The Dissertation Committee for Victoria Christine Cotham Certifies that this is the approved version of the following dissertation:

Mass Spectrometry Combined with Strategic Enzymatic Digestion, Selective Derivatization and Ultraviolet Photodissociation for the Identification and Characterization of Immunoglobulin G Antibodies

Committee:

Jennifer S. Brodbelt, Supervisor

Richard M. Crooks

George Georgiou

Lauren J. Webb

Yan Zhang

Mass Spectrometry Combined with Strategic Enzymatic Digestion, Selective Derivatization and Ultraviolet Photodissociation for the Identification and Characterization of Immunoglobulin G Antibodies

by

Victoria Christine Cotham, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December, 2016

Dedication

To Nicholas, my pillar.

Acknowledgements

What an immensely humbling opportunity it is to thank those who have been a part of this wild and wonderful academic journey. This is by no means an accomplishment based on my efforts alone, and I am deeply indebted to all who have contributed in a myriad of ways. I would first like to acknowledge my parents, Jim and Terry, who instilled the importance of integrity, humility, diligence and a strong work ethic from an early age. These values set the foundation for the scientist I am today, and for that – in addition to their unwavering love and support – I am abundantly grateful. To my three brothers, Matt, Eric and Dillon, thanks for the many years of pestering rooted in love; it helped to develop the thick skin and gumption necessary to make it through a doctorate program. Of course, it is also the support, love, prayers and encouragement of all of my extended family and friends that have made the completion of this work possible.

I would like to thank my graduate research advisor, Dr. Jennifer Brodbelt, for her patience and the freedom to explore my scientific interests. I have developed into a far more independent and creative scientist under her leadership. I am also grateful to Dr. Dan Sykes, my undergraduate advisor, mentor, and the person to whom I attribute my deep love of analytical chemistry. He will forever set the standard as a model leader and educator. Thanks are owed to my collaborators, particularly Andrew Horton, Dr. Yariv Wine and Dr. George Georgiou, for providing many of the antibody samples described in this dissertation, as well as considerable insight regarding the immunological role of these incredible biomolecules. Dr. Justin Sperry from Pfizer and Dr. Sheng Gu, previously at Biogen, are gratefully acknowledged for their generous donations of

monoclonal IgGs that served as excellent models for testing across many stages of method development.

Thanks to those of my fellow Brodbelters, past and present, who willingly shared their expertise, engaged in insightful discussions, and offered constructive criticisms over the years. And to those soldiers who selflessly spent days (and in some cases weeks) troubleshooting LCs, mass spectrometers, and lasers alongside me – please know that you were instrumental (pun intended) to the completion of this work and have my utmost respect and appreciation.

I would be remiss not to also acknowledge my "blue family" and their incredible moral support. I am beyond grateful to have a group of friends who love Nick and I like family and always have our "six". From listening and encouraging in times of great stress and doubt, to referring to me as "The Doctor" or sending pictures of tropical paradises as extra motivation to forge ahead towards graduation – your encouragement, friendships and votes of confidence have been invaluable.

And finally, to my husband, best friend, and biggest champion, Nicholas Cotham, thank you. My life has been enriched by your love, patience, compassion and sacrifice. Thank you for keeping me grounded throughout the compilation of this dissertation and for serving as a constant reminder of how truly blessed I am. I love you now and always.

Mass Spectrometry Combined with Strategic Enzymatic Digestion, Selective Derivatization and Ultraviolet Photodissociation for the Identification and Characterization of Immunoglobulin G Antibodies

Victoria Christine Cotham, Ph.D.

The University of Texas at Austin, 2016

Supervisor: Jennifer S. Brodbelt

Immunoglobulin G (IgG) antibodies represent important analytical targets both for their therapeutic properties and for their critical role in the adaptive immune response. While much of the primary structure is conserved across the IgG class, subtle changes in amino acid sequence and the presence or absence of post-translational modifications can have a profound effect on the function and therapeutic potential of a given antibody. As such, there remains a high demand for versatile analytical tools capable of both identification and complete structural characterization of IgGs. The work presented in this dissertation largely focuses on the development of mass spectrometry-based methods for the improved analysis of antibodies. This was accomplished using strategic enzymatic Brodbeltselectivity for regions of particular diagnostic value or to facilitate comprehensive structural characterization.

A method based on chromophore-mediated 351 nm UVPD was developed as a means to streamline the identification of antibodies in mixtures by enhancing selectively for the third complementarity determining region of the IgG heavy chain (CDR-H3). The hypervariable sequences within this region serve as the primary determinant of antigen binding specificity and thus provide a molecular signature by which to differentiate

unique antibodies. To accomplish this, a highly conserved cysteine residue located in the framework preceding the CDR-H3 region was exploited for selective tagging with an Alexa Fluor 350 (AF350) thiol-selective maleimide. This site-specific tagging combined with strategic enzymatic digestion and 351 nm UVPD allowed selective dissociation of only AF350-labeled peptides for facile discrimination of CDR-H3 sequences within a high-throughput liquid chromatography-tandem mass spectrometry (LC-MS/MS) based workflow.

Two variations of middle-down mass spectrometry based on either restricted Lys-C proteolysis or hinge-selective IdeS digestion combined with 193 nm UVPD were used for the characterization of monoclonal antibodies. Both strategies yielded considerably greater diagnostic sequence information when benchmarked against conventional collision- and electron-based activation methods. The Lys-C proteolysis method was found to have considerable implications for the analysis of serological antibody repertoires owing to its facile implementation into high-throughput proteomic workflows and ability to unambiguously differentiate unique CDR-H3 sequences.

The development and implementation of a front-end dual spray reactor for high-throughput ion/ion-mediated bioconjugation is demonstrated for the enhanced structural characterization of unmodified and post-translationally modified peptide cations by 193 nm UVPD and CID. The ability to generate ion/ion complexes in real-time followed by efficient covalent conversion allowed integration of the dual spray reactor into a high-throughput LC-MS^n workflow for rapid derivatization of peptide mixtures.

Table of Contents

Chapter 1: Introduction 1
1.1 Motivation and Scope of Research 1
1.2 Immunity and the Production of Antigen-Specific Antibodies 2
1.3 Antibody Structure and Function 4
1.4 Therapeutic Antibodies 6
1.5 Mass Spectrometry for Antibody Analysis 9
1.5.1 Bottom-Up MS Analysis 9
1.5.2 Top-Down MS Analysis 11
1.5.3 Middle-Down MS Analysis 11
1.5.4 Polypeptide Fragmentation Nomenclature 12
1.5.5 Collisional Dissociation 14
1.5.6 Ultraviolet Photodissociation (UVPD) 16
1.5.7 Electron Transfer Dissociation (ETD) 19
1.6 Overview of Chapters: 20
1.7 References 24

Chapter 2: Experimental Methods 30
2.1 Overview 30
2.2 Mass Spectrometric Instrumentation 30
2.2.1 Electrospray Ionization (ESI) 30
2.2.2 Dual Source Reactor 32
2.2.3 Linear Ion Trap (LIT) 33
2.2.4 Orbitrap Mass Analyzer 33
2.3 Ion Activation 34
2.3.1 Collisional Activation 34
2.3.2 Ultraviolet Photodissociation (UVPD) 35
2.3.3 Electron Transfer Dissociation (ETD) 36
2.4 Liquid Chromatography (LC) 37
2.4.1 Dionex Ultimate 3000 Nano/Capillary RSLC System 38

2.4.2 Dionex Ultimate 3000 Microbore LC System 39
2.4.3 Eksigent 2D Plus NanoLC 39
2.5 Materials and Reagents 40
2.6 Sample Preparation 40
2.6.1 Single-Chain Antibody Fragment (scAb) Expression and Purification 40
2.6.2 Selective Derivatization and Sample Processing 41
2.6.3 Ion/Ion Reaction-Mediated Peptide Bioconjugation 42
2.6.4 Middle-Down IgG Samples Prepared by Restricted Lys-C Proteolysis 42
2.6.5 Middle-Down IgG Samples Prepared by IdeS Digestion 43
2.7 Automated Tandem Mass Spectrometry Data Interpretation 43
2.7.1 MassMatrix 44
2.7.2 ProSightPC 44
2.7.3 ProSight Lite 45
2.8 References 46

Chapter 3: Selective 351 nm Photodissociation of Cysteine-Containing Peptides for Discrimination of Antigen-Binding Regions of IgG Fragments in Bottom-Up LC-MS/MS Workflows 47
3.1 Overview 47
3.2 Introduction 47
3.3 Experimental 51
3.3.1 Materials and Reagents 51
3.3.2 Single-Chain Antibody Fragment (scAb) Preparation 52
3.3.3 Cysteine Derivatization and Sample Preparation 53
3.3.4 Mass Spectrometry, Liquid Chromatography, and Photodissociation 53
3.3.5 Database Searching 54
3.4 Results and Discussion 55
3.4.1 Selective Modification and 351 nm UVPD of Model Cysteine Peptides 57

3.4.2 Validation of Cysteine-Selective Strategy with Model Protein Digest 63
3.4.3 Selective CDR-H3 Analysis of IgG Fragments 65
3.5 Conclusions 70
3.6 References 72

Chapter 4: Middle-Down 193 nm UVPD for Unambiguous Antibody Identification and its Implications for Immunoproteomic Analysis 74
4.1 Overview 74
4.2 Introduction 75
4.3 Experimental 79
4.3.1 Materials and Reagents 79
4.3.2 Sample Preparation for Middle-Down Analysis 80
4.3.3 Liquid Chromatography and Mass Spectrometry 80
4.3.4 Data Processing 81
4.4 Results and Discussion 82
4.4.1 *In-silico* Proteolysis of a Theoretical Antibody Repertoire 82
4.4.2 UVPD, HCD and ETD Analysis of Lys-C Digested Trastuzumab 86
4.4.3 Middle-Down 193 nm UVPD of Anti-Influenza Antibody Mixture 99
4.5 Conclusions 101
4.6 References 102

Chapter 5: Characterization of Therapeutic Monoclonal Antibodies at the Subunit-Level using Middle-Down 193 nm Ultraviolet Photodissociation 104
5.1 Overview 104
5.2 Introduction 105
5.3 Experimental 109
5.3.1 Materials and Reagents 109
5.3.2 IdeS Digestion and Sample Preparation 109
5.3.3 Liquid Chromatography and Mass Spectrometry 110
5.3.4 Data Processing 111

5.4 Results and Discussion 112
5.4.1 Evaluation of IgG Subunit Separation and Accurate Mass Analysis....112
5.4.2 Optimization of Targeted UVPD for Maximal Sequence Coverage and PTM Localization....118
5.4.3 Benchmarking UVPD against ETD for Subunit Sequence Characterization 127
5.5 Conclusions....132
5.6 References....134

Chapter 6: High-Throughput Bioconjugation for Enhanced 193 nm Photodissociation via Droplet Phase Initiated Ion/Ion Chemistry using a Front-end Dual Spray Reactor 137
6.1 Overview....137
6.2 Introduction....138
6.3 Experimental....140
6.3.1 Chemicals and Materials....140
6.3.2 Implementation of a Front-end Dual Spray Reactor....141
6.3.3 Ion/Ion Chemistry, Mass Spectrometry and Photodissociation 142
6.3.4 High-Throughput Bioconjugation on an LC Timescale 143
6.4 Results and Discussion 143
6.5 Conclusions....158
6.6 References....159

Chapter 7: Modulation of Phosphopeptide Fragmentation via Dual Spray Ion/Ion Reactions using a Sulfonate-Incorporating Reagent 161
7.1 Overview....161
7.2 Introduction....162
7.3 Experimental....165
7.3.1 Chemicals and Materials....165
7.3.2 Solution Phase N-Terminal Derivatization....165
7.3.3 Mass Spectrometry and Front-end Ion/Ion Reactions 165
7.4 Results and Discussion 166

7.4.1 Dual Spray Reactor-Initiated Schiff Base Bioconjugation of Phosphopeptides 168
7.4.2 CID of Unmodified versus FBDSA-Modified Phosphopeptides171
7.4.3 Exploring the Role of the Sulfonate Moiety 176
7.5 Conclusions 183
7.6 References 185
Chapter 8: Conclusions 187
8.1 Summary of Chapters 187
8.2 Future Directions 190
References 192
Vita 209

Chapter 1

Introduction

1.1 Motivation and Scope of Research

Antibodies, whether generated through an immune response or administered therapeutically, function to selectively engage a specific target (antigen) and initiate critical effector mechanisms that confer protective immunity to a host. Monoclonal antibodies (mAbs) have emerged as one of the most effective and rapidly advancing therapeutic modalities used for the treatment of oncogenic, infectious and autoimmune disease. The success of mAb-based biologics arises from their high specificity, favorable pharmacokinetics, and targeted modes of action;[1,2] however, these properties rely critically on antibody structural integrity, both in terms of sequence composition and post-translational modifications (PTMs).[3,4] Despite tremendous inroads in manufacturing and engineering technologies, large scale bioproduction results in heterogeneous molecular compositions that require comprehensive characterization to ensure therapeutic safety and efficacy.[5,6] Alternatively, methods that enable facile differentiation of unique antibodies in highly diverse serum repertoires remain critical for the evolving field of immunoproteomics, which seeks to understand and therapeutically leverage antibody-mediated immunity at the protein level.[7] This objective has be aided by recent developments in next-generation sequencing of immunoglobulin genes (Ig-seq) combined with high-throughput bottom-up mass spectrometry (MS) to identify antibodies elicited during an immune response at functionally relevant concentrations.[8,9] Despite these advances, a number of analytical challenges arising from the intrinsic features of antibodies, such as their high degree of sequence homology, limit the utility of

conventional bottom-up MS approaches for highly complex antibody mixtures and greatly complicate subsequent bioinformatic interpretation.[10]

The work presented in this dissertation focuses on the development of bottom-up and middle-down mass spectrometric methods for the analysis of antibodies using variations of ultraviolet photodissociation (UVPD). Herein, UVPD is advanced as a powerful analytical tool for the comprehensive structural characterization of therapeutic monoclonal antibodies. Furthermore, the utility of combining UVPD and strategic enzymatic digestion is demonstrated for the unambiguous identification of unique antibodies via improved coverage of peptides derived from their diagnostic antigen-binding domains. While not directly applied to highly complex immunoproteomic mixtures at the current stage of development, the latter proof-of-principle studies provide a strong justification for integrating UVPD into future MS based antibody repertoire analyses.

1.2 Immunity and the Production of Antigen-Specific Antibodies

The mammalian immune system employs two primary defense mechanisms, the innate and adaptive response, that act in concert to detect, neutralize and clear foreign pathogens.[11] The innate system serves as the first line of defense and mounts a rapid, but non-specific response that broadly recognizes a conserved set of Pathogen-Associated Molecular Patterns (PAMPs) that are distinct from those expressed by host cells.[12,13] These molecules include bacterial lipopolysaccharides (LPS) and lipopeptides, as well as viral strands of RNA and DNA.[12,13] Once cells of the innate system recognize pathogen invasion, signaling cascades occur that stimulate pro-inflammatory responses, cell-death, and antigen presentation.[14] The latter is responsible for initiating the highly specific adaptive immune response.

The adaptive immune system is further divided into a T cell-mediated cellular component and a B cell-mediated humoral component.[11,15] T cells and B cells each possesses specific receptors (TCRs and BCRs, respectively) that are unique to each clonal population and confer antigen specificity as a result of genetic rearrangements of receptor genes that occur during individual cell development.[15] In B cells, these receptors are membrane-bound immunoglobulins (Ig) that serve as the precursors to secreted antibodies.[16]

During B cell development, Ig variable (V), diversity (D) and joining (J) germline gene segments undergo somatic V(D)J recombination to form the BRCs expressed on naïve, or pre-antigen stimulated B cells.[8,15] Following antigen exposure, naive B cells displaying antigen-recognizing BCRs become activated and undergo clonal expansion and somatic hypermutation within their variable Ig domains to fine-tune antigen specificity in a process referred to as affinity maturation.[8,15] Mature antigen-specific B cells subsequently undergo differentiation into long-lived memory B cells that mount a rapid protective response upon re-exposure to the cognate antigen,[17,18] or into antibody-secreting plasma cells.[8,19] The resulting antigen-specific antibodies released into plasma and extracellular fluids confer protection against cognate antigen through three effector mechanisms: neutralization, opsonization, and complement activation.[20] Neutralization occurs through direct antibody-binding to epitopes of the pathogen to effectively block access to host cells.[20] Alternatively, during opsonization antibodies recruit effector cells to the bound pathogen to initiate functions including antibody-dependent cell-mediated cytotoxicity (ADCC) and antibody-dependent cell-mediated phagocytosis (ADCP).[21,22] Finally, complement activation occurs when antigen-bound antibodies interact with complement protein C1, which in turn activates the complement cascade that ultimately results in the pathogen surface being coated with complement defense proteins that mark

it for destruction via phagocytosis in a process referred to as complement-dependent cytotoxicity (CDC).[23,24]

1.3 ANTIBODY STRUCTURE AND FUNCTION

Secreted immunoglobulins, or antibodies, are composed of two identical heavy chain and two identical light chain polypeptides that form a characteristic "Y" shaped protein structure that is held together by a series of intra- and intermolecular disulfide bonds.[25] Each component chain is further divided into N-terminal variable (V_H and V_L) regions that are responsible for antigen recognition and C-terminal constant (C_H and C_L) regions that interact with effectors cells. Human antibodies are divided into five classes, or isotypes: IgA, IgD, IgE, IgG and IgM, which are distinguished by the structure of their constant regions (α, δ, ε, γ, and μ, respectively) and ultimately their functional role within the immune response.[25,20] Among these, IgG is the predominant isotype produced during B cell activation and constitutes approximately 75% of total serum antibodies in circulation.[25] Moreover, IgG is further divided into four subclasses, IgG1, IgG2, IgG3, and IgG4, that vary slightly in amino acid composition (although they remain >95% homologous), as well as inter-chain disulfide number and connectivity. These structural differences ultimately determine which effector mechanisms are activated, as well as the stability of each IgG subclass. Consequently, the relative abundance, high antigen-specificity, and critical role in nearly all effector mechanisms employed during adaptive immunity make IgG the critical target for mass spectrometry-based analysis of antigen-specific antibody repertoires,[26–29] and currently the only antibody isotype used therapeutically. As such, the remainder of the discussion presented herein will focus on IgG.

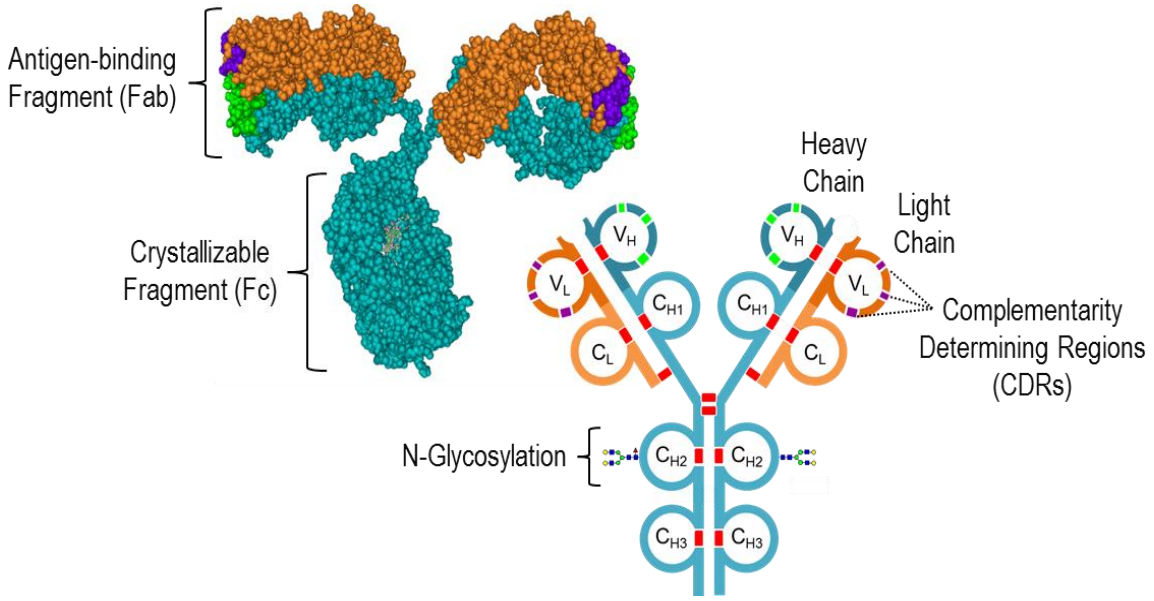


Figure 1.1 The immunoglobulin G (IgG) structure is composed of two identical light chains (orange) and two identical heavy chains (blue) that are further divided into variable (V) and constant (C) regions. The variable regions of each chain contain hypervariable complementarity determining regions (shown in green and purple for the heavy and light chains, respectively). The molecule possesses two functional domains: the antigen-binding fragment (Fab) and the glycosylated crystallizable fragment (Fc).

As demonstrated in **Figure 1.1**, the heavy chains of IgG contain one variable region (V_H) and three constant regions (C_H, C_{H2} and C_{H3}), whereas the light chains contain one variable (V_L) and one constant (C_L) region. The IgG monomer is divided into functional units consisting of two identical fragment antigen-binding (Fab) domains and a single glycosylated fragment crystallizable (Fc) domain, which are responsible for antigen binding and initiation of effector mechanisms, respectively. The variable regions of the Fab contain V, D, and J germline gene segments, which undergo V(D)J recombination in addition to somatic hypermutation,[30] as previously discussed within the context of B cell development. Importantly, the human genome encodes for many V, D,

and J gene segments that can recombined, in theory, to form an incredible number (>10^{13}) of clonally unique antibodies.[20,31] The light chain CDR-L3 is formed in a similar manner, with the exception that only V and J gene segments undergo recombination, thereby significantly limiting its diversity relative to CDR-H3.[30] Conversely, the CDR1 and CDR2 loops of each chain are encoded within the V_H and V_L gene segments and therefore exhibit lower sequence diversity compared to CDRs formed through genetic recombination. However, somatic hypermutation within these regions affords a mechanism by which antigen specificity is fine-tuned through various sequence mutations.[20]

Whereas the variable Fab domain facilitates antigen recognition, the glycosylated Fc domain, composed of the C_{H2} and C_{H3} constant regions, provides a direct link between the adaptive and innate immune systems, as well as the humoral and cellular components of the adaptive immune response.[22] This is mediated by interactions with either Fc gamma receptors (FcγR) expressed on the surface of innate immune effector cells that initiate critical pathogen clearance mechanisms, or with complement protein C1q to initiate the complement pathway. These binding interactions are critically dependent on both the presence and composition of a highly conserved N-linked glycosylation site at asparagine-297 (Asn297) within the C_{H2} constant region.[22]

1.4 THERAPEUTIC ANTIBODIES

Over the past three decades monoclonal antibodies (mAb) have emerged as one of the fastest growing classes of therapeutic modalities within the biopharmaceutical industry. Currently, there are approximately 50 mAb-based products approved in the United State and European Union (EU), most of which possess therapeutic indications for

oncogenic, autoimmune and infectious disease treatment;[32] with many other candidates currently in the clinical development stages.[33] In 2013 alone, the world-wide sales of mAb-based products generated approximately $75 billion in revenue. At the current approval rate these figures are expected to exceed $125 billion in sales by the year 2020.[33,34] The success of mAb-based therapeutics is due to their high target specificity, favorable pharmacokinetic properties and ability to elicit or modulate a desired immune response. This has largely been driven by advances in antibody production and engineering technologies, which has led to more potent therapies with fewer immunogenic side effects.

In theory, antibodies produced from a single B cell clone should possess homogenous structural compositions; however, therapeutic mAbs produced from recombinant DNA technologies are generally complex, heterogeneous and subject to numerous enzymatic and chemical modification during expression, purification and storage that can profoundly influence therapeutic safety and efficacy. Sequence variants are commonly detected in recombinant proteins produced in cell culture, which can arise from mutations at the DNA level, amino acid misincorporation during protein assembly, or miscleavage during post-translational processing.[35,36] Importantly, point mutations within the CDRs have been shown to significantly modulate antigen binding affinity and specificity.[37,38] Other common modifications that can occur during production and storage include variations in the Fc glycan structure, C-terminal processing of the heavy chain, pyroglutamic acid formation, oxidation, and deamidation.[39] These modifications can alter pharmacokinetic properties, reduce therapeutic potency and stimulated deleterious immunogenic responses to varying degrees.[5,6] The potential for adverse effects arising from heterogeneities introduced throughout all stages of production highlight the need for sensitive analytical tools capable of comprehensive mAb structural characterization.

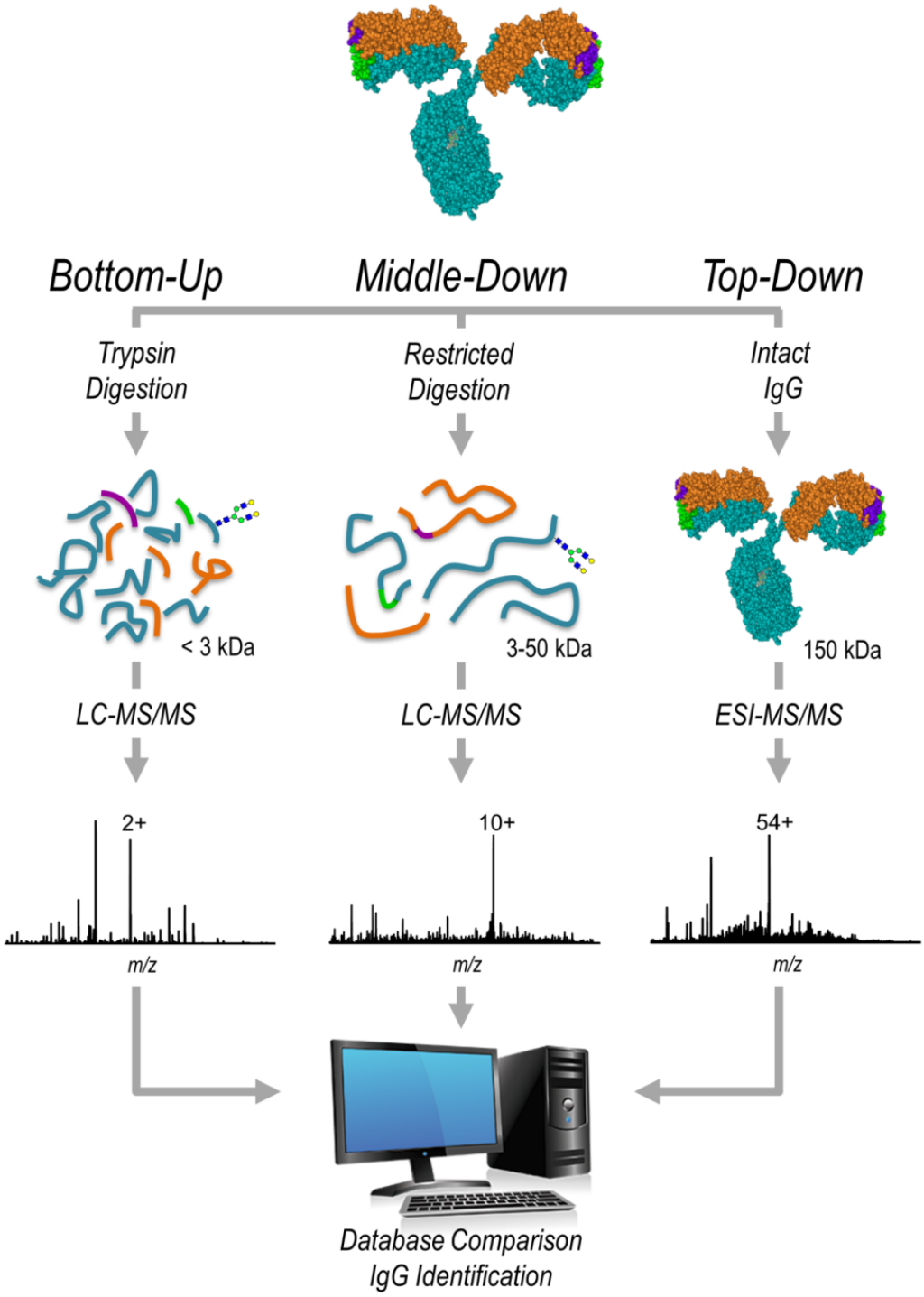


Figure 1.2 Typical mass spectrometry based workflows for IgG analysis.

1.5 MASS SPECTROMETRY FOR ANTIBODY ANALYSIS

Mass spectrometry (MS) has emerged as an essential analytical tool for the structural characterization of therapeutic monoclonal antibodies,[40–42] and more recently for the molecular level deconvolution of antigen-specific antibody repertoires expressed in serum.[9,10,27,26] Advances in MS instrumentation and ion activation methods facilitate detailed and sensitive analysis of antibody amino acid sequence, post-translational modifications, and higher-order structure. As demonstrated in **Figure 1.2**, MS-based strategies for antibody analysis fall into three general categories: bottom-up, middle-down, and top-down.

1.5.1 Bottom-Up MS Analysis

Conventional bottom-up mass spectrometry, which relies on the identification and characterization of antibodies via the analysis of their peptide surrogates, remains the method of choice for the structural characterization of therapeutic mAbs[40,43] and is to date the only strategy employed for MS-based serum antibody proteomics.[10,26,27,29] The utility of the bottom-up workflow largely arises from the fact that peptides are more easily separated, ionized and fragmented using established liquid chromatography-tandem mass spectrometry (LC-MS/MS) based techniques than their intact counterparts.

In the bottom-up approach, antibodies are subjected to enzymatic digestion with trypsin, or a combination of enzymes, to produce small (≤2.5 kDa) proteolytic peptides. The resulting peptide mixture is separated using reverse phase liquid chromatography and the eluting peptides are then analyzed by tandem mass spectrometry (MS/MS) to obtain peptide mass and fragmentation information. The resulting MS/MS data is then used to identify the resulting peptides, as well as characterize any modifications that may be present. This typically involves the use of *in silico* database search algorithms that

compare experimentally obtained MS/MS data with theoretical spectra generated from a candidate sequence database. This approach has proven highly effective for therapeutic IgGs of known primary sequence, but poses a significant technical hurdle for complex mixtures of antibodies derived from serum, for which no *a priori* sequence database is available. This arises from the fact that antibody genes are not encoded in the germline, but are assembled through somatic recombination and hypermutation. Several groups have recently addressed this limitation through the use of high-throughput next-generation DNA sequencing of B cell immunoglobulin variable domains (V genes), or Ig-seq, to generate a sample-specific antibody sequence database for the interpretation of antibody-derived MS/MS data.[26–29] While this paired Ig-seq/bottom-up MS approach has made great strides in terms of handling the complexity of serum antibody repertoires, the high degree of homology intrinsic to the antibody sequence poses a formidable challenge for unambiguous Ig identification using standard decoy-based error modeling employed by most *in silico* algorithms.[10] *De novo* sequencing has also been explored as an alternative approach for mass spectral interpretation of serum-derived antibody mixtures, which overcomes the need for a reference database;[44–46] however, continued development is necessary before these methods can be routinely employed for complex and highly homologous repertoire samples.

While bottom-up methods remain popular, they suffer from several key limitations, particularly with regard to comprehensive structural characterization of mAbs. Sample preparations required for bottom-up analysis tend to be time-consuming and labor-intensive. Moreover, extensive sample handling and proteolysis often introduces artifactual heterogeneities that can obscure intrinsic or manufacturing-related modifications.[40,47] Lastly, incomplete peptide sampling is a common problem owing to the complexity of peptide mixtures generated by proteolytic digestion.[40,47]

1.5.2 Top-Down MS Analysis

The top-down approach omits proteolysis (or chemical reduction) prior to analysis. Structural characterization is therefore accomplished based on accurate mass measurements and MS/MS fragmentation of the intact antibody. This approach is particularly desirable due to the lack of sample preparation required, which maintains high structural integrity, as well as its ability to yield immediate feedback on sequence fidelity and proteoform abundance.[48,49] Despite considerable advances in both ion activation methodologies and high performance MS instrumentation that have increased the scope of top-down strategies in recent years,[49–51] these methods remain non-trivial for the characterization of intact antibodies owing to their size (~150 kDa), structural complexity, and high degree of sequence homology. To date, electron-transfer dissociation (ETD) and electron-capture dissociation (ECD) performed on Orbitrap and FTICR mass spectrometers have shown the greatest promise for top-down sequencing of intact antibodies; however, these methods remain limited to the interrogation of approximately 30-35% of the antibody structure, due to the number of disulfide protected regions within each antibody domain.[52–54] These important structural features currently preclude top-down analysis as a standalone approach for antibody characterization.

1.5.3 Middle-Down MS Analysis

Middle-down mass spectrometry has emerged as a promising intermediate between top-down and bottom-up strategies. This approach typically involves restricted proteolysis with enzymes that target less frequent amino acids to generate peptides within a mass range of 3–20 kDa.[55–59] The larger size of the resulting peptide mixture results in reduced complexity and thus improved MS/MS sampling efficiency. Moreover, restricted digestion generates peptides with more unique sequence character, which can be

exploited for improved differentiation of highly homologous antibody mixtures, or to increase the probability of localizing PTMs or point mutations.[56]

A common variation of the middle-down approach for the analysis of monoclonal antibodies involves the reduction of intermolecular disulfides to produce free heavy chains (~50 kDa) and free light chains (~25 kDa). This strategy can be combined with hinge-selective digestion using proteases such as papain or immunoglobulin G-degrading enzyme from *Streptococcus pyogenes* (IdeS) to generate three distinct subunits (~25 kDa) consisting of the free light chain (Lc), the variable domain fragment of the heavy chain (Fd), and the reduced heavy chain Fc domain (Fc/2).[40,55,60–64] While better adapted to LC-MS/MS analysis compared to intact antibodies, comprehensive characterization of antibodies subunits remains challenging due to inherent limitations in speed and sensitivity of high resolution measurements required for large (≥ 25 kDa), highly charged species within narrow chromatographic elution windows.[65–67]

1.5.4 Polypeptide Fragmentation Nomenclature

The success of mass spectrometric methods for IgG characterization and the differentiation of unique clonotypic antibodies in complex immune mixtures rely on the ability to generate and interpret sequence-specific information from the heavy and light polypeptide chains. Owing to the high degree of sequence homology intrinsic to the IgG scaffold, differentiation of unique antibodies often requires complete sequencing of their hypervariable CDR sequences, of which the CDR-H3 exhibits the greatest sequence diversity and is thus the most diagnostic. As with any tandem mass spectrometry (MS/MS) based protein or peptide sequencing application, this requires controlled fragmentation of the polypeptide backbone. To accomplish this, ion activation methods have been developed that restrict fragmentation to specific bonds along the polypeptide

backbone,[68] for which a systematic nomenclature has been developed to categorize the resulting fragment ions.[69,70] A graphical illustration of this nomenclature and representative product ion formation is shown in **Figure 1.3**.

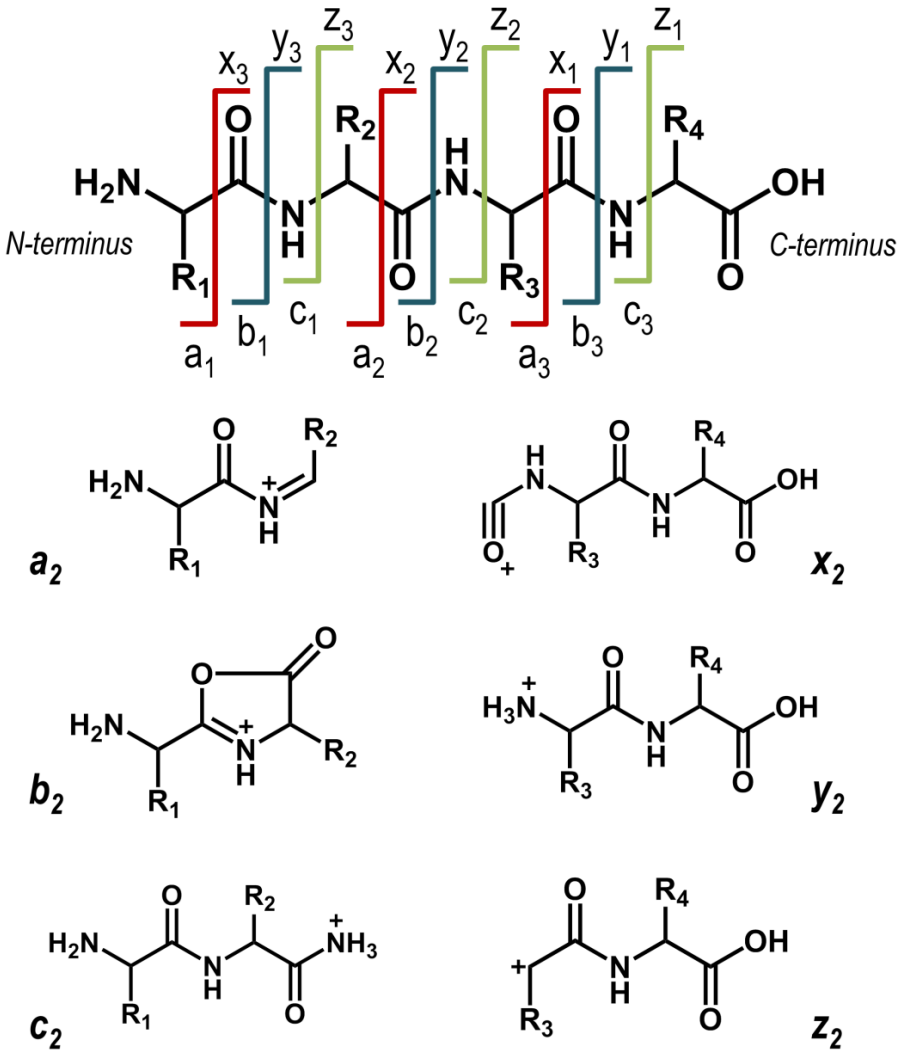


Figure 1.3 Polypeptide fragmentation nomenclature as proposed by Roepstorff *et al.*

Each product ion is defined by the type and location of the bond cleavage across the polypeptide backbone, as well as the terminus (N or C) that is retained. Product ions originating from the N-terminal end of the polypeptide are as classified as *a*, *b* and *c*-type ions, whereas those retaining the C-terminal end are *x*, *y* and *z*-type ions. The two ions produced by cleavage of the same bond are referred to as complementary pairs and include *a/x*, *b/y*, and *c/z* ions. Slow-heating or threshold based activation methods, such as collisional induced dissociation (CID) and some types of photodissociation, result in predominant cleavage at the C–N amide bonds of the polypeptide backbone producing *b*- and *y*-type ions. In contrast, *c*- and *z*-type ions generated from cleavage of the N–C_α bond are characteristic of radical-directed mechanisms. Higher energy activation, such as 157 nm and 193 nm ultraviolet photodissociation (UVPD), also produce *a*- and *x*-type ions arising from C_α–C bond cleavage in addition to *b/y* and *c/z* ions.

1.5.5 Collisional Dissociation

Collision induced dissociation (CID) is the most established and widely utilized ion activation method owing to its robust performance and implementation on virtually all commercial mass spectrometers. CID, as employed in ion trap instruments, is accomplished through resonant excitation using a supplemental AC waveform at the secular frequency of the precursor ion.[71,72] This excitation frequency accelerates the precursor to higher kinetic energy allowing multiple inelastic collisions with inert bath gas molecules present in the trap. Each collision causes conversion of kinetic energy into internal vibrational energy until the dissociation threshold of the precursor is exceeded and fragmentation occurs.[71,72]

The stepwise nature of the activation process restricts CID fragmentation to the lowest energy pathways, which for protonated peptides and proteins is generally observed

as cleavage of the C–N amide bond to produce *b* and *y*-type product ions. However, a number of well-defined variables, including the amino acid composition, charge state and the presence of post-translational modifications can dramatically influence the dissociation behavior of the precursor ion. The mobile proton model, derived from extensive mechanistic studies, provides a qualitative framework by which collision-induced fragmentation can be rationalized.[73,74] This theory posits that ionizing protons are initially localized at the basic sites of the peptide, including the N-terminus or side-chains of arginine, lysine, and histidine. Upon activation, protons become 'mobilized' and can migrate to less-basic sites along the backbone to initiate charge-directed fragmentation. Under sufficiently mobile conditions, N-protonation of the amide bond facilitates sequence-informative cleavage. This pathway is modulated in the presence of proline due to the higher proton affinity of its amide bond, resulting in highly favorable N-terminal to the proline cleavage.[73–76] Alternatively, under proton deficient conditions, fragmentation is governed by charge-remote pathways that lead to preferential cleavage C-terminal to aspartic and glutamic acid residues.[77–79] Collectively, preferential sequence-specific cleavages restrict the utility of CID for characterization peptides and proteins containing these residues. Another shortcoming of CID arises from the slow heating mechanism that governs the activation process, which promotes preferential cleavage of labile PTMs, such as phosphorylation. Chapter 7 of this dissertation directly addresses the latter shortcoming of CID for phosphopeptide analysis using a novel gas-phase bioconjugation technique that stabilizes the labile phosphate group during collisional activation to promote greater sequence informative fragmentation and phosphosite localization.

On the basis of its aforementioned merits, including efficient and predictable fragmentation of proteolytic peptides, CID remains the most commonly employed ion activation method used for the analysis of antibodies in bottom-up workflows. While CID

maintains stable modifications, such as oxidation, deamidation, and N-terminal pyroglutamate, a severe limitation arises from its preferential cleavage of labile modifications that are critical for antibody characterization, such as glycosylation.[80] Of the PTMs commonly found in IgG, the highly conserved N-glycosylation at Asn-297 is generally considered among the most important owing to its critical role in antibody effector functions, stability, and immunogenicity. CID spectra arising from peptides containing this important glycosylation site are often dominated by labile glycosidic bond cleavages that preclude peptide identification.[40,42] To overcome this limitation, CID is often paired with complementary fragmentation techniques, such as electron-based activation, that preserve labile modifications and preferentially cleave the peptide backbone.[81]

1.5.6 Ultraviolet Photodissociation (UVPD)

A laser was first coupled to a mass spectrometer for the purposes of photodissociation of gas phase ions over four decades ago,[82,83] and has since evolved into a powerful and versatile tool for the structural characterization of biological molecules, most notably peptides and proteins.[68,84,85] The photoactivation process relies on energy accumulation and subsequent dissociation through photon absorption, which can occur either non-specifically or in a highly selective chromophore-mediated manner.[68,84,85] The type of UVPD experiment (i.e., non-specific versus chromophore-mediated) is largely dictated by the wavelength selection and the photoabsorption properties of the gas phase ions being interrogated. The most common wavelengths used for UVPD correspond to those generated from pulsed excimer and Nd:YAG lasers, which include 157 nm, 193 nm, 266 nm, 351 nm, and 355 nm (7.9 eV, 6.4 eV, 4.7 eV, and 3.5 eV per photon, respectively).[68,84,85]

UVPD using 157 nm and 193 nm photons has proven highly effective for the comprehensive structural characterization of peptides[86–90] and proteins[91–97] owing to the strong absorption of the polyamide backbone at these wavelengths.[98] The high energy deposition of photons at 157 nm and 193 nm (7.9 eV and 6.4 eV, respectively) promotes electronic excitation, allowing access to new and diverse dissociation pathways that lead to the formation of *a*, *b*, *c*, *x*, *y* and *z* ions.[99–102] Additionally, radical *a* and *x*-ions (*a*• and *x*•) are observed as a result of homolytic C_α–C bond cleavage, which can undergo hydrogen elimination to form *a* and *x* ions or partial loss of amino acid side chains through secondary dissociation to form *v*, *d*, and *w* ions (**Figure 1.4**), which have proven to be useful for the differentiation of isobaric leucine and isoleucine residues.[99] In addition to extensive fragmentation of the polypeptide backbone, UVPD has also been shown to maintain labile modifications, making it well-suited for the identification and localization of biologically relevant post-translational modifications.[89,90,103] These performance attributes prompted our efforts to pursue 193 nm UVPD for the detailed characterization of antibodies as described in Chapters 4 and 5.

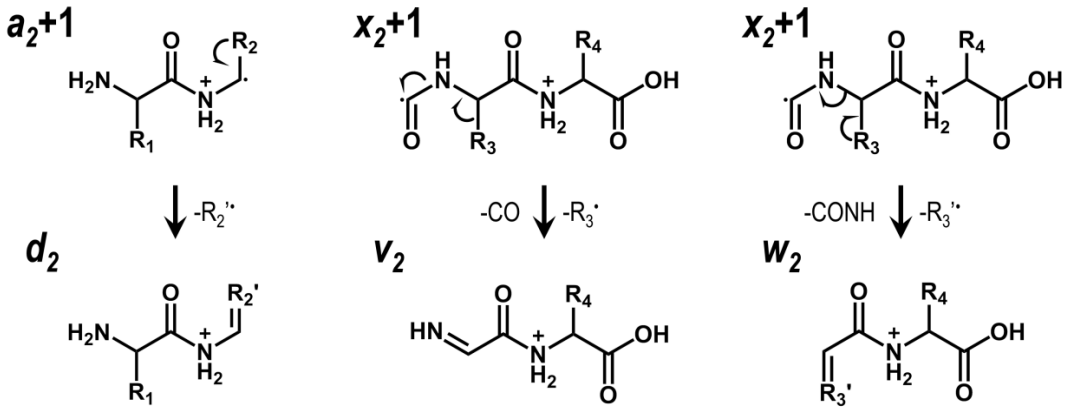


Figure 1.4 Generation of side-chain losses from radical precursor ions (adapted from Reference 98). R' indicates partial loss of side-chain.

As an alternative to the comprehensive and indiscriminant fragmentation observed with 157 nm and 193 nm photodissociation, chromophore-mediated UVPD at wavelengths near 350 nm provide a unique strategy for introducing selectivity into proteomic workflows. In principle, the native polypeptide backbone does not absorb near 350 nm, therefore incorporation of chromogenic moieties that confer photoabsorptivity within this wavelength regime are critical for UVPD. This affords an effective means by which to discriminate chromophore-labeled versus unlabeled peptides within complex mixtures. Several recent applications have demonstrated the utility of site-selective chromophore labeling combined with 351 nm UVPD (XeF excimer laser, 3.5 eV per pulse) for streamlined proteomic analysis. For example, Aponte *et al*. reported a method for targeting tyrosine (Tyr) and histidine (His) containing peptides via a diazonium labeling reaction. Following site-selective derivatization only peptides containing modified Tyr/His residues generated diagnostic fragmentation patterns when analyzed by LC-MS/UVPD. These highly selective MS/MS datasets combined with the low frequency of Tyr and His residues throughout the human proteome vastly diminished the redundancy of *in silico* database searches. Selective chromophore-tagging coupled with 351 nm UVPD has also been exploited for monitoring protein conformational changes based on the relative accessibility of lysine side-chains to an amine-reactive chromogenic probe. Selective dissociation provided a facile means by which to track changes in modified, or solvent exposed residues across ligand-bound and unbound states.[104] Recently, chromophore-mediated UVPD was demonstrated as a powerful tool for high-throughput *de novo* sequencing via spectral simplification. To accomplish this, proteolytic peptides were tagged at their N-termini with a 351 nm active AMCA chromophore and subsequently irradiated with successive laser pulses to eliminate the N-terminal chromophore-containing ions. The resulting MS/MS spectra contained a clean

series of *y* ions for which a novel software platform, UVnovo, was developed to interpret.[105,106] In this dissertation, the development of a selective 351 nm UVPD approach for the discrimination of antigen-binding regions from IgG fragments is described.

1.5.7 Electron Transfer Dissociation (ETD)

Electron-based dissociation methods, such as electron capture dissociation (ECD)[107] and electron transfer dissociation (ETD),[108] have gained traction as powerful alternatives to conventional collisional activation due to their ability to generate more extensive and randomized cleavage of the polypeptide backbone, while also maintaining labile PTMs.[48,109–111] Electron capture dissociation entails the irradiation of multiply charged peptide or protein cations with low-energy electrons, leading to exothermic electron capture and non-ergodic fragmentation of the polypeptide backbone.[112,113] ECD is typically restricted to Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometers, which facilitate simultaneous trapping of electrons and analyte cations within their magnetic fields. Electron transfer dissociation was developed as an ion/ion analogue of ECD that has greatly extended the utility of electron-based activation owning to its facile implementation on ion trap instruments and various hybrid MS platforms.[114] This process involves the transfer of an electron to a multiply charged polypeptide cation via reaction with a radical anion reagent (i.e., fluoranthene, anthracene, azulene). Exothermic electron transfer subsequently initiates backbone cleavage at N-C_α bonds through hydrogen radical migration.[108,113,115] A primary advantage of ECD and ETD is the formation of odd-electron radical species that undergo non-ergodic fragmentation prior to vibrational energy redistribution. This dissociation mechanism overcomes the challenges associated with preferential cleavages and loss of labile PTMs as observed in collisional

activation.[109,116] However, both ECD and ETD exhibit a strong dependency on precursor charge density, which limits their effectiveness as precursor mass-to-charge values increase. Under instances of low charge density, the phenomenon of non-dissociative electron capture/transfer (ECnoD/ETnoD) has been widely reported, in which resulting product ions are held together by non-covalent interactions that impede sequence-informative fragmentation.[117–119]

Recently, electron-based activation methods have gained traction for the middle-down and top-down structural characterization of antibodies.[52–54,59,60,120] Comparable sequence coverages approaching 35% of intact IgG1 have been reported for ETD on an Orbitrap[53] and ECD on a 9.4 T FTICR instrument.[54] While insufficient for complete characterization, these methods far exceed the ~10% total sequence coverage of reduced IgG2 chains obtained with CID on an Orbitrap.[121] Despite the improved fragmentation efficiency of electron-based methods at the intact protein level, the considerable number of disulfide protected regions throughout the antibody structure currently limits their utility for top-down analysis. To overcome this shortcoming, an alternative middle-down approach was recently introduced that relies on the hinge-selective IdeS (immunoglobulin G-degrading enzyme of *Streptococcus pyogenes*) digestion and chemical reduction of disulfide bonds to generate ~25 kDa IgG subunits. Targeted ETD analysis of the resulting subunits resulted in up to 68% sequence coverage when transients were averaged from multiple LC-MS/MS runs collected with varied ETD reaction times.[60]

1.6 OVERVIEW OF CHAPTERS:

The critical role of antibodies, both as therapeutic modalities and as the primary constituents of the humoral immune response, has prompted the development of

analytical methods capable of their detailed characterization. While mass spectrometry-based approaches have led to tremendous progress for both the comprehensive structural characterization of therapeutic mAbs and the differentiation of unique antibodies in complex serological mixtures, there remain opportunities for continued development. The research presented in this dissertation is aimed at improving the versatility and tunability of MS/MS methods for antibody identification and characterization by combining strategic sample preparation with variations of ultraviolet photodissociation.

In Chapter 3, a high-throughput chromophore-mediated 351 nm UVPD method is described to selectively interrogate cysteine-containing peptides as a means to streamline the identification of unique IgGs in homologous antibody mixtures. To accomplish this, a highly conserved cysteine residue located in the framework preceding the CDR-H3 region is exploited for selective derivatization with an Alexa Fluor 350 (AF350) thiol-selective maleimide. When combined with strategic enzymatic digestion and 351 nm UVPD, selective dissociation of AF350-labeled peptides is accomplished for facile discrimination of CDR-H3 sequences.

A middle-down strategy for the improved characterization and differentiation of unique monoclonal antibodies using restricted Lys-C proteolysis and 193 nm UVPD is presented in Chapter 4. The selection of enzymatic digestion using Lys-C is based on its ability to generate peptides spanning the entire length of the diagnostic CDR-H3 region, as demonstrated both experimentally and using *in silico* methods. The merits of this strategy were assessed in the context of a middle-down proteomics experiment combined with next-generation V-gene database searching for the analysis of a simple mixture of anti-influenza monoclonal antibodies.

In Chapter 5, a targeted middle-down UVPD method is demonstrated for the detailed primary sequence analysis and post-translational site localization of therapeutic

monoclonal antibody subunits generated by hinge-selective enzymatic digestion with recombinant immunoglobulin G-degrading enzyme of *Streptococcus pyogenes* (IdeS) followed by chemical reduction. Under optimized conditions, approximately 60% of the IgG sequence is interrogated, in addition to unambiguous glycosylation site localization and extensive coverage of the antigen-binding domains within a single targeted LC-MS/MS experiment. UVPD exhibits improved performance metrics when benchmarked against more conventional ETD activation; however, both methods yield complementary information that is combined to facilitate greater overall subunit characterization.

The development and implementation of a front-end dual spray reactor for high-throughput ion/ion-mediated gas-phase bioconjugation of peptide cations is described in Chapter 6. To demonstrate the utility of the dual spray reactor, peptide cations are subjected to front-end ion/ion reactions with chromogenic 4-formyl-1,3-benezenedisulfonic acid (FBDSA) anions. Subsequent collisional activation of the resulting ion/ion intermediate promotes covalent Schiff base reactions between primary amine sites in the peptide cation and the aldehyde moiety of the FBDSA anion. Resulting Schiff base modified peptides exhibit enhanced 193 nm UVPD efficiencies relative to their unmodified counterparts and yield greater primary sequence information when compared to conventional CID. Moreover, due to the efficiency of the bioconjugation process and its ease-of-integration with liquid chromatography platforms, this strategy is implemented into a LC-MS^n workflow for the rapid derivatization of peptide mixtures.

Chapter 7 extends the dual source methodology presented in Chapter 6 to enhance the collision-induced dissociation of phosphopeptides via rapid Schiff base derivatization with FBDSA. This strategy exploits the strong electrostatic interactions between sulfonate moieties of FBDSA and basic sites within the peptide to facilitate gas-phase

bioconjugation, reduce charge sequestration, and increase the yield of phosphate-retaining sequence ions upon CID.

1.7 REFERENCES

(1) Weiner, L. M.; Surana, R.; Wang, S. *Nat. Rev. Immunol.* **2010**, *10* (5), 317–327.
(2) Chan, A. C.; Carter, P. J. *Nat. Rev. Immunol.* **2010**, *10* (5), 301–316.
(3) Wang, W.; Singh, S.; Zeng, D. L.; King, K.; Nema, S. *J. Pharm. Sci.* **2007**, *96* (1), 1–26.
(4) Walsh, G.; Jefferis, R. *Nat. Biotechnol.* **2006**, *24* (10), 1241–1252.
(5) De Groot, A. S.; Scott, D. W. *Trends Immunol.* **2007**, *28* (11), 482–490.
(6) Harding, F. A.; Stickler, M. M.; Razo, J.; DuBridge, R. B. *mAbs* **2010**, *2* (3), 256–265.
(7) Tjalsma, H.; Schaeps, R. M. J.; Swinkels, D. W. *PROTEOMICS – Clin. Appl.* **2008**, *2* (2), 167–180.
(8) Georgiou, G.; Ippolito, G. C.; Beausang, J.; Busse, C. E.; Wardemann, H.; Quake, S. R. *Nat. Biotechnol.* **2014**, *32* (2), 158–168.
(9) Lavinder, J. J.; Horton, A. P.; Georgiou, G.; Ippolito, G. C. *Curr. Opin. Chem. Biol.* **2015**, *24*, 112–120.
(10) Boutz, D. R.; Horton, A. P.; Wine, Y.; Lavinder, J. J.; Georgiou, G.; Marcotte, E. M. *Anal. Chem.* **2014**, *86* (10), 4758–4766.
(11) Parkin, J.; Cohen, B. *The Lancet* **2001**, *357* (9270), 1777–1789.
(12) Akira, S.; Uematsu, S.; Takeuchi, O. *Cell* **2006**, *124* (4), 783–801.
(13) Mogensen, T. H. *Clin. Microbiol. Rev.* **2009**, *22* (2), 240–273.
(14) Paul, W. E. *Cell* **2011**, *147* (6), 1212–1215.
(15) Nemazee, D. *Nat. Rev. Immunol.* **2006**, *6* (10), 728–740.
(16) Jung, D.; Giallourakis, C.; Mostoslavsky, R.; Alt, F. W. *Annu. Rev. Immunol.* **2006**, *24* (1), 541–570.
(17) McHeyzer-Williams, M.; Okitsu, S.; Wang, N.; McHeyzer-Williams, L. *Nat. Rev. Immunol.* **2012**, *12* (1), 24–34.
(18) Kurosaki, T.; Kometani, K.; Ise, W. *Nat. Rev. Immunol.* **2015**, *15* (3), 149–159.
(19) Batista, F. D.; Harwood, N. E. *Nat. Rev. Immunol.* **2009**, *9* (1), 15–27.
(20) Murphy, K. *Janeway's Immunobiology, Eighth Edition*; Garland Science, 2011.
(21) Nimmerjahn, F.; Ravetch, J. V. *Nat. Rev. Immunol.* **2008**, *8* (1), 34–47.
(22) Borrok, M. J.; Jung, S. T.; Kang, T. H.; Monzingo, A. F.; Georgiou, G. *ACS Chem. Biol.* **2012**, *7* (9), 1596–1602.
(23) Ricklin, D.; Hajishengallis, G.; Yang, K.; Lambris, J. D. *Nat. Immunol.* **2010**, *11* (9), 785–797.
(24) Diebolder, C. A.; Beurskens, F. J.; Jong, R. N. de; Koning, R. I.; Strumane, K.; Lindorfer, M. A.; Voorhorst, M.; Ugurlar, D.; Rosati, S.; Heck, A. J. R.; Winkel, J. G. J. van de; Wilson, I. A.; Koster, A. J.; Taylor, R. P.; Saphire, E. O.; Burton, D. R.; Schuurman, J.; Gros, P.; Parren, P. W. H. I. *Science* **2014**, *343* (6176), 1260–1263.
(25) Schroeder, H. W.; Cavacini, L. *J. Allergy Clin. Immunol.* **2010**, *125* (2), S41–S52.

(26) Wine, Y.; Boutz, D. R.; Lavinder, J. J.; Miklos, A. E.; Hughes, R. A.; Hoi, K. H.; Jung, S. T.; Horton, A. P.; Murrin, E. M.; Ellington, A. D.; Marcotte, E. M.; Georgiou, G. *Proc. Natl. Acad. Sci.* **2013**, *110* (8), 2993–2998.
(27) Cheung, W. C.; Beausoleil, S. A.; Zhang, X.; Sato, S.; Schieferl, S. M.; Wieler, J. S.; Beaudet, J. G.; Ramenani, R. K.; Popova, L.; Comb, M. J.; Rush, J.; Polakiewicz, R. D. *Nat Biotech* **2012**, *30* (5), 447–452.
(28) Sato, S.; Beausoleil, S. A.; Popova, L.; Beaudet, J. G.; Ramenani, R. K.; Zhang, X.; Wieler, J. S.; Schieferl, S. M.; Cheung, W. C.; Polakiewicz, R. D. *Nat. Biotechnol.* **2012**, *30* (11), 1039–1043.
(29) Lavinder, J. J.; Wine, Y.; Giesecke, C.; Ippolito, G. C.; Horton, A. P.; Lungu, O. I.; Hoi, K. H.; DeKosky, B. J.; Murrin, E. M.; Wirth, M. M.; Ellington, A. D.; Dörner, T.; Marcotte, E. M.; Boutz, D. R.; Georgiou, G. *Proc. Natl. Acad. Sci.* **2014**, *111* (6), 2259–2264.
(30) Schatz, D. G. *Immunol. Rev.* **2004**, *200* (1), 5–11.
(31) Schroeder Jr, H. W. *Dev. Comp. Immunol.* **2006**, *30* (1–2), 119–135.
(32) Walsh, G. *Nat. Biotechnol.* **2014**, *32* (10), 992–1000.
(33) Ecker, D. M.; Jones, S. D.; Levine, H. L. *mAbs* **2015**, *7* (1), 9–14.
(34) Aggarwal, S. *Nat. Biotechnol.* **2014**, *32* (1), 32–39.
(35) Zeck, A.; Regula, J. T.; Larraillet, V.; Mautz, B.; Popp, O.; Göpfert, U.; Wiegeshoff, F.; Vollertsen, U. E. E.; Gorr, I. H.; Koll, H.; Papadimitriou, A. *PLOS ONE* **2012**, *7* (7), e40328.
(36) Yang, Y.; Strahan, A.; Li, C.; Shen, A.; Liu, H.; Ouyang, J.; Katta, V.; Francissen, K.; Zhang, B. *mAbs* **2010**, *2* (3), 285–298.
(37) Miyazaki, C.; Iba, Y.; Yamada, Y.; Takahashi, H.; Sawada, J.; Kurosawa, Y. *Protein Eng.* **1999**, *12* (5), 407–415.
(38) Winkler, K.; Kramer, A.; Küttner, G.; Seifert, M.; Scholz, C.; Wessner, H.; Schneider-Mergener, J.; Höhne, W. *J. Immunol.* **2000**, *165* (8), 4505–4514.
(39) Liu, H.; Gaza-Bulseco, G.; Faldu, D.; Chumsae, C.; Sun, J. *J. Pharm. Sci.* **2008**, *97* (7), 2426–2447.
(40) Zhang, Z.; Pan, H.; Chen, X. *Mass Spectrom. Rev.* **2009**, *28* (1), 147–176.
(41) Zhang, H.; Cui, W.; Gross, M. L. *FEBS Lett.* **2014**, *588* (2), 308–317.
(42) Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianférani, S. *Anal. Chem.* **2013**, *85* (2), 715–736.
(43) Lundell, N.; Schreitmüller, T. *Anal. Biochem.* **1999**, *266* (1), 31–47.
(44) Bandeira, N.; Pham, V.; Pevzner, P.; Arnott, D.; Lill, J. R. *Nat. Biotechnol.* **2008**, *26* (12), 1336–1338.
(45) de Costa, D.; Broodman, I.; VanDuijn, M. M.; Stingl, C.; Dekker, L. J. M.; Burgers, P. C.; Hoogsteden, H. C.; Sillevis Smitt, P. A. E.; van Klaveren, R. J.; Luider, T. M. *J. Proteome Res.* **2010**, *9* (6), 2937–2945.
(46) Dekker, L. J. M.; Zeneyedpour, L.; Brouwer, E.; Duijn, M. M. van; Smitt, P. A. E. S.; Luider, T. M. *Anal. Bioanal. Chem.* **2011**, *399* (3), 1081–1091.
(47) Srebalus Barnes, C. A.; Lim, A. *Mass Spectrom. Rev.* **2007**, *26* (3), 370–388.

(48) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121* (4), 806–812.
(49) Zhou, H.; Ning, Z.; E. Starr, A.; Abu-Farha, M.; Figeys, D. *Anal. Chem.* **2012**, *84* (2), 720–734.
(50) Kelleher, N. L.; Thomas, P. M.; Ntai, I.; Compton, P. D.; LeDuc, R. D. *Expert Rev. Proteomics* **2014**, *11* (6), 649–651.
(51) Cui, W.; Rohrs, H. W.; Gross, M. L. *Analyst* **2011**, *136* (19), 3854–3864.
(52) Tsybin, Y. O.; Fornelli, L.; Stoermer, C.; Luebeck, M.; Parra, J.; Nallet, S.; Wurm, F. M.; Hartmer, R. *Anal. Chem.* **2011**, *83* (23), 8919–8927.
(53) Fornelli, L.; Damoc, E.; Thomas, P. M.; Kelleher, N. L.; Aizikov, K.; Denisov, E.; Makarov, A.; Tsybin, Y. O. *Mol. Cell. Proteomics* **2012**, *11* (12), 1758–1767.
(54) Mao, Y.; Valeja, S. G.; Rouse, J. C.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2013**, *85* (9), 4239–4246.
(55) Kleemann, G. R.; Beierle, J.; Nichols, A. C.; Dillon, T. M.; Pipes, G. D.; Bondarenko, P. V. *Anal. Chem.* **2008**, *80* (6), 2001–2009.
(56) Srzentić, K.; Fornelli, L.; Laskay, Ü. A.; Monod, M.; Beck, A.; Ayoub, D.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (19), 9945–9953.
(57) Cannon, J.; Lohnes, K.; Wynne, C.; Wang, Y.; Edwards, N.; Fenselau, C. *J. Proteome Res.* **2010**, *9* (8), 3886–3890.
(58) 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. *Nat. Methods* **2012**, *9* (8), 822–824.
(59) Pang, Y.; Wang, W.-H.; Reid, G. E.; Hunt, D. F.; Bruening, M. L. *Anal. Chem.* **2015**, *87* (21), 10942–10949.
(60) Fornelli, L.; Ayoub, D.; Aizikov, K.; Beck, A.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (6), 3005–3012.
(61) Wang, D.; Wynne, C.; Gu, F.; Becker, C.; Zhao, J.; Mueller, H.-M.; Li, H.; Shameem, M.; Liu, Y.-H. *Anal. Chem.* **2015**, *87* (2), 914–921.
(62) Nicolardi, S.; Deelder, A. M.; Palmblad, M.; van der Burgt, Y. E. M. *Anal. Chem.* **2014**, *86* (11), 5376–5382.
(63) Yan, B.; Valliere-Douglass, J.; Brady, L.; Steen, S.; Han, M.; Pace, D.; Elliott, S.; Yates, Z.; Han, Y.; Balland, A.; Wang, W.; Pettit, D. *J. Chromatogr. A* **2007**, *1164* (1–2), 153–161.
(64) An, Y.; Zhang, Y.; Mueller, H.-M.; Shameem, M.; Chen, X. *mAbs* **2014**, *6* (4), 879–893.
(65) Tran, J. C.; Zamdborg, L.; Ahlf, D. R.; Lee, J. E.; Catherman, A. D.; Durbin, K. R.; Tipton, J. D.; Vellaichamy, A.; Kellie, J. F.; Li, M.; Wu, C.; Sweet, S. M. M.; Early, B. P.; Siuti, N.; LeDuc, R. D.; Compton, P. D.; Thomas, P. M.; Kelleher, N. L. *Nature* **2011**, *480* (7376), 254–258.
(66) Garcia, B. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (2), 193–202.
(67) Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L. *Anal. Chem.* **2011**, *83* (17), 6868–6874.
(68) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.

(69) Roepstorff, P.; Fohlman, J. *Biol. Mass Spectrom.* **1984**, *11* (11), 601.
(70) Biemann, K. *Biol. Mass Spectrom.* **1988**, *16* (1–12), 99–111.
(71) Louris, J. N.; Cooks, R. G.; Syka, J. E. P.; Kelley, P. E.; Stafford, G. C.; Todd, J. F. J. *Anal. Chem.* **1987**, *59* (13), 1677–1685.
(72) McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3* (6), 599–614.
(73) Dongré, A. R.; Jones, J. L.; Somogyi, Á.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118* (35), 8365–8374.
(74) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35* (12), 1399–1406.
(75) Vaisar, T.; Urban, J. *J. Mass Spectrom.* **1996**, *31* (10), 1185–1187.
(76) Bleiholder, C.; Suhai, S.; Harrison, A. G.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 1032–1039.
(77) Tsaprailis, G.; Somogyi, A.; Nikolaev, E. N.; Wysocki, V. H. *Int. J. Mass Spectrom.* **2000**, *195–196*, 467–479.
(78) Gu, C.; Tsaprailis, G.; Breci, L.; Wysocki, V. H. *Anal. Chem.* **2000**, *72* (23), 5804–5813.
(79) Huang, Y.; Triscari, J. M.; Tseng, G. C.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *Anal. Chem.* **2005**, *77* (18), 5800–5813.
(80) Fekete, S.; Guillarme, D.; Sandra, P.; Sandra, K. *Anal. Chem.* **2016**, *88* (1), 480–507.
(81) Desaire, H. *Mol. Cell. Proteomics* **2013**, *12* (4), 893–901.
(82) Dunbar, R. C. *J. Am. Chem. Soc.* **1971**, *93* (18), 4354–4358.
(83) Freiser, B. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1974**, *96* (20), 6260–6266.
(84) Ly, T.; Julian, R. R. *Angew. Chem. Int. Ed.* **2009**, *48* (39), 7130–7137.
(85) Brodbelt, J. S. *Chem. Soc. Rev.* **2014**, *43* (8), 2757–2783.
(86) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. *J. Proteome Res.* **2010**, *9* (8), 4205–4214.
(87) Vasicek, L.; Brodbelt, J. S. *Anal. Chem.* **2010**, *82* (22), 9441–9446.
(88) Robinson, M. R.; Madsen, J. A.; Brodbelt, J. S. *Anal. Chem.* **2012**, *84* (5), 2433–2439.
(89) Robinson, M. R.; Moore, K. L.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (8), 1461–1471.
(90) Fort, K. L.; Dyachenko, A.; Potel, C. M.; Corradini, E.; Marino, F.; Barendregt, A.; Makarov, A. A.; Scheltema, R. A.; Heck, A. J. R. *Anal. Chem.* **2016**, *88* (4), 2303–2310.
(91) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(92) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(93) Cannon, J. R.; Kluwe, C.; Ellington, A.; Brodbelt, J. S. *PROTEOMICS* **2014**, *14* (10), 1165–1173.

(94) O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2014**, *136* (37), 12920–12928.
(95) Cammarata, M. B.; Thyer, R.; Rosenberg, J.; Ellington, A.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2015**, *137* (28), 9128–9135.
(96) Holden, D. D.; McGee, W. M.; Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 1008–1016.
(97) Shaw, J. B.; Robinson, E. W.; Paša-Tolić, L. *Anal. Chem.* **2016**, *88* (6), 3019–3023.
(98) Woody, R. W.; Koslowski, A. *Biophys. Chem.* **2002**, *101–102*, 535–551.
(99) Cui, W.; Thompson, M. S.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (8), 1384–1398.
(100) Kim, T.-Y.; Thompson, M. S.; Reilly, J. P. *Rapid Commun. Mass Spectrom.* **2005**, *19* (12), 1657–1665.
(101) Yoon, S. H.; Chung, Y. J.; Kim, M. S. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (5), 645–655.
(102) Thompson, M. S.; Cui, W.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (8), 1439–1452.
(103) Madsen, J. A.; Ko, B. J.; Xu, H.; Iwashkiw, J. A.; Robotham, S. A.; Shaw, J. B.; Feldman, M. F.; Brodbelt, J. S. *Anal. Chem.* **2013**, *85* (19), 9253–9261.
(104) O'Brien, J. P.; Pruet, J. M.; Brodbelt, J. S. *Anal. Chem.* **2013**, *85* (15), 7391–7397.
(105) Robotham, S. A.; Kluwe, C.; Cannon, J. R.; Ellington, A.; Brodbelt, J. S. *Anal. Chem.* **2013**, *85* (20), 9832–9838.
(106) Robotham, S. A.; Horton, A. P.; Cannon, J. R.; Cotham, V. C.; Marcotte, E. M.; Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (7), 3990–3997.
(107) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120* (13), 3265–3266.
(108) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (26), 9528–9533.
(109) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14* (19), 1793–1800.
(110) Chi, A.; Huttenhower, C.; Geer, L. Y.; Coon, J. J.; Syka, J. E. P.; Bai, D. L.; Shabanowitz, J.; Burke, D. J.; Troyanskaya, O. G.; Hunt, D. F. *Proc. Natl. Acad. Sci.* **2007**, *104* (7), 2193–2198.
(111) Wiesner, J.; Premsler, T.; Sickmann, A. *PROTEOMICS* **2008**, *8* (21), 4466–4483.
(112) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72* (3), 563–573.
(113) Breuker, K.; Oh, H.; Lin, C.; Carpenter, B. K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (39), 14011–14016.
(114) Olsen, J. V.; Schwartz, J. C.; Griep-Raming, J.; Nielsen, M. L.; Damoc, E.; Denisov, E.; Lange, O.; Remes, P.; Taylor, D.; Splendore, M.; Wouters, E. R.; Senko, M.; Makarov, A.; Mann, M.; Horning, S. *Mol. Cell. Proteomics* **2009**, *8* (12), 2759–2769.

(115) Chi, A.; Bai, D. L.; Geer, L. Y.; Shabanowitz, J.; Hunt, D. F. *Int. J. Mass Spectrom.* **2007**, *259* (1–3), 197–203.
(116) McAlister, G. C.; Berggren, W. T.; Griep-Raming, J.; Horning, S.; Makarov, A.; Phanstiel, D.; Stafford, G.; Swaney, D. L.; Syka, J. E. P.; Zabrouskov, V.; Coon, J. J. *J. Proteome Res.* **2008**, *7* (8), 3127–3136.
(117) Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119* (9), 2240–2248.
(118) Breuker, K.; Oh, H.; Horn, D. M.; Cerda, B. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124* (22), 6407–6420.
(119) Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. *Mol. Cell. Proteomics* **2007**, *6* (11), 1942–1951.
(120) Tran, B. Q.; Barton, C.; Feng, J.; Sandjong, A.; Yoon, S. H.; Awasthi, S.; Liang, T.; Khan, M. M.; Kilgour, D. P. A.; Goodlett, D. R.; Goo, Y. A. *J. Proteomics* **2016**, *134*, 93–101.
(121) Bondarenko, P. V.; Second, T. P.; Zabrouskov, V.; Makarov, A. A.; Zhang, Z. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (8), 1415–1424.

Chapter 2

Experimental Methods

2.1 Overview

The methods described herein are aimed at extending the throughput and versatility of tandem mass spectrometric analysis of antibodies through various combinations of strategic enzymatic digestion, site-selective derivatization, and ultraviolet photodissociation. These methods were designed for either extensive fragmentation for improved structural characterization, or to selectively discriminate diagnostic sequence information from a greater pool of uninformative background peptides within the context of a high throughput liquid chromatography-tandem mass spectrometry (LC-MS/MS) based workflow.

2.2 Mass Spectrometric Instrumentation

All experiments described herein were performed using a linear quadrupole ion trap (LIT) mass spectrometer or a Fourier Transform (FT) hybrid linear ion trap-Orbitrap mass spectrometer. Both instruments were equipped with electrospray ionization (ESI) sources, or customized variations thereof, and modified to enable photodissociation.

2.2.1 Electrospray Ionization (ESI)

The gas-phase conversion and ionization of analytes represents the *sine qua non* of mass spectrometry. While a wide variety of techniques have been introduced to generate ions of various forms, the advent of soft ionization methods capable of inducing charging effects while maintaining the integrity of the molecule has revolutionized the field of biological mass spectrometry. Electrospray ionization (ESI)[1,2] is among the most

transformative and widely utilized soft ionization methods owing to its ability to generate multiply charged ions and its ease of integration with orthogonal solution-phase front-end separations (such a liquid chromatography and capillary electrophoresis), making it amenable to high throughput applications using a variety of MS platforms.

The ESI process can be likened to a controlled current electrolytic cell, as demonstrated in **Figure 2.1**. Charge is accumulated at the liquid surface under the influence of an electric field and can also be aided by the addition of solution phase additives (i.e., 0.1-1% formic acid or ammonium hydroxide) that act as proton donors (positive mode) or proton acceptors (negative mode). At sufficiently high voltage conditions (~2-5 kV) a Taylor cone forms at the solution interface that ejects highly charged droplets. These droplets subsequently undergo rapid evaporation that induces numerous fission events driven by Coulombic repulsion in the shrinking droplet. This process continues until desolvated gas-phase ions are formed that can enter the mass spectrometer for detection. Electrospray ionization (ESI) and nano ESI (nESI) were utilized for all experiments described herein using voltages of 1.8–2.0 kV and 3.5–5.0 kV respectively.

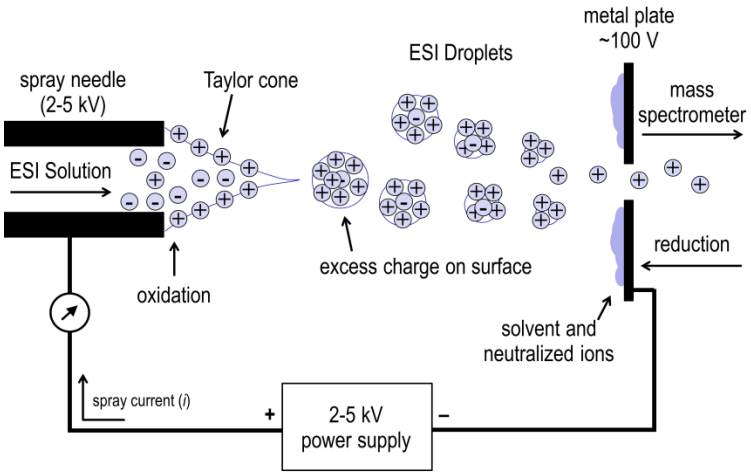


Figure 2.1 Schematic representation of an electrospray ionization process (adapted from Reference 3).

2.2.2 Dual Source Reactor

A dual source reactor was developed in-house to facilitate frontend ion/ion reactions. As illustrated in **Figure 2.2**, the reactor was equipped with two ESI sources (Prosolia Inc., Indianapolis, IN), as described in the previous section 2.2.1, that are coupled to the frontend of the mass spectrometer via a U-shaped railing system that surrounded the MS inlet. The position of each source could be independently adjusted in the x, y and z dimensions relative to the MS inlet. One source was completely integrated with the MS, allowing direct control of polarity and spray voltage through the instrument control software. The spray voltage of the second source was supplied by an external 5 kV high voltage power supply (Stanford Research Systems Inc., Sunnyvale, CA) operated in negative polarity. Nebulizing sheath gas was introduced using an external nitrogen line equipped with a tee fitting and adjustable metering valve to evenly split the gas flow to both sources and allow manipulation of gas flow rates, respectively. The source was use in direct infusion mode or coupled directly to a capillary LC system for high-throughput bioconjugation reactions.

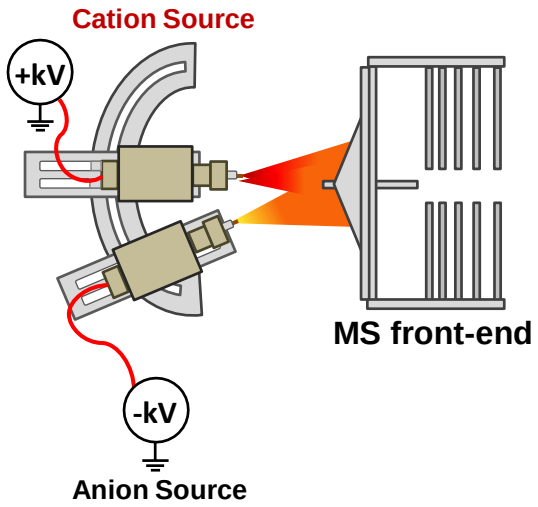


Figure 2.2 Schematic representation of a dual ESI reactor for front-end ion/ion reactions.

2.2.3 Linear Ion Trap (LIT)

A linear ion trap mass analyzer generally consists of a quadrupole with static lens elements placed at the terminal ends of the rods. Opposite RF voltages applied the two pairs of quadrupole rods confine ion motion in the radial direction, whereas DC potentials applied to the terminal ends of the trap confine ions in the axial direction and gate the flow of ions into the trap.[4] Auxiliary AC and broadband waveforms are applied to one pair of rods for ion isolation and activation in a secular frequency dependent manner. Mass analysis is accomplished via sequential ion ejection from the trap using an RF voltage ramp, also referred to as a mass selective instability scan.

All bottom-up analysis and online ion/ion-mediated bioconjugation experiments described herein were performed on a Velos Pro dual linear ion trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA) equipped for CID, HCD and UVPD. The dual LIT configuration consists of a high pressure cell operated at a bath gas (He) pressure of approximately 5 mTorr for high trapping and CID efficiency, and a low pressure cell operated at nominally 0.3 mTorr for mass analysis.

2.2.4 Orbitrap Mass Analyzer

Orbitrap technology was first introduced by Makarov in 2000[5] and has since become a powerful FT-based alternative to Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry owing to its high performance (i.e., resolving power and mass accuracy) and short acquisition times without the need for an expensive superconducting magnet.[6] The analyzer consists of an outer barrel-shaped electrode and an inner spindle-like center electrode. Under the influence of a quadrologarithmic field, ions undergo harmonic oscillations along the axis of spindle electrode at a frequency

proportional to $(m/z)^{-1/2}$, which is detected as image current induced on the barrel electrode.[5]

Middle-down LC-MS/MS analyses from several variations of IgG digests were performed on a modified hybrid linear ion trap Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany)[7,8] equipped with CID, HCD, ETD and UVPD capabilities. To enhance the transfer and detection of large ions, the HCD collision gas pressure in the vacuum chamber containing the Orbitrap mass analyzed was reduce from 10 mTorr to ~5-7 mTorr depending on empirically defined spectral quality. The instrument was modified to enable UVPD in the HCD collision cell as described in greater detail in section 2.3.2.

2.3 ION ACTIVATION

Multiple activation methods were employed for the research presented herein, including 193 nm UVPD, 351 nm UVPD, CID, HCD, and ETD. The set-up, acquisition strategies and relevant parameters for each activation method are detailed in the subsequent sections. Unless otherwise stated, MS^1 full scan spectra were acquired from *m*/z 400–2000 and MS^2 product ion spectra were collected from *m/z* 200–2000 (note that the low *m/z* cutoff (LMCO) varied for CID based on precursor *m/z*) for all LC-MS/MS analyses.

2.3.1 Collisional Activation

Normalized collision energies (NCE) between 25-35% were used for CID and HCD based peptide fragmentation. CID was employed for all collisional activation experiments performed on the Velos Pro dual LIT, whereas HCD was used for all

collisional activation experiments performed on the Orbitrap Elite owing to its superior speed (0.1 ms vs. 10 ms per activation step) and broader *m/z* trapping range. For online bioconjugation experiments, collision-induced covalent conversion of peptide/FBDSA complexes formed during front-end ion/ion reactions was typically performed under "gentle" CID conditions using NCE values ≤18%.

2.3.2 Ultraviolet Photodissociation (UVPD)

Both mass spectrometer platforms used in this work were modified to enable photodissociation as previously described.[7–9] To allow UVPD on the Velos Pro dual linear ion trap, the back flange of the vacuum manifold of the instrument was replaced with a customized viewport flange designed to hold a CaF_2 or fused silica window for transmission of radiation at 193 nm and 351 nm, respectively. An unfocused beam was elevated and aligned coaxially with the dual pressure ion trap using an optical periscope consisting of two 45° mirrors and subsequently passed through a 1.8 mm aperture mounted to the viewport flange followed by a 2 mm aperture on the exit lens of the ion trap. The instrument firmware was modified to allow activation and mass analysis in the low pressure cell in order to minimize beam divergence and maximize the photon flux experienced by the ion cloud. For implementation of UVPD on the Orbitrap Elite the vacuum manifold was modified via the addition of a CaF_2 optical window coaxial to the dedicated HCD collision cell. Additionally, the rods of the ETD reagent ion transfer bent quadrupole were rotated 45° to allow transmission of the laser beam into the HCD cell. The instrument firmware was adjusted to allow simultaneous trapping of the ion cloud and pulse triggering within the HCD cell for a user defined activation period. Coherent ExciStar XS 500 pulsed excimer lasers (Santa Clara, CA) operated at either 193 nm (ArF) or 351 nm (XeF) were used for all experiments and coupled to the mass

spectrometers as described. The maximum repetition rate (500 Hz) was used for all experiments to generate one pulse (5 ns pulse duration) every 2 ms. An external trigger supplied through the auxiliary interface port gated a pulse generator (Berkley Nucleonics Corp., San Rafael, CA) which triggered the laser during the activation period of the instrument scam function.

On the Velos Pro dual linear ion trap, the instrument firmware was modified to enable pulse triggering when the CID NCE was set to zero. The activation time was adjusted in 2 ms intervals to control the number of pulses based on the repetition rate of the laser. Activation *q*-values ranging between 0.1 and 0.125 were used to reduce the low-mass cutoff (LMCO) of the trap. Chromophore selective 351 nm UVPD (3.5 eV per photon) was carried out using ten 3 mJ pulses. Alternatively, 193 nm UVPD (6.4 eV per photon) of Schiff base modified peptides subjected to online on/ion-mediated bioconjugation was performed using one pulse at 2 mJ.

For 193 nm UVPD experiments performed on the Orbitrap Elite mass spectrometer, the instrument firmware was modified to enable UVPD in the HCD cell via a custom UVPD flag. Ions were transferred into the HCD cell using 1% NCE, which allowed efficient ion transfer without sufficient energy to cause undesired fragmentation. All experiments were performed using 1-2 pulses at 2-3 mJ.

2.3.3 Electron Transfer Dissociation (ETD)

Electron transfer dissociation (ETD) experiments were performed on an ETD-enabled hybrid linear-ion trap Orbitrap Elite mass spectrometer. Fluoranthene radical anions generated via a backend chemical ionization source were introduced as the electron-transfer reagent to induce fragmentation of either IgG peptides produced by restricted Lys-C proteolysis or IgG subunits produced by IdeS digestion and chemical

reduction. ETD analysis of IgG subunits was carried out using two variations of targeted acquisition for pre-established precursor ions based on a preceding LC-MS survey scan. For the first approached, the most abundant charge state was isolated using a 20 *m/z* window and continuously activated across the elution profile for each subunit. Alternatively, for the second approach, approximately 3 or 4 less abundant, but more highly charged (greater charge density) precursors were co-isolated using a 150 *m/z* isolation window for targeted broadband ETD acquisition. In both cases, a reaction time of 5 ms was used with a reagent AGC target of 7.5×10^5. To enhance detection of low abundance and large product ions, the pressure in the Orbitrap detector region was reduced so that the change in pressure (Δp) equaled 0.1×10^{-10} (~5 mTorr).

For IgG peptides generated by Lys-C proteolysis, ETD product ion spectra were acquired in a data-dependent manner for the top five most abundant precursors identified in an MS^1 survey scan. ETD was performed with the charge-stated dependent reaction time feature enabled and the reagent AGC target set to 5×10^5. For optimal performance, ETD was only carried out for ions possessing charge states of 4+ or higher.

2.4 LIQUID CHROMATOGRAPHY (LC)

The quality control of therapeutic antibodies using bottom-up or middle-down mass spectrometry requires the complete sampling of all peptides produced upon enzymatic digestion. This relies critically on front-end separations to facilitate the independent interrogation of each peptide. Similarly, effective separation of digests arising from highly homologous IgG mixtures increases the potential for sampling diagnostic peptides from the variable signatures of clonally unique antibodies. Separation of such mixtures is most commonly accomplished using reserve phase liquid

chromatography (RP-LC), which is readily coupled with ESI-MS/MS instrumentation. Reverse phase resins commonly used for bottom-up and middle-down/top-down separations include C18 (for bottom-up) and C8 or C4 (for middle- or top-down), respectively. The LC systems, stationary phases and parameters used for the work described in this dissertation are detailed below.

2.4.1 Dionex Ultimate 3000 Nano/Capillary RSLC System

Separations of tryptic and chymotryptic peptide mixtures were carried out on a Dionex Ultimate 3000 nano LC system configured for preconcentration. The system was equipped with a self-packed New Objective (Woburn, MA) Integrafrit trap column (3.5 cm x 100 μm) and Picofrit nanobore analytical column (15 cm x 75 μm) containing 3 μm Michrom Magic (Auburn, CA) C18 stationary phase. Eluent A was 0.1% formic acid in water and eluent B was 0.1% formic acid in acetonitrile. Peptides were preconcentrated on the trap column with 2% acetonitrile for 5 min at a flow rate of 5 μL/min. For elution of peptides, a linear gradient from 3% to 40% eluent B was used over 70 min for tryptic bovine serum albumin (BSA) digests and 120 min for chymotryptic single-chain antibody fragment (scAb) digests at a flow rate of 0.3 μL/min.

Separation of a model peptide mixture for subsequent front-end ion/ion reactions was accomplished using a Dionex Ultimate 3000 capillary flow system coupled with a homebuilt front-end dual spray reactor (described in section 2.2.2). Peptides were separated on an Agilent ZORBAX 300 Extend-C18 column (0.3 x 150 mm, 3.5 μm particle size) held at a constant temperature of 30°C. Eluent A was 0.1% aqueous formic acid and eluent B was 0.1% formic acid in acetonitrile. A linear gradient was employed from 3% B to 35% B over 30 min at a flow ate 4 μL/min.

2.4.2 Dionex Ultimate 3000 Microbore LC System

Separation of IgG subunits (~25 kDa) produced by IdeS digestion and chemical reduction was accomplished using a Dionex Ultimate 3000 microbore LC system equipped with a Waters XBridge Protein BEH300 C4 column (2.1 x 250 mm, 3.5 μm particle size) held at a constant temperature of 65°C. Eluent A was 0.1% aqueous formic acid and eluent B was 39.9% isopropanol, 60% acetonitrile, and 0.1% formic acid (v/v/v). Two μg of digest was injected directly on column and separated using a steep linear ramp from 5% to 20% B over 2 min followed by a shallow gradient from 20% to 40% B over 28 min at a flow rate of 250 μL/min.

2.4.3 Eksigent 2D Plus NanoLC

Digests of clinical grade IgG1 and a mixture of influenza monoclonal antibodies produced by restricted Lys-C proteolysis were carried out on an Eksigent nano flow LC system configured for preconcentration. The system was equipped with a self-packed New Objective Integrafrit trap column (3 cm x 100 μm) and Picofrit nanobore analytical column (20 cm x 75 μm) containing 5 μm Michrom Magic C8 stationary phase. Eluent A was 0.1% formic acid in water and eluent B was 0.1% formic acid in acetonitrile. Peptides were loaded onto the trap column and preconcentrated for 5 min with 2% acetonitrile containing 0.1% formic acid at a flow rate of 5 μL/min. For separation of single IgG digests, a linear ramp from 4% to 10% B over five minutes and then 10% to 40% B over 55 min was employed at a flow rate of 0.3 μL/min. Separation of digests derived from influenza IgG mixtures was performed under the same gradient conditions, with the exception that the length of separation was doubled to increase the probability of sampling diagnostic variable region peptides required for IgG differentiation.

2.5 MATERIALS AND REAGENTS

Peptides KMVELVHFL, KLVANNTRL, RPPGFSPFR, ASHLGLAR, DRVYIHPFHLVIHN, DAEFRHDSGYQVHHQK, as well as phosphopeptides RQpSVELHSPQSLPR, GGGPApTPKKAKKL, and KKALRRQEpTVDAL were purchased from AnaSpec (Fremont, CA). CDPGYIGSR, AGCKNFFWKTFTSC, SYSMEHFRWG and RRLIEDAEpYAARG-NH_2 were obtained from American Peptide Company (Sunnyvale, CA). DRVYIHPFHL and bovine serum albumin (BSA) were obtained from Sigma-Aldrich (St. Louis, MO). Clinical grade Trastuzumab (Herceptin) IgG1 monoclonal antibody was obtained from Genentech (San Francisco, CA), and Adalimumab (Humira) IgG1 monoclonal antibody was obtained >97% purity from BOC Science (Shirley, NY). Reagents used for site-specific bioconjugation, including 4-formyl-1,3-benzenedisulfonic acid (FBDSA) and AlexaFluor 350 C_5 maleimide were obtained from Sigma-Aldrich and Life Technologies (Grand Island, NY), respectively.

Proteases including trypsin, Lys-C and chymotrypsin were purchased from Promega (Madison, WI). Recombinant immunoglobulin G-degrading enzyme from *Streptococcus pyogenes* (IdeS/FabRICATOR, Genovis) was obtained from Bulldog Bio, Inc. (Portsmouth, NH). All other solvents, chemicals and reagents were obtained from Sigma-Aldrich, Thermo Fisher Scientific or EMD Millipore (Temecula, CA).

2.6 SAMPLE PREPARATION

2.6.1 Single-Chain Antibody Fragment (scAb) Expression and Purification

Single-chain antibody fragments were prepared are previously described by Rani *et al*.[10] with the following modifications: antibody fragments were expressed as scAbs by inserting single-chain variable fragment (scFv)-encoding genes of 2 anti-HA33 [HA33

(33 kDa), part of the botulinum neurotoxin complex[11]] variants (#3 and #4) into pMopac16 vector, a pAK400 derivative[12] in which the scFv was fused in frame to a C-terminal human kappa light chain constant domain and hexahistidine tag. Antibody fragments were expressed in *Escherichia coli* BL21. Individual colonies were inoculated into 4 mL TB media with 50 μg/mL ampicillin and 2% glucose and were grown overnight at 37°C. Overnight cultures were used to inoculate 400 mL TB media with 50 μg/mL ampicillin, and cells were grown at 37°C until an $o.d._{600}$ of 0.6 was reached. Cultures were brought to 25°C and protein expression was induced with 1 mM IPTG. After 5 h incubation at 25°C, cells were collected by centrifugation, and the protein was purified from osmotic shock fraction, as previously described.[10,13] The purity of isolated scAb was verified by gel electrophoresis on a 4-20% SDS-PAGE gel (NuSep, Lawrenceville, GA) stained with Coomassie blue.

2.6.2 Selective Derivatization and Sample Processing

Protein-level modification of cysteine residues in BSA and scAb samples was accomplished by first reducing disulfide bonds in the presence of excess tris(2-carboxyethyl)phosphine (TCEP) at 55°C for 60 min. Free thiols were then reacted in the presence of 45 mM Alexa Fluor 350 maleimide for 3 h, or with 15 mM iodoacetamide (IAM) for 45 min. Alkylation reactions were carried out at room temperature in the dark. Modified BSA and scAbs were then digested overnight at 37°C using a 20:1 ratio of protein to trypsin or chymotrypsin, respectively.

Phosphopeptides were subjected to N-terminal derivatization via 4-sulfophenyl isothiocyanate (SPITC) and carbamylation. SPITC reactions were carried out by reacting 20 μL of reagent stock solution prepared by dissolving 1 mg of SPITC in 100 μL of 1x PBS (pH 7.4) with 10 nmol of peptide for 30 minutes at 55°C. N-terminal carbamylation

was accomplished via incubation of 10 nmol of peptide with 8 M urea in 50 mM Tris-HCl (pH 8) at 80°C for 4 h. The reaction products were immediately desalted on C18 to terminate further reaction.

2.6.3 Ion/Ion Reaction-Mediated Peptide Bioconjugation

Peptide cations were subjected to front-end ion/ion reactions with 4-formyl-1,3-benezenedisulfonic acid (FBDSA) anions using the dual source reactor detailed in section 2.2.2. To accomplish this peptide cations were produced either from direct infusion or LC separation at flow rates of 1.5-4 μL/min using spray voltage of 1.5-1.75 kV. Reagent anions were simultaneously generated by direct infusion of 1-2 mM FBDSA at 3 μL/min produced using spray voltages ranging between -2 and -3 kV. Peptide/FBDSA complexes formed by ion/ion reaction were then isolated and subjected to low energy CID (NCE ≤18%) for conversion to covalent Schiff base modified forms.

2.6.4 Middle-Down IgG Samples Prepared by Restricted Lys-C Proteolysis

IgG samples were prepared in 50 mM Tris-HCl (pH 8) containing 2M urea and digested at 37°C with Lys-C for two hours using a 1:75 enzyme-to-substrate ratio. Digestion was carried out prior to reduction and alkylation to promote the formation of larger peptides (i.e., a greater number of missed cleavages). The resulting digests were diluted to a final urea concentration of 0.5 M in Tris-HCl and sequentially reduced in the presence of 5 mM DTT for 30 min at 37°C and alkylated with 25 mM IAM at room temperature, in the dark, for 30 min. Immediately following alkylation, IgG digests were acidified with 0.5% formic acid and desalted on a solid phase extraction (SPE) column containing C8 resin.

2.6.5 Middle-Down IgG Samples Prepared by IdeS Digestion

Acetone-precipitated monoclonal IgG1 antibodies were resuspended at 5 μg/μL in IdeS cleavage buffer (50 mM sodium phosphate, 150 mM NaCl, pH 6.6) and subsequently digested with one unit of IdeS per microgram of IgG at 37˚C for 30 minutes. Following digestion, IgGs were denatured with 4 M urea and reduced in the presence of 30 mM TCEP for 30 minutes at room temperature. To quench the reaction and prevent disulfide bond reformation, the sample was acidified with 1% formic acid (FA). Immediately prior to analysis the sample was diluted to 1 μg/μL in 0.1% FA.

2.7 AUTOMATED TANDEM MASS SPECTROMETRY DATA INTERPRETATION

LC-MS/MS experiments often generate thousands of mass spectra within a single dataset, of which only a fraction can be meaningfully correlated to proteolytic peptides. The most common approach for interpreting large volumes of MS/MS spectra in a time and computationally effective manner is through comparison with *in-silico* generated theoretical fragmentation patterns generated from candidate protein sequence databases. Once potential spectral matches are identified, they can then be validated through rigorous statistical measures to facilitate confident identifications based on precursor mass and fragmentation ion data. The strength of these searches is therefore largely dependent on mass accuracy, sequence coverage across the peptide backbone, and the ability of the algorithm to interpret the resulting spectra. *In silico* searching has proven extremely effective for conventional proteomics; however, these methods are limited to the interpretation of datasets for which *a priori* genome data is readily available. Until recently, this limitation precluded the use of *in silico* algorithms to interpret mass spectral data arising from antibody repertoires, for which there is little germline data available due to incredible sequence diversity of immunoglobulins generated by somatic recombination

and hypermutation. This has been partially circumvented with the advent of Ig-seq technology, which enables the generation of an individualized database of immunoglobulin variable gene (V-gene) sequences constructed by next-generation DNA sequencing of mature B cells. In section 2.7.2 an Ig-seq donor V-gene database was used in combination with the ProSightPC search algorithm for the identification of middle-down sized peptides arising from influenza monoclonal antibodies. Additionally, all search algorithms and search parameters used for MS/MS data interpretation are detailed in the subsequent sections.

2.7.1 MassMatrix

The MassMatrix database search algorithm (version 2.4.0) was used for *in silico* interpretation of CID and 351 nm UVPD fragmentation datasets arising from tryptic BSA and chymotryptic scAb digestions. Searches were performed against forward and reverse candidate sequences generated from a bovine proteome FASTA database (UniProtKB) modified to include single-chain antibody fragment (scAb) sequences. Peptide matching was based on *b*- and *y*-type product ions. All data was collected on an LIT mass spectrometer and therefore searched using default low resolution parameters including a precursor ion tolerance of ±1.8 Da and a fragment ion tolerance of ±0.8 Da. Peptide hits were filtered by pp and pp_2 thresholds of 5 and a pp_{tag} threshold of 2.0.

2.7.2 ProSightPC

Automated database searching of middle-down LC-MS/MS data derived from Lys-C digested IgGs was performed using ProSightPC 4.0 (Thermo Fisher Scientific) equipped with UVPD searching functionalities to account for 9 ion types: *a*, *a*+1, *b*, *c*, *x*, *x*+1, *y*, *y*-1 and *z*. Spectra were converted to neutral monoisotopic masses using the Xtract

algorithm (Thermo Fisher Scientific) using precursor and fragment ion signal-to-noise thresholds of 7 and 3, respectively. Data derived from single monoclonal IgGs was searched against a custom database containing ten closely related therapeutic IgG sequences obtained from the DrugBank resource database (http://www.drugbank.ca/). Alternatively, UVPD data from a mixture of anti-influenza IgGs discovered in post-vaccinated donor serum was searched against a custom V-gene database consisting of 17 V_L and 14,499 V_H sequences constructed from next generation sequencing of V_L and V_H genes from B cells of the same donor. Searches were performed in absolute mass mode with a precursor mass tolerance of 2.2 Da and a strict fragment ion tolerance of 5 ppm.

2.7.3 ProSight Lite

All scans collected for a given precursor or group of precursors during targeted LC-MS/MS analysis of IgG subunits were combined to generate an averaged Thermo.RAW file for each subunit. The averaged UVPD spectra were deconvolved using the Xtract algorithm (Thermo Fisher Scientific) using a fragment ion signal-to-noise threshold of 3 and searched using ProSight Lite with UVPD searching enabled.

2.8 REFERENCES

(1) Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. *Science* **1989**, *246* (4926), 64–71.

(2) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Mass Spectrom. Rev.* **1990**, *9* (1), 37–70.

(3) Cech, N. B.; Enke, C. G. *Mass Spectrom. Rev.* **2001**, *20* (6), 362–387.

(4) Douglas, D. J.; Frank, A. J.; Mao, D. *Mass Spectrom. Rev.* **2005**, *24* (1), 1–29.

(5) Makarov, A. *Anal. Chem.* **2000**, *72* (6), 1156–1162.

(6) Michalski, A.; Damoc, E.; Lange, O.; Denisov, E.; Nolting, D.; Müller, M.; Viner, R.; Schwartz, J.; Remes, P.; Belford, M.; Dunyach, J.-J.; Cox, J.; Horning, S.; Mann, M.; Makarov, A. *Mol. Cell. Proteomics* **2012**, *11* (3), O111.013698.

(7) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.

(8) Vasicek, L. A.; Ledvina, A. R.; Shaw, J.; Griep-Raming, J.; Westphall, M. S.; Coon, J. J.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 1105–1108.

(9) Gardner, M. W.; Smith, S. I.; Ledvina, A. R.; Madsen, J. A.; Coon, J. J.; Schwartz, J. C.; Stafford, G. C.; Brodbelt, J. S. *Anal. Chem.* **2009**, *81* (19), 8109–8118.

(10) Rani, M.; Bolles, M.; Donaldson, E. F.; Van Blarcom, T.; Baric, R.; Iverson, B.; Georgiou, G. *J. Virol.* **2012**, *86* (17), 9113–9121.

(11) Ito, H.; Sagane, Y.; Miyata, K.; Inui, K.; Matsuo, T.; Horiuchi, R.; Ikeda, T.; Suzuki, T.; Hasegawa, K.; Kouguchi, H.; Oguma, K.; Niwa, K.; Ohyama, T.; Watanabe, T. *FEMS Immunol. Med. Microbiol.* **2011**, *61* (3), 323–331.

(12) Krebber, A.; Bornhauser, S.; Burmester, J.; Honegger, A.; Willuda, J.; Bosshard, H. R.; Plückthun, A. *J. Immunol. Methods* **1997**, *201* (1), 35–55.

(13) Hayhurst, A.; Happe, S.; Mabry, R.; Koch, Z.; Iverson, B. L.; Georgiou, G. *J. Immunol. Methods* **2003**, *276* (1–2), 185–196.

Chapt[1]er 3

Selective 351 nm Photodissociation of Cysteine-Containing Peptides for Discrimination of Antigen-Binding Regions of IgG Fragments in Bottom-Up LC-MS/MS Workflows[*]

3.1 OVERVIEW

Despite tremendous inroads in the development of more sensitive LC-MS/MS strategies for mass spectrometry-based proteomics, there remains a significant need for enhancing the selectivity of MS/MS-based workflows for streamlined analysis of complex biological mixtures. Here, a novel LC-MS/MS platform based on 351 nm ultraviolet photodissociation (UVPD) is presented for the selective analysis of cysteine-peptide subsets in complex protein digests. Cysteine-selective UVPD is mediated through the site-specific conjugation of reduced cysteine residues with a 351 nm active chromogenic Alexa Fluor 350 (AF350) maleimide tag. Only peptides containing the AF350 chromophore undergo photodissociation into extensive arrays of b- and y-type fragment ions, thus providing a facile means for differentiating cysteine-peptide targets from convoluting peptide backgrounds. Using this approach in addition to strategic proteolysis, the selective analysis of diagnostic heavy chain complementarity determining regions (CDRs) of single-chain antibody (scAb) fragments is demonstrated.

3.2 INTRODUCTION

Over the past several decades, tandem mass spectrometry (MS/MS) has played a pivotal role in expanding the depth and breadth of proteomics research.[1–5] Most notably,

[*]Cotham, V. C.; Wine, Y.; Brodbelt, J. S. *Anal. Chem.* **2013**, 85, 5577-5585.
V.C.C. designed all experiments and completed the labeling reactions and mass spectral analysis.

workflows based on bottom-up liquid chromatography tandem mass spectrometry (LC-MS/MS) have emerged as the primary analytical technique for large-scale characterization of complex protein mixtures, such as those originating from whole cell lysates[6,7] or the blood serum proteome.[8,9] The vast majority of these workflows rely on non-selective activation methods that generate informative and comprehensive mass spectral datasets that are interpreted bioinformatically to enable protein identification, detection of post-translational modifications (PTMs), and structural elucidation.[5,10] Collision induced dissociation (CID)[11,12] and more recently introduced electron-based dissociation (electron capture dissociation (ECD) and electron transfer dissociation (ETD))[13–15] are arguably the most widely employed activation strategies for accomplishing these goals.[16–18]

Despite the merits of using indiscriminate activation processes for proteomic analysis, the exorbitant, and often times redundant, amount of data generated by these methods places a substantial computational burden on bioinformatic databases used for spectral interpretation.[10,19] This often results in long search times and high false discovery rates (FDRs), particularly for low resolution instruments where mass accuracy and resolving power cannot be exploited for improved spectral matching.[20] Furthermore, these non-selective methods greatly reduce the ability to efficiently differentiate diagnostic subsets of information from a greater pool of less-informative data. One common approach to increase selectivity entails the use of data-dependent methods.[21] Data-dependent approaches, in which the most abundant precursor m/z values are excluded from repeated MS/MS interrogation, help streamline data collection but the data discrimination process is linked to precursor abundances rather than structure- or sequence-specific features. Moreover, the common use of 3 *m/z* precursor selection

windows means that MS/MS spectra may be composites of several precursors, thus further confounding spectral interpretation and peptide identification.

The issues surrounding indiscriminate MS/MS methods are highlighted when considered in the context of current proteomic approaches for the characterization of the antigen-specific immunoglobulin repertoire of polyclonal serum antibodies produced during an immune response. The high degree of sequence homology shared between complex mixtures of unique high-specificity IgGs presents a particularly daunting analytical challenge for traditional bottom-up methods and hinges on the ability to sequence low abundance peptides originating from the complementarity determining regions (CDR1, CDR2, and CDR3) of the heavy and light chain variable domains. In particular, heavy chain CDR3 (CDR-H3) peptides, which contain hypervariable amino acid sequences caused by V(D)J gene recombination and somatic hypermutation,[22,23] serve as diagnostic molecular signatures that enable differentiation of IgGs,[24,25] but remain difficult to analyze in the presence of more abundant homologous peptides. Moreover, conventional database searches, which rely on strong statistical correlations between experimental MS/MS spectra and those generated *in silico* from known protein sequences, are not well-suited for antibody-based applications as a result of the high degree of homology between IgG sequences leading to an excessively high number of false identifications.[25] Recent efforts to overcome these challenges have focused on reducing sample complexity via affinity enrichment of antigen-specific IgGs in combination with next generation DNA sequencing to create refined variable gene sequence databases.[25,26] Although these strategies have demonstrated improved deconvolution of the serum antibody response, methods that afford more discriminate analysis of diagnostic CDR peptides in bottom-up workflows remain desirable.

In the ongoing efforts to incorporate greater selectivity into mass spectrometric strategies, photodissociation has cultivated considerable interest as an alternative dissociation method due to its versatility, tunable energy deposition and potential for streamlining tandem MS workflows for complex mixtures.[27–30] While infrared multiphoton photodissociation (IRMPD) has exhibited some capacity for selective screening, this method is largely limited to the differentiation of phosphorylated peptides in protein digests.[31,32] Alternatively, ultraviolet photodissociation (UVPD) has demonstrated significant attributes as an ion activation method due to its capability for fast, high energy deposition and tunable dissociation behavior based on wavelength selection and the incorporation of appropriate chromophores. This last point has proven particularly effective for enhancing the selectivity of UVPD-based techniques through chromophore-mediated dissociation at wavelengths that are minimally absorbed by native peptides. For example, Julian et al. have developed a number of 266 nm UVPD methods that rely on the strategic attachment of a photolabile antenna to drive site-specific radical induced dissociation. This method has been applied to allow facile identification of phosphorylation sites[33,34] and determination of the presence and location of cysteine residues.[35,36] Previously, our group has demonstrated the utility of 355 nm UVPD for applications including the manipulation of product ion distributions for *de novo* sequencing through terminal chromophore labeling,[37] as well as the identification of peptide cross-links via selective photodissociation of a bis-aryl hydrazone crosslinker.[38] Recently, Lemoine et al. introduced a novel photodissociation-based technique, referred to as photo-selected reaction monitoring (photo-SRM), which combines both traditional SRM with selective chromophore-mediated dissociation at visible wavelengths (473 nm).[39] Using site-specific labeling of cysteine residues they were able to show targeted analysis of modified cysteine-containing peptides from complex plasma protein digests.[40]

The present study incorporates both site-specific conjugation of cysteine residues with a commercially available thiol reactive Alexa Fluor 350 chromophore and selective photodissociation at 351 nm. The use of a fast repetition rate excimer laser provides a significant advantage over selective 355 nm UVPD methods previously reported[37,38] in that its affords rapid activation on a timescale amenable to LC-MS/MS. Efficient and exclusive dissociation of Alexa Fluor 350 modified peptides is observed upon irradiation at 351 nm, yielding an extensive array of *b* and *y*-type ions, while unmodified precursors neither absorb nor dissociate. Therein, this selective strategy enables both facile identification of cysteine-containing peptides and effectively eliminates convoluting MS/MS spectra for streamlined data analysis. Following validation of this approach using a standard protein digest, its application is further extended to the selective analysis of the diagnostic third heavy chain complementarity determining region (CDR-H3) of IgG single-chain antibody fragments (scAbs). To accomplish this goal, a highly conserved cysteine residue located N-terminal to the CDR-H3 region of interest[41,42] was targeted for site-specific labeling. Combined with the use of appropriate proteases and selective 351 nm UVPD, high sequence coverage of diagnostic CDR-H3-containing peptides was observed and differentiation of IgG scAbs is accomplished.

3.3 EXPERIMENTAL

3.3.1 Materials and Reagents

Model cysteine-containing peptides CDPGYIGSR and AGCKNFFWKTFTSC were purchased from American Peptide Company (Sunnyvale, CA). Bovine serum albumin (BSA), iodoacetamide, glutathione and urea were obtained from Sigma Aldrich (St. Louis, MO). Mass spectrometry grade trypsin and sequencing grade chymotrypsin

were purchased from Promega (Madison, WI) and thiol-reactive Alexa Fluor 350 C_5 maleimide was obtained from Life Technologies (Grand Island, NY). Tris(2-carboxyethyl)phosphine (TCEP) solution, Slide-A-Lyzer dialysis cassettes (10 kDa MWCO), and PepClean C18 spin columns were purchased from Pierce (Rockford, IL). All other buffer components and solvents were obtained from Fisher Scientific (Fairlawn, NJ).

3.3.2 Single-Chain Antibody Fragment (scAb) Preparation

Single-chain antibody fragments were prepared as previously described[43] with the following modifications: antibody fragments were expressed as scAbs by inserting the scFv encoding genes of 2 anti-HA33 (HA33 (33 kDa) - part of the botulinum neurotoxin complex[44]) variants (#3 and #4) into pMopac16 vector, a pAK400 derivative[45] in which the scFv was fused in frame to a C-terminal human kappa light chain constant domain and hexahistidine tag. Antibody fragments were expressed in *E. coli* BL21. Individual colonies were inoculated into 4 ml TB media with 50 μg/ml ampicillin and 2% glucose, and were grown overnight at 37°C. Overnight cultures were used to inoculate 400 ml of TB media with 50 μg/ml ampicillin and the cells were grown at 37°C until OD_{600} 0.6 was reached. Cultures were transferred to 25°C and protein expression was induced with 1 mM IPTG. After a 5 hour incubation at 25°C, cells were collected by centrifugation and protein was purified from the osmotic shock fraction as described previously.[43,46] The purity of isolated scAb was verified by gel electrophoresis on a 4-20% SDS-PAGE gel (NuSep, Lawrenceville, GA) stained with Coomassie blue.

3.3.3 Cysteine Derivatization and Sample Preparation

Disulfide bonds were reduced using a 5x molar excess of TCEP solution (500 mM in water) for 1 hour at 55˚C in the presence of 8 M urea (in 1x PBS, pH 7.4). Following reduction, protein solutions were diluted 2-fold with 1x PBS to lower the urea concentration to 4 M immediately prior to site-selective conjugation of reduced cysteine residues. Alexa Fluor C_5 maleimide stock solution (45 mM in water) was added at a 10x molar excess over the protein concentration and the reaction was allowed to proceed for 3 hours in the dark at room temperature under gentle mixing conditions. Conjugation with Alexa Fluor 350 maleimide results in a mass shift of 478 Da per modified cysteine residue. Standard alkylation was carried out by adding iodoacetamide to a final concentration of 15 mM in solution and allowed to react at room temperature for 45 minutes in the dark. Reactions were quenched with an excess of glutathione. Modified BSA was dialyzed overnight against 1x PBS to remove urea and excess reagent prior to enzymatic digestion with trypsin. Alternatively, modified scAbs were buffer exchanged into 100 mM Tris-HCl, 10 mM $CaCl_2$ (pH 8) for conditions compatible with chymotrypsin digestion. Modified BSA and scAbs were digested overnight at 37˚C using a 20:1 ratio of protein to trypsin or chymotrypsin, respectively.

3.3.4 Mass Spectrometry, Liquid Chromatography, and Photodissociation

All experiments were undertaken on a Thermo Scientific Velos Pro dual-pressure linear ion trap mass spectrometer (San Jose, CA) outfitted with a Coherent ExciStar XeF excimer laser operated at 351 nm. The laser setup was similar to that previously described,[47,48] with the exception that a quartz window was used for transmission of 351 nm photons. Modified peptides were analyzed by direct infusion ESI-MS using a spray voltage of 4 kV and a heated capillary temperature of 200ºC. Protein digests were

separated on a Dionex UltiMate 3000 RSLCnano system (Sunnyvale, CA) configured for on-line preconcentration using a New Objective Integrafrit trap column (3.5 cm x 100 μm) and Picofrit nanobore analytical column with integrated emitter (15 cm x 75 μm). Both columns were packed in-house with 3 μm Michrom Magic C18 stationary phase (Auburn, CA). 1 μL of digest prepared at 1 μM in 0.1% formic acid was injected onto the trap column and preconcentrated with 2% acetonitrile with 0.1% formic acid at a flow rate of 5 μL/min for 5 minutes. For separation on the analytical column eluent A consisted of 0.1% formic acid in water and eluent B was 0.1% formic acid in acetonitrile. A linear gradient from 3% to 40% B over 70 min at 0.3 μL/min was used for all tryptic BSA digests. Due to the considerably greater number of peptides generated using the lower specificity chymotrypsin protease, scAb digests were subjected to a 120 minute gradient using the same conditions previously stated.

LC-MS/MS data was collected using data-dependent acquisition in which the first event was the full mass scan (*m/z* range of 400 – 2000) followed by 10 consecutive isolation and activation (UVPD or CID) events of the most abundant ions detected in the full mass scan. For all 351 nm UVPD experiments, the q_z-value was set to 0.1 to reduce the low mass cut-off to below *m/z* 150 and precursor ions were irradiated with 10 pulses at a repetition rate of 500 Hz using a power of 3 mJ/pulse. The q_z-value was increased to 0.25 and a normalized collision energy (NCE) of 35% was applied during a 10 ms activation period for all comparative CID experiments.

3.3.5 Database Searching

MassMatrix database search algorithm (version 2.4.0) was used for *in silico* interpretation of all MS/MS data.[49] All LC-MS/MS RAW files generated in the Thermo Xcalibur software (version 2.2) were converted to the mzXML file format compatible

with the MassMatrix algorithm. MS/MS data for BSA digests was searched against the reference bovine proteome database (UniProtKB). This same database was modified to incorporate both scAb sequences for interpretation of MS/MS data from scAb digests. A precursor ion mass tolerance of ± 1.8 Da and a fragment ion tolerance of ± 0.8 Da were used for all searches. Oxidation of methionine (+15.9949 Da) and conjugation of cysteine residues with Alexa Fluor 350 maleimide in its standard (+478.1284 Da) and hydrated (+496.1389 Da) forms were searched as dynamic modifications and peptides consisting of fewer than four amino acids were filtered out. All database search results were verified manually. In some cases, peptides not identified by MassMatrix were manually interpreted based on MS/MS fragmentation and the presence of a unique reporter ion at m/z 296 generated by cleavage of the amide bond within the Alexa Fluor 350 maleimide tag.

3.4 RESULTS AND DISCUSSION

The workflow of the LC-MS/351 nm UVPD platform used in this study for the selective-analysis of cysteine-containing peptides is shown in **Figure 3.1**. In principle, native peptides exhibit minimal gas phase absorption near 350 nm, therefore incorporation of chromogenic moieties that enhance photoabsorptivity are critical for successful UVPD within this wavelength regime. The present study uses a fast repetition rate 351 nm excimer laser to achieve selective photodissociation of cysteine-containing peptides on a timescale amenable to chromatographic separation of complex protein digests. As shown in **Figure 3.1**, the disulfide bonds in the proteins were reduced, and the free cysteines were subsequently alkylated using the Alexa Fluor maleimide reagent in lieu of the conventional iodoacetamide utilized in most bottom-up proteomic methods.

The proteins were proteolyzed, and the resulting digest was analyzed by nanoLC-MS with UVPD as the activation mode. Only the Alexa Fluor-modified cysteine peptides absorb, undergo photodissociation, and give diagnostic fragmentation patterns.

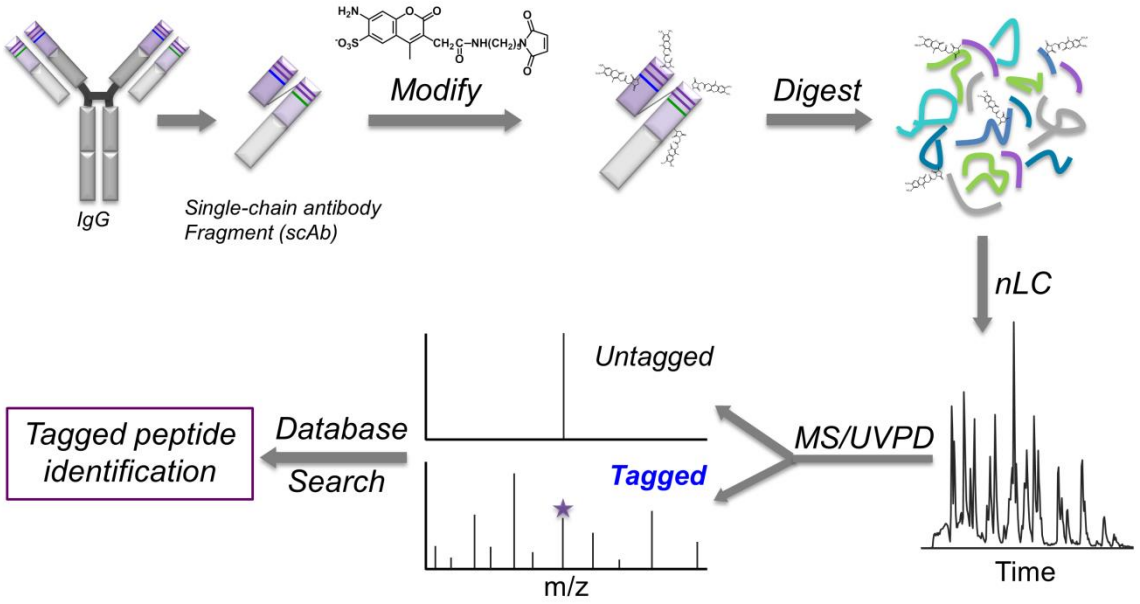


Figure 3.1 LC-MS/MS workflow based of 351 nm UVPD for the selective analysis of cysteine-containing peptides in complex mixtures. Proteins are subjected to site-specific conjugation at cysteine residues with a chromogenic Alexa Fluor 350 maleimide tag. Modified protein digests are separated by nanoLC and activated by 351 nm UVPD which promotes selective photodissociation of Alexa Fluor 350 modified peptides.

Cysteine residues are compelling analytical targets for this study due in part to the intrinsically high nucleophilicity of the free sulfhydryl functionality contained in their side chain, making them attractive for site-specific incorporation of chromophores necessary for selective 351 nm activation. Furthermore, cysteines are responsible for less

than 1.4% of the total amino acid composition, but are present in nearly all proteins of well-characterized proteomes (i.e. 97% in the human proteome), making them ideal targets for streamlined protein identification via selective analysis of the unique cysteine peptide subset. Lastly, it has been shown that cysteine is a highly conserved residue located at the N-terminal position of CDR3 sequences of monoclonal antibody heavy chains. CDR-H3 sequences exhibit the greatest amino acid diversity of all hypervariable antibody domains and are predominantly responsible for antigen recognition. As such, these diagnostic regions provide a unique molecular signature by which antigen-specific IgGs can be identified.[25] Thus, when combined with the appropriate proteases, this conserved cysteine residue provides a fixed target for selective chromophore-mediated dissociation of diagnostic CDR-H3 sequences to enable efficient characterization and differentiation of IgGs. As shown in the following sections, we evaluate the selectivity of our 351 nm UVPD platform against conventional bottom-up LC-MS/MS of protein and IgG digests to assess its analytical potential for streamlined proteomic and immunogenic analysis.

3.4.1 Selective Modification and 351 nm UVPD of Model Cysteine Peptides

The utility of the cysteine-selective UVPD strategy is highly dependent on the ability to efficiently conjugate cysteine residues in a protein (or protein mixture) to a chromophore that affords strong absorption at 351 nm. Solution phase absorbance data are shown in **Figure 3.2** for unmodified and AF350-modified forms of the cysteine-containing tripeptide glutathione. The low absorbance profile of unmodified glutathione is highly representative of native proteolytic peptides, which do not contain active chromophores near 350 nm. By incorporating the AF350 chromophore into the peptide structure via derivatization of the cysteine residue, absorption was dramatically increased

over a wavelength range well-suited for 351 nm activation. Although gas-phase and solution absorption maxima are not necessarily identical, the high absorbance from 325 to 375 nm in solution offered a sufficiently broad range to guide our choice of Alexa Fluor 350 equipped with a cysteine-reactive moiety for our gas-phase study.

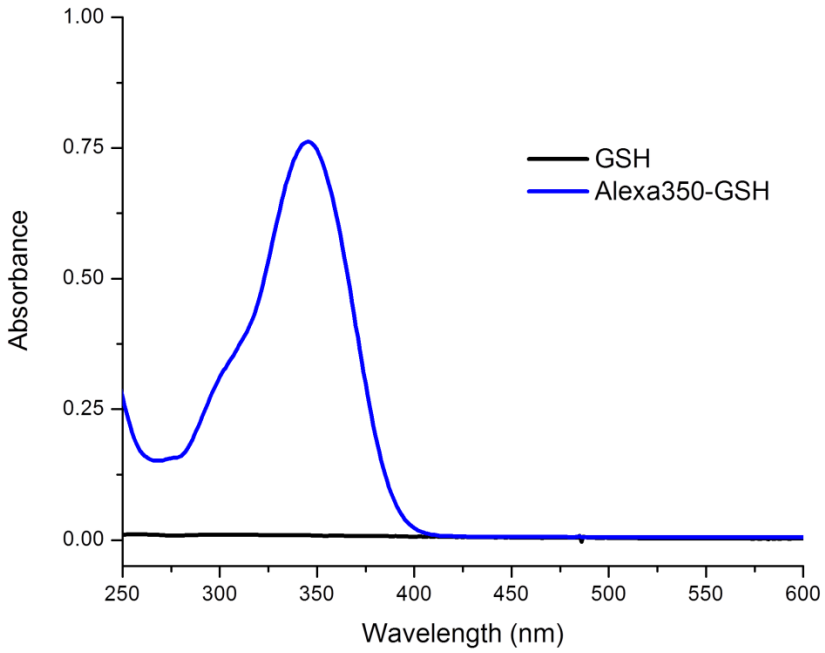


Figure 3.2 UV/VIS absorbance curves of native and Alexa Fluor 350 maleimide modified glutathione (GSH) in water.

Model peptides CDPGYIGSR and AGCKNFFWKTFTSC were selected to assess the efficiency of site-specific modification at cysteine residues with Alexa Fluor 350 maleimide, as well as to evaluate the feasibility of promoting selective photodissociation at 351 nm. Reduction and subsequent derivatization of cysteine residues with the chromogenic AF350 tag exhibited complete conversion of peptides to their modified forms as demonstrated in the representative ESI spectrum shown in **Figure 3.3** for AF350

labeled CDPGYIGSR. Although the AlexaFluor moiety contains an ionizable sulfonic acid group, the Alexa Fluor-modified peptides exhibited the same charge states, and similar distribution of charge states and ionization efficiencies as observed for iodoacetamide-alkylated peptides, and thus there is no evidence that the sulfonic acid group is deprotonated in the gas phase.

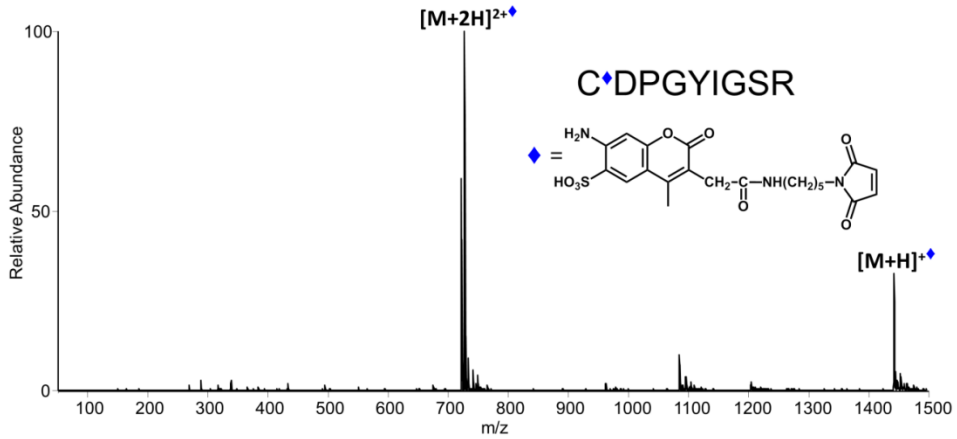


Figure 3.3 ESI mass spectrum of Alexa Fluor 350 maleimide modified CDPGYIGSR. A mass shift of 478 Da corresponding to Alexa Fluor 350 conjugation of the cysteine residue is denoted by (♦).

Conjugation at the cysteine side chain was confirmed by both CID and UVPD, as shown in **Figure 3.4** for AF350 modified AGCKNFFWKTFTSC. Both spectra exhibited abundant and identical arrays of *b*- and *y*-type fragment ions with mass shifts consistent with AF350-maleimide modification at both Cys3 and Cys14. The multiple sites of modification were further verified by consecutive losses of 80 Da from the precursor ion, which correspond to the sulfonic acid moieties from two separate AF350 labels. This characteristic loss proved effective in not only identifying conjugated peptides, but also determining the number of modifications along the peptide backbone. One noteworthy

attribute of using UVPD over conventional CID for the activation of AF350-peptide conjugates arises from access to a lower m/z trapping limit during UVPD because the effectiveness of UVPD is not mediated by the rf voltage applied to the trap, as is the case for CID. As exemplified by the UVPD mass spectrum shown in **Figure 3.4b**, the detection of informative low mass ions, particularly a diagnostic reporter ion at *m/z* 296 generated from cleavage of the amide bond within the AF350 maleimide tag provided an additional means to pinpoint modified cysteine peptides in mixtures and afforded greater confidence in subsequent peptide identifications.

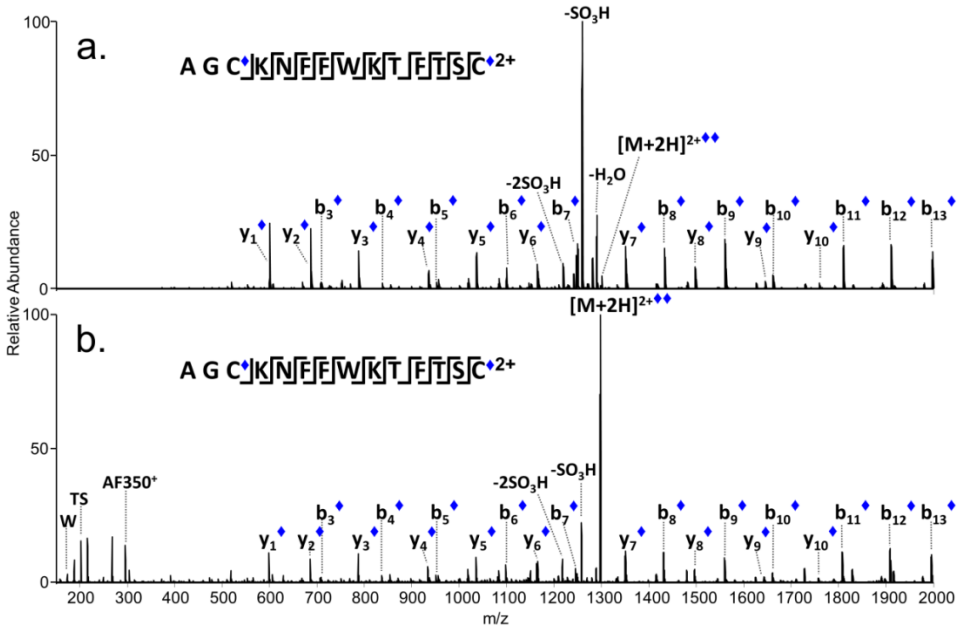


Figure 3.4 MS/MS spectra based on (a) CID and (b) 351 nm UVPD (10 pulses, 3 mJ) of AGCKNFFWKTFTSC 2+ modified at both cysteine residues with Alexa Fluor 350 maleimide (♦).

The remarkable overlap in CID and UVPD spectra for the modified cysteine-containing peptide (**Figure 3.4**) indicated similarities in the activation processes and

internal energy deposition; however, unlike resonant CID, which caused exclusive excitation of the isolated precursor, the non-resonant nature of UVPD promoted activation of both precursor and primary fragment ions retaining the chromogenic tag. This potential for secondary activation enabled greater tunability of energy deposition based on laser parameter selection (i.e., laser power and number of pulses). In this study, optimal photoactivation was accomplished using ten 5 ns laser pulses at a constant energy of 3 mJ per pulse. These conditions were selected based on three criteria: i) adequate dissociation of the precursor, ii) abundance of informative fragment ions, and iii) activation on a timescale comparable to conventional strategies used in LC-MS/MS workflows.

To demonstrate the concept of selective 351 nm photodissociation, the peptide CDPGYIGSR was alkylated by either iodoacetamide (IAM) or chromogenic Alexa Flour 350 maleimide, infused, and activated by 351 nm UVPD. The overlaid isolation spectra for both forms of the doubly charged peptide are shown in **Figure 3.5a** during the laser "off" state. Upon triggering the laser to deliver 10 pulses of 351 nm photons (laser "on" state, **Figure 3.5b**), the AF350-modified precursor diminished by approximately 97% of its initial intensity, confirming that this ion was highly responsive to photoirradiation. Alternatively, the IAM-alkylated precursor remained unaffected by UV irradiation. This point was illustrated more explicitly by comparing the behavior of each precursor ion in the respective single ion chronograms (SIC) during alternating laser on/laser off states (**Figure 3.5c**). The ion intensity of the IAM-alkylated precursor showed no dependence on the state of the laser as demonstrated by a negligible change in intensity over the course of the laser on/off switching experiment, whereas the ion intensity of the AF350-modified precursor showed immediate depletion during the laser on periods.

Furthermore, the UVPD spectrum of the AF350-modified peptide obtained during the laser on state yielded diagnostic *b*- and *y*-type fragment ions (**Figure 3.5d**), whereas that of the corresponding IAM-alkylated peptide showed only the presence of the intact precursor (i.e. no dissociation, data not shown).

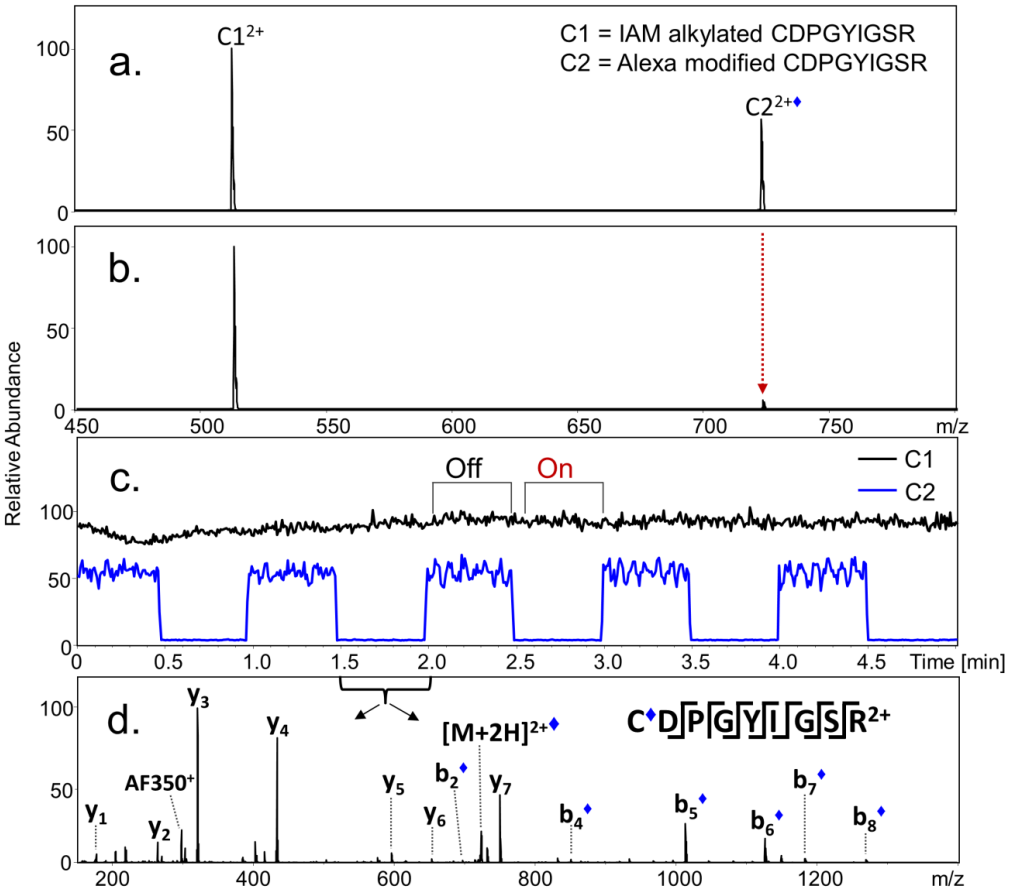


Figure 3.5 Overlaid and normalized isolation spectra of Alexa Fluor 350 (♦) and IAM modified CDPGYIGSR 2+ under (a) laser OFF and (b) laser ON conditions. (c) Single ion chronograms (SIC) of the 2+ charge state of modified and unmodified peptide are shown under alternating laser ON/OFF states in 30 s intervals and (d) the UVPD spectrum of Alexa modified peptide during laser ON conditions.

3.4.2 Validation of Cysteine-Selective Strategy with Model Protein Digest

The selective 351 nm UVPD method was extended to a considerably more complex tryptic protein digest to assess the feasibility of cysteine-selective analysis within the context of a typical LC-MS/MS based proteomic workflow. Bovine serum albumin (BSA) was selected as a model protein due to the significant number of cysteine residues (35) contained throughout the protein sequence, thus providing a large pool of analytical targets. LC-MS/MS analysis based on CID and 351 nm UVPD was performed for AF350 maleimide-conjugated tryptic BSA to compare the relative peptide populations identified by both MS/MS methods. The discriminatory power of selective 351 nm UVPD on a chromatographic timescale is illustrated by the representative UVPD spectra shown in **Figure 3.6** of non-cysteine containing HPYFYAPELLYYANK and AF350-modified DDPHACYSTVFDK tryptic BSA peptides, respectively. The non-cysteine peptide (identified by CID) exhibited no dissociation upon activation by 351 nm UVPD, whereas extensive sequence information was obtained for the AF350-modified peptide. While conventional CID yielded approximately 80% sequence coverage of BSA through the identification of over 70 peptides, a database search of the UVPD data returned an exclusive list of 24 peptides, all cysteine-containing, accounting for 42% of the BSA sequence and 34 of the possible 35 (97%) cysteine residues. As demonstrated in **Figure 3.7**, examination of the extracted ion chromatogram (XIC) of the *m/z* 296 reporter ion provided a facile means to track cysteine-peptide elution in the resulting LC-MS/UVPD dataset.

The results from the model protein study provided validation of selective cysteine-peptide analysis in a conventional LC-MS/MS workflow. Importantly, 351 nm UVPD demonstrated the potential for streamlined database searching of the cysteine-peptide subset in a complex protein digest compared to indiscriminant CID, which

generated a significant number of convoluting non-cysteine peptide identifications. Furthermore, these results demonstrated the ability to confidently identify a protein based exclusively on the unique cysteine-peptide data generated by 351 nm UVPD.

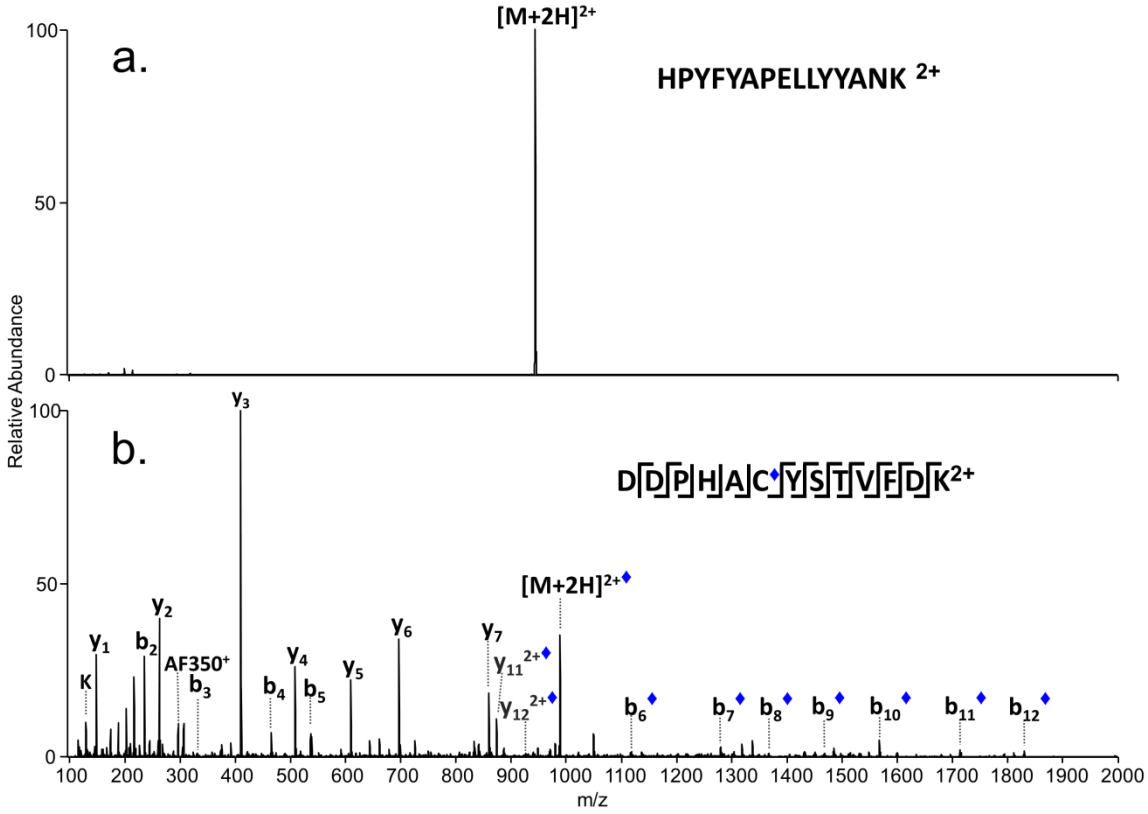


Figure 3.6 351 nm UVPD mass spectra of tryptic BSA peptides (a) HPYFYAPELLYYANK 2+ and (b) Alexa Fluor 350 conjugated DDPHACYSTVFDK 2+ following 351 nm photoirradiation with 10 pulses at 3 mJ.

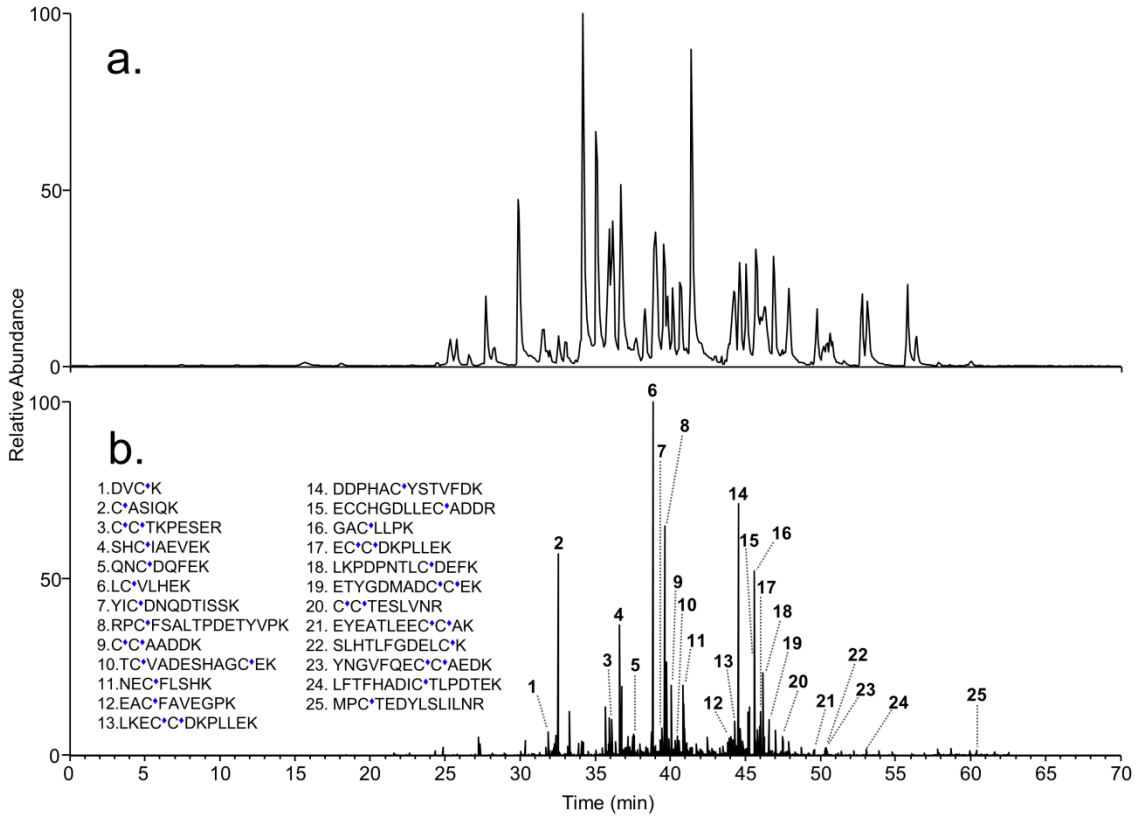


Figure 3.7 Alexa Fluor 350 maleimide modified tryptic BSA digest subjected to LC-MS/351 nm UVPD: (a) Base peak chromatogram and (b) extracted ion chromatogram (XIC) for m/z 296 reporter ion showing peaks associated with identified AF350 modified tryptic cysteine-containing peptides.

3.4.3 Selective CDR-H3 Analysis of IgG Fragments

To demonstrate the analytical merit of this strategy in a biological context, cysteine-selective 351 nm UVPD was used to facilitate selective characterization of the highly diagnostic hypervariable CDR-H3 sequences of single-chain antibody fragments (scAb). Two scAbs sharing over 97% sequence homology with identical CDR-H3 sequences and each containing a total of six cysteine residues (**Figure 3.8**) were subjected to site-specific conjugation with AF350 maleimide. Modified scAbs were strategically digested with chymotrypsin due in part to the low frequency of tryptic sites

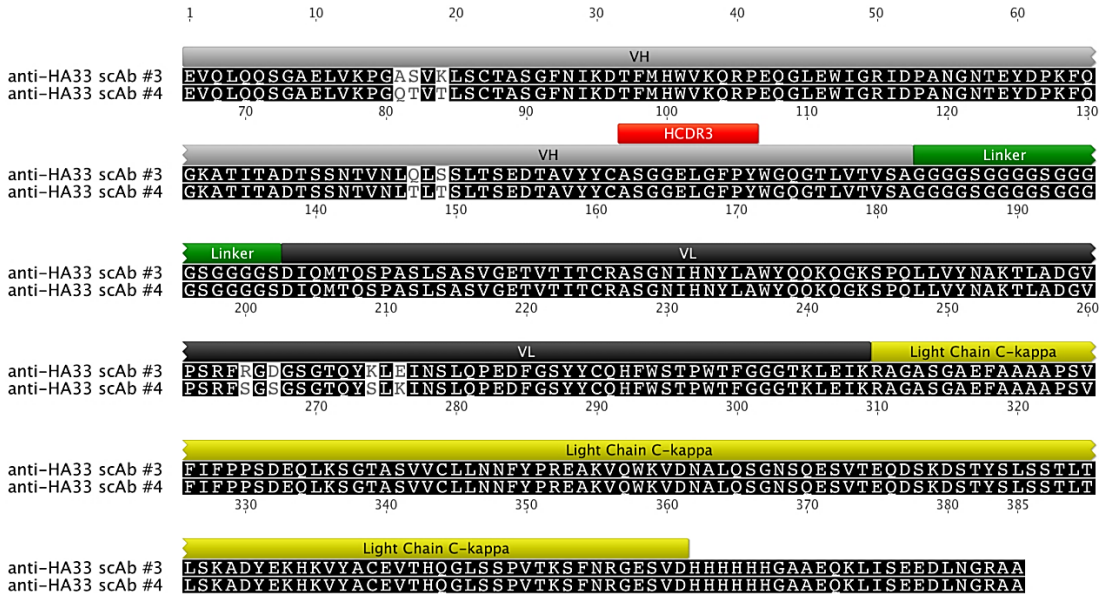


Figure 3.8 Sequence alignment of single-chain antibody fragments (scAbs). Sequence variations are highlighted in white.

in the regions surrounding the CDR-H3 sequences, but more importantly due to the potential for generating CDR-H3 peptides containing a highly conserved cysteine residue in the heavy chain framework to facilitate selective photodissociation. Digests were subsequently analyzed by LC-MS/MS based on CID and 351 nm UVPD. **Table 3.1** provides a comparative summary of peptides identified using both activation platforms. As shown, the CID data resulted in over 40 peptide identifications from each scAb, yielding sequence coverages of 92% and 89% for scAbs #3 and #4, respectively. Alternatively, the 351 nm UVPD data generated selective identification of only modified cysteine-containing peptides, thereby reducing the amount of convoluting peptide information by more than 60%. With the exception of CQHFW, all cysteine-containing peptides identified by both MS/MS strategies were found exclusively in their modified forms, thereby suggesting nearly complete derivatization of cysteine residues. Interestingly, AF350-modified cysteine peptides YCASGGEL and CQHFW were

identified from the 351 nm UVPD datasets for each scAb, but were not identified from the corresponding CID datasets, perhaps suggesting greater sensitivity of selective UVPD towards AF350-modified peptides.

	Scab #3		Scab #4	
Peptide	CID	UVPD	CID	UVPD
EVQLQQSGAELVKPGASVKL	√	-	-	-
EVQLQQSGAELVKPGQTVTL	-	-	√	-
QQSGAELVKPGASVKL	√	-	√	-
QQSGAELVKPGQTVTL	-	-	√	-
VKPGASVKLSC**TASGF	√	√	-	-
SC*TASGF	√	√	√	√
SC**TASGF	√	√	√	√
NIKDTFMHW + Oxi M (7)	√	-	√	-
MHWVKQRPEQGLEW + Oxi M (1)	√	-	√	-
VKQRPEQGLEW	√	-	√	-
IGRIDPANGNTEY	√	-	√	-
IGRIDPANGNTEYDPKFQGKATITADTSSNTVNL	√	-	√	-
DPKFQGKATITADTSSNTVNL	√	-	√	-
QGKATITADTSSNTVNL	√	-	√	-
QGKATITADTSSNTVNLQL	√	-	-	-
QGKATITADTSSNTVNLTL	-	-	√	-
QLSSLTSEDTAVY	√	-	-	-
TLTSLTSEDTAVY	-	-	√	-
YC*ASGGEL	-	√	-	√
YC*ASGGELGFPY	√	√	√	√
C*ASGGELGFPY	√	√	√	√
WGQGTL	√	-	√	-
WGQGTLVTVSAGGGGSGGGGSGGGGSGGGGSDIQMTQSPASL	√	-	√	-
VTVSAGGGGSGGGGSGGGGSGGGG-SDIQMTQSPASL	√	-	√	-
SASVGETVTITC*RASGNIHNY	√	√	√	√
SASVGETVTITC**RASGNIHNY	√	√	√	√
LAWY	√	-	√	-
YQQKQGKSPQLL	√	-	√	-
YQQKQGKSPQLLVY	√	-	√	-
NAKTLADGVPSRF	√	-	√	-
KLEINSLQPEDFGSYY	√	-	-	-
SLKINSLQPEDFGSYY	-	-	√	-
CQHFW	√	-	√	-
C*QHFW	-	√	-	√
WSTPW	√	-	√	-
TFGGGTKL	√	-	√	-
EIKRAGASGAEF	√	-	√	-
AAAAPSVF	√	-	√	-
IFPPSDEQLKSGTASVVC*L	√	√	√	√
IFPPSDEQLKSGTASVVC*LL	√	√	√	√
KSGTASVVC*L	√	√	√	√
KSGTASVVC*LL	√	√	√	√
LNNFYPREAKVQW	√	-	√	-
KVDNALQSGNSQESVTEQDSKDSTY	√	-	√	-
SLSSTL	√	-	√	-
TLSKADYEKHKVY	√	-	-	-
AC*EVTHQGL	√	√	√	√
AC**EVTHQGL	√	√	√	√
AC*EVTHQGLSSPVTKSF	-	√	-	-
SSPVTKSF	√	-	√	-
ISEEDLNGRAA	√	-	√	-
NGRAA	√	-	√	-

Table 3.1 Summary of peptides identified by LC-MS/MS based on CID and 351 nm UVPD of chymotrypsin digested Alexa Fluor 350 maleimide conjugated scAbs.

UVPD also facilitated differentiation of scAbs despite their high degree of sequence homology and limited number of cysteine residues in areas containing sequence variations. This was accomplished by means of selective 351 nm activation of the unique scAb #3 peptide VKPGASVKLSCTASGF, as shown in **Figure 3.9**. We speculate that the corresponding scAb #4 peptide was not identified by either CID or UVPD due to ionization suppression resulting from sequence mutation K19T, which reduced the number of charge-carrying sites along the peptide backbone. Although selective UVPD afforded a facile approach to withdraw peptide subset information from scAb digests (**Figure 3.10**), the true impact of this method and its implications for streamlined antibody characterization are demonstrated by the UVPD spectrum provided in **Figure 3.10c** for the chymotryptic CDR-H3 peptide YCASGGELGFPY. As shown, by targeting the highly conserved cysteine residue contained in the framework preceding the diagnostic CDR-H3 (see **Figure 3.8**) for site-specific conjugation, in addition to strategic digestion, 100% sequence coverage of the diagnostic CDR-H3 was achieved using the cysteine-selective LC-MS/351 nm UVPD workflow. The base peak chromatogram for the chymotryptic digest of Alexa Fluor 350-modified scAb #3 shows the profile of all eluting peptides (**Figure 3.10a**), and the extracted ion chromatogram (XIC) of the UVPD reporter ion (*m/z* 296) shown in **Figure 3.10b** allowed the elution of all of the cysteine-peptides to be readily pinpointed within the complex mixture, thus streamlining the identification of key peptides of interest. The presence of excess Alexa Fluor 350 maleimide that was not effectively removed from the sample during the dialysis or desalting step also generated the reporter ion at *m/z* 296 and is responsible for several of the unlabeled peaks in the XIC.

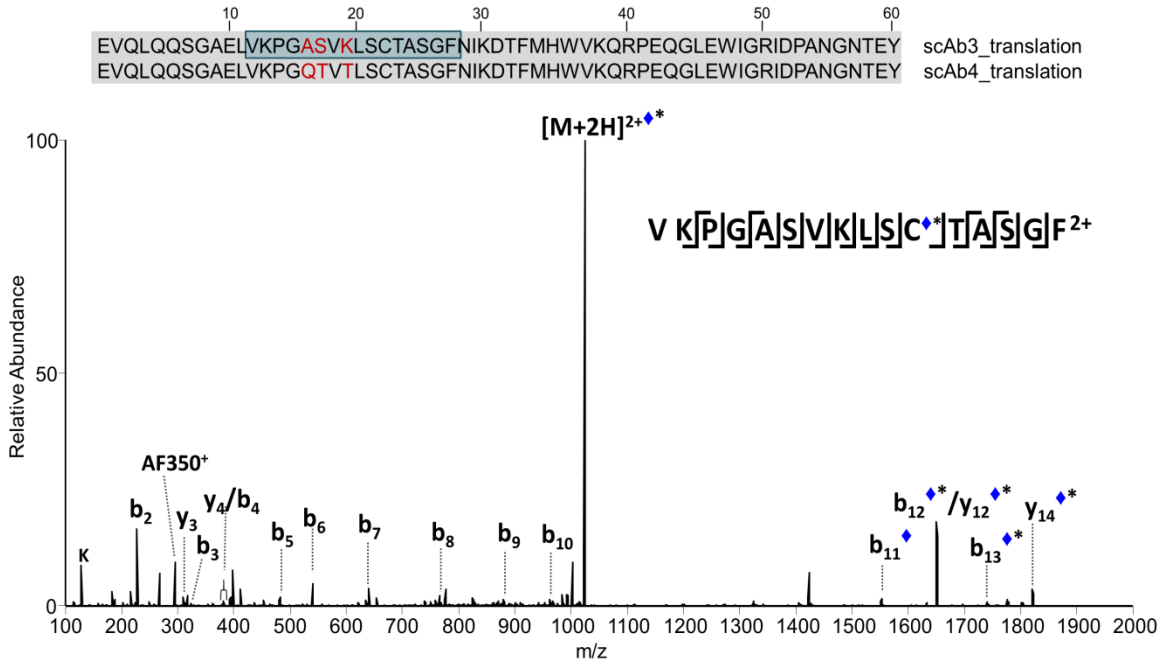


Figure 3.9 Sequence alignment of the first 60 amino acid residues from scAb #3 and scAb #4 are shown, respectively. Sequence variations are highlighted in red. The blue shaded region corresponds to the unique scAb #3 peptide identified by cysteine-selective LC-MS/MS as shown in the 351 nm UVPD spectrum. The presence of the Alexa Fluor 350 maleimide tag is indicated by (♦) and the hydrated form is shown as (♦*)

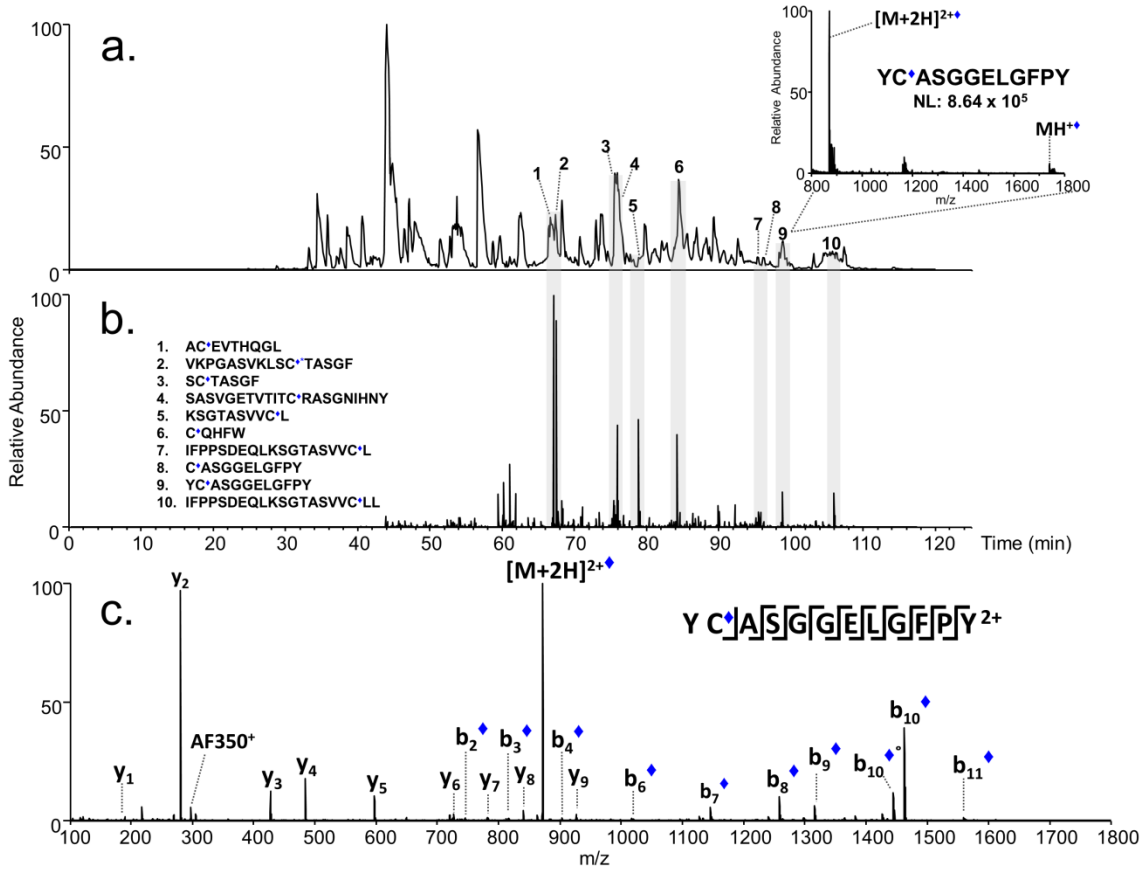


Figure 3.10 (a) Base peak chromatogram for a chymotrypsin digest of Alexa Fluor 350-modified scAb #3 and (b) corresponding XIC for reporter ion at m/z 296. Labeled peaks correspond to modified cysteine-containing peptides identified by 351 nm UVPD. The inset shows the full scan mass spectrum of the modified chymotryptic CDR-H3 peptide YCASGGELGFPY identified by 351 nm UVPD as shown in (c). The presence of Alexa Fluor 350 maleimide is denoted by (♦).

3.5 CONCLUSIONS

When combined with strategic proteolysis, cysteine-selective UVPD was shown to be highly effective for improved characterization of the diagnostic CDR-H3 sequences in antibody digests. The gas-phase photoabsorptivity of native cysteine-containing peptides was significantly enhanced at 351 nm by means of site-specific conjugation of a chromogenic Alexa Fluor 350 moiety for the purpose of selective chromophore-mediated

UVPD. Modified cysteine peptides underwent rapid and extensive photodissociation to yield diagnostic sequence ions, while unmodified peptides exhibited no response upon photoirradiation. The use of a fast repetition rate laser facilitated cysteine-selective UVPD on a timescale amenable to chromatographic separation, thus allowing facile differentiation and streamlined data collection of the cysteine-containing peptide subset of complex mixtures in the context of a conventional bottom-up LC-MS/MS workflow. This method proved successful for the analysis of the diagnostic heavy chain complementarity determining regions of single-chain antibody fragments, ones containing a highly conserved cysteine residue located N-terminal to the CDR-H3 region. Ongoing work to develop a novel thiol-reactive chromogenic tag that provides the added advantage of cysteine-peptide enrichment in addition to selective photodissociation at 351 nm is currently underway. We anticipate that this type of tag will further improve the selectivity of the workflow because the enrichment step should discriminate against unmodified (non-cysteine containing) peptides and alleviate both the congestion of digests and the chances that unmodified peptides will saturate the ESI signal. A complementary enrichment step should be particularly useful for the analysis of low abundance peptides, thus facilitating the analysis of considerably more complex proteomic and immunogenic samples

3.6 REFERENCES

(1) Yates, J. R. *J. Mass Spectrom.* 1998, *33* (1), 1–19.
(2) Aebersold, R.; Mann, M. *Nature* 2003, *422* (6928), 198–207.
(3) Patterson, S. D.; Aebersold, R. H. *Nat. Genet.* 2003, *33*, 311–323.
(4) Domon, B.; Aebersold, R. *Science* 2006, *312* (5771), 212–217.
(5) Mann, M.; Kelleher, N. L. *Proc. Natl. Acad. Sci.* 2008, *105* (47), 18132–18138.
(6) Peng, J.; Gygi, S. P. *J. Mass Spectrom.* 2001, *36* (10), 1083–1091.
(7) Walther, T. C.; Mann, M. *J. Cell Biol.* 2010, *190* (4), 491–500.
(8) Hood, B. L.; Malehorn, D. E.; Conrads, T. P.; Bigbee, W. L. *Methods Mol. Biol. Clifton NJ* 2009, *520*, 107–128.
(9) Coombes, K. R.; Morris, J. S.; Hu, J.; Edmonson, S. R.; Baggerly, K. A. *Nat. Biotechnol.* 2005, *23* (3), 291–292.
(10) Nesvizhskii, A. I.; Vitek, O.; Aebersold, R. *Nat. Methods* 2007, *4* (10), 787–797.
(11) Laskin, J.; Futrell, J. H. *Mass Spectrom. Rev.* 2003, *22* (3), 158–181.
(12) McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* 1992, *3* (6), 599–614.
(13) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* 2004, *101* (26), 9528–9533.
(14) Mikesh, L. M.; Ueberheide, B.; Chi, A.; Coon, J. J.; Syka, J. E. P.; Shabanowitz, J.; Hunt, D. F. *Biochim. Biophys. Acta BBA - Proteins Proteomics* 2006, *1764* (12), 1811–1822.
(15) Zubarev, R. A. *Curr. Opin. Biotechnol.* 2004, *15* (1), 12–16.
(16) Molina, H.; Matthiesen, R.; Kandasamy, K.; Pandey, A. *Anal. Chem.* 2008, *80* (13), 4825–4835.
(17) Molina, H.; Horn, D. M.; Tang, N.; Mathivanan, S.; Pandey, A. *Proc. Natl. Acad. Sci.* 2007, *104* (7), 2199–2204.
(18) Vasicek, L.; O'Brien, J. P.; Browning, K. S.; Tao, Z.; Liu, H.-W.; Brodbelt, J. S. *Mol. Cell. Proteomics* 2012, *11* (7).
(19) Reiter, L.; Claassen, M.; Schrimpf, S. P.; Jovanovic, M.; Schmidt, A.; Buhmann, J. M.; Hengartner, M. O.; Aebersold, R. *Mol. Cell. Proteomics* 2009, *8* (11), 2405–2417.
(20) Elias, J. E.; Haas, W.; Faherty, B. K.; Gygi, S. P. *Nat. Methods* 2005, *2* (9), 667–675.
(21) Stahl, D. C.; Swiderek, K. M.; Davis, M. T.; Lee, T. D. *J. Am. Soc. Mass Spectrom.* 1996, *7* (6), 532–540.
(22) Kim, S.; Davis, M.; Sinn, E.; Patten, P.; Hood, L. *Cell* 1981, *27* (3 Pt 2), 573–581.
(23) Honjo, T.; Habu, S. *Annu. Rev. Biochem.* 1985, *54* (1), 803–830.
(24) de Costa, D.; Broodman, I.; VanDuijn, M. M.; Stingl, C.; Dekker, L. J. M.; Burgers, P. C.; Hoogsteden, H. C.; Sillevis Smitt, P. A. E.; van Klaveren, R. J.; Luider, T. M. *J. Proteome Res.* 2010, *9* (6), 2937–2945.
(25) Wine, Y.; Boutz, D. R.; Lavinder, J. J.; Miklos, A. E.; Hughes, R. A.; Hoi, K. H.; Jung, S. T.; Horton, A. P.; Murrin, E. M.; Ellington, A. D.; Marcotte, E. M.; Georgiou, G. *Proc. Natl. Acad. Sci.* 2013.

(26) Cheung, W. C.; Beausoleil, S. A.; Zhang, X.; Sato, S.; Schieferl, S. M.; Wieler, J. S.; Beaudet, J. G.; Ramenani, R. K.; Popova, L.; Comb, M. J.; Rush, J.; Polakiewicz, R. D. *Nat Biotech* 2012, *30* (5), 447–452.
(27) Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* 2011, *22* (2), 197–206.
(28) Ly, T.; Julian, R. R. *Angew. Chem. Int. Ed.* 2009, *48* (39), 7130–7137.
(29) Reilly, J. P. *Mass Spectrom Rev* 2009, *28* (3), 425–447.
(30) Brodbelt, J. S.; Wilson, J. J. *Mass Spectrom Rev* 2009, *28* (3), 390–424.
(31) Crowe, M. C.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* 2004, *15* (11), 1581–1592.
(32) Crowe, M. C.; Brodbelt, J. S. *Anal. Chem.* 2005, *77* (17), 5726–5734.
(33) Diedrich, J. K.; Julian, R. R. *J. Am. Chem. Soc.* 2008, *130* (37), 12212–12213.
(34) Diedrich, J. K.; Julian, R. R. *Anal. Chem.* 2011, *83* (17), 6818–6826.
(35) Diedrich, J. K.; Julian, R. R. *Anal Chem* 2010, *82* (10), 4006–4014.
(36) Diedrich, J. K.; Julian, R. R. *Anal. Bioanal. Chem.* 2012, *403* (8), 2269–2277.
(37) Wilson, J. J.; Brodbelt, J. S. *Anal Chem* 2007, *79* (20), 7883–7892.
(38) Gardner, M. W.; Brodbelt, J. S. *Anal Chem* 2009, *81* (12), 4864–4872.
(39) Enjalbert, Q.; Simon, R.; Salvador, A.; Antoine, R.; Redon, S.; Ayhan, M. M.; Darbour, F.; Chambert, S.; Bretonnière, Y.; Dugourd, P.; Lemoine, J. *Rapid Commun. Mass Spectrom.* 2011, *25* (22), 3375–3381.
(40) Enjalbert, Q.; Girod, M.; Simon, R.; Jeudy, J.; Chirot, F.; Salvador, A.; Antoine, R.; Dugourd, P.; Lemoine, J. *Anal. Bioanal. Chem.* 1–11.
(41) Morea, V.; Tramontano, A.; Rustici, M.; Chothia, C.; Lesk, A. M. *J. Mol. Biol.* 1998, *275* (2), 269–294.
(42) Kiss, C.; Fisher, H.; Pesavento, E.; Dai, M.; Valero, R.; Ovecka, M.; Nolan, R.; Phipps, M. L.; Velappan, N.; Chasteen, L.; Martinez, J. S.; Waldo, G. S.; Pavlik, P.; Bradbury, A. R. M. *Nucleic Acids Res.* 2006, *34* (19), e132.
(43) Rani, M.; Bolles, M.; Donaldson, E. F.; Van Blarcom, T.; Baric, R.; Iverson, B.; Georgiou, G. *J. Virol.* 2012, *86* (17), 9113–9121.
(44) Ito, H.; Sagane, Y.; Miyata, K.; Inui, K.; Matsuo, T.; Horiuchi, R.; Ikeda, T.; Suzuki, T.; Hasegawa, K.; Kouguchi, H.; Oguma, K.; Niwa, K.; Ohyama, T.; Watanabe, T. *FEMS Immunol. Med. Microbiol.* 2011, *61* (3), 323–331.
(45) Krebber, A.; Bornhauser, S.; Burmester, J.; Honegger, A.; Willuda, J.; Bosshard, H. R.; Plückthun, A. *J. Immunol. Methods* 1997, *201* (1), 35–55.
(46) Hayhurst, A.; Happe, S.; Mabry, R.; Koch, Z.; Iverson, B. L.; Georgiou, G. *J. Immunol. Methods* 2003, *276* (1–2), 185–196.
(47) Gardner, M. W.; Vasicek, L. A.; Shabbir, S.; Anslyn, E. V.; Brodbelt, J. S. *Anal. Chem.* 2008, *80* (13), 4807–4819.
(48) Gardner, M. W.; Smith, S. I.; Ledvina, A. R.; Madsen, J. A.; Coon, J. J.; Schwartz, J. C.; Stafford, G. C.; Brodbelt, J. S. *Anal. Chem.* 2009, *81* (19), 8109–8118.
(49) Xu, H.; Freitas, M. A. *PROTEOMICS* 2009, *9* (6), 1548–1555.

Chapter 4

Middle-Down 193 nm UVPD for Unambiguous Antibody Identification and its Implications for Immunoproteomic Analysis

4.1 Overview

Mass spectrometry has emerged as a powerful tool within the growing field of immunoproteomics, which aims to understand antibody-mediated immunity at the molecular level based on the direct analysis of serological antigen-specific antibody repertoires. To date, these methods have relied on the use of high resolution bottom-up proteomic strategies that require effective sampling and characterization of low abundance peptides derived from the antigen binding domains of clonally unique antibodies within complex mixtures of highly homologous peptide backgrounds. Herein, we describe a method that uses restricted Lys C enzymatic digestion to increase the average mass of proteolytic IgG peptides ($\geq$ 4.5 kDa), thereby enhancing the uniqueness of their amino acid composition and reducing the total sample complexity. When combined with 193 nm ultraviolet photodissociation, improved characterization of the antibody sequence is accomplished relative to conventional collision- and electron-based activation methods. Moreover, we demonstrate the utility of pairing this middle-down UVPD strategy with next-generation V-gene database searching for unambiguous differentiation of unique antibodies in mixtures, thus demonstrating its analytical potential for MS-based serological antibody repertoire analysis.

4.2 INTRODUCTION

A potent humoral immune response hinges on a diverse and highly dynamic antibody repertoire that confers protection through efficient recognition and targeted neutralization of antigenic threats.[1,2] Antibodies, which exist as either membrane-bound B cell receptors (BCRs) or as their secreted protein analogues, are composed of two identical units of paired heavy (H) and light (L) immunoglobulin (Ig) chains that are functionally divided into N-terminal variable domains (V_H and V_L) and C-terminal constant domains (C_H and C_L).[3] In general, the immunoglobulin scaffold exhibits a high degree of sequence homology between individual antibodies, with the exception of three hypervariable complementarity determining regions (CDRs) that are embedded within the frameworks of each V_H and V_L domain.[4] Collectively, the CDRs and their adjacent frameworks form a unique antigen-binding pocket that is distinct to antibodies derived from each clonal population of B cells.[3,4] The potential for substantial sequence diversity within the CDRs is the product of somatic recombination and junctional diversification of germline gene segments that encode for these regions during B cell development.[5] Together these processes give rise to a pre-immune repertoire that is estimated to consist of $>10^8$ distinct antibodies in humans, thus providing the mechanism for broad antigen-recognition that is critical for effective host defense.[4,6]

Upon antigen challenge, the immune repertoire is modulated in an antigen-specific manner through exclusive activation and clonal expansion of naïve B cells expressing antibodies with sufficient antigen affinity.[7] Over time the repertoire undergoes further affinity maturation via somatic hypermutation (SHM) within the variable domains to generate a polyclonal antibody population that is highly specific for the target antigen.[7,8] Currently, there is enormous interest in methods capable of unbiased identification and characterization of such affinity matured antibodies, particularly when

raised at functionally relevant levels against deleterious disease antigens, owing to their intrinsic therapeutic value for use as putative biomarkers and biotherapeutic agents, as well as to guide in vaccine development.[9–11]

High-resolution mass spectrometry (MS) has emerged as a powerful tool to directly monitor serological antibody repertoire composition and dynamics under disease or vaccine-induced states.[12–21] To date, these studies have relied on bottom-up proteomic strategies that entail enzymatic digestion with trypsin or multiple proteases and analysis by liquid chromatography–tandem mass spectrometry (LC-MS/MS). While such methods are well established for conventional proteome analysis,[22] antibody repertoires present several formidable technical hurdles that must be overcome for successful immunoproteomic implementation. For example, no *a priori* sequence database is available for mass spectral interpretation since antibody genes are not directly encoded in the germline, but instead undergo extensive diversification and hypermutation.[13,15] This shortcoming has been largely circumvented by recent integration of MS-based approaches with next-generation DNA sequencing of mature B cell Ig variable (V-gene) domains, or Ig-seq technology, to facilitate the construction of individualized antibody sequence databases.[12,17–21] Despite this improvement, however, deconvolution of antibody repertoires from complex proteolytic mixtures remains non-trivial in the presence of abundant and highly homologous peptide backgrounds arising from conserved immunoglobulin frameworks.[19] As a result, unambiguous antibody identification is contingent on the ability to both sample and fully characterize low abundance peptides originating from unique hypervariable antigen-binding domains.[19] This limitation can be partially mitigated by reducing the complexity of the peptide population, which is often accomplished in proteomics through selective enrichment or depletion of certain types of peptides.[23] Another option, termed the middle-down approach, relies on restrictive

enzymatic digestion to generate fewer and larger peptides than conventional trypsin proteolysis,[24–26] and offers a compelling alternative to bottom-up methods for antibody repertoire analysis. Several groups have already demonstrated the merits of utilizing middle-down approaches for enhanced characterization of therapeutic monoclonal antibodies (mAbs), both in terms of peptide-level sequence coverage and detection of post-translational modifications (PTMs).[27–31] Notably, Tsybin and co-workers showed that enzymatic digestion of a mixture of therapeutic mAbs with secreted aspartic acid protease 9 (Sap9) yielded a greater fraction of peptides that contained full-length or multiple hypervariable CDR sequences to improve the confidence of antibody identification compared to conventional bottom-up analysis.[30] This study represented an important step towards advancing middle-down proteomics for the evaluation of increasingly complex antibody mixtures and motivated our interest in expanding this strategy for integration with more challenging V-gene database searching against an expanded set of closely related sequences to better assess its merits for antibody repertoire analysis. Moreover, the conventional activation methods used to interrogate these longer peptides, including both low- and high-energy collision induced dissociation (CID and HCD, respectively) and electron-transfer dissociation (ETD), do not consistently afford complete inter-residue coverage of the CDRs,[30] which may lead to ambiguous spectral matches during repertoire analysis.[19]

Recently, ultraviolet photodissociation (UVPD) has been established as a versatile alternative ion activation strategy that is amenable to a broad range of proteomics applications.[32–41] In particular, our group has shown that UVPD at 193 nm offers considerable performance gains relative to conventional activation methods for the characterization of whole proteins and large polypeptides thereof on chromatographic timescales,[29,40] thus prompting our efforts to evaluate its utility for middle-down

immunoproteomic analysis. Herein, we describe a high-throughput middle-down strategy for antibody identification and characterization that exploits the combination of restricted lys-C enzymatic digestion to increase the average mass of proteolytic peptides (≥ 4.5 kDa) with the sequencing power of 193 nm UVPD. The overall workflow is illustrated in **Figure 4.1**. Unlike other less specific proteases that have been used to produce comparably sized peptides for middle-down antibody analysis, such as Sap9 and pepsin,[30,31] lys-C cleaves with high selectivity at lysine side-chains to generate a predictable and lower complexity peptide population for streamlined bioinformatic interpretation. Moreover, we compare the performance of 193 nm ultraviolet photodissociation (UVPD) to that of more conventional collision- and electron-based activation methods for the sequence characterization of these large antibody peptides, with a particular emphasis on coverage of the diagnostic hypervariable CDR sequences. Finally, the utility of this method was assessed within the context of a middle-down immunoproteomics experiment combined with next-generation V-gene database searching for the analysis of a mixture of anti-influenza monoclonal antibodies.

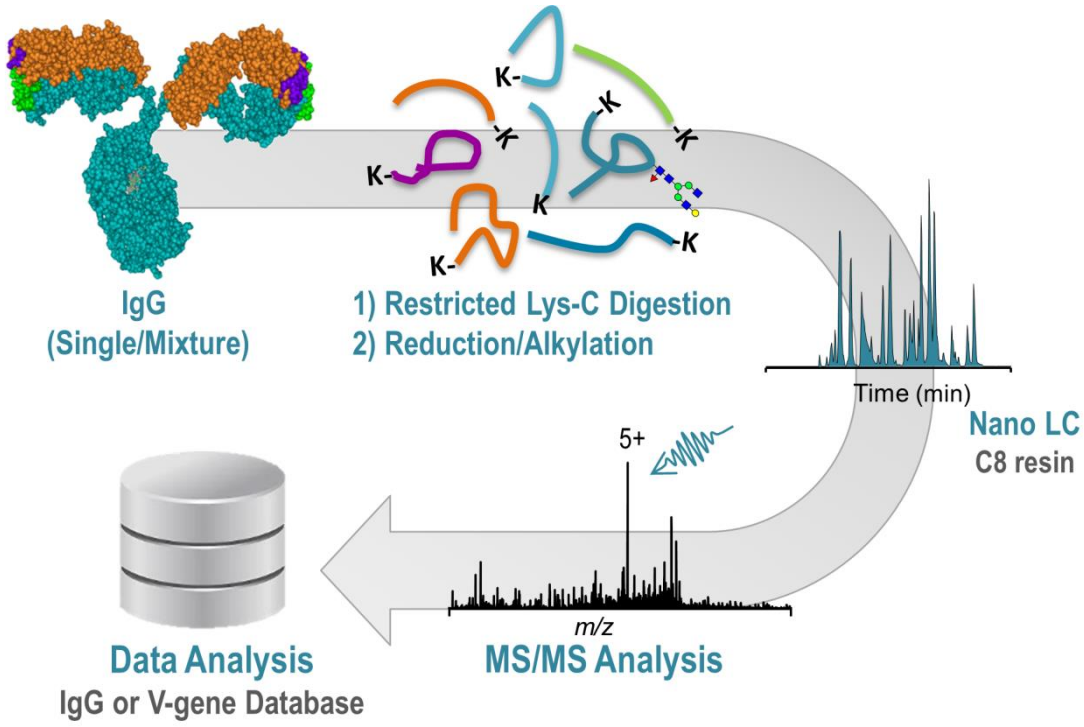


Figure 4.1 Schematic representation of workflow for middle-down analysis of antibodies based on restricted proteolytic digestion with lys-C. Digestion is carried out prior to disulfide reduction and alkylation steps to promote a greater number of missed cleavages as a means to bias the resulting peptide population toward higher mass.

4.3 EXPERIMENTAL

4.3.1 Materials and Reagents

Clinical grade Trastuzumab (Herceptin) monoclonal IgG1 used for initial method development was donated by Genentech. Influenza monoclonal IgGs were provided by the Georgiou lab (University of Texas at Austin) and were synthesized as described by Lee *et. al*.[42] Proteomics-grade r-LysC was obtained from Promega (Madison, WI). LC-MS grade solvents were obtained from EMD Millipore (Temecula, CA). All other

reagents and buffer components were acquired from Sigma-Aldrich (St. Louis, MO). Integrafrit columns (360 μm O.D. x 100 μm I.D.) and picofrit columns (360 μm O.D. x 75 μm I.D. x 30 μm emitter tip I.D.) were purchased from New Objective (Woburn, MA).

4.3.2 Sample Preparation for Middle-Down Analysis

Intact Intact Trastuzumab and an equimolar mixture of three purified anti-influenza IgGs previously discovered in post-vaccinated donor serum were resuspended in 50 mM Tris-HCl (pH 8) containing 2M urea and digested at 37°C with LysC for two hours using a 1:75 enzyme-to-substrate ratio. The resulting digests were diluted to a final urea concentration of 0.5 M in Tris-HCl and sequentially reduced in the presence of 5 mM DTT for 30 min at 37°C and alkylated with 25 mM IAM at room temperature, in the dark, for 45 min. Immediately following alkylation, IgG digests were acidified with 0.5% formic acid and desalted on a solid phase extraction (SPE) column containing C8 resin (Agilent) prior to LC-MS/MS analysis.

4.3.3 Liquid Chromatography and Mass Spectrometry

Proteolytic IgG peptides were separated by reversed phase chromatography on an Eksigent 2D Plus nanoLC system (Redwood, CA) configured for preconcentration using an Integrafrit trap column (3 cm) and a picofrit analytical column (20 cm) packed in-house with 5μm Michrom Magic C8 stationary phase (Auburn, CA). Eluent A was 0.1% aqueous formic acid and eluent B was 0.1% formic acid (v/v) in acetonitrile. Peptides were loaded onto the trap column and preconcentrated for 5 min in aqueous solvent containing 2% acetonitrile and 0.1% formic acid at a flow rate of 5 μL/min. For single IgG digests separation was carried out over 60 minutes using a linear gradient from 4-10% eluent B over the first 5 minutes and further increased to 40% eluent B over the last

55 min at 0.3 μL/min. For the mixture of IgGs discovered in donor serum, separation was carried out over 120 min to increase the probability of sampling diagnostic peptides derived from the CDR regions.

The LC system was coupled via a nano electrospray ionization source (Nanospray Flex ion source, Thermo Scientific) maintained at 1.8 kV to a Thermo Fisher Orbitrap Elite mass spectrometer (Bremen, Germany) outfitted with a 193 nm ArF excimer laser (Coherent ExciStar XS) to allow UVPD in the HCD cell as previously described.[39,43] For all experiments, MS^1 spectra were collected at a resolving power of 120K (at *m/z* 400) and MS^2 product ion spectra were acquired for the top five most abundant precursors by averaging three microscans at 120K resolution. The automatic gain control (AGC) target for MS^2 was set to 1E6 with a maximum injection time of 300 ms using a 5 *m/z* isolation width. To bias data acquisition towards larger peptides, ions 4+ and higher were selected for activation. Monoisotopic precursor selection and dynamic exclusion were enabled using a 30s exclusion window. For UVPD experiments, precursor ions were transferred to the HCD cell for activation and spectra were acquired using two laser pulses at 3 mJ per pulse. HCD was carried out using normalized collision energy (NCE) of 30% while ETD was performed using a reaction time of either 25 ms or 50 ms with a reagent AGC target of 5E5 and a maximum injection time of 100 ms.

4.3.4 Data Processing

RAW files were deconvolved using the Xtract algorithm (Thermo Fisher Scientific) embedded within ProSightPC 3.0 to obtain neutral monoisotopic mass information using a precursor ion signal-to-noise (S/N) threshold of 7 and a fragment ion S/N threshold of 3. All data was searched using ProSightPC 4.0 equipped with nine ion type (*a*, *a*+1, *b*, *c*, *x*, *x*+1, *y*, *y*-1, and *z*) search capabilities. Data from single IgG analyses

were searched against a custom database containing closely related therapeutic IgG sequences obtained from the DrugBank resource database (http://www.drugbank.ca/). UVPD data from the mixture of anti-influenza IgGs discovered in post-vaccinated donor serum was searched against a database consisting of 14,499 V_H sequences constructed from next generation sequencing of V_H genes from B cells of the same donor. All searches, with the exception of those used to identify glycoforms in trastuzumab, were performed in Absolute Mass mode using a 2.2 Da precursor mass window based on peptide monoisotopic mass and a 5 ppm product ion tolerance. Glycoform searches were conducted using a loose precursor window of 2000 Da and filtered based on known N-linked glycan masses.

4.4 RESULTS AND DISCUSSION

4.4.1 *In-silico* Proteolysis of a Theoretical Antibody Repertoire

The success of immunoproteomic methods for the evaluation of serological antibody repertoires hinges on the ability to sample and confidently identify diagnostic peptides arising from the antigen-binding domains of unique monoclonal antibodies. As demonstrated in the sequence alignment of multiple immunoglobulin heavy chain variable domains (V_H) shown in **Figure 4.2b**, the heavy chain complementarity determining region 3 (CDR-H3) exhibits the greatest diversity in both length and amino acid sequence of all hypervariable antigen-binding regions. As such, this region functions as the primary determinant of antigen recognition and provides a unique molecular signature by which to differentiate IgGs. To this end, we sought to strategically exploit highly conserved lysine residues in adjacent frameworks (**Figure 4.2**) to generate peptides containing the full-length CDR-H3. To evaluate the occurrence of lys-C

proteolytic sites (K) flanking this target region, V and J gene sequence logos, as shown in **Figure 4.2a**, were prepared from IMGT germline gene sequences (www.IMGT.org). Based on these data, lysine represents a strong candidate cleavage site for generating middle-down sized CDR-H3 peptides. The relatively high number of conserved lys-C cleavage sites that occur in the V_H sequence N-terminal to the region of interest also suggest that a considerable portion of the resulting peptide population will overlap with tryptic length distributions. While this is non-ideal from a purely middle-down method development perspective, it may offer an additional opportunity to strategically reduce sample complexity using a mass-biased partitioning approach, such as that previously reported by Cannon *et. al*.,[44] to improve the sampling efficiency of CDR-H3-containing peptides in MS-based immunoproteomic analyses.

Using the IgSimulator software,[45] a set of artificial V_H sequences were constructed to simulate the relative performance of trypsin and lys-C for generating suitable peptides for proteomic identification of the CDR-H3. Sequences were translated, filtered to remove any with stop codons or out-of-frame J regions, and digested *in silico* at each site C-terminal to either lysine (Lys-C) or arginine and lysine (trypsin). Other proteases do not consistently produce peptides spanning the CDR-H3 region and were therefore not considered. **Figure 4.3** shows the expected mass distribution plots for the resulting theoretical peptides produced using each of the two proteases (lys-C and trypsin) assuming no missed cleavage sites. The inset table shows counts of CDR-H3-containing peptides from both synthetic IgSimulator V_H data (used for the mass distribution plots) and empirical data derived from next-generation sequencing of V_H genes of peripheral B cells obtained from post-influenza vaccinated donor serum as described elsewhere.[42] The theoretical and empirical data demonstrate consistent shifts toward higher mass peptides possessing greater unique sequence information following lys-C digestion. Importantly,

more than 50% of the resulting CDR-H3-containing peptide population falls within an ideal middle-down size regime of 5-12 kDa, making this a compelling strategy for immunoproteomic analysis.

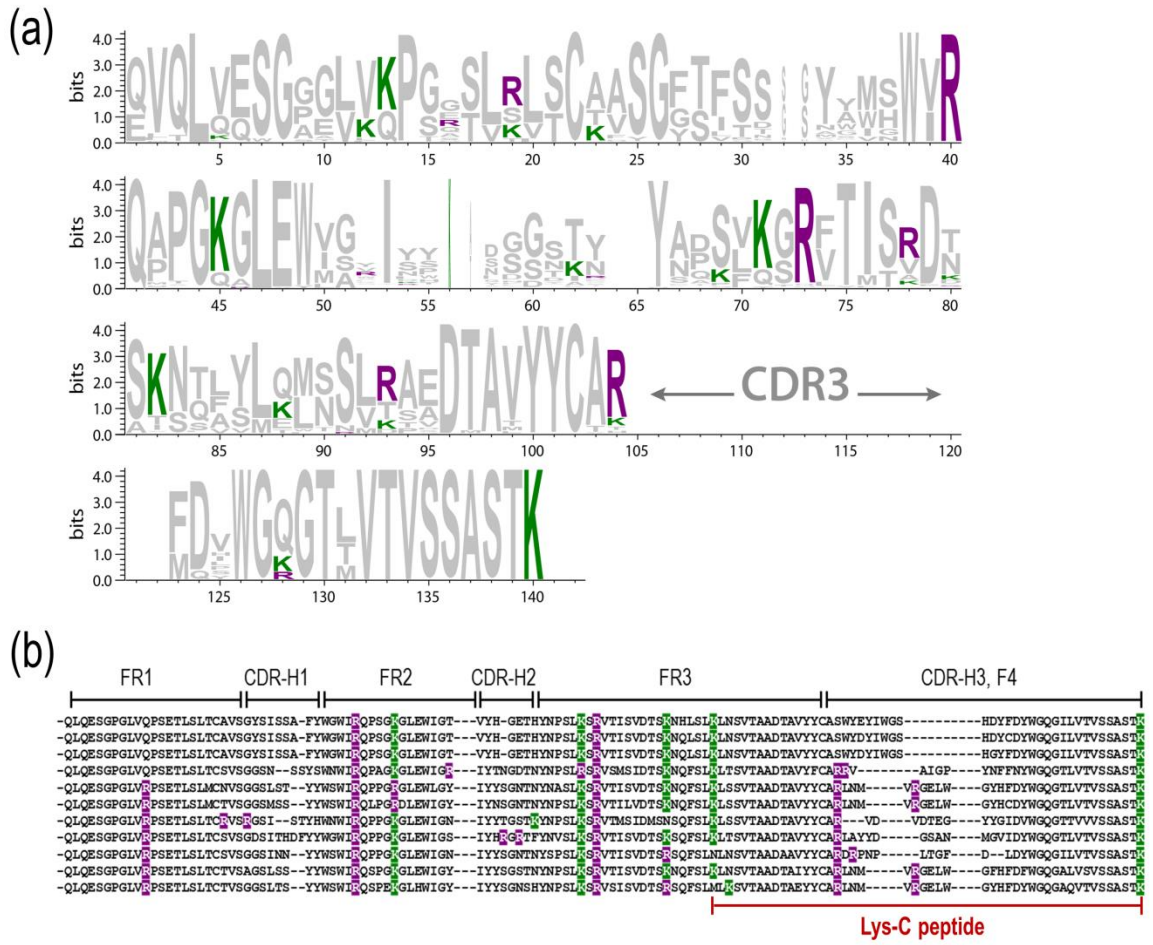
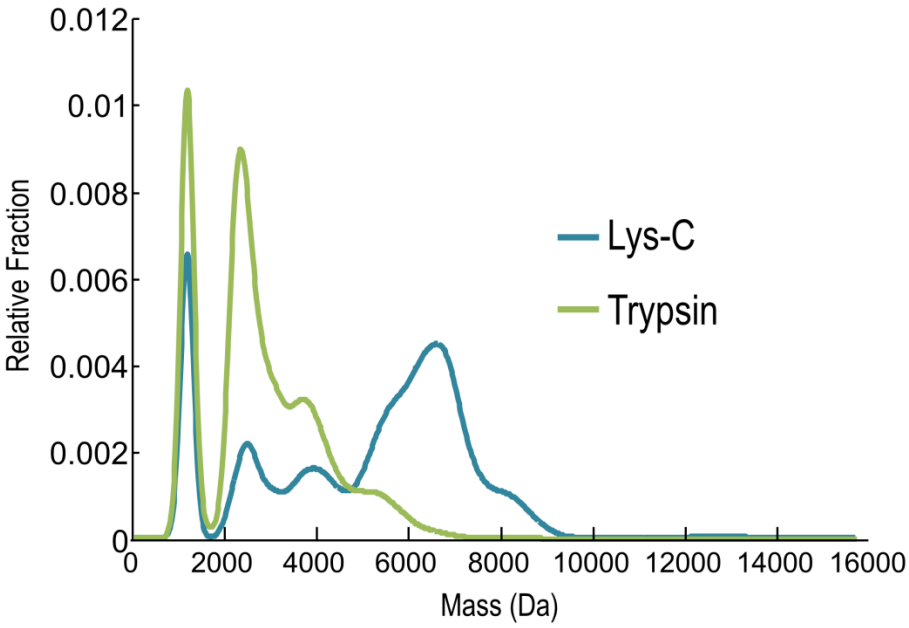


Figure 4.2 (a) V_H sequence logo constructed from IMGT germline V and J gene segments and (b) sequence alignment of V_H sequences from multiple unique antibodies. Hypervariable complementarity determining regions (CDRs) are embedded within conserved frameworks (FRs). Lysine (K) and arginine (R) residues are shown in green and purple, respectively. Based on highly conserved cleavage sites, lys-C often generates a full-length CDR-H3-containing peptide as indicated in red.



		CDR-H3-Containing Peptides		Peptide Count by Mass Range (Da)			
Sample	**Protease**	**Total**	**Unique**	**0 - 3k**	**3k - 5k**	**5k - 12k**	**>12k**
Igsim	Lys C	75318	61911	18747	12356	43838	377
	Trypsin	75318	41060	49120	21526	4672	0
flu_V11	Lys C	53000	43288	7868	17848	27028	256
	Trypsin	53000	22743	35864	16902	234	0

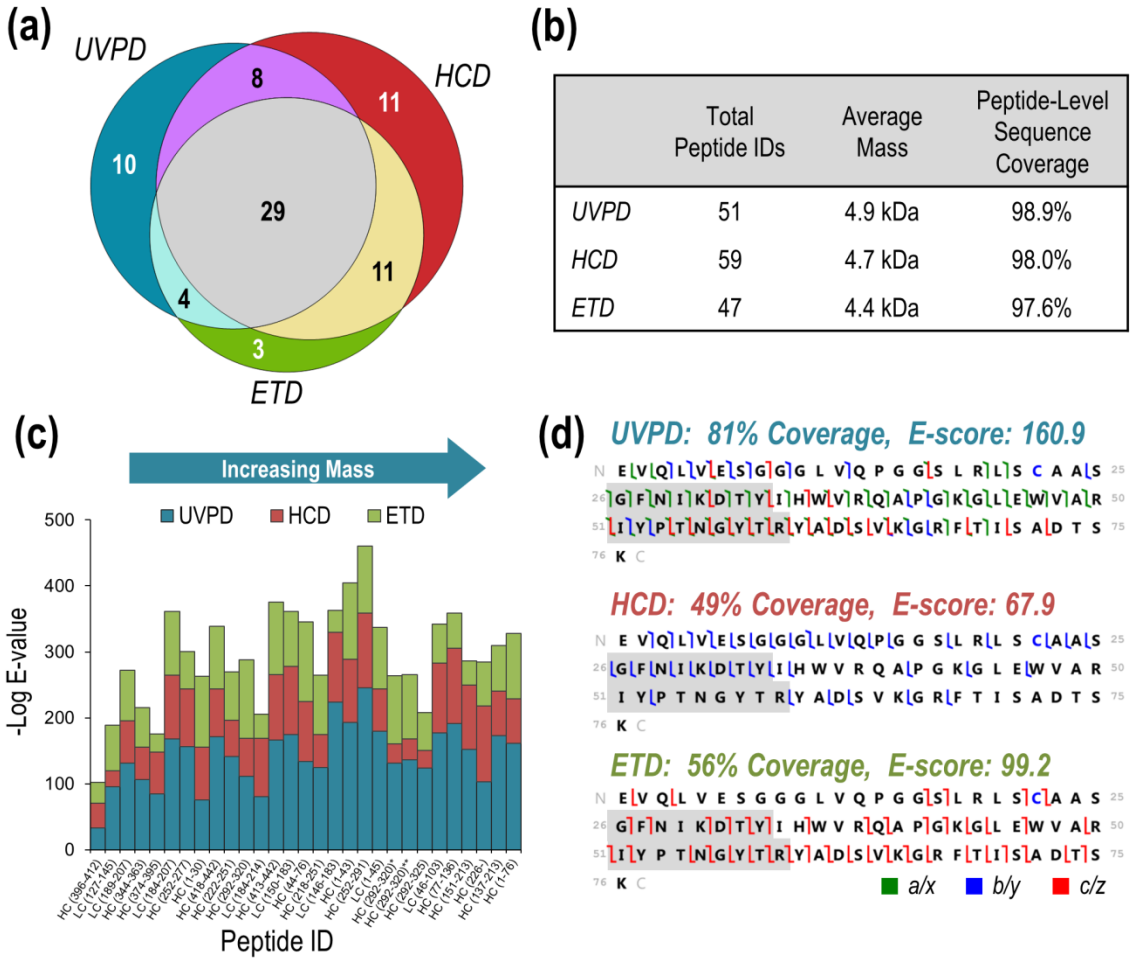
Figure 4.3 CDR-H3 peptide mass distribution plots for simulated V_H sequences generated by *in-silico* trypsin and Lys-C digestion. The inset table shows counts of theoretical CDR-H3-containing peptides following Lys-C or trypsin digestion of simulated (Igsim) V_H sequences and empirical next-generation sequenced V_H genes of peripheral B cells obtained from post-influenza vaccinated donor serum.

4.4.2 UVPD, HCD and ETD Analysis of Lys-C Digested Trastuzumab

Based on the *in silico* results described above, lys-C digestion addresses the first of two critical variables required for deconvolution of complex antibody mixtures: the sampling of peptides derived from the antigen-binding regions (CDRs). The second involves the unambiguous characterization of these diagnostic sequences based on informative MS/MS spectra. Recently, UVPD has proven to outperform other conventional activation methods (e.g. HCD, ETD) for the characterization of proteins and large proteolytic peptides on chromatographic timescales owing to its fast and high energy deposition, which gives rise to extensive and diverse fragmentation (*a*, *b*, *c*, *x*, *y*, and *z*-type ions) throughout the polypeptide backbone.[29,40] Thus, UVPD is a compelling option to address this latter variable for middle-down sized peptides produced upon lys-C proteolysis.

Trastuzumab was selected as a model antibody to benchmark our middle-down UVPD strategy owing to its well-defined sequence, which closely reflects the theoretical and empirical repertoire trends described above. Peptide masses were experimentally biased using restricted two-hour lys-C digestion at the intact (unreduced) level to promote the occurrence of missed cleavages (i.e., fewer exposed lysine residues resulting from extensive disulfide bonding) and increase the average length of resultant peptides. Reduction and alkylation were carried out immediately following digestion to eliminate undesirable crosslinks arising from the presence of intrinsic disulfide bonds and to ensure the formation of readily interpretable linear peptides. The resulting mixture was then chromatographically separated on a reversed phase C8 column and analyzed by UVPD, HCD and ETD on an Orbitrap Elite mass spectrometer operated in high resolution/high mass accuracy mode for both precursor and product ion detection. **Figure 4.4a,b** provides comparative summaries of the peptide populations confidently identified based

on separate UVPD, HCD and ETD datasets. All three methods performed comparably with respect to peptide-level mapping of trastuzumab, yielding ≥ 97% sequence coverage of the heavy and light chains (**Figure 4.5**). These results are consistent with the production of a predictable peptide mixture based on selective cleavage at lysines that, on average, falls within a suitable middle-down size regime (≥ 4.4 kDa, **Figure 4.4b**) and is amenable to high-throughput LC-MS/MS analysis. However, while all three activation methods returned similar peptide-level sequence coverages as a whole, considerable variability was observed in the peptide identifications upon which these coverages are based (**Figure 4.4a**). On average, only 56% of peptides identified from each individual run were confidently identified by the other two activation methods used, despite exceptional run-to-run chromatographic reproducibility. Further evaluation of these peptide populations revealed inconsistencies in the identifications of low abundance peptides that contained partial or complete sequence overlap with dominant peptide forms, but varied based on their number of missed cleavages. For example, UVPD and HCD identified three overlapping forms of light chain N-terminal peptides: LC(1-39), LC(1-45), and LC(1-103) (where "LC" represents the light chain and the numbers in parentheses indicate the amino acid positions) containing up to three internal lysine sites; whereas ETD only returned hits to the dominant peptide LC(1-45). Many of these variably truncated peptides (a negative byproduct of incomplete digestion) contained redundant information, as evidenced by the fact that approximately 90% sequence coverage of the heavy and light chain was obtained based exclusively on the 29 overlapping peptides identified by all three activation methods (**Figure 4.6**). While the majority of these overlapping peptides were present in high abundance (**Figure 4.7**), additional optimization of lys-C digestion parameters may help to limit sample dilution and complexity arising from undesirable peptide redundancy.



	Total Peptide IDs	Average Mass	Peptide-Level Sequence Coverage
UVPD	51	4.9 kDa	98.9%
HCD	59	4.7 kDa	98.0%
ETD	47	4.4 kDa	97.6%

Figure 4.4 Comparison of UVPD, HCD and ETD for middle-down analysis of lys-C digested trastuzumab in terms of (a) overlap in identified peptide populations and (b) average peptide mass and total peptide-level sequence coverage obtained by each activation method. The relative E-scores for the 29 overlapping peptides identified by all three activation methods are compared in (c) and ion cleavage maps obtained for an 8.2 kDa CDR-H1 and CDR-H2-containing peptide are shown in (d). The hypervariable regions are shaded in gray

UVPD:

>Trastuzumab Light Chain *(100% coverage)*
DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISS
LQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL
QSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

>Trastuzumab Heavy Chain *(98% coverage)*
EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISADTSKN
TAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFP
EPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPP
CPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSV
LTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWE
SNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

HCD:

>Trastuzumab Light Chain *(98% coverage)*
DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISS
LQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL
QSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

>Trastuzumab Heavy Chain *(98% coverage)*
EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISADTSKN
TAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFP
EPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPP
CPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSV
LTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWE
SNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

ETD

>Trastuzumab Light Chain *(98% coverage)*
DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISS
LQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL
QSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

>Trastuzumab Heavy Chain *(97% coverage)*
EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISADTSKN
TAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFP
EPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPP
CPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSV
LTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWE
SNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

Figure 4.5 Summary of peptide-level sequence coverage of the light and heavy chain of trastuzumab obtained by UVPD, HCD, and ETD, respectively.

UVPD, HCD and ETD Peptide Overlap:

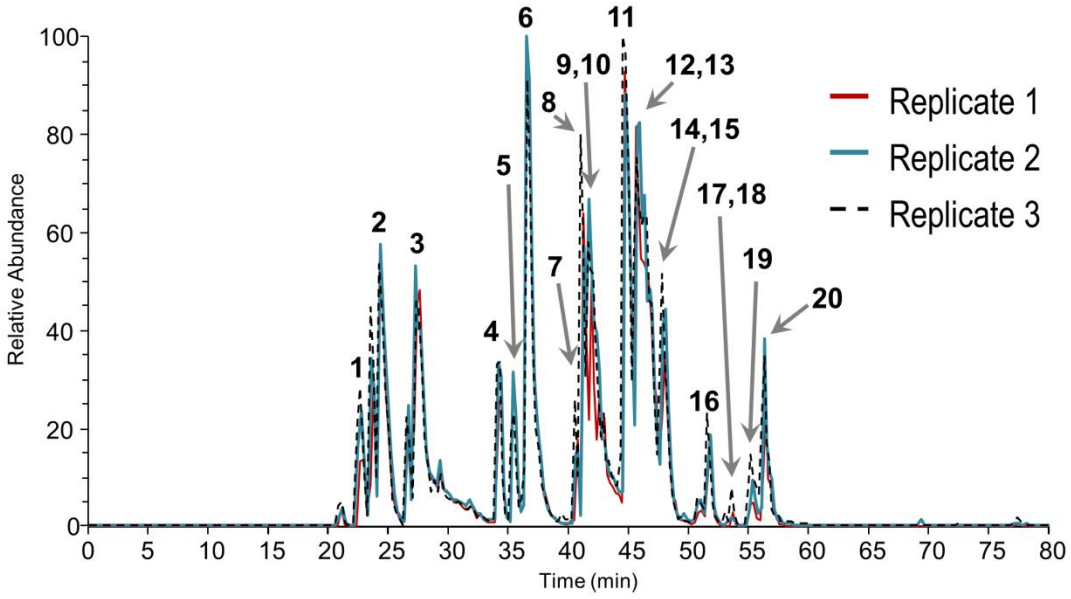
>Trastuzumab Light Chain *(89% Coverage)*

DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPS
RFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLT
LSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

>Trastuzumab Heavy Chain *(90% Coverage)*

EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRY
ADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSS
ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSS
GLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLGG
PSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYN
STYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREE
MTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRW
QQGNVFSCSVMHEALHNHYTQKSLSLSPGK

Figure 4.6 Summary of peptide-level sequence coverage of trastuzumab based on peptides identified by all three activation methods: UVPD, HCD and ETD.



#	Sequence	Mass (Da)	Charge State
1	HKVYACEVTHQGLSSPVTK	2140.07	4
2	ADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	3596.68	5
3	GQPREPQVYTLPPSREEMTK	2342.17	4
4	SRWQQGNVFSCSVMHEALHNHYTQK	3043.39	5
5	LTVDKSRWQQGNVFSCSVMHEALHNHYTQK	3599.72	5
6	VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK	4160.00	4
7	DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPK	4876.42	6
8	GLEWVARIYPTNGYTRYADSVK	2558.29	4
9	GLEWVARIYPTNGYTRYADSVKGRFTISADTSK	3721.89	5
10	DTLMISRTPEVTCVVVDVSHEDPEVK	2954.44	4
11	EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGK	4581.35	6
12	TKPREEQYN**STYRVVSVLTVLHQDWLNGK	5066.40	5
13	THTCPPCPAPELLGGPSVFLFPPKPK	2843.46	4
14	DTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK	4613.24	5
15	EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISADTSK	8285.23	7
16	GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK	7864.79	7
17	NTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSSASTKGPSVFPLAPSSK	6558.07	5
18	LLIYSASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTK	6475.11	5
19	RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAK	4208.17	5
20	STSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTK	8014.97	6

Figure 4.7 Base peak chromatograms for consecutive injections of restricted lys-C digested trastuzumab demonstrating high run-to-run reproducibility.

To better evaluate the performance of UVPD against HCD and ETD for the characterization lys-C generated trastuzumab peptides, the highest ranking spectral matches for each of the 29 peptides found by all three activation methods were compared

based on total number of matching fragment ions, percent coverage of inter-residue positions, and relative E-scores (**Table 4.1**). Note that the E-score represents the inverse log of the expectation value (or E-value), which is a confidence metric that describes the probability at which a set of fragment ions match to a given sequence by random chance alone. As demonstrated in **Figure 4.8**, UVPD consistently gives rise to a greater number of matched fragment ions per peptide (**Figure 4.8a**), an outcome attributed to the diverse array of fragmentation channels accessible to this high energy activation process. To more accurately reflect the gain in sequence-informative fragmentation obtained for each peptide upon UVPD, the total number of matched fragment ions was normalized to peptide length (**Figure 4.8b**). Based on these data, fragmentation per residue is enhanced by an average ratio of 2.3:1:1 for UVPD relative to HCD and ETD alike, allowing more extensive and uniform coverage of the peptide backbone that approaches single amino acid resolution (**Figure 4.8b**), as well as results in higher confidence identifications across the entire range of lys-C generated peptides. This latter point is illustrated in **Figure 4.4c**, which compares the relative peptide E-scores derived from the UVPD, HCD and ETD spectra for the 29 peptides. The relative size of the bars associated with each activation method scale with the confidence of the identification. With few exceptions, UVPD dramatically outperformed HCD and ETD with respect to peptide confidence across the entire mass distribution, resulting in an average E-score that roughly doubled that of the more conventional methods (i.e., 143 versus 74 and 77 for HCD and ETD, respectively).

Mass (Da)	Chain (residues)	Sequence	E-score			Inter-residue Coverage (%)			# Matching Fragment Ions		
			ETD	HCD	UVPD	ETD	HCD	UVPD	ETD	HCD	UVPD
1872.91	HC (396-412)	TTPPVLDSDGSFFLYSK	31.9	37.8	32.9	75	88	94	14	16	36
2125.06	LC (127-145)	SGTASVVCLLNNFYPREAK	68.9	24.0	95.8	94	78	100	28	15	48
2140.07	LC (189-207)	HKVYACEVTHQGLSSPVTK	76.7	64.3	131.2	94	83	100	30	28	62
2342.17	HC (344-363)	GQPREPQVYTLPPSREEMTK	60.4	49.3	105.9	79	74	89	24	20	60
2543.12	HC (374-395)	GFYPSDIAVEWESNGQPENNYK	27.6	63.5	84.4	57	81	95	14	27	52
2746.34	LC (184-207)	ADYEKHKVYACEVTHQGLSSPVTK	95.9	96.6	168.4	96	100	100	39	41	82
2954.44	HC (252-277)	DTLMISRTPEVTCVVVDVSHEDPEVK	57.1	86.9	156.6	76	96	100	27	37	83
3029.56	HC (1-30)	EVQLVESGGGLVQPGGSLRLSCAASGFNIK	107.6	79.8	75.8	83	79	90	46	31	50
3043.39	HC (418-442)	SRWQQGNVFSCSVMHEALHNHYTQK	94.6	73.0	171.1	96	83	96	39	31	87
3333.63	HC (222-251)	SCDKTHTCPPCPAPELLGGPSVFLFPPKPK	73.9	55.1	140.9	69	72	97	35	30	77
3459.79	HC (292-320)	TKPREEQYNSTYRVVSVLTVLHQDWLNGK	119.6	57.6	111.0	96	79	93	50	33	73
3596.68	LC (184-214)	ADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	36.4	88.5	80.3	63	93	77	27	48	52
3599.72	HC (413-442)	LTVDKSRWQQGNVFSCSVMHEALHNHYTQK	109.6	99.4	166.1	90	100	97	49	51	89
3618.70	LC (150-183)	VDNALQSGNSQESVTEQDSKDSTYSLSSTLTL SK	83.5	103.0	174.7	67	94	97	37	54	110
3721.89	HC (44-76)	GLEWVARIYPTNGYTRYADSVKGRFTISADTSK	120.4	90.9	134.0	91	88	97	52	43	77
3786.89	HC (218-251)	VEPKSCDKTHTCPPCPAPELLGGPSVFLFPPK PK	89.9	50.1	124.6	67	58	88	38	29	70
4160.00	LC (146-183)	VQWKVDNALQSGNSQESVTEQDSKDSTYSLS STLTLSK	33.8	105.5	223.7	32	89	100	15	44	119
4581.35	HC (1-43)	EVQLVESGGGLVQPGGSLRLSCAASGFNIKDT YIHWVRQAPGK	115.9	95.2	193.3	79	83	93	53	46	110
4613.22	HC (252-291)	DTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD GVEVHNAK	102.0	112.7	245.7	72	95	100	48	62	137
4876.42	LC (1-45)	DIQMTQSPSSLSASVGDRVTITCRASQDVNTAV AWYQQKPGKAPK	93.2	64.2	179.5	82	66	89	48	35	106
4904.32	HC (292-320)	TKPREEQYN*STYRVVSVLTVLHQDWLNGK	103.3	29.7	131.0	96	61	89	46	17	83
5066.38	HC (292-320)	TKPREEQYN**STYRVVSVLTVLHQDWLNGK	97.3	31.6	136.7	93	68	89	44	19	83
5774.70	HC (292-325)	TKPREEQYN**STYRVVSVLTVLHQDWLNGKEY KCK	57.9	26.5	123.9	61	61	91	28	20	79
6475.11	LC (46-103)	LLIYSASFLYSGVPSRFSGSRSGTDFTLTISSLQ PEDFATYYCQQHYTTPPTFGQGTK	59.1	105.1	177.4	35	65	89	29	53	113
6558.07	HC (77-136)	NTAYLQMNSLRAEDTAVYYCSRWGGDGFYAM DYWGQGTLVTVSSASTKGPSVFPLAPSSK	53.0	114.2	191.2	29	81	93	26	51	125
6712.31	HC (151-213)	DYFPEPVTVSWNSGALTSGVHTFPAVLQSSGL YSLSSVVTVPSSSLGTQTYICNVNHKPSNTK	37.4	97.1	152.2	27	66	82	21	50	103
7438.66	HC (226-291)	THTCPPCPAPELLGGPSVFLFPPKPKDTLMISR TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN AK	66.4	114.9	103.2	52	77	85	39	62	103
8014.97	HC (137-213)	STSGGTAALGCLVKDYFPEPVTVSWNSGALTS GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTY ICNVNHKPSNTK	68.4	67.6	173.1	49	41	74	40	31	125
8285.23	HC (1-76)	EVQLVESGGGLVQPGGSLRLSCAASGFNIKDT YIHWVRQAPGKGLEWVARIYPTNGYTRYADSV KGRFTISADTSK	99.2	68.0	161.0	56	49	81	56	39	115

*G0F glycoform (+1444.5339 Da)
**G1F glycoform (+1606.5867 Da)

Table 4.1 Comparison of performance metrics for 29 lys-C generated trastuzumab peptides identified by UVPD, HCD and ETD. Peptides are ordered by increasing mass and compared in terms of E-value/E-score, percent inter-residue coverage, and total number of matching fragment ions.

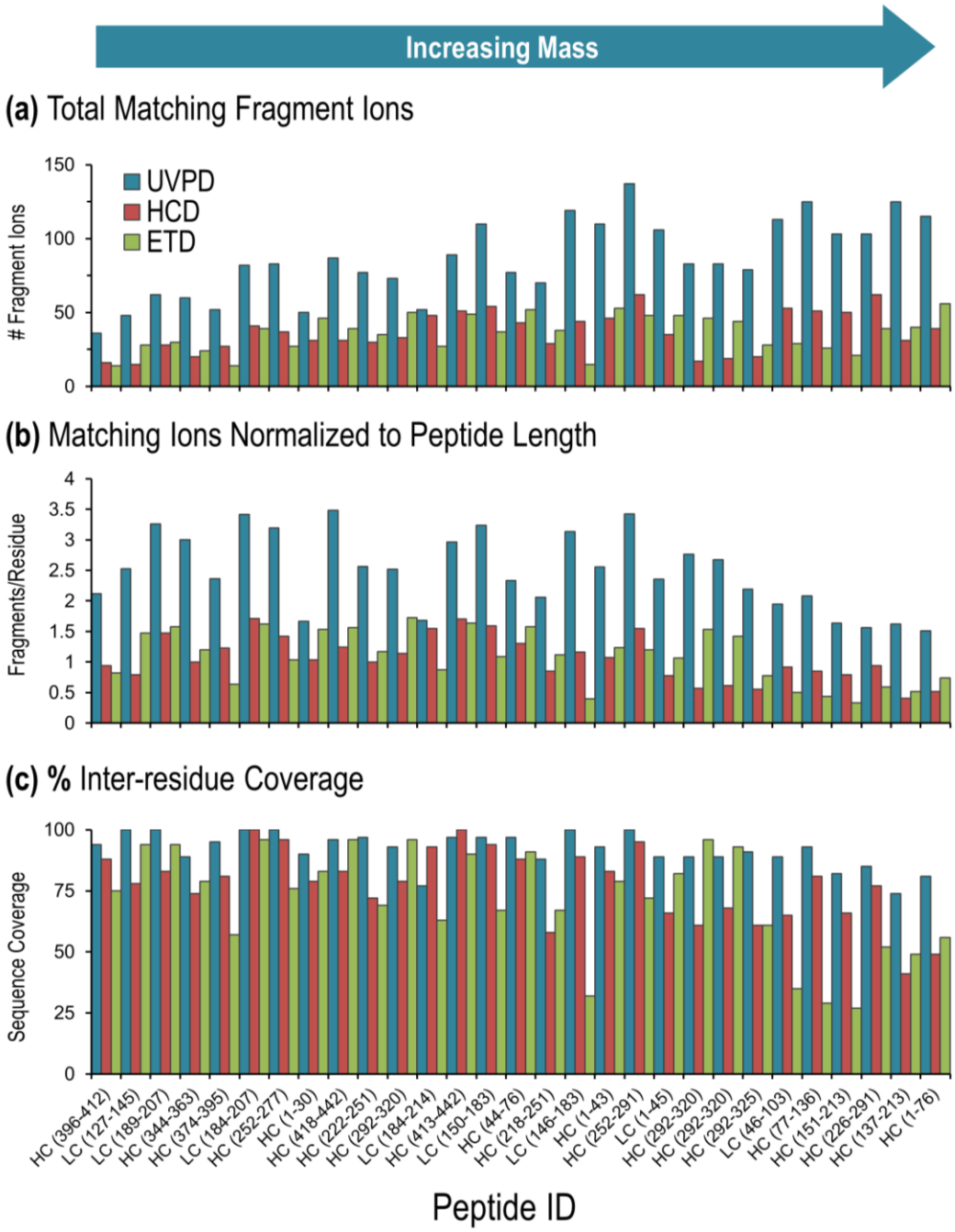


Figure 4.8 Histograms comparing the (a) total number of matching fragment ions, (b) number of fragment ions normalized to peptide length, and (c) percent coverage of inter-residue sites for lys-C generated trastuzumab peptides identified by UVPD (2 pulse, 3 mJ), HCD (30% NCE), and ETD (25 ms and 50 ms).

While the ability to obtain high confidence peptide identifications on the whole is compelling, a better assessment of the analytical utility of the middle-down UVPD strategy for high-throughput antibody identification is based on its capacity for comprehensive characterization of hypervariable CDR sequences that are critical for the differentiation of unique antibodies. Of the 29 peptides identified by all three activation methods, four peptides exceeding 4.8 kDa in mass were found to encompass all hypervariable regions of the heavy (CDR-H1, CDR-H2, and CDR-H3) and light (CDR-L1, CDR-L2, and CDR-L3) chains of trastuzumab. **Figure 4.4d** shows a comparison of the resulting fragment ion maps for one 8.2 kDa peptide, HC(1-76), encompassing both the heavy chain CDR-H1 and the CDR-H2 (both shown in gray) based on high-throughput analysis by UVPD, HCD, and ETD, respectively. As demonstrated in both the sequence maps (**Figure 4.4d**) and corresponding mass spectra (**Figure 4.9**), UVPD yielded extensive and pairwise fragmentation across the peptide backbone, giving rising to a rich array of product ion types (*a*, *b*, *c*, *x*, *y*, and *z*) that enabled the unambiguous characterization of both hypervariable regions. Alternatively, both HCD and ETD showed more limited fragmentation of the peptide backbone and incomplete inter-residue coverage of the CDR regions, resulting in lower confidence identification relative to UVPD (**Figure 4.4d**).

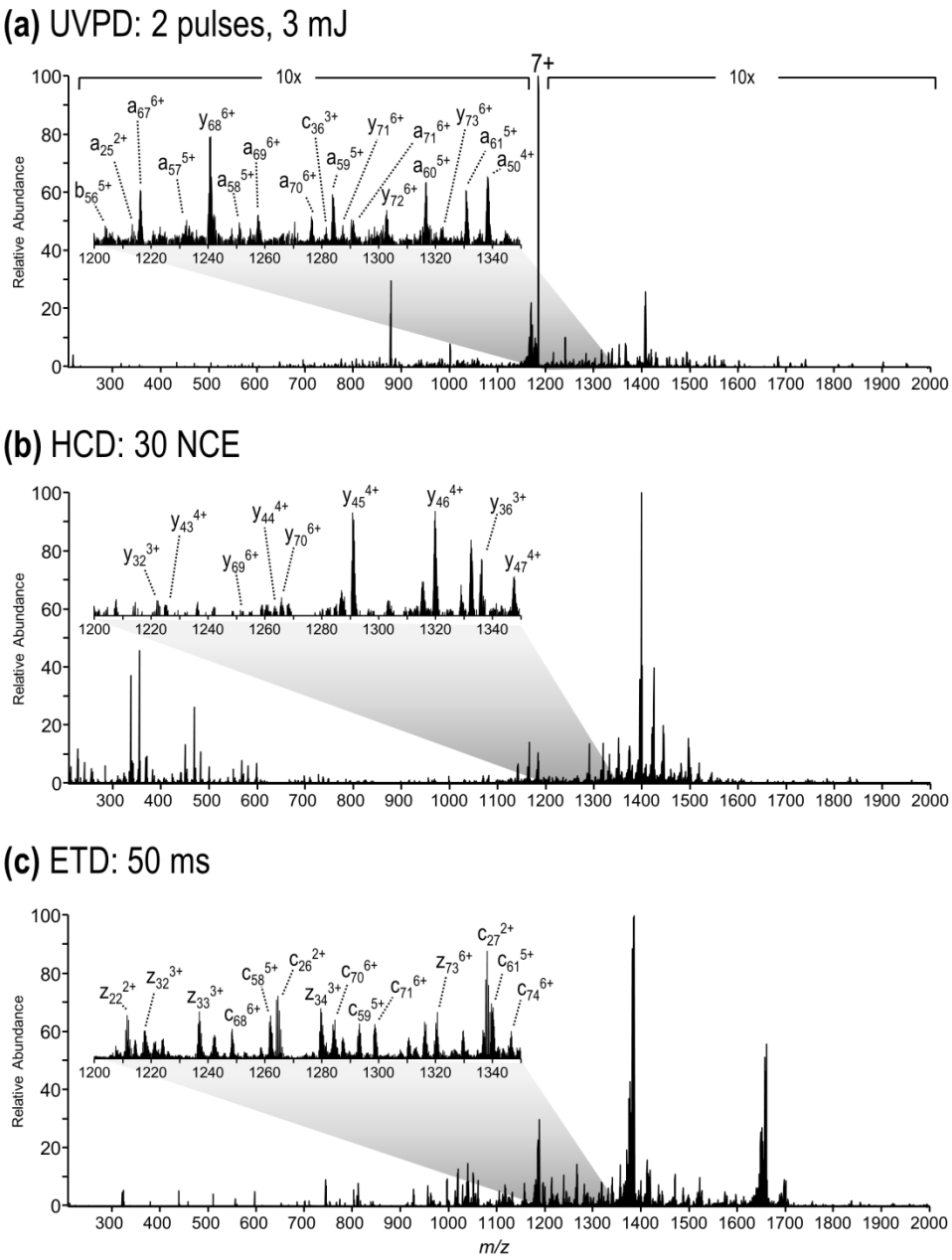


Figure 4.9 Comparative MS/MS spectra (120K at *m/z* 400) for the N-terminally derived CDR-H1 and CDR-H2-containing lys-C peptide from trastuzumab, HC(1-76), activated by (a) UVPD using 2 pulses at 3 mJ, (b) HCD using an NCE of 30%, and (c) ETD at 50 ms reaction time. (M_r of peptide = 8285.23 Da, 7+ charge state). The insets demonstrate the product ion diversity observed for each activation strategy. The sequences, inter-residue coverages, and E-scores are shown in Figure 2d.

As previously stated, the CDR-H3 is commonly used in MS-based immunoproteomics as a unique molecular signature by which to differentiate antibody clonotypes due to its nature as the most diverse hypervariable region and primary determinant of antigen recognition.[10] Ultimately, the diagnostic value of the CDR-H3 and the occurrence of highly conserved lysine sites flanking this region prompted our use of restricted lys-C digestion to generate peptides that encompassed this entire hypervariable region, making it a particularly important target by which to evaluate our middle-down strategy. **Figure 4.10** demonstrates the successful implementation of strategic lys-C proteolysis to generate an ideal middle-down sized CDR-H3-containing peptide (6.5 kDa) that was readily identified using all three activation methods compared herein. Similar to the results obtained for the CDR-H1/CDR-H2-containing peptide described above, UVPD resulted in complete and pairwise coverage of the diagnostic CDR-H3 (**Figure 4.10a**). HCD performed comparably to UVPD with respect to coverage throughout the CDR-H3 (**Figure 4.10b**), but yielded lower overall coverage of the peptide backbone (**Figure 4.10e**). ETD performed poorly for the characterization of this peptide overall; however, it still provided reasonable coverage of the CDR-H3 region (**Figure 4.10c**). Again, the extensive fragmentation afforded by UVPD allowed improved characterization and more confident identification of this diagnostic peptide relative to more conventional activation methods as evidenced by comparisons of the number of matched ions, the percentage of inter-residue coverage, and the peptide E-scores, all shown in **Figure 4.10d-f**.

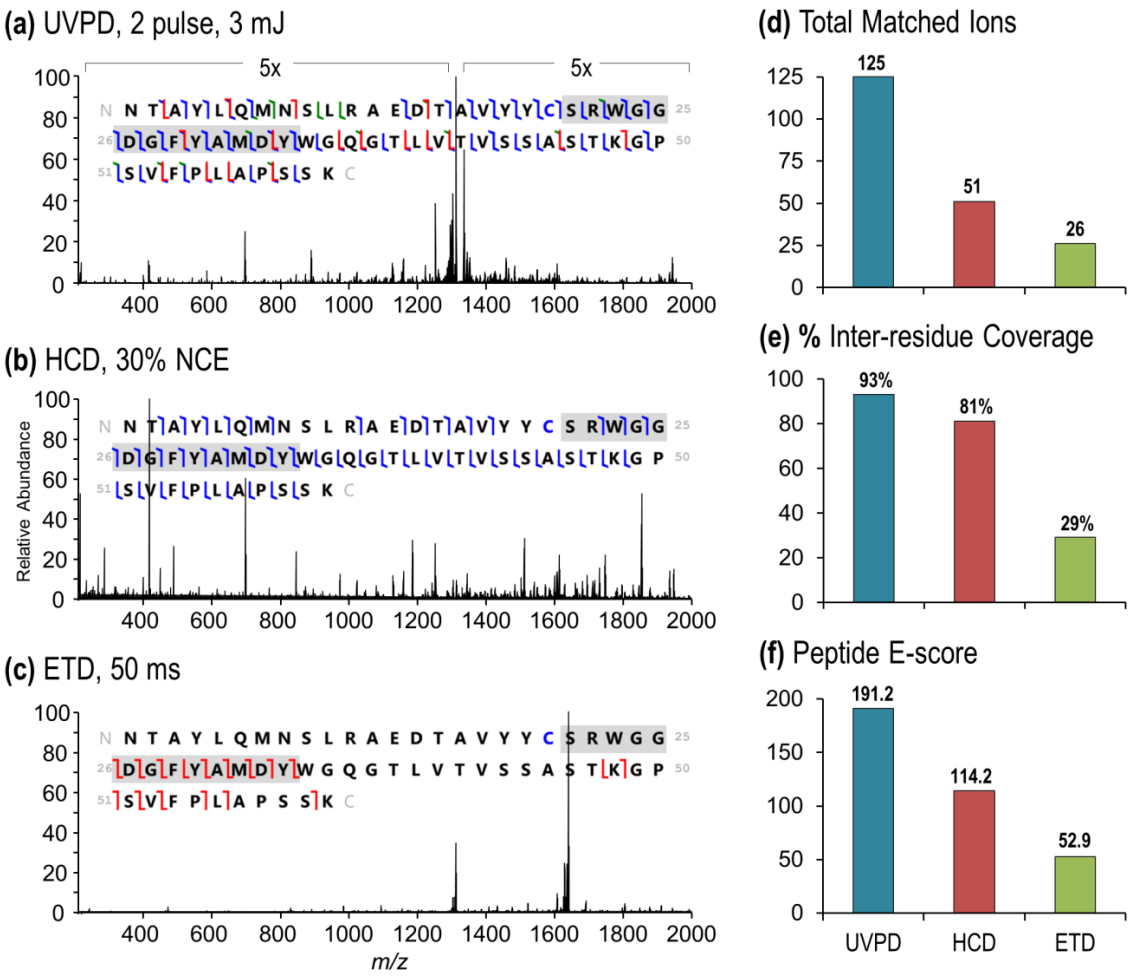
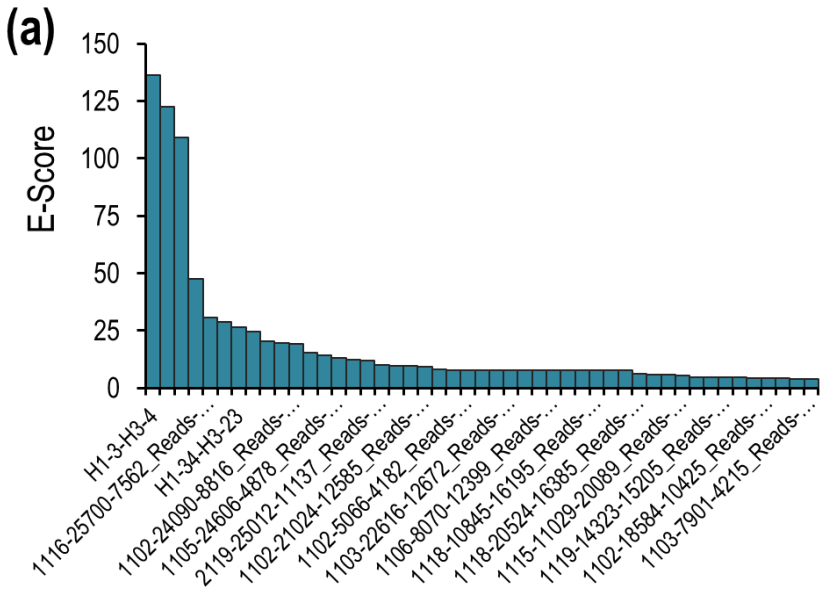


Figure 4.10 MS/MS spectra (120K at m/z 400) and associated fragment ion maps for a full-length CDR-H3-containing Lys-C peptide (one missed cleavage) from trastuzumab activated by (a) UVPD using 2 pulses at 3 mJ, (b) HCD using an NCE of 30%, and (c) ETD at 50 ms reaction time. (M_r of peptide = 6558.07 Da, 5+ charge state). Shown for each of the activation methods are (d) the total number of matched fragment ions, (e) the percent coverage of possible inter-residue cleavage sites, and (f) the peptide confidence reported in terms of the relative E-score..

4.4.3 Middle-Down 193 nm UVPD of Anti-Influenza Antibody Mixture

Owing to the superior performance of UVPD over more conventional HCD and ETD for the characterization of large CDR-containing peptides, as demonstrated in initial benchmarking experiments with trastuzumab, we next sought to evaluate the potential for integration of the middle-down UVPD strategy into an immunoproteomics workflow. For these proof-of-principle experiments a mixture of three anti-influenza monoclonal antibodies, previously identified in the serum repertoire of a vaccinated donor,[42] was analyzed. In the interest of evaluating the discriminatory power of this method in an unbiased manner, no sequence information was provided for the antibody mixture. The LC-MS/UVPD data was then searched against a V_H-gene database constructed from approximately 15,000 heavy chain sequences previously obtained by next generation sequencing of V_H genes from peripheral B cells of the donor from which the antibody mixture was derived.[42] The resulting peptide hits were then ranked by E-score (highest to lowest) using an applied E-value threshold of 1E-4, and filtered to remove multiple hits to the same V_H sequence entry. This process ensures that only the most confident peptide hit for each unique V_H sequence entry is reported. Hits to 47 unique V_H sequences remained using the filtering criteria employed as summarized in the E-score distribution plot shown in **Figure 4.11**. Based on these data, hits to three unique V_H entries exhibited markedly higher E-score values than the other 44 hits (**Figure 4.11)**. In all cases these hits correspond to peptides containing complete or partial CDR-H3 sequences. A representative example is provided in **Figure 4.12** which shows the UVPD spectrum and resulting fragmentation map for the highest scoring peptide match to a full-length CDR-H3 peptide. The extensive coverage afforded by UVPD resulted in 100% coverage of the diagnostic hypervariable sequence allowing for unambiguous antibody identification.

.



(b) Top 5 Hits to Unique Ig V_H Sequences

Protein Description	E-score	# Matching Fragments	PPM Error
flu rep_V2-Full-VH_H1-3-H3-4	136.1	89	3.37
flu rep_V2-Full-VH_H1-2-S	122.6	101	1.45
flu rep_V2-Full-VH_Bris-4-Tex-2	109.2	76	1.56
flu rep_FluZone_V2-Full-VH_1116-25700-7562	47.4	27	1.73
flu rep_FluZone_V2-Full-VH_2119-10115-10218	30.6	29	1894.48

Figure 4.11 (a) Comparison of E-score distributions for top ranking peptide hits to unique Ig V_H sequences based on LC-MS/UVPD analysis of a Lys-C digested anti-influenza three IgG mixture searched against a donor V_H gene sequence database, and (b) a summary of ProSight PC search results for the top five highest ranking hits to unique Ig V_H sequences.

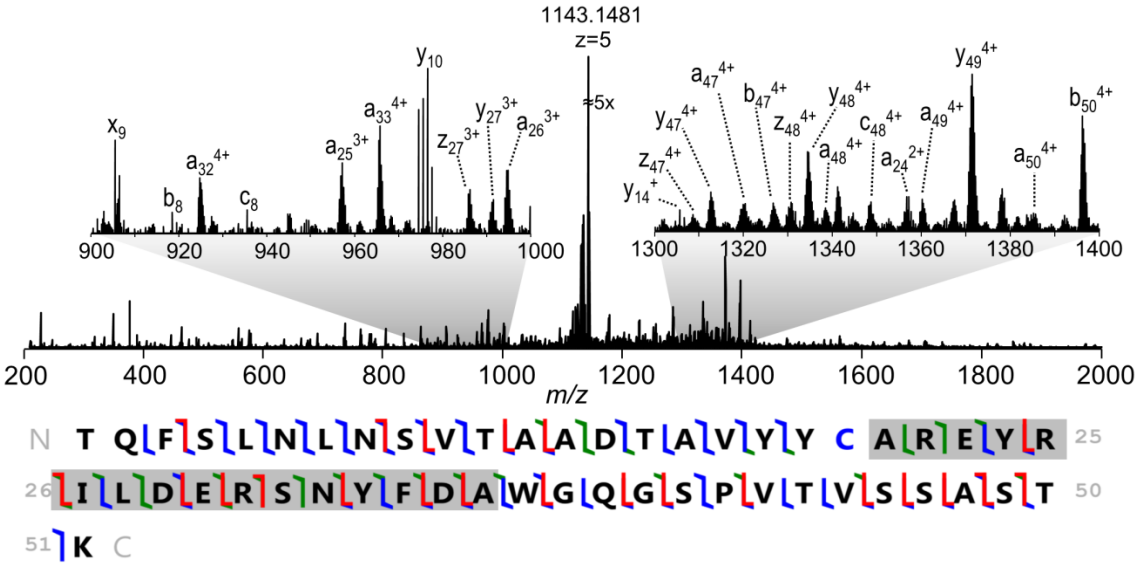


Figure 4.12 UVPD spectrum and resulting fragmentation map for an unambiguously identified full-length CDR-H3 peptide (M_r = 5708.71 Da, 5+ charge state, Ig V_H = H1-2-S) from an anti-influenza monoclonal IgG mixture searched against a donor V_H sequence database. The hypervariable region is shaded in gray.

4.5 CONCLUSIONS

The middle-down workflow presented herein exploits highly conserved lysine residues flanking the diagnostic CDR-H3 to generate ideal middle-down sized peptides (3-10 kDa) that can be used to uniquely identify antibodies in mixtures. The extensive fragmentation afforded by 193 nm UVPD allows comprehensive sequencing of large CDR-H3-containing peptides for unambiguous IgG identifications. Moreover, we demonstrated the potential utility of integrating middle-down UVPD with Ig-seq V_H database searching for enhanced antibody serum proteomics.

4.6 REFERENCES

(1) Batista, F. D.; Harwood, N. E. *Nat. Rev. Immunol.* **2009**, *9* (1), 15–27.
(2) Nutt, S. L.; Hodgkin, P. D.; Tarlinton, D. M.; Corcoran, L. M. *Nat. Rev. Immunol.* **2015**, *15* (3), 160–171.
(3) Murphy, K. *Janeway's Immunobiology, Eighth Edition*; Garland Science, 2011.
(4) Schroeder, H. W.; Cavacini, L. *J. Allergy Clin. Immunol.* **2010**, *125* (2), S41–S52.
(5) Tonegawa, S. *Nature* **1983**, *302* (5909), 575–581.
(6) Briney, B. S.; Crowe, J. E. J. *B Cell Biol.* **2013**, *4*, 42.
(7) Goodnow, C. C.; Vinuesa, C. G.; Randall, K. L.; Mackay, F.; Brink, R. *Nat. Immunol.* **2010**, *11* (8), 681–688.
(8) Gitlin, A. D.; Shulman, Z.; Nussenzweig, M. C. *Nature* **2014**, *509* (7502), 637–640.
(9) Wilson, P. C.; Andrews, S. F. *Nat. Rev. Immunol.* **2012**, *12* (10), 709–719.
(10) Georgiou, G.; Ippolito, G. C.; Beausang, J.; Busse, C. E.; Wardemann, H.; Quake, S. R. *Nat. Biotechnol.* **2014**, *32* (2), 158–168.
(11) Robinson, W. H. *Nat. Rev. Rheumatol.* **2015**, *11* (3), 171–182.
(12) Lavinder, J. J.; Wine, Y.; Giesecke, C.; Ippolito, G. C.; Horton, A. P.; Lungu, O. I.; Hoi, K. H.; DeKosky, B. J.; Murrin, E. M.; Wirth, M. M.; Ellington, A. D.; Dörner, T.; Marcotte, E. M.; Boutz, D. R.; Georgiou, G. *Proc. Natl. Acad. Sci.* **2014**, *111* (6), 2259–2264.
(13) de Costa, D.; Broodman, I.; VanDuijn, M. M.; Stingl, C.; Dekker, L. J. M.; Burgers, P. C.; Hoogsteden, H. C.; Sillevis Smitt, P. A. E.; van Klaveren, R. J.; Luider, T. M. *J. Proteome Res.* **2010**, *9* (6), 2937–2945.
(14) VanDuijn, M. M.; Dekker, L. J. M.; Zeneyedpour, L.; Smitt, P. A. E. S.; Luider, T. M. *J. Biol. Chem.* **2010**, *285* (38), 29247–29253.
(15) Dekker, L. J. M.; Zeneyedpour, L.; Brouwer, E.; Duijn, M. M. van; Smitt, P. A. E. S.; Luider, T. M. *Anal. Bioanal. Chem.* **2011**, *399* (3), 1081–1091.
(16) Maat, P.; VanDuijn, M.; Brouwer, E.; Dekker, L.; Zeneyedpour, L.; Luider, T.; Smitt, P. S. *J. Autoimmun.* **2012**, *38* (4), 354–360.
(17) Cheung, W. C.; Beausoleil, S. A.; Zhang, X.; Sato, S.; Schieferl, S. M.; Wieler, J. S.; Beaudet, J. G.; Ramenani, R. K.; Popova, L.; Comb, M. J.; Rush, J.; Polakiewicz, R. D. *Nat Biotech* **2012**, *30* (5), 447–452.
(18) Wine, Y.; Boutz, D. R.; Lavinder, J. J.; Miklos, A. E.; Hughes, R. A.; Hoi, K. H.; Jung, S. T.; Horton, A. P.; Murrin, E. M.; Ellington, A. D.; Marcotte, E. M.; Georgiou, G. *Proc. Natl. Acad. Sci.* **2013**, *110* (8), 2993–2998.
(19) Boutz, D. R.; Horton, A. P.; Wine, Y.; Lavinder, J. J.; Georgiou, G.; Marcotte, E. M. *Anal. Chem.* **2014**, *86* (10), 4758–4766.
(20) Ogishi, M.; Yotsuyanagi, H.; Moriya, K.; Koike, K. *Sci. Rep.* **2016**, *6*, 29532.
(21) Lavinder, J. J.; Horton, A. P.; Georgiou, G.; Ippolito, G. C. *Curr. Opin. Chem. Biol.* **2015**, *24*, 112–120.
(22) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R. *Chem. Rev.* **2013**, *113* (4), 2343–2394.

(23) Fang, X.; Zhang, W.-W. *J. Proteomics* **2008**, *71* (3), 284–303.
(24) Cannon, J.; Lohnes, K.; Wynne, C.; Wang, Y.; Edwards, N.; Fenselau, C. *J. Proteome Res.* **2010**, *9* (8), 3886–3890.
(25) 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. *Nat. Methods* **2012**, *9* (8), 822–824.
(26) Tsiatsiani, L.; Heck, A. J. R. *FEBS J.* **2015**, *282* (14), 2612–2626.
(27) Kleemann, G. R.; Beierle, J.; Nichols, A. C.; Dillon, T. M.; Pipes, G. D.; Bondarenko, P. V. *Anal. Chem.* **2008**, *80* (6), 2001–2009.
(28) Fornelli, L.; Ayoub, D.; Aizikov, K.; Beck, A.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (6), 3005–3012.
(29) Cotham, V. C.; Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (7), 4004–4013.
(30) Srzentić, K.; Fornelli, L.; Laskay, Ü. A.; Monod, M.; Beck, A.; Ayoub, D.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (19), 9945–9953.
(31) Pang, Y.; Wang, W.-H.; Reid, G. E.; Hunt, D. F.; Bruening, M. L. *Anal. Chem.* **2015**, *87* (21), 10942–10949.
(32) Reilly, J. P. *Mass Spectrom Rev* **2009**, *28* (3), 425–447.
(33) Ly, T.; Julian, R. R. *Angew. Chem. Int. Ed.* **2009**, *48* (39), 7130–7137.
(34) Brodbelt, J. S. *Chem. Soc. Rev.* **2014**, *43* (8), 2757–2783.
(35) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.
(36) Moon, J. H.; Yoon, S. H.; Kim, M. S. *Rapid Commun. Mass Spectrom.* **2005**, *19* (22), 3248–3252.
(37) Girod, M.; Sanader, Z.; Vojkovic, M.; Antoine, R.; MacAleese, L.; Lemoine, J.; Bonacic-Koutecky, V.; Dugourd, P. *J. Am. Soc. Mass Spectrom.* **2014**, *26* (3), 432–443.
(38) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. *J. Proteome Res.* **2010**, *9* (8), 4205–4214.
(39) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(40) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(41) Shaw, J. B.; Robinson, E. W.; Paša-Tolić, L. *Anal. Chem.* **2016**, *88* (6), 3019–3023.
(42) Lee, J.; Boutz, D. R.; Joyce, M. G.; Vollmers, C.; Chromikova, V.; Leung, K.; Horton, A. P.; DeKosky, B. J.; Lee, C. H.; Lavinder, J. J.; Georgiou, G. *Nat. Med. (publication pending)*.
(43) Vasicek, L. A.; Ledvina, A. R.; Shaw, J.; Griep-Raming, J.; Westphall, M. S.; Coon, J. J.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 1105–1108.
(44) Cannon, J. R.; Edwards, N. J.; Fenselau, C. *J. Mass Spectrom.* **2013**, *48* (3), 340–343.
(45) Safonova, Y.; Lapidus, A.; Lill, J. *Bioinformatics* **2015**, *31* (19), 3213–3215.

Chapter 5

Characterization of Therapeutic Monoclonal Antibodies at the Subunit-Level using Middle-Down 193 nm Ultraviolet Photodissociation*

5.1 Overview

Monoclonal antibodies (mAbs) are a rapidly advancing class of therapeutic glycoproteins that possess wide clinical utility owing to their biocompatibility, high antigen specificity, and targeted immune stimulation. These therapeutic properties depend greatly on the composition of the immunoglobulin G (IgG) structure, both in terms of primary sequence and post-translational modifications (PTMs); however, large-scale production in cell culture often results in heterogeneous mixtures that can profoundly affect clinical safety and efficacy. This places a high demand on analytical methods that afford comprehensive structural characterization of mAbs to ensure their stringent quality control. Here we report the use of targeted middle-down 193 nm ultraviolet photodissociation (UVPD) to provide detailed primary sequence analysis and PTM site localization of therapeutic monoclonal antibody subunits (~25 kDa) generated upon digestion with recombinant immunoglobulin G-degrading enzyme of *Streptococcus pyogenes* (IdeS) followed by chemical reduction. Under optimal conditions, targeted UVPD resulted in approximately 60% overall coverage of the IgG sequence, in addition to unambiguous glycosylation site localization and extensive coverage of the antigen-binding complementarity determining regions (CDRs) in a single LC-MS/MS experiment. Moreover, we exploited the tunable energy deposition afforded by UVPD, as well as the complementary nature of UVPD and ETD to obtain deeper sequencing and

*Cotham, V. C.; Brodbelt, J. S. Anal. Chem. 2016, 88, 4004-4013.
V.C.C. designed and conducted all experiments.

greater overall characterization of IgG subunits. Overall, this targeted UVPD approach represents a promising new strategy for the comprehensive characterization of antibody-based therapeutics.

5.2 INTRODUCTION

Since the first therapeutic monoclonal antibody (mAb) entered the clinic nearly three decades ago, the number of mAb-based products approved by the US Food and Drug Administration (FDA) and the European Medicine Agency (EMA) has increased by approximately 50-fold.[1,2] Moreover, upwards of 300 mAb candidates and their derivatives are currently in the clinical development stages, making them among the most rapidly advancing biotherapeutic modalities in the pharmaceutical industry.[2,3] The success of antibody-based biologics, particularly those used in oncogenic and autoimmune disease treatment, arise from their ability to elicit or modulate a desired immune response through efficient and highly specific interactions with a given target.[4–8] These acute interactions and the resulting downstream therapeutic efficacy are regulated by mAb structural integrity with regard to primary sequence and the presence and abundance of post-translational modifications (PTMs), as well as the impact of these variables on higher order structure.[9] Importantly, heterogeneities introduced via production in cell culture or during purification and storage procedures can alter pharmacokinetic properties, reduce therapeutic potency and in severe cases stimulate deleterious immunogenic responses.[10,11] Such modifications include sequence mutations, differential glycosylation, and heavy chain C-terminal processing, as well as varying levels of deamidation and oxidation.[12,13] Consequently, methods that facilitate detailed characterization of antibody primary sequence in addition to PTM identification and site

localization are critical to ensure mAb safety and efficacy. This objective remains challenging, however, due to the size (~150 kDa) and structural complexity of the tetrameric immunoglobulin G (IgG) scaffold from which all approved mAb therapeutics are derived. The IgG structure consists of two identical heavy chains (HC, ~50 kDa) and two identical light chains (LC, ~25 kDa) that are joined through a series of intra- and intermolecular disulfides.[14] Each chain is further divided into variable (V) and constant (C) domains that are responsible for antigen-specific binding and initiation of cell-mediated effector mechanisms, respectively.[15,16]

Advances in sensitivity, resolution, and throughput have established mass spectrometry (MS) as the primary analytical tool used for the characterization and quality control (QC) of antibody therapeutics within the pharmaceutical industry.[17,18] Bottom-up peptide mapping workflows, which involve reduction, alkylation and enzymatic digestion into constituent peptides followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis, are routinely used to monitor mAb primary sequence across all stages of production.[19] Although well-established both in terms of peptide-level separations and MS/MS activation methods, bottom-up strategies tend to be labor-intensive and suffer from limitations arising from incomplete peptide sampling and artifactual heterogeneities introduced via sample handling and proteolysis that can obscure the differentiation of manufacturing-related modifications.[17,20] To circumvent these shortcomings, tandem-MS based methods that restrict or completely bypass enzymatic digestion have gained traction for mAb characterization.[21–29] The latter, or top-down approach, is particularly desirable due to the lack of sample preparation required, which translates to maintenance of high structural integrity, in addition to its ability to provide immediate feedback on sequence fidelity and proteoform abundance based on intact mass measurement.[30,31] However, both the size and structural complexity of IgGs

render high-throughput MS/MS analysis at the top-down level non-trivial using conventional activation methods and current state-of-the-art instrumentation.[26–28,32] To date, electron-transfer dissociation (ETD) and electron-capture dissociation (ECD) have emerged as the methods of choice for top-down sequencing of mAbs due to their enhanced protein-level fragmentation efficiency compared to low energy, slow heating methods, such as collision-induced dissociation (CID).[26–28] Despite this improved efficiency, electron-based methods are limited to the interrogation of approximately 30-35% of the antibody structure, due in part to the number of disulfide protected regions across each IgG domain.[27,28] These important structural features currently preclude top-down MS/MS analysis as a standalone method for comprehensive mAb characterization.

Middle-down mass spectrometry has surfaced as a compelling alternative that combines the strengths of both bottom-up and top-down methods and mitigates their shortcomings. This approach generally involves restricted enzymatic digestion to generate large peptides (~3-20 kDa) that are amenable to high resolution LC-MS/MS analysis.[21,24,33–35] IgG subdomain analysis is a common variation of the middle-down approach accomplished via chemical or electrochemical reduction of intermolecular disulfides to produce free heavy chains (~50 kDa) and free light chains (~25 kDa) or in combination with selective cleavage near the hinge region with proteases such as papain or IdeS to generate three distinct ~25 kDa subunits consisting of the free light chain (Lc), the heavy chain variable domain (Fd), and the heavy chain Fc monomer (Fc/2).[17,21–23,36–38] While better suited for high-throughput analysis compared to their intact counterparts, characterization of antibody subunits remains challenging due to inherent limitations in the speed and sensitivity of high resolution measurements required for large (≥ 25 kDa), highly charged species within narrow chromatographic elution windows.[39–41] Fornelli and co-workers addressed these issues on an Orbitrap Elite instrument by utilizing a survey

LC-MS run to first establish accurate mass measurements of IdeS generated IgG subunits, followed by subsequent targeted LC-MS/MS analyses based on broadband ETD of multiple highly charged precursors.[22] Using this targeted approach, acquisition of a greater number of scans was possible to generate averaged spectra with improved signal-to-noise, resulting in approximately 30-50% sequence coverage of IgG subunits from a single LC-MS/ETD run. At a slight cost in throughput, sequence coverages were improved to 60-70% when transients were averaged from 6-10 independent LC-MS/ETD runs collected with varied ETD reaction times.[22] Wang *et al*. demonstrated only marginally lower sequence coverage of the Lc subunit using a single LC-MS/MS analysis based on broadband higher energy collisional activation (HCD).[23] However, since this middle-down approach did not incorporate enzymatic digestion of the heavy chain, a direct comparison between HCD and ETD performance for the Fd and Fc/2 subunits is currently lacking.

Recently, our group has demonstrated substantial performance gains in the characterization of intact proteins using 193 nm ultraviolet photodissociation (UVPD) and hybrid activation methods thereof relative to conventional collision- and electron-based strategies.[42–45] The high energy deposition afforded by 193 nm UV photoabsorption facilitates interrogation of nearly every inter-residue site to provide unparalleled identification and site localization of sequence variants and PTMs in intact proteins.[42,44] Moreover, the utility of UVPD for whole protein characterization on chromatographic timescales has also been demonstrated.[43,45] These performance attributes make UVPD an ideal candidate to address the increasing demand for analytical methods that provide efficient and detailed characterization of antibody therapeutics to ensure product quality and clinical safety. Herein, we present a middle-down strategy that capitalizes on the high energy deposition and tunability of UV photoactivation to achieve unprecedented

sequence coverage of monoclonal antibody subunits within a single LC-MS/MS experiment. Modulation of the activation parameters, both in terms of pulse number and pulse energy allows for greater control over the extent of coverage of terminal and interior regions of the subunit sequence. By combining the information from four separate targeted UVPD experiments using variable laser parameters, subunit sequence coverages as high as 85% were achieved.

5.3 EXPERIMENTAL

5.3.1 Materials and Reagents

Clinical grade recombinant monoclonal antibodies of the IgG1 subclass were used for all experiments. Trastuzumab (Herceptin) was kindly donated by Genentech. Adalimumab (Humira) was purchased at >97% purity from BOC Sciences (Shirley, NY). Lyophilized recombinant immunoglobulin G-degrading enzyme from *Streptococcus pyogenes* (IdeS/FabRICATOR, Genovis) was obtained from Bulldog Bio, Inc. (Portsmouth, NH). All solvents and mobile phase additives were purchased in LC-MS grade purity. Water, acetonitrile, urea and all other buffer components were obtained from Sigma-Aldrich (St. Louis, MO). Isopropyl alcohol and formic acid were purchased from Fisher Scientific (Fairlawn, NJ). Tris(2-carboxyethyl)phosphine (TCEP) solution was obtained from Pierce (Rockford, IL).

5.3.2 IdeS Digestion and Sample Preparation

Sample preparation was adapted from that previously described by Fornelli *et al.*[22] Briefly, acetone-precipitated monoclonal antibodies were resuspended at 5 μg/μL in IdeS cleavage buffer (50 mM sodium phosphate, 150 mM NaCl, pH 6.6) and subjected to hinge-selective cleavage with one unit of IdeS per microgram of IgG at 37˚C for 30

minutes to produce the F(ab')2 and Fc fragments. Following digestion, these fragments were denatured with 4 M urea and reduced in the presence of 30 mM TCEP for 30 minutes at room temperature to produce the Fc/2, Lc and Fd IgG subunits (~25 kDa). To quench the reaction and prevent disulfide bond reformation, the sample was acidified with 1% formic acid (FA). Immediately prior to analysis the sample was diluted to 1 μg/μL in 0.1% FA.

5.3.3 Liquid Chromatography and Mass Spectrometry

Chromatographic separation of IgG subunits was accomplished using a Dionex Ultimate 3000 microbore liquid chromatography system (Sunnyvale, CA) equipped with a Waters XBridge Protein BEH300 C4 column (2.1mm x 250 mm, 3.5 μm particle size) heated to 65˚C. Mobile phase A consisted of 0.1% formic acid in water, and mobile phase B consisted of 39.9% IPA, 60% ACN, and 0.1% formic acid. 2 μg of IgG subunit mixture was injected on-column and separated using a steep linear ramp from 5% B to 20% B over 2 minutes followed by a shallow linear gradient from 20% B to 40% B over 28 minutes at a flow rate of 250 μL/min. The LC system was coupled via an electrospray ionization (ESI) source to a Thermo Fisher Orbitrap Elite mass spectrometer (Bremen, Germany) outfitted with a 193 nm ArF excimer laser (Coherent ExciStar XS) to allow UVPD in the HCD cell as previously described.[42,46] For all experiments, an ESI source voltage of 3.75 kV, S-lens rf level of 70%, and heated capillary temperature of 350˚C were used. The sheath and auxiliary gas flow was maintained at 35 and 5 arbitrary units, respectively. The HCD collision gas pressure in the vacuum chamber containing the Orbitrap mass analyzer was reduced so that the change in pressure (Δp) equals 0.1×10^{-10} (~5 mTorr) as previously described to enhance the detection of low abundance and large product ions.[22,42] Prior to acquisition of MS/MS data, a survey LC-MS run was performed

to establish the chromatographic elution window and charge state distributions for each IgG subunit. All MS^1 spectra were acquired using a mass range of *m/z* 400-2000 and resolving power of 120,000 at *m/z* 400. For all MS^2 experiments, targeted LC-MS/MS programs were developed in which spectra were acquired using a mass range of *m/z* 205-2000 with 240,000 resolution at *m/z* 400, and each FTMS scan was the result of 20 microscans. For UVPD, an automatic gain control (AGC) target of 1.0×10^6 was used and the most abundant charge state was continuously activated across the elution profile for each subunit using either one or two 5 ns laser pulses at a range of 1 to 2.5 mJ per pulse. In the UVPD set-up, the laser is neither focused nor collimated. ETD data was acquired using a reaction time of 5 ms with a reagent AGC target of 7.5×10^5 and a maximum anion injection time of 50 ms. For direct comparison with UVPD under the same conditions, ETD activation was performed on the most abundant charge state of each subunit using a 20 *m/z* isolation window. To compare the performance of UVPD with more optimal ETD conditions, ETD was also carried out using multiple highly charged precursors via a wide isolation window (150 *m/z*) as described by Fornelli *et al*.[22]

5.3.4 Data Processing

MS^1 spectra collected across the entire chromatographic elution profile for each IgG subunit were combined to generate a single averaged MS^1 spectrum that was subsequently deconvolved using the Xtract algorithm (Thermo Fisher Scientific) with a S/N ratio of 3 to determine the monoisotopic mass of each subunit. For MS/MS data, all scans for a given precursor or range of precursors were averaged and a new Thermo .RAW file was generated for each subunit. The averaged UVPD spectra were deconvolved with Xtract to obtain monoisotopic mass information for all product ions with a S/N ratio of 3 or higher and then analyzed using ProSight Lite with UVPD

searching enabled in addition to Protein Prospector-assisted manual interpretation to search for 10 ion types: *a, a+1, b, c, x, x+1, y, y-1, Y* and *z*. ETD spectra were searched for *c*, *z,* and *y*-type product ions. Cleavage sites were assigned using a 10 ppm tolerance.

5.4 RESULTS AND DISCUSSION

5.4.1 Evaluation of IgG Subunit Separation and Accurate Mass Analysis

Trastuzumab and adalimumab used in this study are two therapeutic monoclonal IgG1 antibodies that possess well-characterized primary sequences and post-translational modifications to allow for reliable benchmarking of our targeted middle-down UVPD strategy. As summarized in **Figure 5.1**, the first part of the workflow utilizes hinge-selective digestion with IdeS, which cleaves at the conserved di-glycine motif of the heavy chain (Xxx-L-L-G/G-P-S-xxX), followed by reduction of inter- and intra-chain disulfide bonds to produce the Fc/2, Lc, and Fd antibody subunits. Complete sample preparation was accomplished in approximately one hour using mild reaction conditions previously shown to minimize processing-derived artifacts, such as oxidation and carbamylation that are common to lengthier bottom-up digestion procedures.[22,38] The overlaid total ion chromatograms (TIC) shown in **Figure 5.2** for IdeS digested trastuzumab (**Figure 5.2a**) and adalimumab (**Figure 5.2b**) demonstrate efficient conversion of intact IgG into the three constituent subdomains of interest. Moreover, the TICs demonstrate baseline resolution of the subunit mixture with high chromatographic reproducibility for replicate injections of the same antibody using reversed-phase C4 separation. These criteria are critical for the integration of subunit-specific targeted MS/MS activation across discreet elution windows used in the second half of the workflow (**Figure 5.1**).

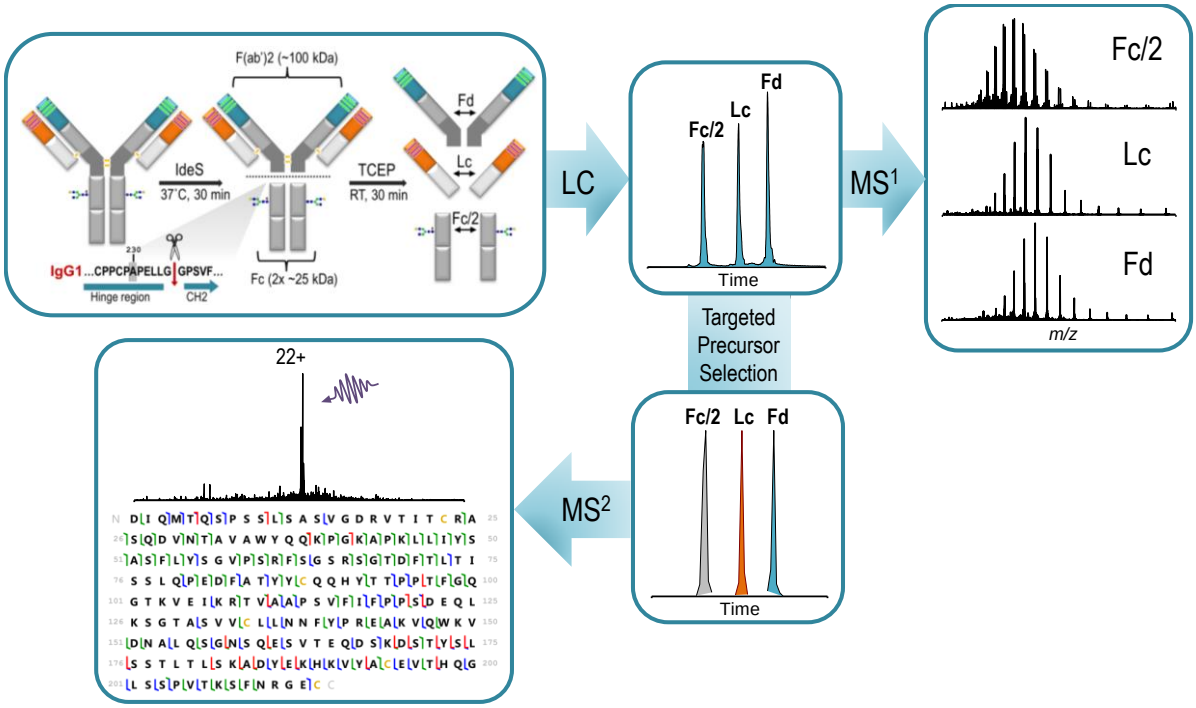


Figure 5.1 Schematic representation of sample preparation and analysis workflow. IgG subunits are first produced from IdeS digestion and TCEP reduction. High resolution LC-MS[1] analysis using 120K resolution (at *m/z* 400) provides accurate mass measurements of subunits in addition to elution profiles and charge state distributions necessary for targeted MS/MS activation.

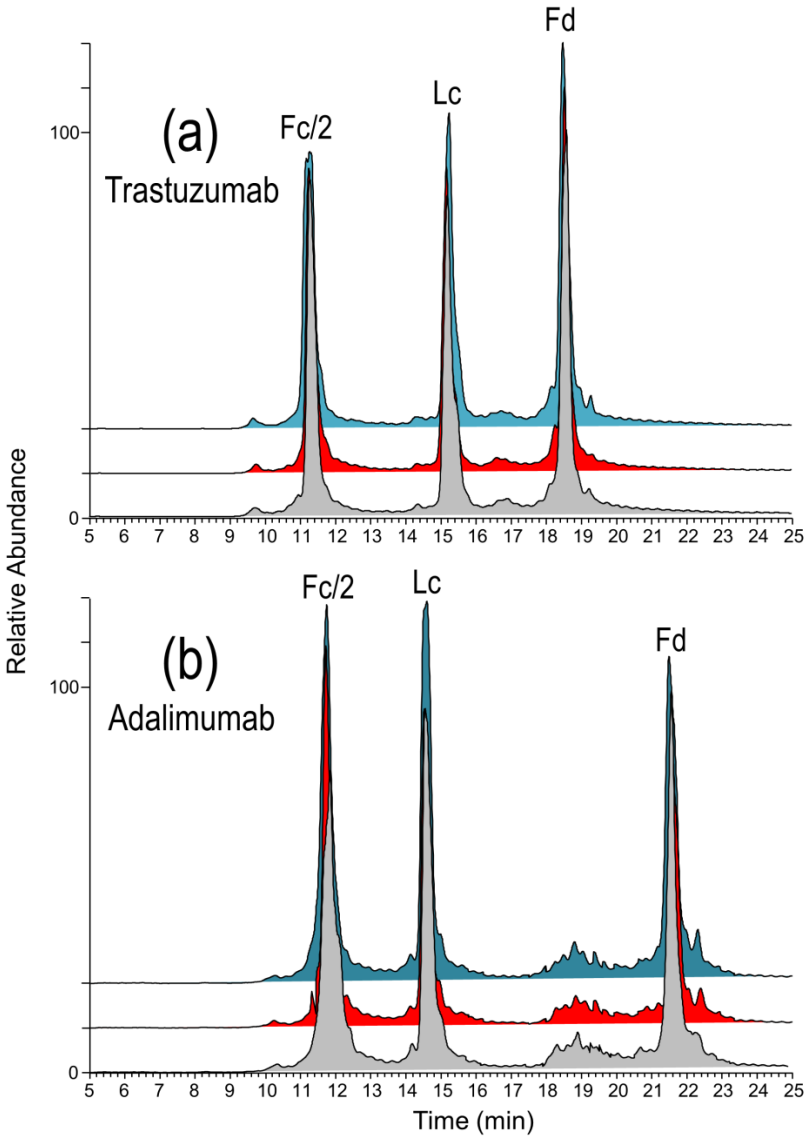


Figure 5.2 Stacked total ion chromatograms for triplicate LC-MS analyses of IdeS-derived (a) trastuzumab and (b) adalimumab subunits: Fc/2, Lc, and Fd, respectively, with baseline chromatographic resolution.

Despite greater than 90% sequence identity between full length trastuzumab and adalimumab, the majority of the amino acid variability occurs within the antigen-binding complementarity determining regions (CDRs) of the Lc and Fd subdomains (**Figure 5.3**) resulting in considerable variations in both chromatographic and ESI profiles that must be empirically defined for respective mAbs. This was accomplished using a strategy similar to that described by Fornelli *et al.*,[22] in which a high resolution LC-MS survey run was first performed to provide immediate feedback regarding subunit elution and charge state distribution. As demonstrated in **Figure 5.4**, MS^1 acquisition at a resolving power of 120,000 (*m/z* 400) affords isotopic resolution of the charge state envelope and monoisotopic mass determination with sub 4 ppm mass accuracy for each subunit within a single LC-MS experiment. High resolution accurate mass analysis also revealed the presence of two commonly observed post-translational modifications on the heavy chain-derived Fc/2 subunit: N-linked glycosylation and C-terminal lysine truncation. Glycan microheterogeneity was observed as consecutive mass shifts consistent with either fucose or hexose saccharide additions. Masses corresponding to the G0, G0F, G1F and G2F glycoforms were identified for trastuzumab (**Figure 5.4a**), whereas adalimumab exhibited only the G0F and G1F glycovariants (**Figure 5.4d**). Furthermore, each Fc/2 variant showed a conserved loss of 128 Da from the theoretical subunit mass, indicative of complete processing of the heavy chain C-terminal lysine. Both of these modifications have important implications on the regulation of cell-mediated effector functions and are thus essential for the complete characterization of IgG.[16,47] While accurate mass information provides insight regarding the presence of probable structural features based on *a priori* knowledge of common mAb modifications, the MS^1 data alone does not facilitate unambiguous localization of these modifications to specific residues.

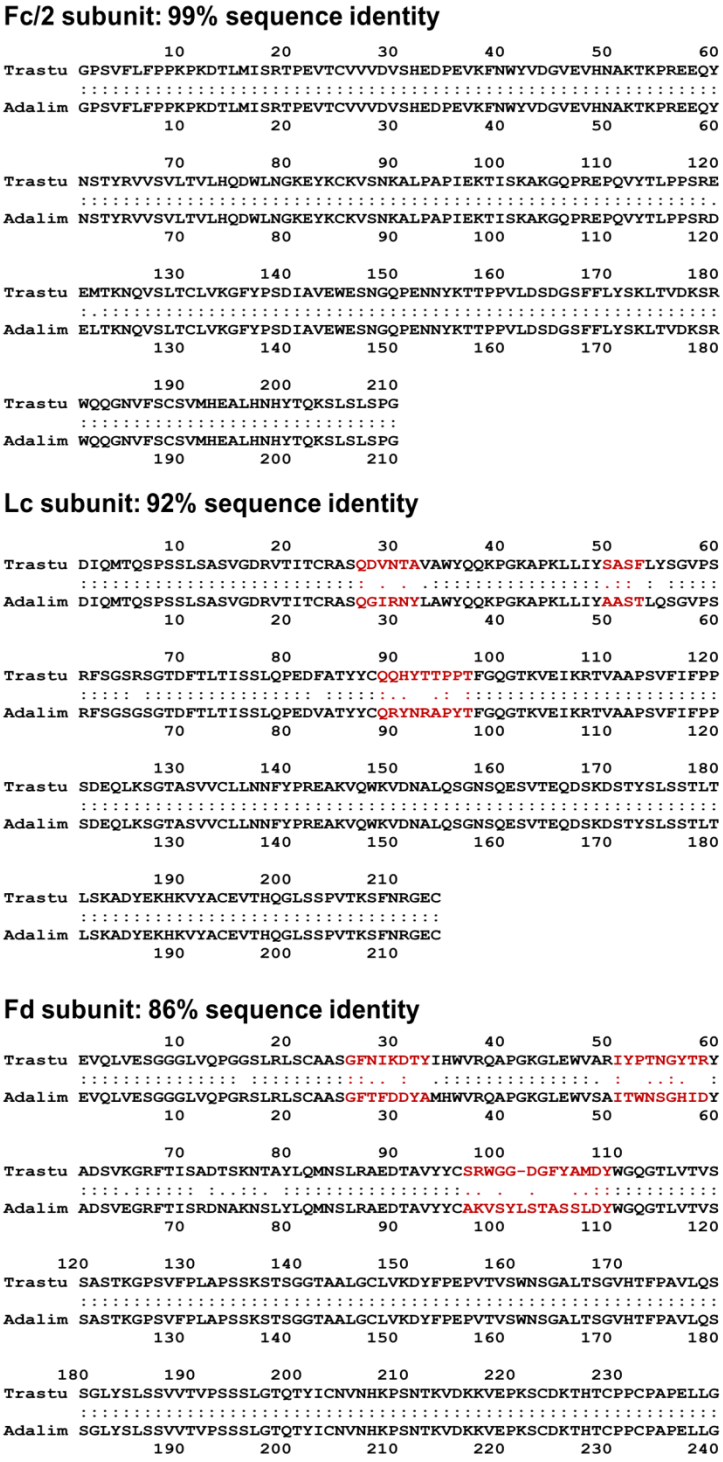


Figure 5.3 Sequence alignment of trastuzumab and adalimumab Fc/2, Lc, and Fd subunits. Hypervariable CDRs are shown in red.

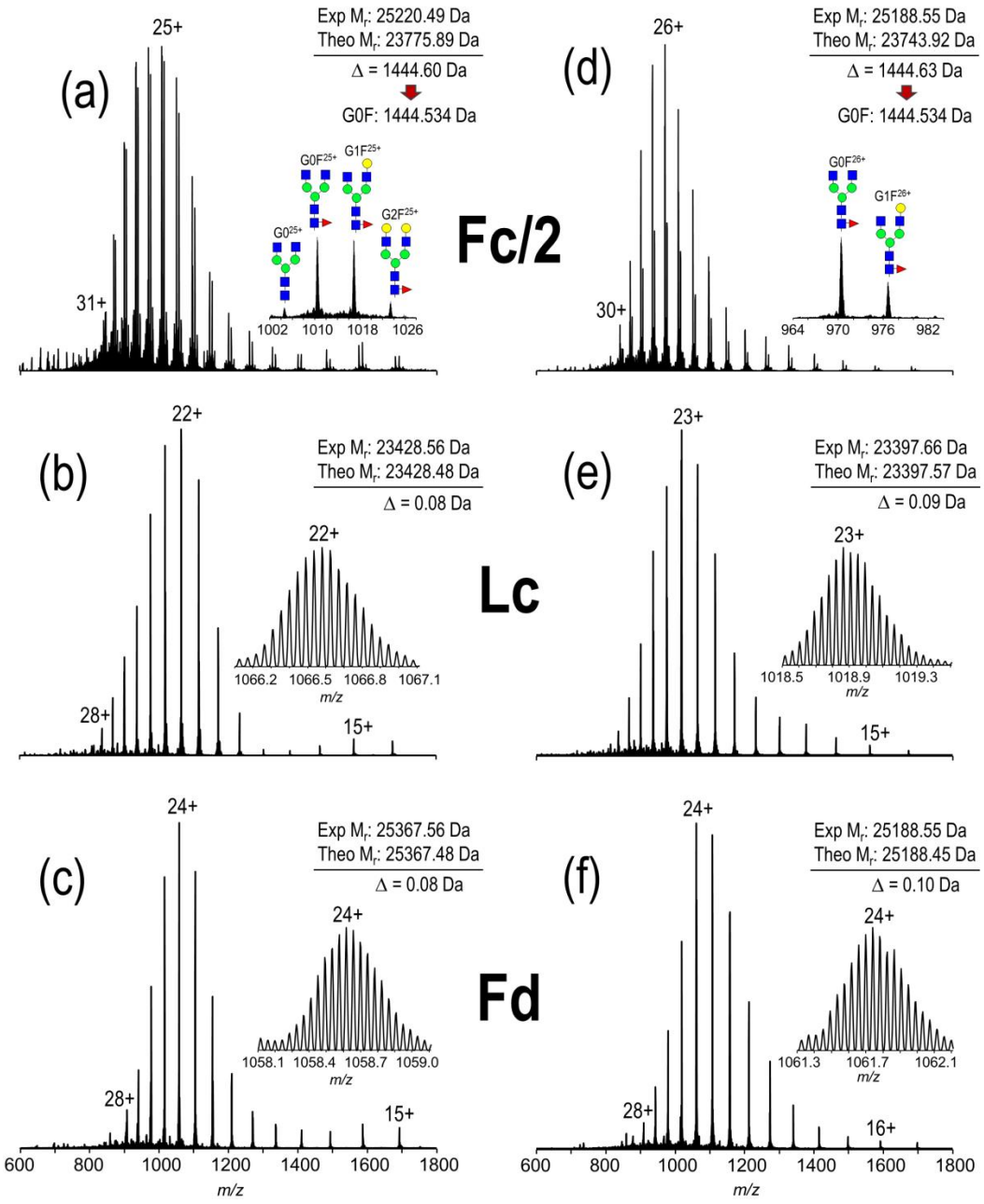


Figure 5.4 ESI mass spectra for the Fc/2, Lc, and Fd subunits of trastuzumab (a-c) and adalimumab (d-f), respectively, collected at 120K resolution (at m/z 400). The insets for the Fc/2 subdomains demonstrate the glycoform heterogeneity in each IgG based on accurate mass measurement. Trastuzumab exhibited the G0, G0F, G1F, and G2F glycoforms (a), whereas adalimumab exhibited the G0F and G1F variants only. The insets for all other subunits demonstrate the isotope distribution for the most abundant charge state.

5.4.2 Optimization of Targeted UVPD for Maximal Sequence Coverage and PTM Localization

Characterization of whole proteins or large polypeptides thereof on chromatographic timescales has been shown to rely critically on the nominal resolving power and number of spectral averages acquired during activation.[22,45,48] A recent study in our lab, which utilized 193 nm UVPD for the high-throughput interrogation of intact proteins, showed that optimal characterization, both in terms of the number of inter-residue cleavages throughout the primary sequence and site localization of PTMs, was facilitate by the use of maximal resolving power.[45] This was attributed to improved deconvolution of rich UVPD spectra that contain extensive arrays of highly charged product ions and instances of overlapping isotopic distributions that were otherwise obscured at lower resolution settings.[45] To address the latter variable, Fornelli and co-workers employed targeted MS/MS acquisition to maximize the number of scans collected during middle-down ETD analysis to generate more informative spectra with enhanced single-to-noise (S/N).[22] For the present study, we hypothesized that combining these data acquisition strategies with the sequencing power of 193 nm UVPD would afford greater depth of characterization of therapeutic IgG subunits within the context of a high-throughput workflow.

Optimization of laser parameters for UVPD in terms of pulse number and energy per pulse was carried out for trastuzumab with the goal of maximizing sequence coverage of all subunits within a single targeted LC-MS/MS experiment, in addition to obtaining fragmentation-level confirmation and site localization of PTMs. A limited set of test conditions was evaluated based on the degree of coverage, in terms of the ratio of observed versus total number of inter-residue positions, for a model modified protein of similar size (alpha casein, ~24 kDa, data not shown) and consisted of combinations of

one or two laser pulses with energies ranging from 1 to 2.5 mJ per pulse. To achieve the highest MS^2 sensitivity and spectral S/N, we exploited the charge-independent nature of 193 nm UVPD[42,49] to continuously activate the most abundant charge state for each subunit (25+, 22+ and 24+ charge states for the Fc/2, Lc, and Fd, respectively) across the elution window defined by a preceding LC-MS^1 survey run. Composite MS^2 spectra generated by merging 60 to 80 microscans acquired at 240K resolution during targeted activation were then used to evaluate which set of laser conditions promoted optimal IgG characterization.

As previously mentioned, Fc glycosylation is a key regulator of antibody effector functions, including but not limited to antibody-dependent cell-mediated cytotoxicity (ADCC) and phagocytosis (ADCP), via the modulation of IgG-Fc interactions with respective cell surface receptors.[16] Due to its critical role in the immune response, most therapeutic IgGs possess a highly conserved site of N-linked glycosylation at asparagine 297 (Asn-297) of the heavy chain, which consists of a heptasaccharide core ($GlcNAc_4Man_3$) that is often variably modified during protein engineering and bioproduction processes.[50,51] Both the type and relative abundance of glycan microheterogeneities can have a profound effect on the pharmacological properties and therapeutic efficacy of mAb-based drugs. Thus an important metric by which to evaluate UVPD for the characterization of IgG is its ability to retain intact glycan and glycosylation site information. While photodissociation at 157 nm and 193 nm have shown improvement over collisional activation methods at preserving labile glycosidic bonds at the peptide level,[52,53] this has yet to be investigated at the protein or subunit level. A representative example of UVPD performance for the characterization of the glycosylated Fc/2 subunit of trastuzumab is shown in **Figure 1**. Due to its greatest overall abundance in the MS^1 spectrum, the 25+ charge state of the G0F glycoform (see top right

inset of **Figure 5.5**) was selected for repeated photoactivation. The resulting UVPD spectrum was first searched against theoretical fragment ion masses corresponding to the unmodified Fc/2 sequence using a strict 10 ppm mass tolerance (**Figure 5.5** bottom left). The identification of N-terminally derived product ions (*a*, *b*, *c*) showed an abrupt stop at the Asn-61 position, which corresponds to the glycosylated residue of interest (**Figure 5.5** middle). A similar loss of matched C-terminal product ions (*x*, *y*, *z*) beyond Asn-61 was also observed; albeit this loss appears less dramatic due to a lower frequency of bi-directional fragmentation in terminally-located regions, consistent with previously reported UVPD results for proteins of similar size.[44] It should be noted here that C-terminal product ions were only identified when the terminal residue was removed, thus confirming lysine processing of the Fc/2. Upon the addition of 1444.53 Da at the Asn-61 position to reflect the mass of the G0F glycan, 40 additional product ions were identified and the total sequence coverage was boosted from 50% to 62% (**Figure 5.5** bottom right). As shown in the middle panel of **Figure 5.5**, consecutive *a*-ions allow unambiguous site localization of the G0F glycan and provide unequivocal evidence for the generation of diagnostic fragment ions that retain intact glycan information.

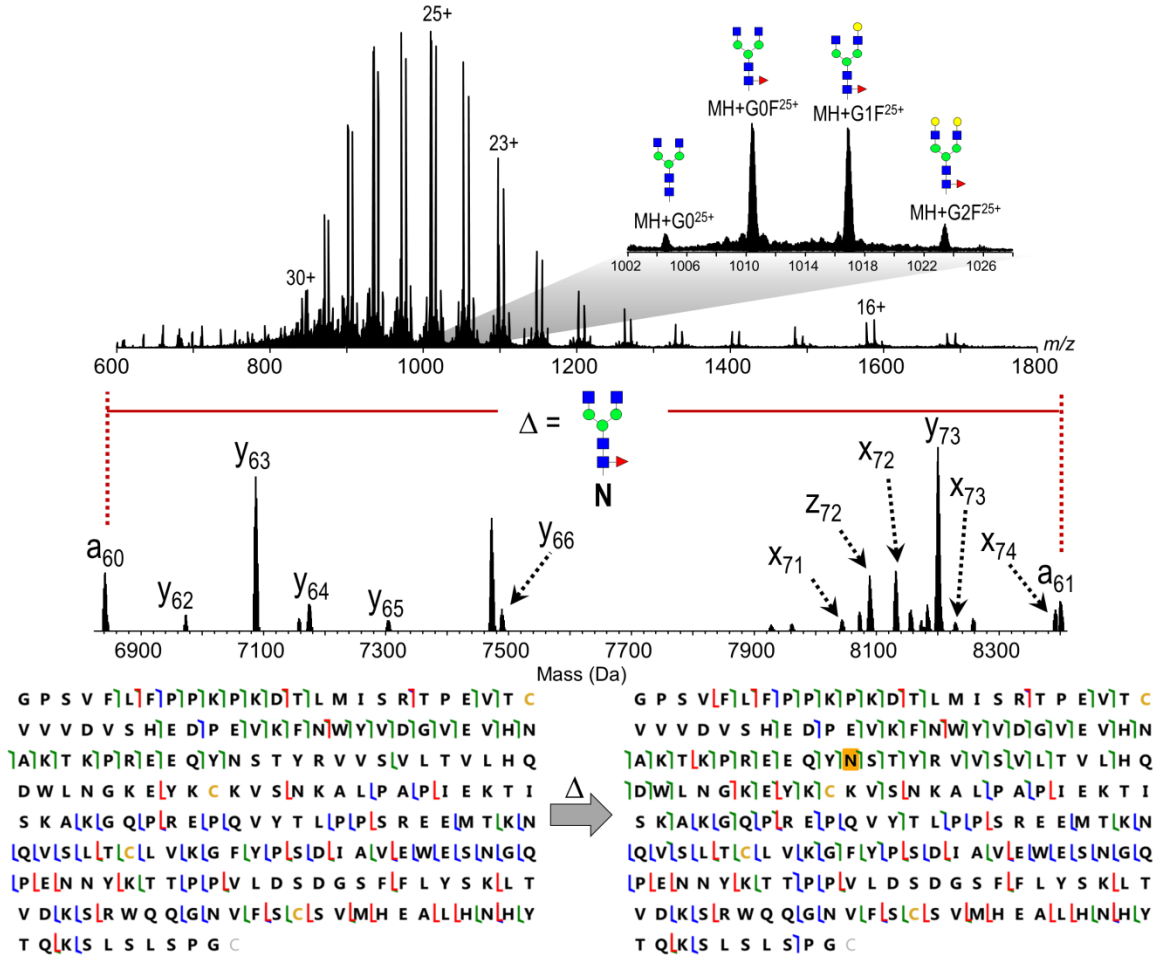


Figure 5.5 MS[1] spectrum of the Fc/2 subunit of trastuzumab showing mass shifts consistent with glycoform microheterogeneity (top). The inset demonstrates consecutive saccharide additions to the core N-linked glycan structure. The zoomed region of the deconvolved 193 nm UVPD mass spectrum (middle) of the 25+ charge state of the G0F glycoform shows a mass shift consistent with the intact glycan structure between consecutive *a* ions. An abrupt stop in matched N-terminally derived ions allows for unambiguous glycan site localization (bottom).

Once it was confirmed that Fc/2 fragmentation was not prohibitively biased by preferential cleavage of the glycan moiety, UVPD performance was further evaluated based on the total sequence coverage per subunit as a function of laser parameter selection (**Figure 5.6**). In general, two pulses of 193 nm photons yielded a greater degree of sequence informative fragmentation compared to single pulse activation, likely due to enhanced energization of the polypeptide backbone leading to more efficient photodissociation. With regard to energy per pulse, both the Fc/2 and Lc subunits exhibited improved coverage as laser power was increased, while the opposite was true for the Fd subdomain. This is readily demonstrated by comparing subunit sequence coverages obtained for individual targeted LC-MS/UVPD experiments performed using dual pulse activation at either 1 or 2.5 mJ per pulse. As summarized in **Figure 5.6**, a laser power of 1 mJ/pulse promoted 58%, 59% and 53% sequence coverage of the Fc/2, Lc, and Fd subunits, respectively. These values rose to 65% and 66% for the Fc/2 and Lc subunits at the higher energy setting of 2.5 mJ/pulse, whereas Fd coverage fell to 42%. These results likely reflect "over"-dissociation of the Fd backbone by way of enhanced production of secondary fragment ions or internal ions at increasingly energetic activation conditions. This outcome for the Fd polypeptide can be rationalized based on its greater length (i.e., 239 amino acids (aa) versus 210 aa and 214 aa of the Fc/2 and Lc, respectively) taken together with the fact that the amide backbone serves as the chromophore at 193 nm.[54] Thus, as the length of the amino acid chain increases, the magnitude of the absorption cross-section also increases, resulting in more efficient photoactivation.

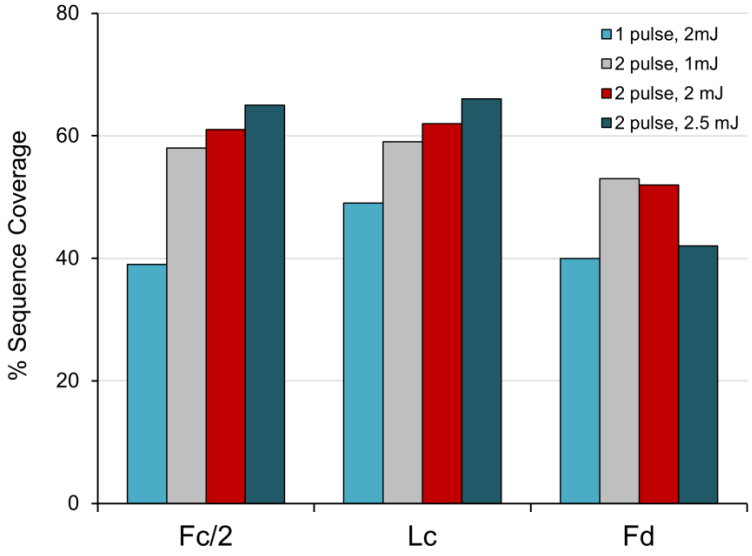


Figure 5.6 Sequence coverage observed as a function of laser parameter selection used for targeted UVPD of the most abundant precursor of the Fc/2 (25+), Lc (22+) and Fd (24+) subunits of trastuzumab.

To explore this phenomenon in greater detail, the fragmentation information obtained by UVPD at each set of activation parameters was plotted as histograms showing the summed N-terminal (*a*, *b*, *c*) and C-terminal (*x*, *y*, *z*) product ion abundances arising at all inter-residue sites across the Fd backbone (**Figure 5.7**). Changes in fragmentation, both in terms of cleavage location and ion abundance support the hypothesis of increasing secondary dissociation at elevated laser power. **Figure 5.7d** is particularly illustrative of this point as the histogram is essentially devoid of sequence information from the interior of the polypeptide and instead shows bias towards smaller terminally-derived product ions that are expected to undergo less efficient secondary dissociation due to their inherently lower photoabsorption cross-sections. Conversely, greater access to interior regions of the Fd sequence was afforded by lowering the laser power to effectively modulate energy deposition in favor of the formation of large energetically stable primary fragment ions (**Figure 5.7a-c**). This strategy was particularly useful for obtaining sequence information in the complementarity determining regions

(CDRs) of the IgG variable domains, which are critical indicators of antigen-binding specificity and mAb therapeutic efficacy. Consequently, coverage of these highly diagnostic regions derived from the heavy chain was used as the final metric by which to evaluate UVPD performance for IgG characterization. As shown in **Figure 5.7**, dual pulse activation at both 1 mJ/pulse (**Figure 5.7b**) and 2 mJ/pulse (**Figure 5.7c**) yielded nearly complete fragmentation in these diagnostic regions (shown in gray), while also affording the greatest overall coverage of the Fd subunit.

Based on the criteria of total sequence coverage, PTM site localization, and CDR sequence confirmation, optimal characterization of IgG subunits was accomplished when UVPD was carried out using dual pulse activation at 2 mJ/pulse. Under these conditions, a single targeted LC-MS/UVPD experiment resulted in 62% sequence coverage of both the Fc/2 and Lc subunits of trastuzumab, and 52% coverage of the larger Fd subunit (**Figure 5.7**). Importantly, this level of characterization is comparable to that reported for ETD upon combining the MS/MS data from multiple runs completed at varying ETD reaction times,[22] thus demonstrating significant gains in throughput using this targeted UVPD strategy. To emphasize the utility of tunable energy deposition for sequence characterization by UVPD, we combined the data from all four LC-MS/MS runs carried out using the laser conditions previously described. The resulting fragment ion maps shown in **Figure 5.8a-c** for the Fc/2 (G0F glycoform), Lc and Fd subunits of trastuzumab, respectively, reveal unprecedented characterization of IgG subunits owing to greater control over energy deposition and subsequent fragmentation. As summarized in **Figure 5.8d**, sequence coverages were dramatically increased to values above 80% for the Fc/2 and Lc subunits, and just under 70% for the Fd subunit. Moreover, all CDRs from both the light and heavy chain variable domains were fully sequenced (these regions are shown in gray in **Figures 5.8b** and **5.8c**).

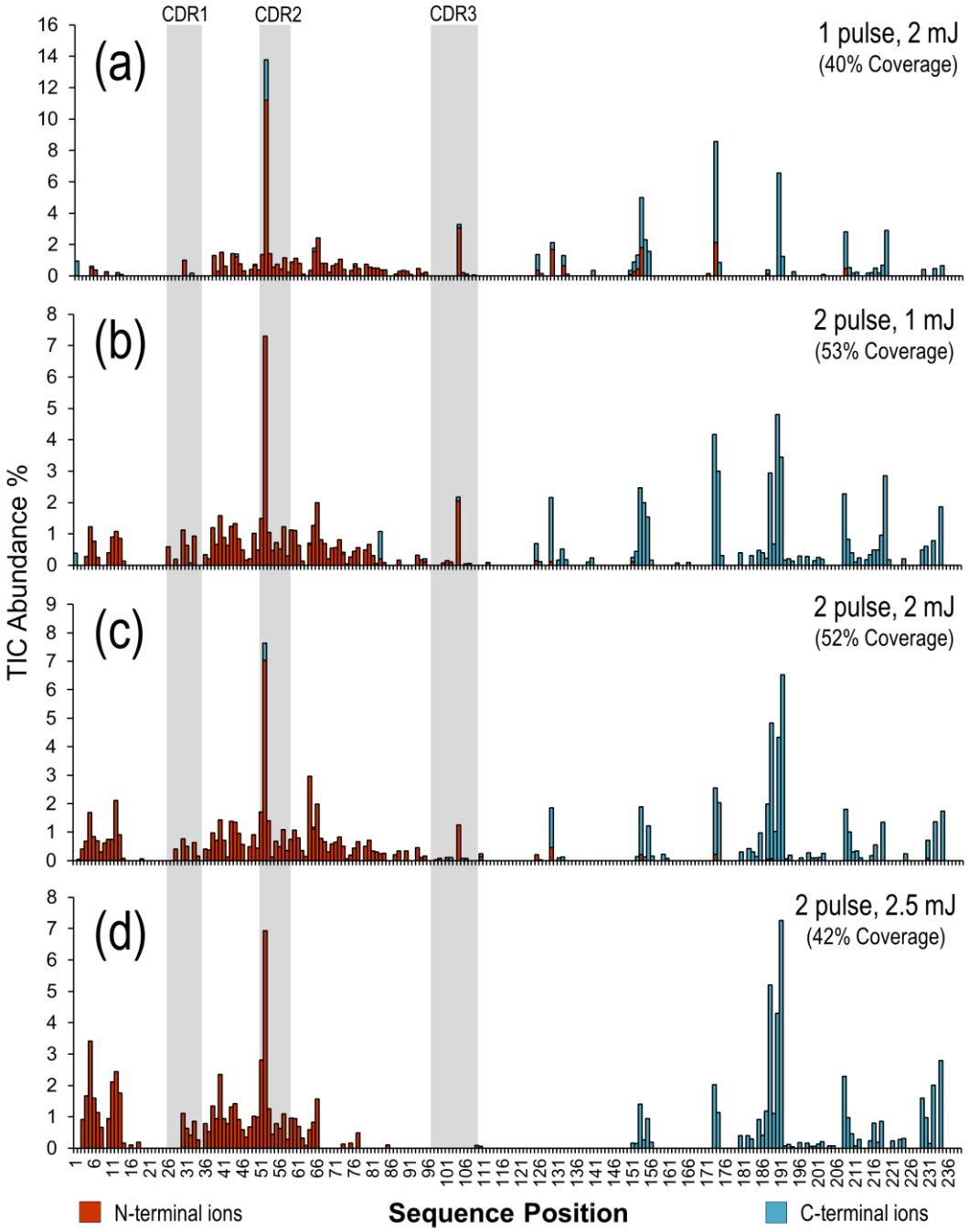


Figure 5.7 Histograms showing the summed N-terminal (*a, b, c*) and C-terminal (*x, y, z*) fragment ion abundances originating from cleavages at each inter-residue position across the Fd subunit of trastuzumab as a function of pulse number (1 or 2 pulses) and pulse energy (nominally 1 mJ, 2 mJ, or 2.5 mJ) used for UV photoactivation. The sequence positions corresponding to the CDRs are shaded in gray.

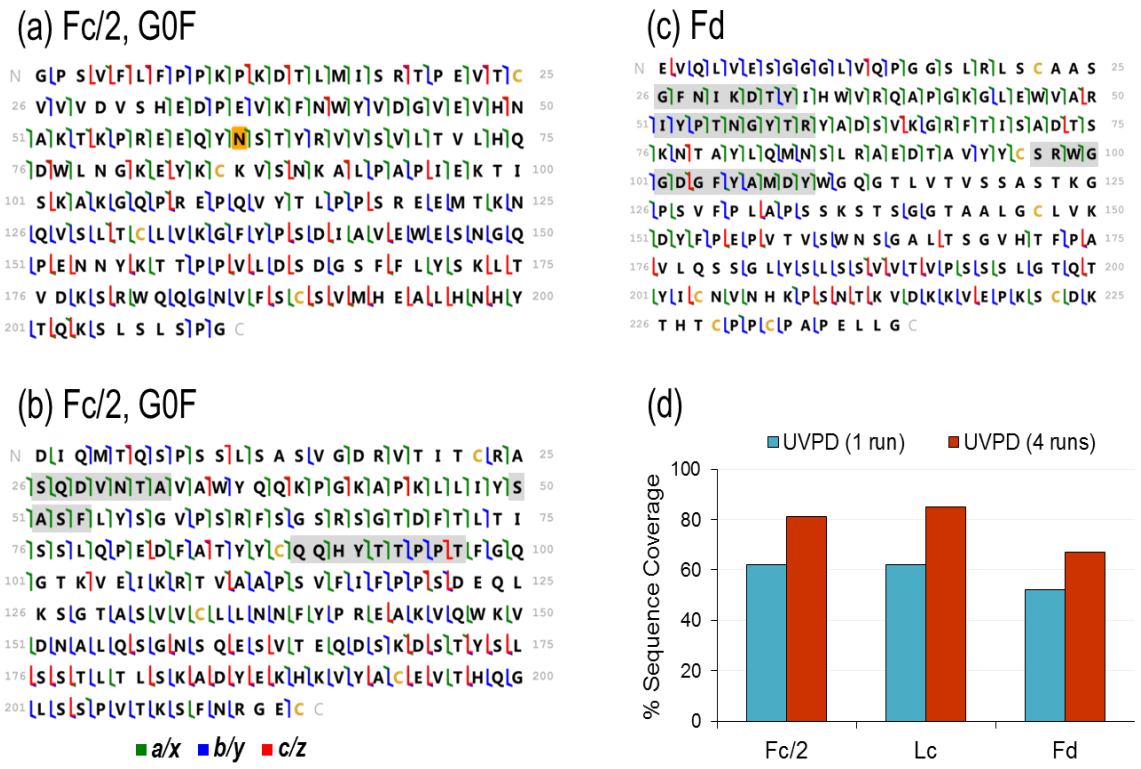


Figure 5.8 (a-c) Fragmentation maps for trastuzumab subunits generated from combining the fragment ion information obtained from four independent UVPD runs collected using varied laser parameters. The CDR sequences are shown in gray. (d) Comparison of sequence coverage obtained for a single LC-MS/UVPD run using fixed laser conditions versus that obtained from combining UVPD data from four runs collected with varied laser parameters. The following charge states were isolated for UVPD: +25 for Fc/2, +22 for Lc, and + 24 for Fd.

5.4.3 Benchmarking UVPD against ETD for Subunit Sequence Characterization

To rule out the possibility that greater sequence coverage of IgG subunits obtained with UVPD relative to the leading ETD-based method[22] arises from differences in instrument performance or the acquisition parameters used, a direct comparison of both activation strategies was carried out on our Orbitrap system. To minimize bias, ETD was performed in two fashions based on the observations of Fornelli *et al*.[22] using either: 1) narrow 20 *m/z* isolation of the most abundant charge state (identical to UVPD) or alternatively, 2) wide 150 *m/z* isolation and broadband activation of multiple highly charged precursors. All ETD experiments were carried out using a 5 ms reaction time and UVPD was accomplished using previously optimized activation conditions. The lower complexity of ETD mass spectra, containing mainly *c/z* ions, relative to the higher complexity of UVPD mass spectra (*a*, *b*, *c*, *x*, *y*, *z*-ions, respectively) mitigates the need for acquisition at 240K resolution; however, all data was collected at this resolving power to facilitate spectral comparisons. Performance metrics were evaluated based on the total number of positionally unique N-terminal and C-terminal fragments produced, as well as the overall sequence coverage obtained for each subunit of trastuzumab and adalimumab within single targeted LC-MS/MS analyses. In all cases, UVPD outperformed both variations of ETD as summarized in **Figure 5.9**. As expected, ETD exhibited improved fragmentation efficiency under conditions of greater charge density, such as those arising from simultaneous isolation and activation of multiple highly charged precursors; however, UVPD still produced between 25-30% more positionally unique fragment ions. Note that this increase does not directly correlate with changes in sequence coverage because N-terminal and C-terminal ion pairs arising from cleavage at the same sequence position were not accounted for. Collectively, these results point to enhanced conversion of the precursor into sequence-informative product ions by UVPD, likely due to its ability

to better disrupt non-covalent interactions that persist in the gas phase and lead to non-dissociative electron transfer by ETD, as previously described.[55,56] Differences in the locations in which these MS^2 events occur (i.e., UVPD in the HCD cell versus ETD in the linear ion trap) introduce variations in product ion transfer that may also contribute to the differences observed in the resulting fragment ion populations; although this effect was minimized through optimization of pressure conditions for improved transfer efficiencies. While UVPD leads to greater overall sequence informative fragmentation within a single experiment, the analytical merit of combining the unique information from both UVPD and broadband ETD was evaluated to assess the complementarity of the fragment ion populations produced by both activation methods. As demonstrated in the lower half of **Figure 5.9** based on the fourth bar in the clustered bar graph, combining the sequence information from separate UVPD and ETD spectra yields considerable gains in coverage, going from 59% for UVPD (on average for the three subunits) and 46% for ETD (on average for the three subunits) to a net coverage of 74% for trastuzumab and similarly 72% for adalimumab. An illustrative example of the complementary nature of these activation methods is demonstrated in the fragment ion maps for the Fc/2 subunit of trastuzumab shown in **Figure 5.10**. Notably, in numerous regions throughout the sequence, where one activation method produced sparse coverage the other yields considerably greater sequence information (i.e., see coverage for regions including residues 25-34 and 62-71 in **Figure 5.10**), thereby greatly extending the level of characterization.

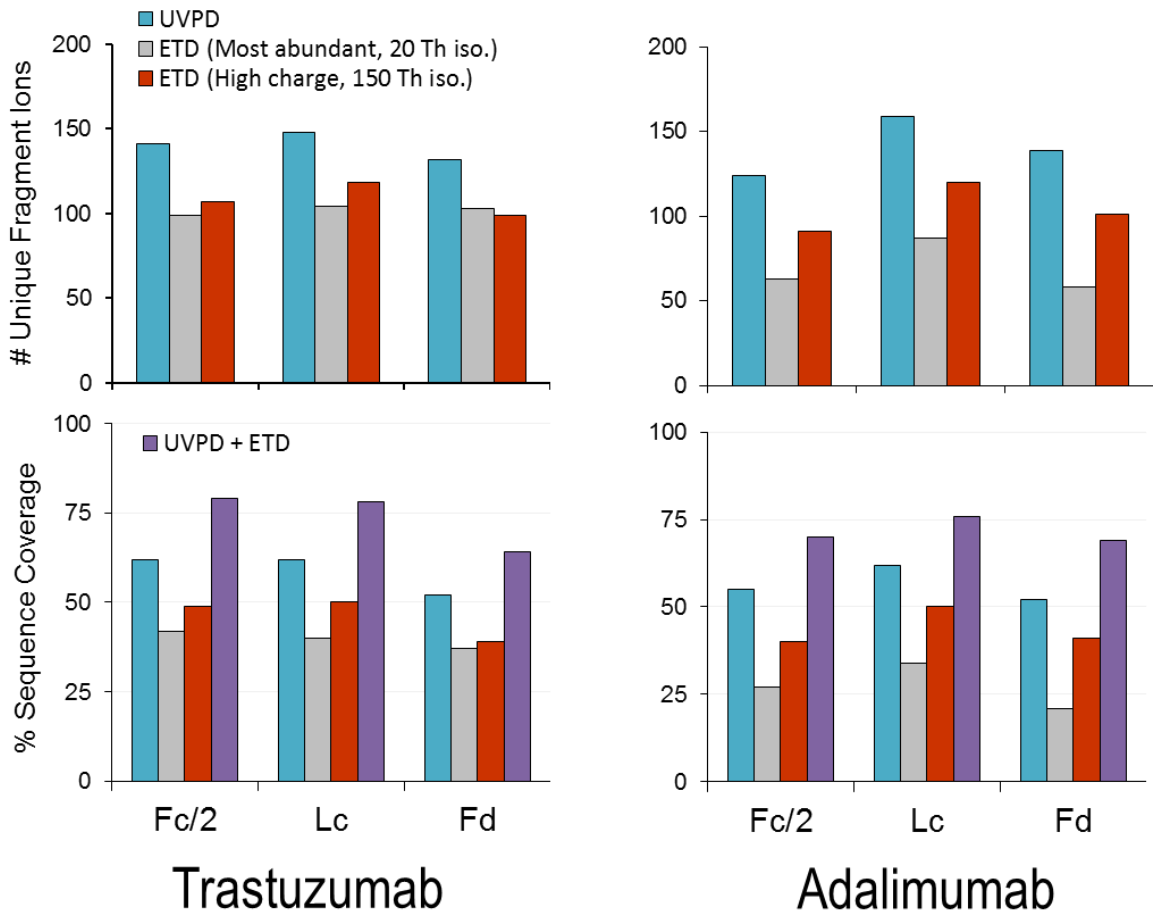


Figure 5.9 Evaluation of unique fragment ions and sequence coverage obtained from single targeted LC-MS/MS analyses of trastuzumab and adalimumab subunits based on UVPD (20 *m/z* isolation, most abundant charge state), ETD (5 ms reaction time) using single precursor isolation (20 *m/z* isolation, the most abundant charge state), and ETD (5 ms reaction time) using multiple precursor isolation (150 Th isolation, high charge states). The most abundant charge states for trastuzumab were as follows: +25 for Fc/2, +22 for Lc, and +24 for Fd. The isolation range used for multiple precursor isolation included: +25 to +31 for Fc/2, +24 to +28 for Lc, and +25 to +29 for Fd. The most abundant charge states for adalimumab were as follows: +26 for Fc/2, +23 for Lc, and +24 for Fd. The isolation range used for multiple precursor isolation included: +26 to +30 for Fc/2, +25 to +27 for Lc, and +25 to +28 for Fd. The fourth bar (purple) in the lower bar graphs shows the net sequence coverage for combined product ion information from UVPD and broadband ETD.

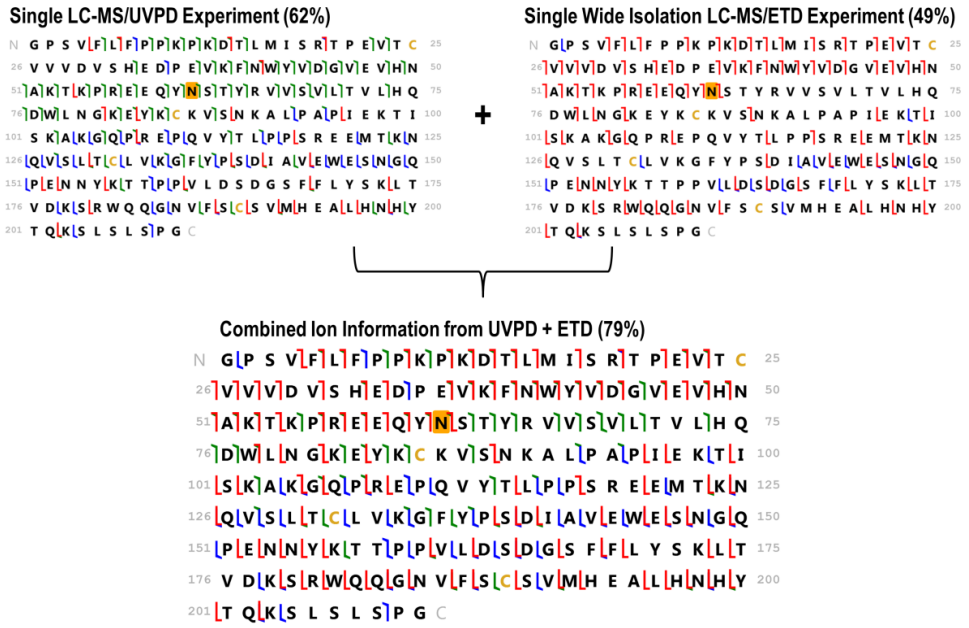


Figure 5.10 Fragment ion maps for the Fc/2 subunit of trastuzumab generated by a single LC-MS/MS experiment based on UVPD (top left) and broadband ETD (top right) and a composite map (bottom) generated by combining the fragment ion information from both experiments.

Representative fragmentation spectra and associated sequence ion maps are shown in **Figure 5.11** for UVPD, narrow isolation- and wide isolation ETD of the Fd subunit of trastuzumab. UVPD exhibited both the greatest overall product ion density and ion-type diversity, resulting in a rich array of *a*, *b*, *c, x*, *y*, and *z*-type ions as indicated in both the insets and fragment ion maps provided (**Figure 5.11a**). As expected, ETD produced predominantly *c* and *z* ions, as well as a slight secondary contribution of *y*-type ions (**Figures 5.11b-c**). Importantly, better coverage of the terminal regions of the Fd sequence by UVPD allowed improved sequencing of the N-terminally located variable domain CDR1 and CDR2, and similar coverage of the CDR3 compared to both variations of ETD; however, ETD does contribute information from several unique inter-residue sites, most notably in the diagnostic CDR3, which are not observed by UVPD, thus further emphasizing the complementary nature of these activation techniques.

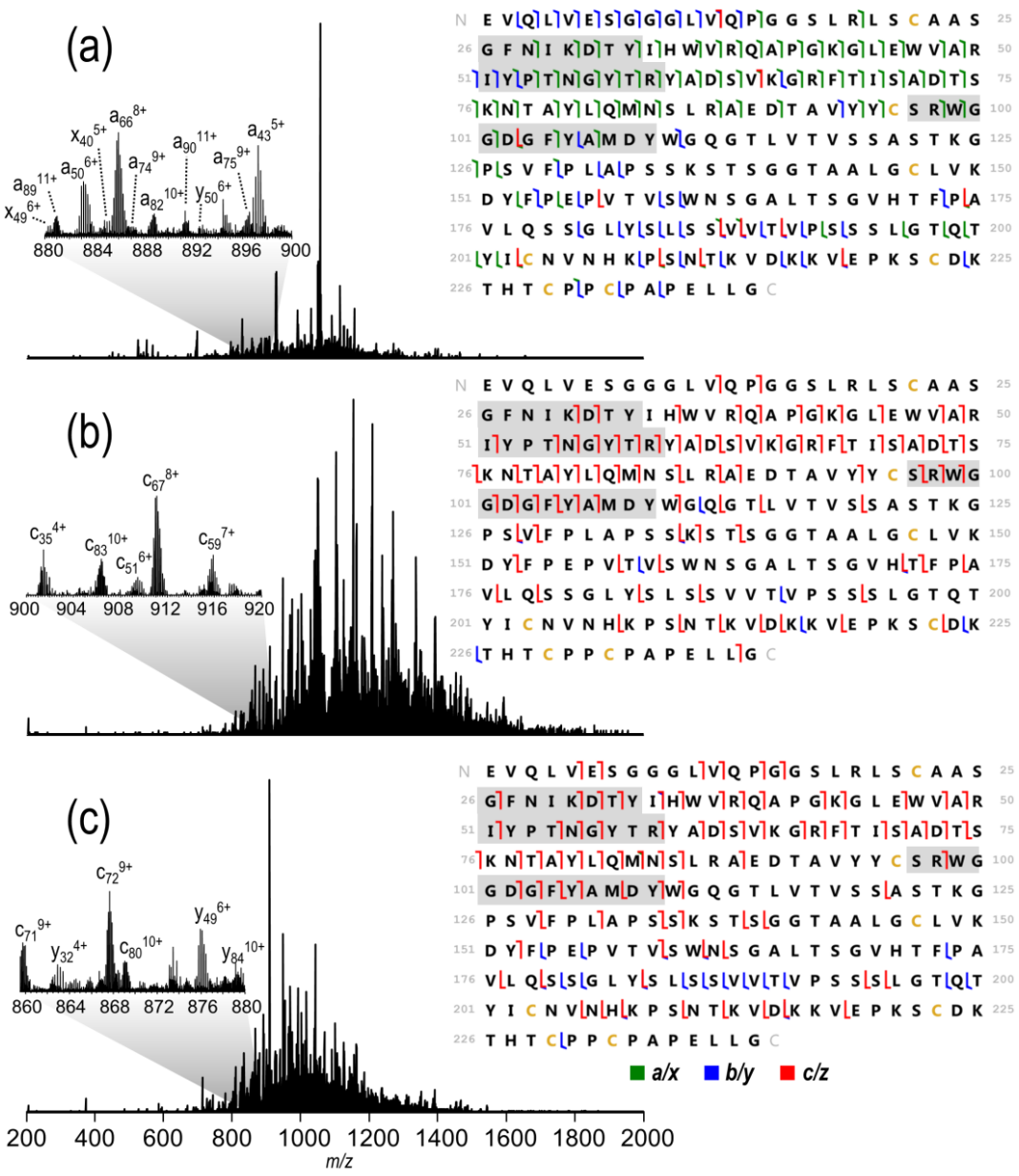


Figure 5.11 Comparative MS/MS spectra and associated fragmentation maps for the Fd subunit of trastuzumab following isolation of the most abundant precursor ion (24+) obtained using (a) UVPD (2 pulses, 2 mJ), and (b) ETD (5 ms reaction time), respectively, and (c) ETD (5 ms reaction time) with wide isolation (150 *m/z*) centered at the 27+ charge state. The insets demonstrate the product ion diversity observed for each activation strategy. The CDRs are shaded in gray. All spectra are shown as the combination of multiple scans collected at 240K resolution across the elution profile for a single targeted LC-MS/MS analysis.

5.5 CONCLUSIONS

The utility of targeted middle-down 193 nm photodissociation for improved characterization of therapeutic monoclonal antibody subunits within an LC-MS/MS workflow was demonstrated. The rapid and high energy activation afforded by UVPD resulted in cleavage at a greater number of inter-residue positions for all subunits compared to ETD performed for single charge states or spanning a range of charge states, while also maintaining the integrity of labile modifications, as demonstrated for the Fc/2 glycan. Collectively, this allowed confident glycosylation site localization in the Fc/2 heavy chain constant domain and confirmation of diagnostic CDR sequences in the variable portions of the Lc and Fd subdomains. Moreover, a degree of control over the extent of secondary dissociation was demonstrated via modulation of pulse number and laser power used for UVPD. This can be used strategically to improve coverage or enhance product ion signal-to-noise from specific regions of the backbone. For example, if a particular modification is expected to occur near one terminus of a subunit, higher energy photoactivation conditions might be preferable in order to bias fragmentation towards the production of smaller, high intensity terminal product ions. The resulting data from this kind of "customized" UVPD run can be used independently or in combination with data from other runs performed using varied pulse conditions, or with complementary broadband ETD to obtain more comprehensive coverage of the IgG sequence, as shown in this study.

While this work demonstrates an advantage to using UVPD over ETD for subunit characterization on an Orbitrap Elite mass spectrometer, we anticipate that ongoing development of newer generation Orbitrap platforms will continue to improve performance metrics for both activation methods described herein. Moreover, the use of

hybrid MS^2 techniques, such as EThcD[57,58] and ETUVPD[43] may offer compelling advantages to further improve subunit characterization within single LC-MS/MS experiments via the simultaneous generation of product ions that are both complementary and unique to each activation type, as well as enhanced conversion of both initial precursors and charge-reduced precursors into diagnostic fragment ions of analytical value.

5.6 REFERENCES

(1) Aggarwal, S. *Nat. Biotechnol.* **2014**, *32* (1), 32–39.
(2) Ecker, D. M.; Jones, S. D.; Levine, H. L. *mAbs* **2015**, *7* (1), 9–14.
(3) Nelson, A. L.; Dhimolea, E.; Reichert, J. M. *Nat. Rev. Drug Discov.* **2010**, *9* (10), 767–774.
(4) Weiner, L. M.; Surana, R.; Wang, S. *Nat. Rev. Immunol.* **2010**, *10* (5), 317–327.
(5) Scott, A. M.; Wolchok, J. D.; Old, L. J. *Nat. Rev. Cancer* **2012**, *12* (4), 278–287.
(6) Chan, A. C.; Carter, P. J. *Nat. Rev. Immunol.* **2010**, *10* (5), 301–316.
(7) Beck, A.; Haeuw, J.-F.; Wurch, T.; Goetsch, L.; Bailly, C.; Corvaïa, N. *Discov. Med.* **2010**, *10* (53), 329–339.
(8) Hughes, B. *Nat. Rev. Drug Discov.* **2010**, *9* (9), 665–667.
(9) Wang, W.; Singh, S.; Zeng, D. L.; King, K.; Nema, S. *J. Pharm. Sci.* **2007**, *96* (1), 1–26.
(10) De Groot, A. S.; Scott, D. W. *Trends Immunol.* **2007**, *28* (11), 482–490.
(11) Harding, F. A.; Stickler, M. M.; Razo, J.; DuBridge, R. B. *mAbs* **2010**, *2* (3), 256–265.
(12) Walsh, G.; Jefferis, R. *Nat. Biotechnol.* **2006**, *24* (10), 1241–1252.
(13) Kozlowski, S.; Swann, P. *Adv. Drug Deliv. Rev.* **2006**, *58* (5–6), 707–722.
(14) Janeway, C. A.; Travers, P.; Walport, M.; Shlomchik, M. J. *Immunobiology*, 5th ed.; Garland Science, 2001.
(15) Gabrielli, E.; Pericolini, E.; Cenci, E.; Ortelli, F.; Magliani, W.; Ciociola, T.; Bistoni, F.; Conti, S.; Vecchiarelli, A.; Polonelli, L. *PLoS ONE* **2009**, *4* (12), e8187.
(16) Borrok, M. J.; Jung, S. T.; Kang, T. H.; Monzingo, A. F.; Georgiou, G. *ACS Chem. Biol.* **2012**, *7* (9), 1596–1602.
(17) Zhang, Z.; Pan, H.; Chen, X. *Mass Spectrom. Rev.* **2009**, *28* (1), 147–176.
(18) Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianférani, S. *Anal. Chem.* **2013**, *85* (2), 715–736.
(19) Lundell, N.; Schreitmüller, T. *Anal. Biochem.* **1999**, *266* (1), 31–47.
(20) Srebalus Barnes, C. A.; Lim, A. *Mass Spectrom. Rev.* **2007**, *26* (3), 370–388.
(21) Kleemann, G. R.; Beierle, J.; Nichols, A. C.; Dillon, T. M.; Pipes, G. D.; Bondarenko, P. V. *Anal. Chem.* **2008**, *80* (6), 2001–2009.
(22) Fornelli, L.; Ayoub, D.; Aizikov, K.; Beck, A.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (6), 3005–3012.
(23) Wang, D.; Wynne, C.; Gu, F.; Becker, C.; Zhao, J.; Mueller, H.-M.; Li, H.; Shameem, M.; Liu, Y.-H. *Anal. Chem.* **2015**, *87* (2), 914–921.
(24) Srzentić, K.; Fornelli, L.; Laskay, Ü. A.; Monod, M.; Beck, A.; Ayoub, D.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (19), 9945–9953.
(25) Bondarenko, P. V.; Second, T. P.; Zabrouskov, V.; Makarov, A. A.; Zhang, Z. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (8), 1415–1424.
(26) Tsybin, Y. O.; Fornelli, L.; Stoermer, C.; Luebeck, M.; Parra, J.; Nallet, S.; Wurm, F. M.; Hartmer, R. *Anal. Chem.* **2011**, *83* (23), 8919–8927.

(27) Fornelli, L.; Damoc, E.; Thomas, P. M.; Kelleher, N. L.; Aizikov, K.; Denisov, E.; Makarov, A.; Tsybin, Y. O. *Mol. Cell. Proteomics* **2012**, *11* (12), 1758–1767.
(28) Mao, Y.; Valeja, S. G.; Rouse, J. C.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2013**, *85* (9), 4239–4246.
(29) Zhang, J.; Liu, H.; Katta, V. *J. Mass Spectrom.* **2010**, *45* (1), 112–120.
(30) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121* (4), 806–812.
(31) Zhou, H.; Ning, Z.; E. Starr, A.; Abu-Farha, M.; Figeys, D. *Anal. Chem.* **2012**, *84* (2), 720–734.
(32) Zhang, Z.; Shah, B. *Anal. Chem.* **2007**, *79* (15), 5723–5729.
(33) Cannon, J.; Lohnes, K.; Wynne, C.; Wang, Y.; Edwards, N.; Fenselau, C. *J. Proteome Res.* **2010**, *9* (8), 3886–3890.
(34) 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. *Nat Meth* **2012**, *9* (8), 822–824.
(35) Pang, Y.; Wang, W.-H.; Reid, G. E.; Hunt, D. F.; Bruening, M. L. *Anal. Chem.* **2015**, *87* (21), 10942–10949.
(36) Nicolardi, S.; Deelder, A. M.; Palmblad, M.; van der Burgt, Y. E. M. *Anal. Chem.* **2014**, *86* (11), 5376–5382.
(37) Yan, B.; Valliere-Douglass, J.; Brady, L.; Steen, S.; Han, M.; Pace, D.; Elliott, S.; Yates, Z.; Han, Y.; Balland, A.; Wang, W.; Pettit, D. *J. Chromatogr. A* **2007**, *1164* (1–2), 153–161.
(38) An, Y.; Zhang, Y.; Mueller, H.-M.; Shameem, M.; Chen, X. *mAbs* **2014**, *6* (4), 879–893.
(39) Tran, J. C.; Zamdborg, L.; Ahlf, D. R.; Lee, J. E.; Catherman, A. D.; Durbin, K. R.; Tipton, J. D.; Vellaichamy, A.; Kellie, J. F.; Li, M.; Wu, C.; Sweet, S. M. M.; Early, B. P.; Siuti, N.; LeDuc, R. D.; Compton, P. D.; Thomas, P. M.; Kelleher, N. L. *Nature* **2011**, *480* (7376), 254–258.
(40) Garcia, B. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (2), 193–202.
(41) Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L. *Anal. Chem.* **2011**, *83* (17), 6868–6874.
(42) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(43) Cannon, J. R.; Holden, D. D.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (21), 10970–10977.
(44) Cannon, J. R.; Kluwe, C.; Ellington, A.; Brodbelt, J. S. *PROTEOMICS* **2014**, *14* (10), 1165–1173.
(45) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(46) Vasicek, L. A.; Ledvina, A. R.; Shaw, J.; Griep-Raming, J.; Westphall, M. S.; Coon, J. J.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 1105–1108.

(47) Bremer, E. T. van den; Beurskens, F. J.; Voorhorst, M.; Engelberts, P. J.; Jong, R. N. de; Boom, B. G. van der; Cook, E. M.; Lindorfer, M. A.; Taylor, R. P.; Berkel, P. H. van; Parren, P. W. *mAbs* **2015**, *7* (4), 672–680.
(48) Michalski, A.; Damoc, E.; Lange, O.; Denisov, E.; Nolting, D.; Müller, M.; Viner, R.; Schwartz, J.; Remes, P.; Belford, M.; Dunyach, J.-J.; Cox, J.; Horning, S.; Mann, M.; Makarov, A. *Mol. Cell. Proteomics* **2012**, *11* (3), O111.013698.
(49) Holden, D. D.; McGee, W. M.; Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 1008–1016.
(50) Roy, J. *Trends Glycosci. Glycotechnol.* **2009**, *21* (118), 105–117.
(51) Solá, D. R. J.; Griebenow, K. *BioDrugs* **2012**, *24* (1), 9–21.
(52) Zhang, L.; Reilly, J. P. *J. Proteome Res.* **2009**, *8* (2), 734–742.
(53) Ko, B. J.; Brodbelt, J. S. *Int. J. Mass Spectrom.* **2015**, *377*, 385–392.
(54) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.
(55) Breuker, K.; Oh, H.; Horn, D. M.; Cerda, B. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124* (22), 6407–6420.
(56) Riley, N. M.; Westphall, M. S.; Coon, J. J. *Anal. Chem.* **2015**, *87* (14), 7109–7116.
(57) Frese, C. K.; Altelaar, A. F. M.; van den Toorn, H.; Nolting, D.; Griep-Raming, J.; Heck, A. J. R.; Mohammed, S. *Anal. Chem.* **2012**, *84* (22), 9668–9673.
(58) Brunner, A. M.; Lössl, P.; Liu, F.; Huguet, R.; Mullen, C.; Yamashita, M.; Zabrouskov, V.; Makarov, A.; Altelaar, A. F. M.; Heck, A. J. R. *Anal. Chem.* **2015**, *87* (8), 4152–4158.

Chapter 6

High-Throughput Bioconjugation for Enhanced 193 nm Photodissociation via Droplet Phase Initiated Ion/Ion Chemistry using a Front-end Dual Spray Reactor*

6.1 OVERVIEW

Fast on-line chemical derivatization of peptides with an aromatic label for enhanced 193 nm ultraviolet photodissociation (UVPD) is demonstrated using a dual electrospray reactor implemented on the front-end of a linear ion trap (LIT) mass spectrometer. The reactor facilitates the intersection of protonated peptides with a second population of chromogenic 4-formyl-1,3-benzenedisulfonic acid (FBDSA) anions to promote real-time formation of ion/ion complexes at atmospheric pressure. Subsequent collisional activation of the ion/ion intermediate results in Schiff base formation generated via reaction between a primary amine in the peptide cation and the aldehyde moiety of the FBDSA anion. Utilizing 193 nm UVPD as the subsequent activation step in the MS^3 workflow results in acquisition of greater primary sequence information relative to conventional collision induced dissociation (CID). Furthermore, Schiff base modified peptides exhibit on average a 20% increase in UVPD efficiency compared to their un-modified counterparts. Due to the efficiency of covalent labeling achieved with the dual spray reactor, we demonstrate that this strategy can be integrated into a high-throughput LC-MS^n workflow for rapid derivatization of peptide mixtures.

---

*Cotham, V. C.; Shaw, J. B.; Brodbelt, J. S. *Anal. Chem.* **2015**, 87, 9396-9402.
V.C.C. designed and conducted all experiments.

6.2 INTRODUCTION

Bioconjugation techniques have long been used to extend the versatility of mass spectrometry (MS) for proteomic analysis.[1,2] The number of reactive moieties incorporated into polymeric amino acid chains makes them amenable to a diverse array of site-selective chemistries, which have been exploited in both MS and MS^n modes for purposes such as enhancing ionization efficiencies,[3,4] incorporating isotopic labels for quantification,[5–7] modulating fragmentation,[8–10] and promoting selective dissociation.[11–13] Despite their utility, derivatization reactions are often the rate-limiting step in MS-based workflows since most require off-line solution phase chemistry prior to mass spectrometric analysis. Recently, this shortcoming has been addressed by several compelling strategies that utilize the mass spectrometer as a tool to facilitate rapid functional group derivatization analogous to that performed in bulk solution. These strategies apply the principles of two fundamentally distinct chemical platforms: 1) microdroplet chemistry[14] and 2) gas phase ion/ion chemistry.[15] Other approaches, including the use of theta capillaries and a sheath gas reaction mode, have been used to modulate electrospray ionization in an on-line manner via fast mixing of droplets or exposure of droplets to gaseous acids or bases.[16–20]

MS-based microdroplet chemistry exploits the interfacial region between the ion source and the vacuum inlet of the mass spectrometer to promote rapid chemical modification during the ionization process.[14,21–26] Cooks and co-workers demonstrated that the confined volume of a charged evaporating droplet acted as a microreactor for heterogeneous spray mixtures and promotes accelerated bond formation relative to bulk solution.[23–25] The enhanced reaction kinetics were attributed to extremes in pH and concentration within the shrinking droplet environment, as well as increased collision frequencies at atmospheric pressure.[23] This phenomenon was most prominent at sub-

nanospray volumes, such as that of secondary droplets formed during reactive desorption electrospray ionization (rDESI).[24] Strategies incorporating on-line microdroplet derivatization have primarily been limited to small molecule applications, with only a select few studies adapting these methods for peptide or protein analysis.[14] These exceptions include the in-situ crosslinking of primary amines in single peptides with bis(sulfosuccinimidyl) suberate (BS3) using reactive DESI[25] and online dithiothreitol (DTT) reduction of disulfides in intact proteins via reactive electrospray-assisted laser desorption/ionization (rELDI).[27] Although these techniques established the feasibility of on-line microdroplet-assisted modification of several functionalities in polymeric amino acid chains, the potential for applying these methods to high-throughput proteomic mixtures remains largely unexplored.

Gas phase ion/ion reactions represent the second major MS-based platform for on-line bioconjugation of peptides and proteins. These strategies move the chemical reaction step from the atmospheric pressure interface into the vacuum chamber of the mass spectrometer.[15] Implementation of such methods requires the use of sophisticated instrumentation with bipolar trapping capabilities for mutual storage of populations of oppositely charged reagent and analyte ions.[28] To date, a number of covalent chemistries have been used to selectively functionalize moieties of polypeptide chains in the gas phase, including Schiff base and NHS-ester modification of primary amines,[29–31] carbodiimide derivatization of carboxylic acids,[32] as well as directed peptide bond formation via sulfo-NHS ester N-terminal coupling.[33] Unlike analogous condensed phase reactions, which offer limited control over reacting species and are thus subject to undesirable side products, *in vacuo* ion/ion reactions afford fine-tuned selection of reactant ions and a high degree of conversion into the products of interest.[15] Despite these merits, complete gas-phase transformation often comes at the cost of long activation

times (>100 ms) that are not amenable to on-line chromatographic separations, thereby limiting analysis to low complexity samples.

In this Chapter we demonstrate a method that combines the principles of droplet phase and gas phase ion/ion chemistry to promote rapid bioconjugation using a front-end dual spray reactor. The reactor is designed to be readily coupled to LC platforms for high-throughput proteomic applications that were either not possible or unexplored using previously described droplet phase or gas phase strategies alone. To demonstrate the utility of the reactor for droplet-phase initiated ion/ion reactions, we borrow from covalent Schiff base chemistry that has been extensively explored in the gas phase.[29,30,34,35] We expand on this further by taking advantage of the chromophore addition at peptide N-termini following Schiff base formation with 4-formyl-1,3-benzenedisulfonic acid (FBDSA) to enhance photodissociation at 193 nm.[36]

6.3 EXPERIMENTAL

6.3.1 Chemicals and Materials

Model peptides KMVELVHFL, KLVANNTRL, RPPGFSPFR, ASHLGLAR, DRVYIHPFHLVIHN and DAEFRHDSGYQVHHQK were purchased from AnaSpec Inc. (Fremont, CA). SYSMEHFRWG was purchased from American Peptide Company (Sunnyvale, CA). DRVYIHPFHL and 4-formyl-1,3-benzenedisulfonic acid (FBDSA) were obtained from Sigma-Aldrich (St. Louis, MO). Peptides were used without purification. All other solvents were purchased from Fisher Scientific (Fairlawn, NJ).

6.3.2 Implementation of a Front-end Dual Spray Reactor

Design and fabrication of the dual spray reactor mounting system was performed in house according to the illustration shown in **Figure 6.1**. The reactor was equipped with two electrosonic spray ionization (ESSI) sources (Prosolia Inc., Indianapolis, IN) that were integrated on the front-end of the mass spectrometer by way of a free-standing mount attached to a U-shaped rail that surrounded the MS inlet. The curved rail allowed for independent adjustment of the angle between the two sources in addition to the angle of each source relative to the axis of the mass spectrometer. Both sources were fixed to sliding supports that enabled adjustment of the y-dimensional distance between the emitter tip and the MS inlet. These supports were further mounted to a precision 1D translational stage for positional control in the z-direction. The first source was completely integrated, allowing direct control of polarity and spray voltage through the instrument control software. The spray voltage of the second source was supplied by an external 5 kV high voltage power supply (Stanford Research Systems Inc., Sunnyvale, CA) operated in negative polarity. Nebulizing sheath gas was introduced using an external nitrogen line equipped with a tee fitting and adjustable metering valve to evenly split the gas flow to both sources and allow manipulation of gas flow rates, respectively.

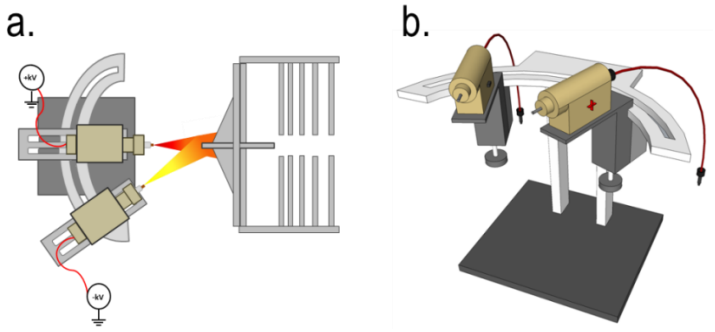


Figure 6.1 (a) Aerial view of dual source reactor mounted at the front-end of a mass spectrometer and (b) free-standing design for facile adaptation to multiple instrument platforms.

6.3.3 Ion/Ion Chemistry, Mass Spectrometry and Photodissociation

All experiments were conducted on a Thermo Velos Pro dual linear ion trap mass spectrometer (San Jose, CA) outfitted with a GAM EX5 or Coherent Excistar XS 500 Hz ArF excimer laser operated at 193 nm. The back flange of the mass spectrometer was modified to allow introduction of the laser beam coaxial to the dual cell linear ion trap through a CaF2 window and 2 mm stainless steel aperture, as previously described.[37] In all experiments, peptide cations were generated by positive mode electrospray ionization of 5 μM peptide working solutions prepared in equal parts water and acetonitrile containing 0.1% formic acid and infused at rate of 1.5 μL/min. For ion/ion reactions, a second ion population consisting of a large excess of reagent anions was produced by negative mode electrospray of 1 mM FBDSA in 50:50 water/methanol at an infusion rate of 3 μL/min. The peptide source was positioned on axis to the inlet of the mass spectrometer; whereas the reagent ion source was position approximately 45 degrees off axis. Complex formation was accomplished at atmospheric pressure by intersecting analyte sprays of opposite polarity in the reaction region prior to the capillary inlet. The spray voltage for the cation source was held constant at 1.75 kV, while the anion spray voltage was varied between -2.5 and -3 kV depending on optimal complex formation. Electrostatic adducts formed in the overlapping sprays were isolated and subjected to low energy collisional activation using a normalized collision energy (NCE) between 10-20% to promote loss of water and formation of the stable Schiff base imine product. Isolation and MS^3 activation of the Schiff base was carried out using either collisional induced dissociation (CID) or 193 nm ultraviolet photodissociation (UVPD). MS^3 experiments based on UVPD were performed using a single 2 mJ, 5 ns laser pulse. A q-value of 0.125 was used to extend the low m/z cutoff. UVPD efficiencies were calculated as previously described.[36]

6.3.4 High-Throughput Bioconjugation on an LC Timescale

The dual spray reactor was coupled to a Dionex Ultimate 3000 (Sunnyvale, CA) capillary flow system for all high-throughput bioconjugation experiments. Peptides were separated on an Agilent ZORBAX 300 Extend-C18 column (150 x 0.3 mm, 3.5 μm particle size) held at a constant temperature of 30˚C. Eluent A consisted of 0.1% formic acid in water and eluent B was 0.1% formic acid in acetonitrile. A linear gradient from 3% to 35% B over 30 min at 4 μL/min was used. An auxiliary syringe method was programed to infuse FBDSA at a rate of 3 μL/min throughout the course of separation. The spray voltage for the cation and anion source was held constant at 1.75 kV and -2.5 kV, respectively. Rapid derivatization was accomplished using an automated data-dependent neutral loss (DDNL) MS^3 program similar to that described by Gygi *et al.*[38] The top five most abundant ions in the full MS scan were subjected to low energy collisional activation (18% NCE, isolation width of 5 Th, *q*-value 0.15) for 50 ms. MS^3 UVPD (1 pulse, 2 mJ, *q*-value of 0.125) was triggered if a neutral loss product was detected at -18 Da (-18 Th, -9 Th and -6 Th to account for the 1+, 2+ and 3+ charge states, respectively) and was within the top 3 most abundant ions in the MS^2 spectrum. All data was filtered such that grouped MS^2 and MS^3 scans were manually interrogated.

6.4 RESULTS AND DISCUSSION

Covalent modification in the gas phase has been shown to occur via stable, long-lived electrostatic intermediates formed between polypeptide cations and bi-functional reagent anions.[15,29,31] Within this context, bi-functionality refers to the ability of reagent anions to engage in stabilizing non-covalent interactions that favor complex formation over competing proton transfer, as well as facilitate functional group derivatization.[15] Han and McLuckey reported the first example of gas phase bioconjugation within the confines

of an electrodynamic ion trap mass spectrometer using the aldehyde-containing reagent 4-formyl-1,3-benzenedisulfonic acid (FBDSA) to covalently modify primary amines of peptide cations via Schiff base ion/ion chemistry.[29] The resulting derivatized peptides were found to yield more informative fragmentation upon collisional activation compared to their unmodified counterparts. Similarly, gas phase FBDSA-based chemistry has been used to enhance collisional dissociation in the negative ion mode for charge inverted Schiff base modified peptides.[34,35,39] These seminal studies demonstrated the feasibility of online ion/ion mediated bioconjugation, as well as its utility for improving the structural characterization of modified species. Despite the improved throughput of this online approach relative to orthogonal in-solution chemistry, the complex scan functions and average reaction times (50-1000 msec) required for formation of electrostatic intermediates *in vacuo*, followed by covalent conversion and subsequent activation to obtain structurally relevant information, have rendered these methods not well adapted for chromatographic timescales. These limitations have prompted our efforts to develop a streamlined approach that utilizes a front-end dual spray reactor to initiate ion/ion reactions in the droplet phase prior to introduction into the mass spectrometer. To benchmark the performance of the dual spray reactor, we employed aforementioned Schiff base reactions between peptide cations and FBDSA anions for comparison with previously reported gas phase ion/ion covalent chemistry.

The process for dual spray assisted bioconjugation of peptides with FBDSA is summarized in **Figure 6.2**. The reactor utilizes two ESI sources biased at opposite polarity to simultaneously generate overlapping populations of peptide cations and FBDSA anions. This configuration was designed to maximize collisions between reactive species in the droplet, pseudo-droplet, and gas phase at the high pressure interface of the

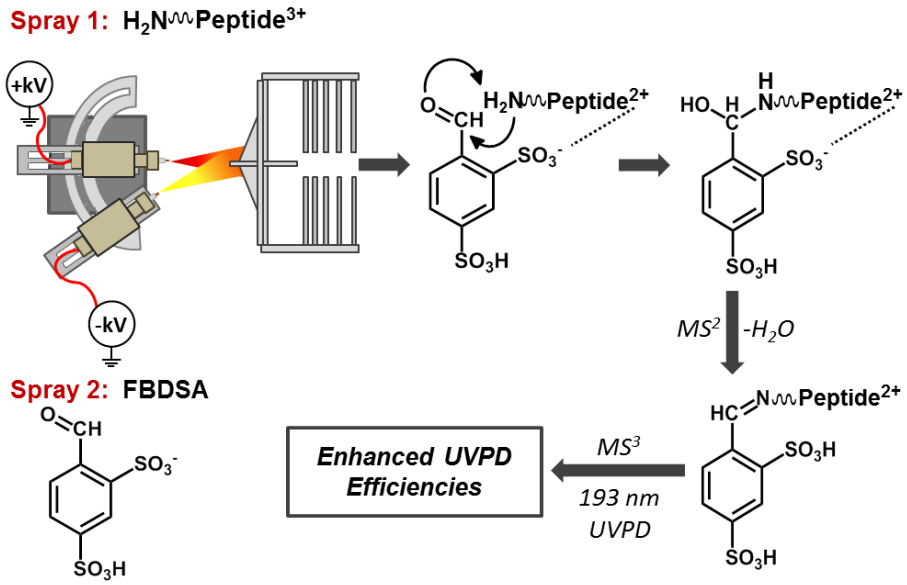


Figure 6.2 Schematic summary of dual spray initiated bioconjugation of a peptide cation with an FBDSA anion for enhanced ultraviolet photodissociation.

mass spectrometer, similar in concept to the Y-shaped reactor inlet used to merge ions from two independently biased sources in early ion/ion proton transfer reactions,[40] or that used in extractive electrospray (EESI) configurations.[41,42] The role of the sulfonate groups of FBDSA is two-fold: first, they provide acidic sites that are readily deprotonated under negative electrospray conditions to form 1- and 2- anions (**Figure 6.3a**), and secondly they engage in stabilizing acid-base interactions with multiply charged peptide cations to form charge-reduced non-covalent complexes that persist into the gas phase. This process is demonstrated in **Figure 6.3**, which compares the MS^1 spectrum of the peptide DRVYIHPFHLVIHN before (**Figure 6.3b**) and after dual spray infusion with anionic FBDSA (**Figure 6.3c**). The major products in the post ion/ion reaction spectrum are charge-reduced relative to the unmodified peptide and arise from partial neutralization of multiply charged peptide cations (up to five sites of protonation) with singly and doubly deprotonated FBDSA anions. This is clearly indicated by the absence of the 5+ charge

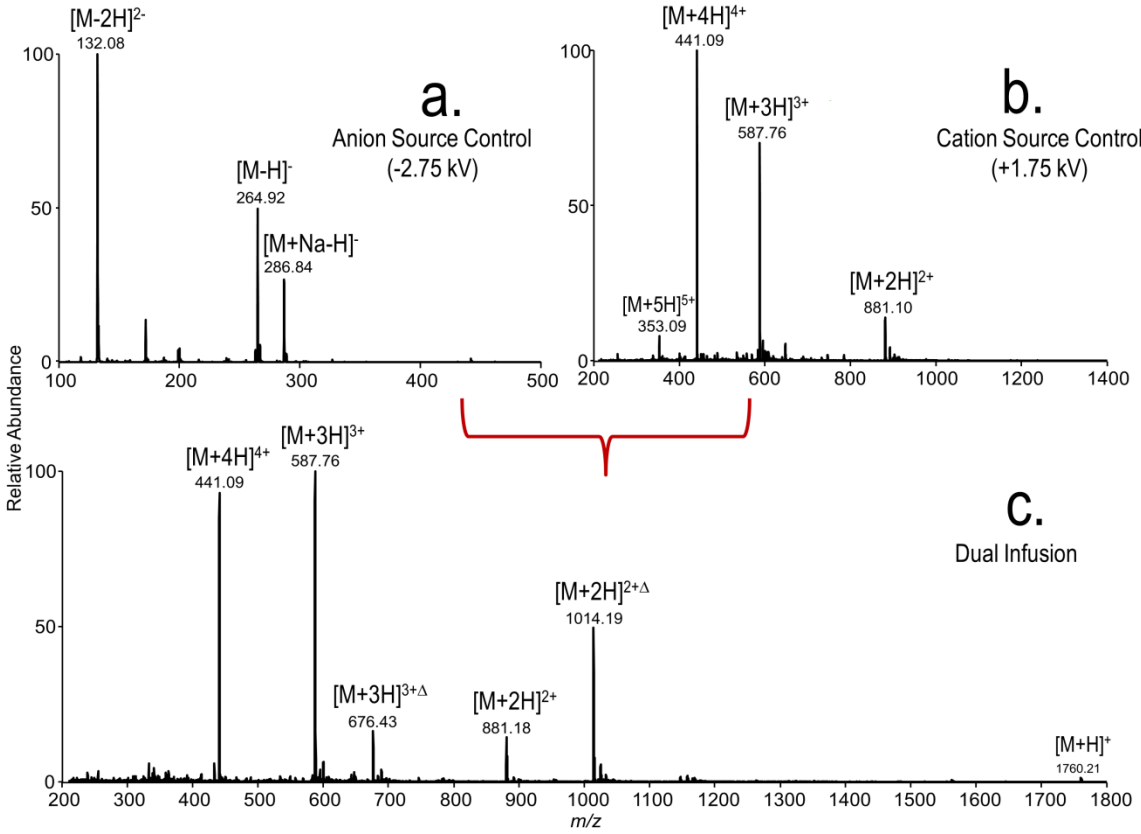


Figure 6.3 (a) Negative mode ESI control spectrum for FBDSA and (b) positive mode ESI control spectrum for DRVYIHPFHLVIHN. (c) Dual spray spectrum for simultaneously infused DRVYIHPFHLVIHN (+) and FBDSA (-). Electrostatic DRVYIHPFHLVIHN/FBDSA complexes are denoted by Δ.

state and attenuation of the relative abundances of the 3+ and 4+ charge states in the resulting dual spray spectrum. Moreover, new products are observed at *m/z* values consistent with the formation of 2+ and 3+ peptide/FBDSA complexes as indicated by a mass shift of +266 Da (denoted by the "Δ" symbol) relative to the unmodified peptide. As shown in **Figure 6.4**, a direct relationship was observed between the magnitude of the anion source voltage and the relative contribution of complexes to the total ion current of the dual spray spectrum. The absence of peptide/FBDSA complexes when the voltage is

set to zero (akin to ion/molecule reaction conditions), in addition to the observed shift towards more abundant complex formation as the anion source voltage is stepped to increasingly negative potentials confirms that the reaction occurs exclusively through an ion/ion mediated pathway.

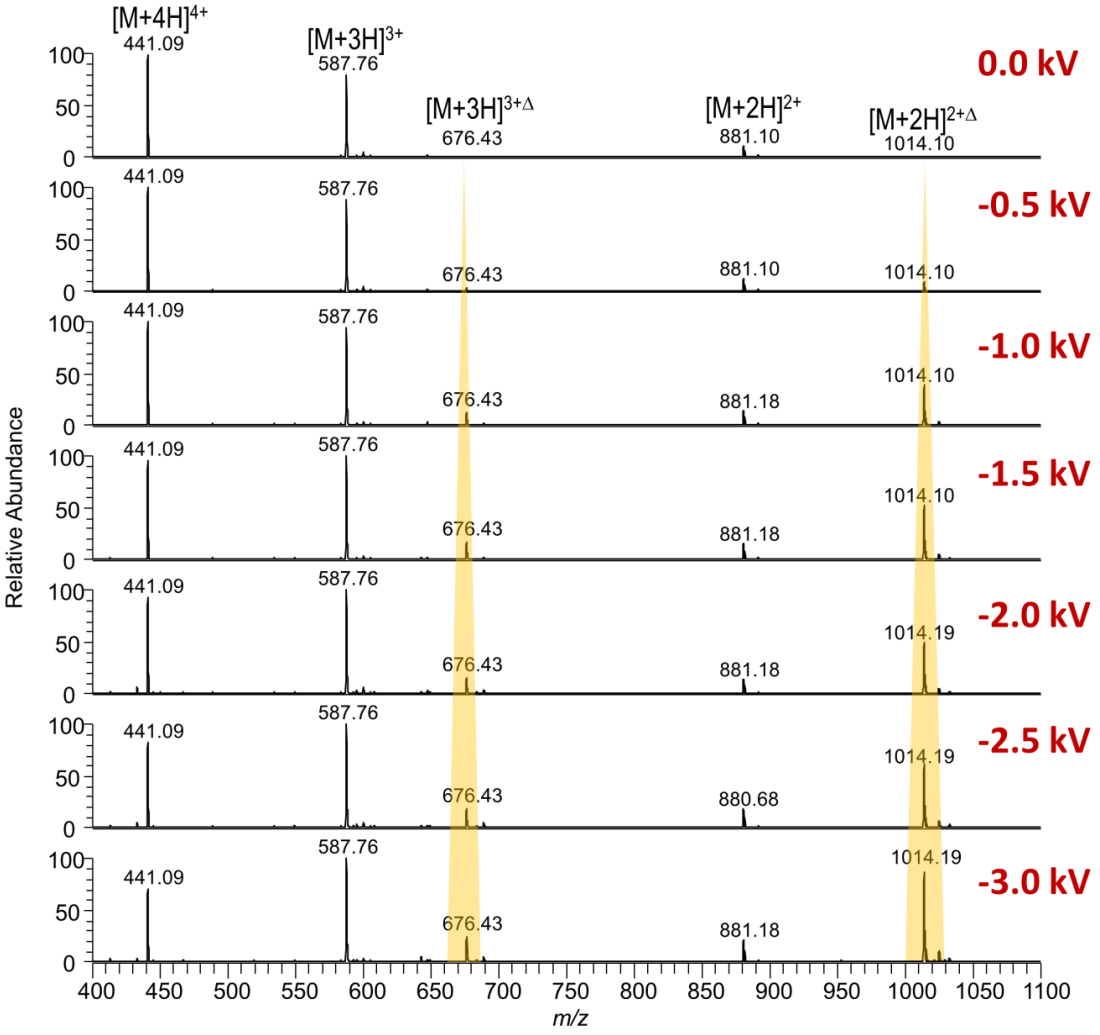


Figure 6.4 Anion source voltage optimization: normalized abundance of unmodified DRVYIHPFHLVIHN and DRVYIHPFHLVIHN/FBDSA complex as a function of anion source voltage.

An example of an optimized dual spray spectrum is shown in **Figure 6.5** for FBDSA-reacted DRVYIHPFHLVIHN. Here, signal from the non-covalent complex accounts for approximately 37% of the total ion current, thus demonstrating the efficiency of the front-end ion/ion reaction. It should be noted that competing proton transfer reactions may also contribute to the observed attenuation of peptide charge state; however, an approach for isolating and quantifying this contribution relative to complex formation is beyond the scope of this work.

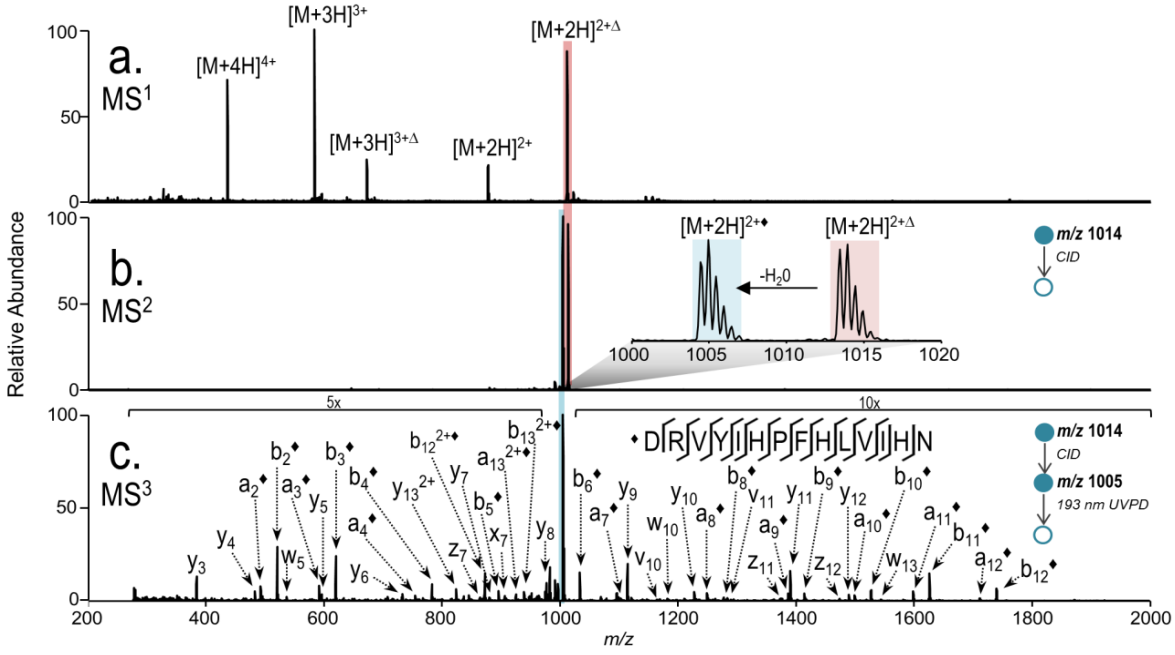


Figure 6.5 Process for on-line modification of peptides with FBDSA using a dual spray reactor. (a) Electrostatic complexes are formed at atmospheric pressure between multiply charged peptide cations and FBDSA anions. These ion/ion intermediates are denoted with a "Δ" superscript. (b) Low energy collisional activation of the intermediate ion/ion species promotes Schiff base formation via the concerted formation of an imine bond and loss of a water molecule, resulting in a chromophore-labeled Schiff base product (♦) with a mass shift of -18 Da relative to the electrostatic complex. (c) 193 nm UVPD of the labeled peptide exhibits extensive backbone fragmentation.

Following optimization of the front-end ion/ion reaction, the second step of the process requires isolation and gentle collisional activation of the electrostatic complex to promote covalent bioconjugation. This input of energy into the system is likely required to overcome the activation barrier necessary for nucleophilic attack of the FBDSA aldehyde by the free N-terminus of the peptide. Since water represents the major byproduct of imine formation, the dominant water loss product observed in the MS^2 spectrum (**Figure 6.5b**) is highly indicative of covalent Schiff base derivatization. An additional ion activation event results in an MS^3 spectrum that confirms Schiff base formation (**Figure 6.5c**). Unlike previously reported *in vacuo* methods, which accomplish MS^3 by using a second collisional activation step, we integrate 193 nm ultraviolet photodissociation (UVPD) as an alternative activation method to both verify Schiff base formation and achieve comprehensive structural characterization of resulting modified peptides. The use of 193 nm UVPD is strategic for several reasons: 1) UVPD can be accomplished in a much shorter activation period than required for CID, thus improving the throughput of the analysis, 2) in many cases, UVPD provides more extensive coverage of the peptide backbone compared to CID,[43] and finally 3) N-terminal modification of peptides with chromogenic labels, such as FBDSA, has been shown to enhance UVPD efficiencies,[36] thereby offering an additional metric by which to evaluate covalent attachment. The exceptional sequence coverage of the peptide backbone afforded by 193 nm UVPD is demonstrated in **Figure 6.5c**. The combined information obtained by the extensive array of complementary N- and C-terminal ions (in this case *a*/*b* and *x*/*y*/*z*, respectively) allows for unambiguous localization of Schiff base modification at the peptide N-terminus. This lack of ambiguity arises from the fact that all N-terminally derived ions exhibit a conserved mass shift of +248 Da (denoted in the spectrum by the addition of "♦" to the ion labels), whereas the entire set of

complementary C-terminal product ions remain unmodified. The differentiation of several isobaric leucine and isoleucine residues across the peptide backbone is also possible by unique *v*- and *w*-type side-chain ions produced upon UV activation. Moreover, the non-resonant nature of UVPD allows access to a lower m/z trapping limit, whereas CID suffers from a low mass cutoff (LMCO) restriction imposed by the RF amplitude applied to the trap. By overcoming this limitation, UVPD provides greater depth of coverage for high *m/z* (low charge) precursor ions commonly observed following the formation of charge-reduced peptide/FBDSA complexes during front-end ion/ion reactions.

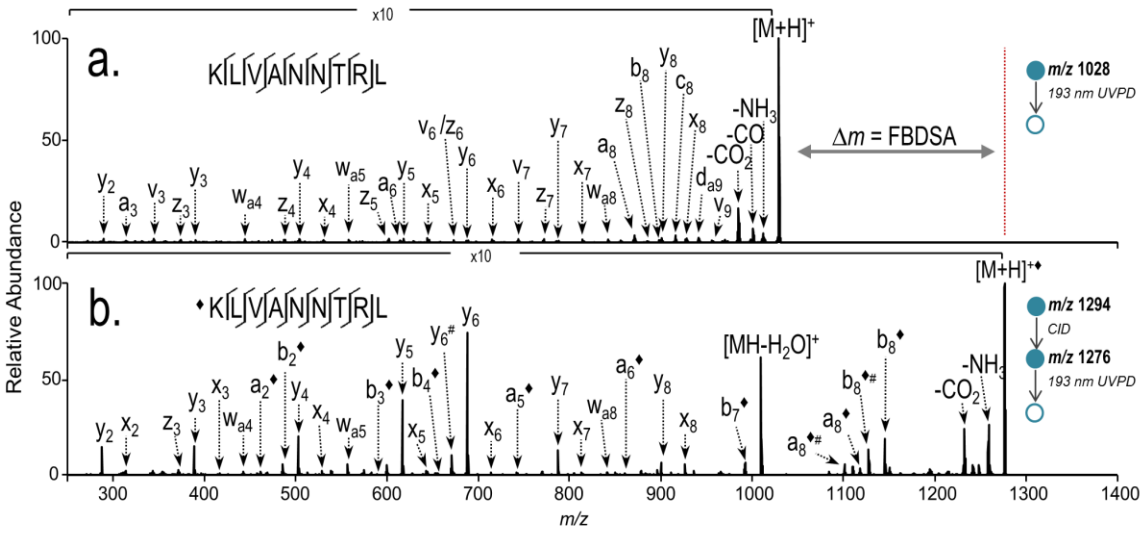


Figure 6.6 Comparison of 193 nm UVPD efficiencies before and after Schiff base modification of KLVANNTRL (1+). (a) MS^2 UVPD mass spectrum for unlabeled peptide and (b) MS^3 UVPD mass spectrum following on-line derivatization using the dual spray reactor.

Previous work in our lab demonstrated enhanced 193 nm photodissociation of peptides containing native aromatic residues,[43] as well as for those whose intrinsic

aromaticity was increased through N-terminal derivatization with chromogenic labels.[36] Based on these findings, we anticipated that front-end ion/ion mediated N-terminal modification with chromogenic FBDSA would enhance the UVPD dissociation behavior of peptides. To test this hypothesis, UVPD was evaluated for a series of model peptides in both the unmodified and FBDSA-labeled states. Notably, spectral changes were most prominent for peptides lacking intrinsic aromatic residues such as in the case of KLVANNTRL (**Figure 6.6**). Despite complete backbone coverage of the unmodified peptide, the fragment ion abundance only accounts for approximately 17% of the total ion intensity, therefore indicating rather poor dissociation of the singly charged precursor (**Figure 6.6a**). Alternatively, photoactivation of the Schiff base modified form of this peptide showed nearly 27% increase in the resulting photodissociation efficiency (**Figure 6.6b**), as indicated by a significant increase in fragment ion intensity relative to the surviving precursor. Similar results were obtained for all peptides investigated in this study, with Schiff base modification accounting for an average enhancement in photodissociation efficiency of approximately 20% (**Figure 6.7**). This value is highly consistent with previous findings for peptides modified with SPITC and PPITC chromogenic N-terminal labeles.[36] To verify that this change in dissociation behavior is the direct result of enhanced gas-phase photoabsorption as opposed to changes in the critical energies of the modified peptides, an energy variable collisional activation analysis was performed on unmodified and Schiff base labeled KLVANNTRL (1+). **Figure 6.8** shows the normalized precursor ion intensity for both forms of KLVANNTRL (1+) as a function of increasing collision energy. The high degree of overlap in the variable CID profiles suggests that FBDSA predominantly enhances the photoexcitation energy, while having no appreciable impact on the critical energy of the peptide. Despite improved dissociation efficiency, decreased UVPD sensitivity is

possible for FBDSA-labeled peptides due to incomplete reaction conversion to the Schiff base product.

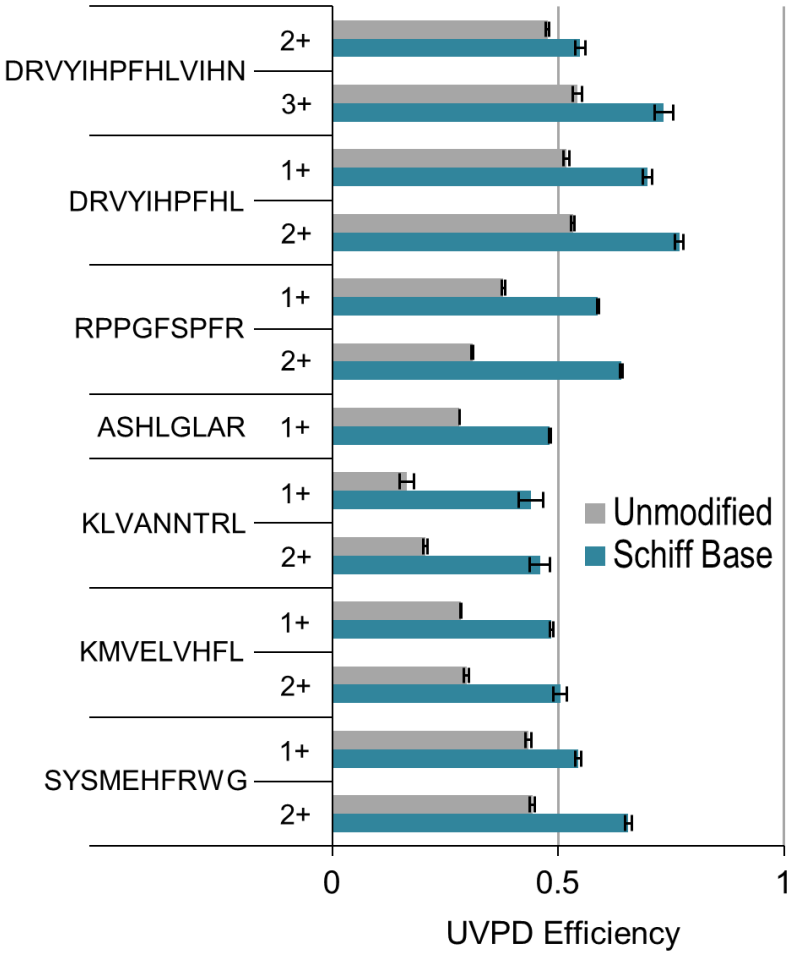


Figure 6.7 Comparison of 193 nm UVPD efficiencies for a set of model peptides before and after Schiff base modification.

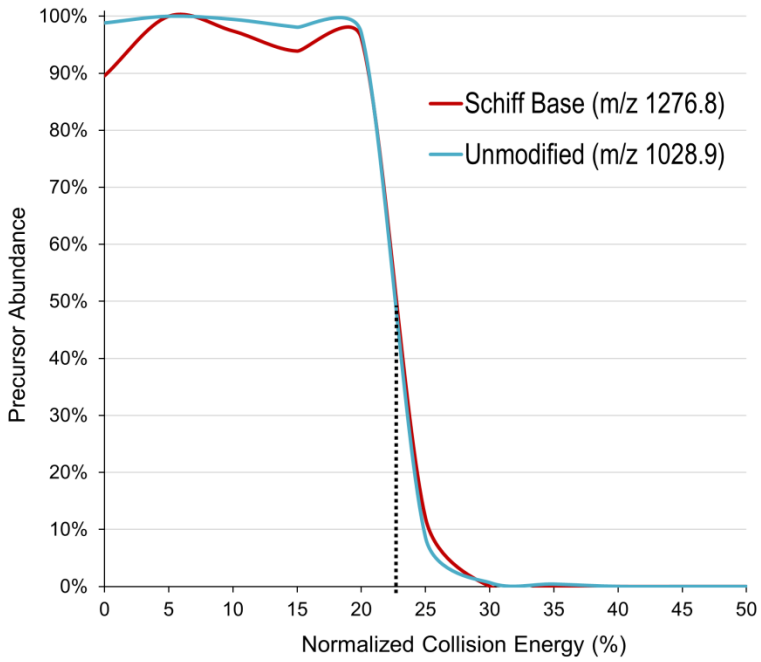


Figure 6.8 Variable energy collision analysis of unmodified and Schiff base labeled KLVANNTRL (1+). Normalized precursor abundances are plotted as a function of increasing collision energy. The dashed line represents the point of 50% precursor dissociation.

Finally, we wanted to assess the feasibility of coupling the dual spray reactor to a liquid chromatograph in order to show the first example of high-throughput ion/ion mediated bioconjugation on a chromatographic timescale. For these proof-of-principle studies, a five peptide mixture was utilized to simulate the complexity of a small, single protein digest. Unlike direct infusion experiments, which allow source parameters to be independently adjusted in order to maximize complex abundance, integration of the reactor into a high-throughput workflow requires using a fixed set of source parameters to generate sufficiently high levels of ion/ion complex for selection during automated data-dependent acquisition. **Figure 6.9** provides a comparison of spectra for each peptide following LC-MS analysis of the mixture under control (**Figure 6.9a-e**) and dual source

reaction conditions (**Figure 6.9f-j**). In all cases, spectral differences between the samples arise from the presence of electrostatic complexes. Interestingly, the degree of complex formation appears greatest for early eluting peptides where the mobile phase composition is predominantly aqueous. This corresponds to larger droplet sizes on the basis of the higher surface tension of water relative to organic solvents, and thus more dilute droplet conditions. Contrary to microdroplet reaction theory, which states that accelerated rates of reaction are observed as droplet size decreases,[24] our results point to a greater emphasis on stabilizing the electrostatic interactions for ion/ion mediated reactions as opposed to increasing the relative concentration of reactants per unit volume. Once formed and transferred into gas phase, the extent to which complexes are converted to covalent Schiff base products is highly dependent on the amount of time allotted for the reaction to proceed in the MS^2 step via low energy collisional activation.). Alternatively, rapid activation at high collision energies (NCE>25%) resulted in dominant loss of FBDSA (data not shown). This observation indicates that covalent conversion occurs optimally under slow heating conditions as opposed to fast energy transfer. Efficient conversion within narrow elution windows is critical for compatibility with front-end separations, and the extent to which this conversion occurs has an immediate impact on the quality of subsequent UVPD spectra necessary for characterization of Schiff base labeled peptides. For all non-covalent complexes observed in these experiments, a reaction time of 50 ms combined with 18% normalized collision energy (NCE) resulted in predominant conversion to Schiff base products. These MS^2 spectra were easily distinguished based on the abundance of the water loss product; whereas activation at 18% NCE was sufficiently high to promote some extent of backbone cleavage for unreacted peptides (data not shown).

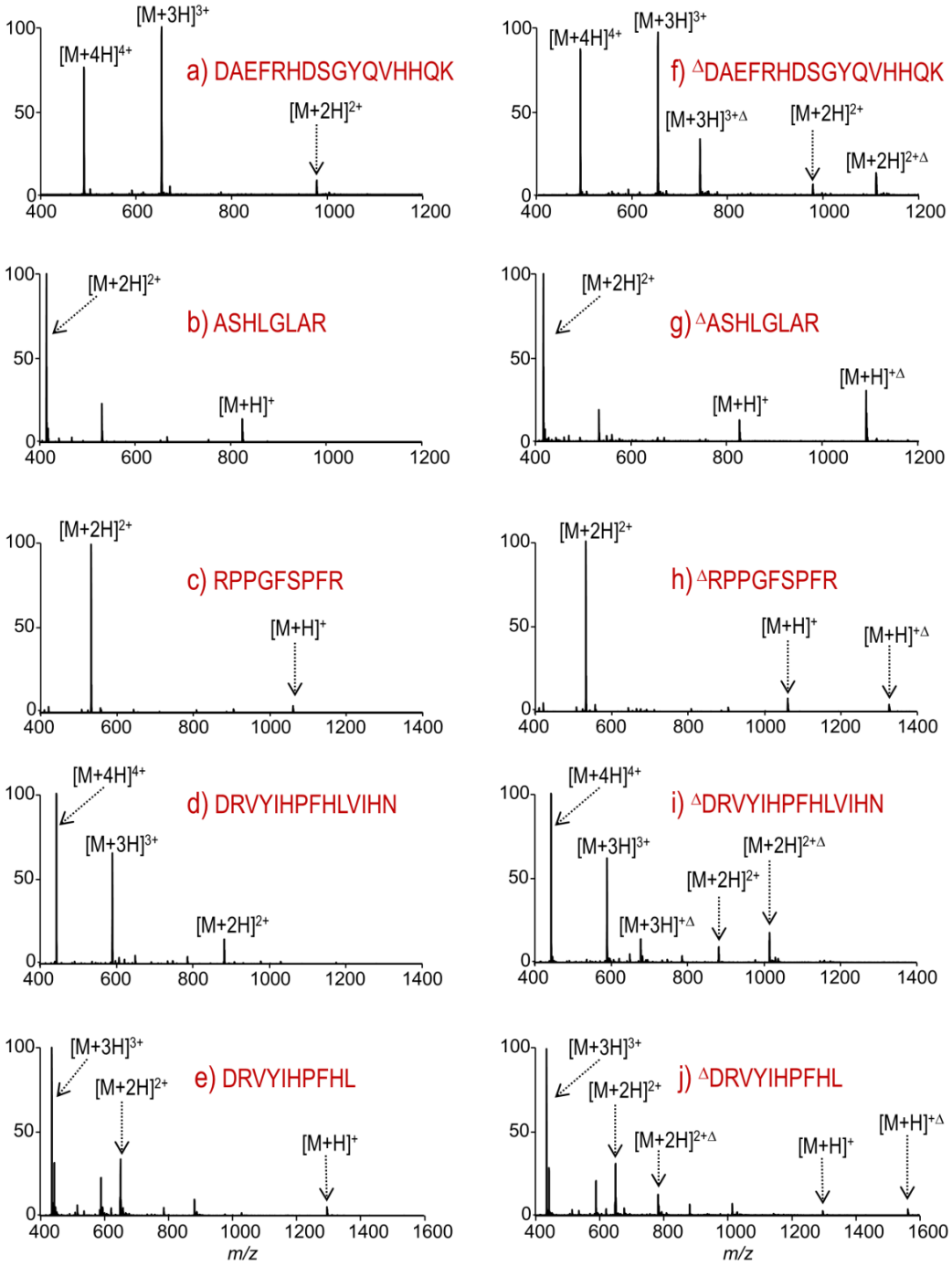


Figure 6.9 (a-e) Control and (f-j) FBDSA-reacted LC mass spectra for each peptide of a five peptide mixture. For all data shown, the peptide source voltage was held constant at 1.75 kV. Dual source initiated ion/ion reactions were carried out using an anion source voltage of -2.5 kV. Note that spectra are shown in order of elution.

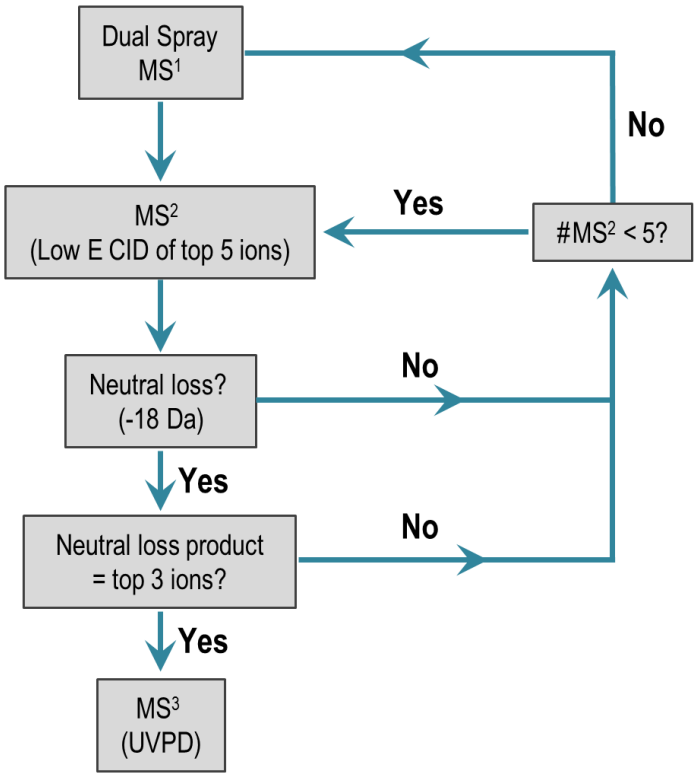


Figure 6.10 Data-dependent neutral loss MS^3 decision tree for online derivatization using an LC-MS3 platform.

The ability to establish a set of dual source parameters capable of promoting ion/ion mediated covalent derivatization for all peptides in a mixture provided confidence that online bioconjugation followed by fast UVPD characterization of modified peptides was possible using an automated LC-MS^n approach. To test this strategy, a data-dependent neutral loss MS^3 ($DDNLMS^3$) method was developed (**Figure 6.10**) in which the top five most abundant ions in the MS^1 spectrum were subjected to MS^2 using the aforementioned activation parameters. If the product ion corresponding to the neutral loss of water was within the top three most abundant ions in the low energy CID spectrum, a

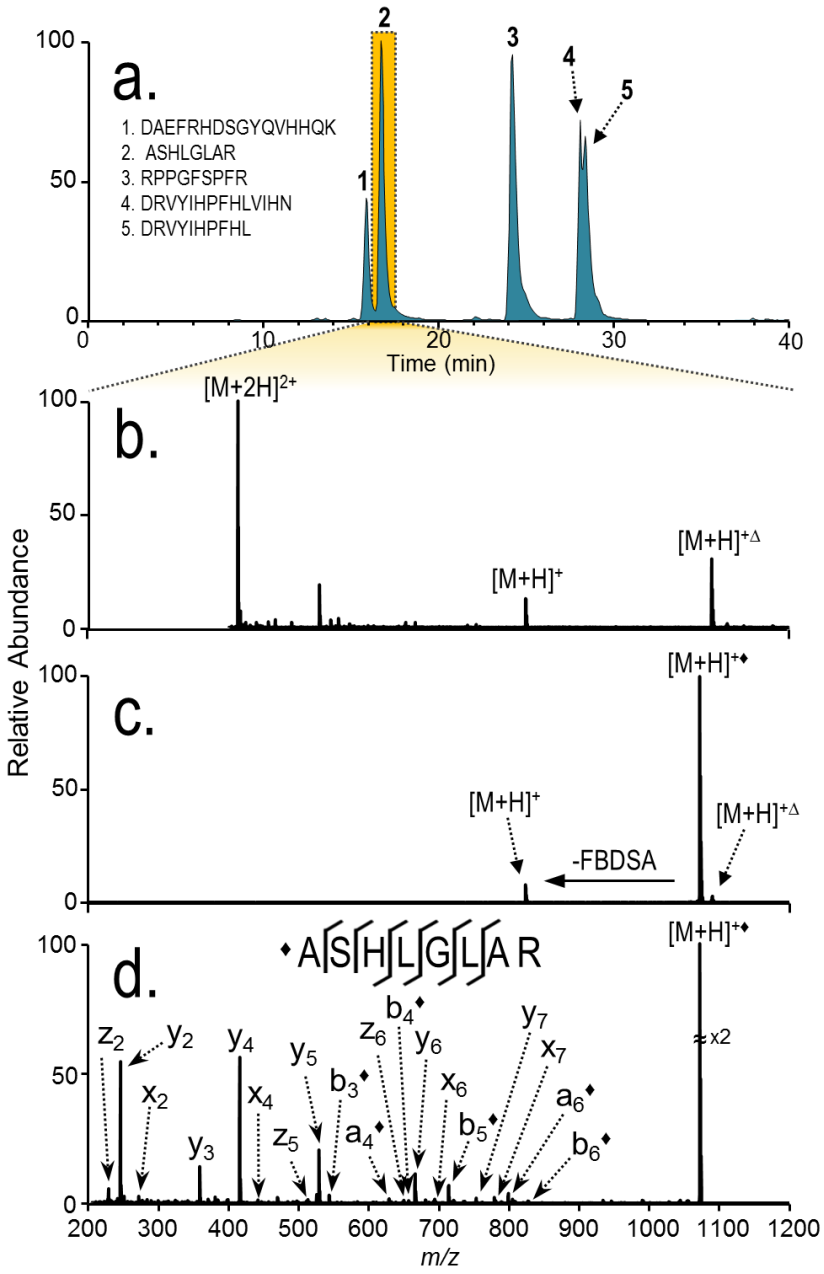


Figure 6.11 Online derivatization of a peptide mixture using the dual spray reactor combined with an LC-MS based data-dependent neutral loss MS^3 scan program. (a) Base peak chromatogram for the peptide mixture. (b) MS^1 scan during the elution of ASHLGLAR showing a mixture of unreacted peptide and peptide/FBDSA complexes (Δ). The electrostatic complex is selected by the DDNLMS^3 program to undergo (c) low energy CID (NCE = 18%). A neutral loss product is detected at -18 Da (♦), which triggers (d) 193 nm UVPD of the Schiff base product.

subsequent 193 nm UVPD activation event was initiated. Automation of the complete ion/ion mediated bioconjugation process is shown via consecutive scan events in **Figure 6.11** for the peptide ASHLGLAR corresponding to the chromatographic peak at 17 minutes of the elution profile (**Figure 6.11a**). The 1+ charge state of the electrostatic complex appears as the second most abundant ion in the MS^1 spectrum (**Figure 6.11b**), allowing efficient selection during data-dependent acquisition. Collisional activation of the complex results in a dominant neutral loss of water (**Figure 6.11c**), thus triggering 193 nm UVPD (**Figure 6.11d**). The array of fragment ions generated provides extensive structural characterization of the modified peptide. Moreover, greater than 20% improvement in UVPD efficiency was observed relative to the unmodified analog, thus showing excellent agreement of the high-throughput LC-MS^n results with those obtained via dual source infusion set-up.

6.5 CONCLUSIONS

The development and implementation of a front-end dual electrospray reactor used for online ion/ion mediated bioconjugation was demonstrated for rapid Schiff base derivatization of peptide cations with FBDSA anions. Dual spray-assisted covalent chemistry was found to be highly consistent with analogous reactions performed *in vacuo*; however, by shifting the ion/ion reaction step to the interface of the mass spectrometer, both the throughput and adaptability of this approach was streamlined for integration into chromatographic workflows. Furthermore, we showed the advantage of utilizing fast online reactions to enhance the photodissociation efficiencies of peptides using 193 nm photoactivation as well as to improve the structural characterization of modified peptides.

6.6 REFERENCES

(1) Hermanson, G. T. *Bioconjugate Techniques*; Academic Press, 2013.
(2) Stephanopoulos, N.; Francis, M. B. *Nat. Chem. Biol.* **2011**, *7* (12), 876–884.
(3) Mirzaei, H.; Regnier, F. *Anal. Chem.* **2006**, *78* (12), 4175–4183.
(4) Pashkova, A.; Moskovets, E.; Karger, B. L. *Anal. Chem.* **2004**, *76* (15), 4550–4557.
(5) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17* (10), 994–999.
(6) Wiese, S.; Reidegeld, K. A.; Meyer, H. E.; Warscheid, B. *PROTEOMICS* **2007**, *7* (3), 340–350.
(7) Ulbrich, A.; Merrill, A. E.; Hebert, A. S.; Westphall, M. S.; Keller, M. P.; Attie, A. D.; Coon, J. J. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (1), 6–9.
(8) Keough, T.; Youngquist, R. S.; Lacey, M. P. *Proc. Natl. Acad. Sci.* **1999**, *96* (13), 7131–7136.
(9) Madsen, J. A.; Brodbelt, J. S. *Anal. Chem.* **2009**, *81* (9), 3645–3653.
(10) Robinson, M. R.; Madsen, J. A.; Brodbelt, J. S. *Anal. Chem.* **2012**, *84* (5), 2433–2439.
(11) Cotham, V. C.; Wine, Y.; Brodbelt, J. S. *Anal. Chem.* **2013**, *85* (11), 5577–5585.
(12) Diedrich, J. K.; Julian, R. R. *Anal Chem* **2010**, *82* (10), 4006–4014.
(13) Vasicek, L.; O'Brien, J. P.; Browning, K. S.; Tao, Z.; Liu, H.-W.; Brodbelt, J. S. *Mol. Cell. Proteomics* **2012**, *11* (7), 1–10.
(14) Espy, R. D.; Wleklinski, M.; Yan, X.; Cooks, R. G. *TrAC Trends Anal. Chem.* **2014**, *57*, 135–146.
(15) Prentice, B. M.; McLuckey, S. A. *Chem. Commun.* **2013**, *49* (10), 947–965.
(16) Mortensen, D. N.; Williams, E. R. *Anal. Chem.* **2014**, *86* (18), 9315–9321.
(17) Mortensen, D. N.; Williams, E. R. *Anal. Chem.* **2014**.
(18) Fisher, C. M.; Kharlamova, A.; McLuckey, S. A. *Anal. Chem.* **2014**, *86* (9), 4581–4588.
(19) Miladinović, S. M.; Fornelli, L.; Lu, Y.; Piech, K. M.; Girault, H. H.; Tsybin, Y. O. *Anal. Chem.* **2012**, *84* (11), 4647–4651.
(20) Kharlamova, A.; McLuckey, S. A. *Anal. Chem.* **2011**, *83* (1), 431–439.
(21) Grimm, R. L.; Hodyss, R.; Beauchamp, J. L. *Anal. Chem.* **2006**, *78* (11), 3800–3806.
(22) Kim, H. I.; Kim, H.; Shin, Y. S.; Beegle, L. W.; Jang, S. S.; Neidholdt, E. L.; Goddard, W. A.; Heath, J. R.; Kanik, I.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2010**, *132* (7), 2254–2263.
(23) Müller, T.; Badu-Tawiah, A.; Cooks, R. G. *Angew. Chem. Int. Ed.* **2012**, *51* (47), 11832–11835.
(24) Girod, M.; Moyano, E.; Campbell, D. I.; Cooks, R. G. *Chem. Sci.* **2011**, *2* (3), 501–510.
(25) Badu-Tawiah, A. K.; Li, A.; Jjunju, F. P. M.; Cooks, R. G. *Angew. Chem. Int. Ed.* **2012**, *51* (37), 9417–9421.

(26) Chen, H.; Wortmann, A.; Zhang, W.; Zenobi, R. *Angew. Chem. Int. Ed.* **2007**, *46* (4), 580–583.
(27) Peng, I. X.; Ogorzalek Loo, R. R.; Shiea, J.; Loo, J. A. *Anal. Chem.* **2008**, *80* (18), 6995–7003.
(28) Xia, Y.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **2011**, *19* (2), 173–189.
(29) Han, H.; McLuckey, S. A. *J. Am. Chem. Soc.* **2009**, *131* (36), 12884–12885.
(30) Stutzman, J. R.; Luongo, C. A.; McLuckey, S. A. *J. Mass Spectrom.* **2012**, *47* (6), 669–675.
(31) Mentinova, M.; McLuckey, S. A. *J. Am. Chem. Soc.* **2010**, *132* (51), 18248–18257.
(32) Prentice, B. M.; Gilbert, J. D.; Stutzman, J. R.; Forrest, W. P.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (1), 30–37.
(33) McGee, W. M.; McLuckey, S. A. *Proc. Natl. Acad. Sci.* **2014**, *111* (4), 1288–1292.
(34) Hassell, K. M.; Stutzman, J. R.; McLuckey, S. A. *Anal. Chem.* **2010**, *82* (5), 1594–1597.
(35) Stutzman, J. R.; McLuckey, S. A. *Anal. Chem.* **2012**, *84* (24), 10679–10685.
(36) Vasicek, L.; Brodbelt, J. S. *Anal. Chem.* **2010**, *82* (22), 9441–9446.
(37) Gardner, M. W.; Smith, S. I.; Ledvina, A. R.; Madsen, J. A.; Coon, J. J.; Schwartz, J. C.; Stafford, G. C.; Brodbelt, J. S. *Anal. Chem.* **2009**, *81* (19), 8109–8118.
(38) Villén, J.; Beausoleil, S. A.; Gygi, S. P. *PROTEOMICS* **2008**, *8* (21), 4444–4452.
(39) Stutzman, J. R.; Hassell, K. M.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2012**, *312*, 195–200.
(40) Ogorzalek Loo, R. R.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1992**, *3* (7), 695–705.
(41) Chen, H.; Venter, A.; Cooks, R. G. *Chem. Commun.* **2006**, No. 19, 2042–2044.
(42) Law, W. S.; Wang, R.; Hu, B.; Berchtold, C.; Meier, L.; Chen, H.; Zenobi, R. *Anal. Chem.* **2010**, *82* (11), 4494–4500.
(43) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. *J. Proteome Res.* **2010**, *9* (8), 4205–4214.

Chapter 7

Modulation of Phosphopeptide Fragmentation via Dual Spray Ion/Ion Reactions using a Sulfonate-Incorporating Reagent*

7.1 OVERVIEW

The labile nature of phosphoryl groups has presented a long-standing challenge for the characterization of protein phosphorylation via conventional mass spectrometry-based bottom-up proteomics methods. Collision-induced dissociation (CID) causes preferential cleavage of the phospho-ester bond of peptides, particularly under conditions of low proton mobility, and results in the suppression of sequence-informative fragmentation that often prohibits phosphosite determination. In the present study, the fragmentation patterns of phosphopeptides are improved through ion/ion-mediated peptide derivatization with 4-formyl-1,3-benezenedisulfonic acid (FBDSA) anions using a dual spray reactor. This approach exploits the strong electrostatic interactions between the sulfonate moieties of FBDSA and basic sites to facilitate gas-phase bioconjugation and to reduce charge sequestration and increase the yield of phosphate-retaining sequence ions upon CID. Moreover, comparative CID fragmentation analysis between unmodified phosphopeptides and those modified online with FBDSA or in solution via carbamylation and 4-sulfophenyl isothiocyanate (SPITC) provided evidence for sulfonate interference with charge-directed mechanisms that result in preferential phosphate elimination. Our results indicate the prominence of charge-directed neighboring group participation reactions involved in phosphate neutral loss, and the implementation of ion-ion reactions

*Cotham, V. C.; McGee, W. M.; Brodbelt, J. S. *Anal. Chem.* **2016**, 88, 8158-8165.
V.C.C. designed and conducted all experiments.

in a dual spray reactor set-up provides a means to disrupt the interactions by competing hydrogen-bonding interactions between sulfonate groups and the side-chains of basic residues.

7.2 INTRODUCTION

Protein phosphorylation is a highly dynamic post-translational modification (PTM) that plays a central role in the signaling and regulatory machinery that mediate nearly all cellular processes including transcription, differentiation, cell cycle progression, and metabolism.[1–3] These processes are controlled through the coordinated interplay of protein kinases and phosphatases that modulate the function of target proteins by transiently altering their phosphorylation states at serine, threonine and tyrosine sites.[4] Moreover, aberrant phosphorylation arising from the dysregulation of this activity has been linked to the onset and progression of numerous neurodegenerative, oncogenic and metabolic diseases.[5–7] Consequently, the molecular-level characterization of protein phosphorylation is essential for the comprehensive understanding of complex mechanisms governing cell health and disease and offers critical insight for the development of new therapeutics.

Mass spectrometry (MS) has emerged as the analytical method of choice for the identification and characterization of phosphorylated proteins on both the individual and global scale.[8,9] However, despite exceptional speed and sensitivity, common MS-based approaches suffer from several key impediments arising from the intrinsic biological and chemical properties of phosphoproteins and their peptide constituents. For example, phosphorylation often occurs at substoichiometric levels that are below the sampling depth of most bottom-up data-dependent driven workflows.[10,11] Enrichment strategies

such as immobilized metal affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC) have helped to overcome this limitation by selectively increasing the relative abundance of phosphorylated targets within full MS survey scans.[9] While this additional step facilitates improved detection, subsequent tandem mass spectrometric (MS/MS) analysis to obtain sequence and phosphosite information by direct fragmentation of selected phosphopeptides is often inhibited by the higher gas phase lability of the phospho-ester bond relative to the polypeptide backone.[12] Collision-induced dissociation (CID) remains the most established and widely utilized ion activation method;[13] however, the slow heating mechanism that governs ion dissociation promotes cleavage at the most labile sites, thereby inducing preferential neutral loss of the phosphate group and suppression of diagnostic sequence and phosphosite-informative fragmentation.[12,14,15] This outcome has proven particularly problematic under conditions of low proton mobility where hydrogen bonding interactions between basic side-chains and the phosphate group facilitate nearly exclusive charge-directed neutral loss of the phosphate.[12,16] This shortcoming has prompted the use of alternative activation strategies that provide more informative MS/MS spectra for phosphopeptide characterization, including electron-driven approaches (ETD and ECD),[17,18] higher-energy collisional activation (HCD),[19] ultraviolet photodissociation (UVPD),[20–23] and several combinations thereof (i.e., ETcaD,[24] EThcD,[25] ETUVPD[26]).

In addition to alternative activation methods, chemical and enzymatic strategies that modify the intrinsic properties of phosphopeptides to make them more suitable for MS/MS interrogation have also been reported, albeit at the cost of more extensive sample preparation and experimental complexity. Approaches based on β-elimination of phosphoryl groups from phosphoserine and phosphothreonine residues followed by Michael addition with a nucleophilic reactant have been used in a diversity of protein

phosphorylation studies.[27–32] The purpose of this type of strategy is generally two-fold: 1) removal of the CID labile phosphate group to generate more informative MS/MS spectra and 2) incorporation of novel chemistry that can be exploited for streamlined phosphopeptide analysis.[28,31–33] Despite these merits, one major drawback to β-elimination-based approaches arises from their lack of selectivity toward phosphotyrosine residues.[34] Other strategies aimed at eliminating or minimizing the hydrogen bonding interactions that lead to preferential phosphate cleavage have also shown success for generating more informative CID spectra with the added benefit of being universally applicable to all phosphopeptide types (S/T/Y).[35–38] This has been accomplished via selective derivatization of either the phosphate moiety[35,36] or basic side-chains of the peptide,[37] or alternatively by the complete enzymatic removal of basic residues.[38]

Recently, we described a method for the high-throughput bioconjugation of peptide cations via pseudo-droplet phase initiated ion/ion reactions using a front-end dual spray reactor.[39] This previous work relied on well-characterized gas-phase ion/ion-mediated covalent chemistry using 4-formyl-1,3-benzenedisulfonic acid (FBDSA) anions[40–42] to both facilitate rapid derivatization on a timescale compatible with chromatographic separations and to increase the intrinsic photoabsorption cross-section of peptides for enhanced photodissociation at 193 nm.[39] Herein, we demonstrate that this same chemistry can also be leveraged to modulate the collisional dissociation behavior of basic phosphopeptides in real time for improved sequence coverage and phosphosite localization relative to their unmodified counterparts. In a manner similar to the removal or selective derivatization of basic sites, this method relies on the preferential formation of noncovalent interactions between the sulfonate moieties of FBDSA and the basic sites within the peptide to overcome or partially disrupt the mechanisms leading to preferential phosphate loss by collisional activation.

7.3 EXPERIMENTAL

7.3.1 Chemicals and Materials

Phosphopeptides RQpSVELHSPQSLPR, GGGPApTPKKAKKL, and KKALRRQEpTVDAL were purchased from AnaSpec Inc. (Fremont, CA). RRLIEDAEpYAARG-NH_2 was purchased from American Peptide Company (Sunnyvale, CA). 4-formyl-1,3-benzenedisulfonic acid (FBDSA), 4-sulfophenyl isothiocyanate (SPITC), urea and all other solvents were obtained from Sigma-Aldrich (St. Louis, MO). Peptides and reagents were used without further purification.

7.3.2 Solution Phase N-Terminal Derivatization

Aliquots of phosphopeptides lacking internal lysine residues were subjected to N-terminal derivatization via 4-sulfophenyl isothiocyanate (SPITC) and carbamylation in the presence of excess urea. SPITC modification was accomplished by reacting 20 μL of stock solution (1 mg of SPITC in 100 μL of 1x PBS, pH 7.4) with 10 nmol of peptide for 30 minutes at 55°C. Carbamylation reactions were carried out via incubation with 8 M urea in 50 mM Tris-HCl (pH 8) at 80°C for 4 h. N-terminally labeled peptides were then desalted on C18 spin columns (Life Technologies, Grand Island, NY), evaporated to dryness and resuspended in 50:50 water/methanol for infusion.

7.3.3 Mass Spectrometry and Front-end Ion/Ion Reactions

All experiments were conducted on a Thermo Scientific Velos Pro dual linear ion trap mass spectrometer (San Jose, CA) equipped with a front-end dual spray reactor as previously described.[39] Briefly, the reactor was equipped with two electrospray ionization (ESI) sources (Prosolia Inc., Indianapolis, IN) mounted on a U-shaped railing system that surrounded the front-end of the mass spectrometer. The first source was fully integrated to allow for direct control of spray voltage and polarity in the Thermo Tune Plus control

software, whereas the spray voltage of the second source was supplied via an external 5 kV dual polarity high voltage power supply (Stanford Research Systems Inc., Sunnyvale, CA).

For standard infusion experiments, the dual spray reactor was operated in single source mode in a manner comparable to conventional ESI. Phosphopeptide cations were generated by positive mode electrospray ionization of 5 μM working solutions infused at rate of 1.5 μL/min using a spray voltage of 1.5 kV. Alternatively, ion/ion reactions were carried out using dual source mode, during which the second source was simultaneously operated to generate a second population of reagent anions via negative mode ESI of 2 mM FBDSA prepared in 50:50 water/methanol infused at a rate of 3 μL/min. Electrostatic ion/ion complexes were formed at or near atmospheric pressure in the region of spray overlap prior to the inlet of the mass spectrometer. Anion source voltage was varied between -2.0 and -2.75 kV to achieve optimal complex formation and spray stability. Schiff base reactions were performed by collisionally activating the electrostatic phosphopeptide/FBDSA complexes with low normalized collision energy (NCE = 10-18%) to overcome the activation barrier for imine formation. Covalent Schiff base products were then isolated and subjected to MS^3 collision-induced dissociation (CID) to generate diagnostic product ions using 20-30% NCE and a *q*-value of 0.25.

7.4 RESULTS AND DISCUSSION

Gas-phase ion/ion-mediated Schiff base derivatization with 4-formyl-1,3-benezenedisulfonic acid (FBDSA) is an established method for rapid covalent transformation of peptides within the context of a tandem mass spectrometry-based experiment.[39–43] This chemistry proceeds via the formation of long-lived electrostatic

complexes arising from noncovalent interactions between the sulfonate moieties of FBDSA reagent anions and protonated sites of peptide cations.[40] Subsequent collisional activation of these ion/ion intermediates promotes nucleophilic attack on the FBDSA aldehyde by an unprotonated primary amine in the substrate peptide, resulting in concerted dehydration and imine bond formation.[40] Although well-defined for unmodified substrates, no studies to date have explored this chemistry with peptides containing labile post-translational modifications. Thus, we sought to evaluate both the feasibility and analytical utility of gas-phase derivatization with FBDSA for a series of phosphorylated peptides containing modified serine, threonine, and tyrosine residues. We hypothesized that this Schiff base reaction could provide a means to modulate hydrogen-bonding interactions in phosphopeptides and minimize phosphate cleavage, and at the same time allow implementation of this method in an on-line fashion via a dual spray reactor set-up.

Peptides that contain multiple basic sites pose a particularly difficult challenge for conventional CID-based characterization owing to the immobilization of charges caused by the presence of multiple basic residues that sequester ionizing protons. Previous mechanistic studies indicate that the protonated basic residues form strong hydrogen bonding interactions with the phosphate group.[16] This consequently lowers the energy barrier for charge-directed mechanisms that lead to preferential neutral loss of phosphate and suppression of sequence-informative fragmentation upon collisional activation, thus further exacerbating the phosphate loss problem prevalent for MS/MS analysis of phosphopeptides.[16] An illustrative example of this phenomenon is shown in the CID product ion spectrum of doubly charged KKALRRQEpTVDAL (**Figure 7.1**). The uninformative loss of H_3PO_4 from the precursor accounts for approximately 80% of the total product ion signal, while the remaining 20% arises from the b_8^{2+} through b_{12}^{2+} ions

split between their phosphate retained and neutral loss forms. All diagnostic product ions contain the N-terminus, which is consistent with proton sequestration at basic arginine (R) or lysine (K) side-chains that occur near the N-terminal region of the peptide. As underscored by this example, the impediments associated with charge-immobilization make this a compelling gas-phase environment in which to probe the effects of FBDSA incorporation on phosphopeptide fragmentation.

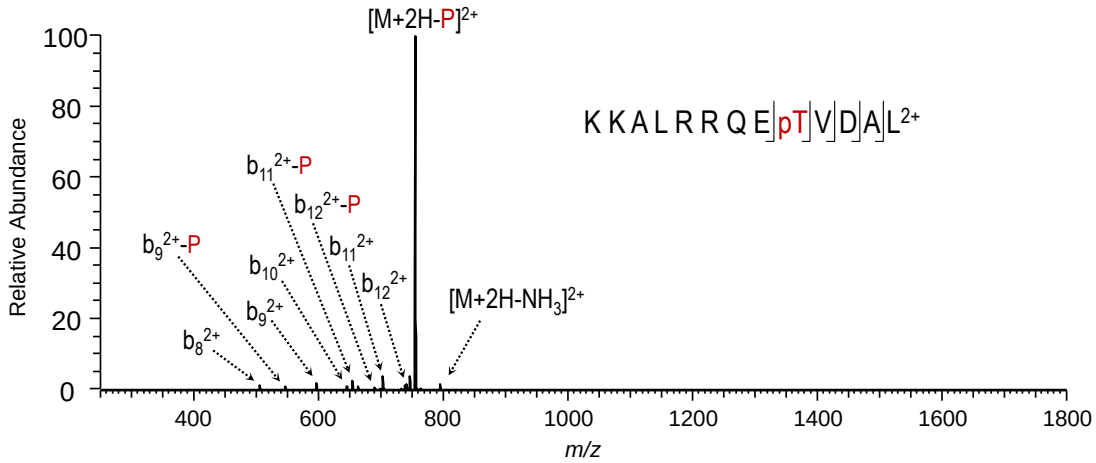


Figure 7.1 CID product ion spectrum of the 2+ charge state of KKALRRQEpTVDAL. Neutral loss of phosphate is indicated by "-P" in the product ion label.

7.4.1 Dual Spray Reactor-Initiated Schiff Base Bioconjugation of Phosphopeptides

The formation of long-lived phosphopeptide/FBDSA complexes was accomplished in real-time using a front-end dual spray reactor as previously described.[39] Briefly, the reactor utilized two oppositely biased ESI sources to simultaneously generate overlapping populations of phosphopeptide cations and FBDSA anions in the high pressure region prior to the inlet of the mass spectrometer. The anionic reagent of interest, FBDSA, is negatively charged and thus has the potential to cause neutralization of peptides during formation of ion-ion complexes. To ensure successful detection of

ion/ion complexes in the positive mode, the phosphopeptide substrates each contained at least two positive charge-bearing residues (i.e., arginine or lysine). This process is demonstrated in **Figure 7.2**, which compares the MS^1 spectrum of KKALRRQEpTVDAL before (**Figure 7.2a**) and after (**Figure 7.2b**) interaction with FBDSA anions. Two highly abundant ions consistent with the 1+ and 2+ charge states of the charge-reduced KKALRRQEpTVDAL/FBDSA complex (denoted by the addition of "Δ" in the label) are observed exclusively in the post-reaction spectrum at *m/z* 1874 and 937, respectively. Together the complexed species account for approximately 60% of the total analyte signal, thus demonstrating high ion/ion reaction efficiency in the region of overlap between the dual sprays prior to transmission into the mass spectrometer. These results suggest that the presence of phosphorylated side-chains do not have a prohibitive effect on FBDSA binding despite possible competition with the phosphate moiety to form stabilizing noncovalent interactions with the protonated basic sites of the peptide.[16,44]

Once formed, ion/ion complexes were isolated in the linear ion trap and subjected to low energy CID to initiate covalent conversion to products, as demonstrated for the doubly charged complex of KKALRRQEpTVDAL/FBDSA (**Figure 7.2c**). Unlike collisional activation of the unreacted phosphopeptide, which is dominated by phosphate neutral loss (**Figure 7.1**), the most energetically favored pathway of the electrostatic complex results in dehydration with complete retention of the phosphate group and formation of the covalent Schiff base product. This is reflected by the single product ion observed in the MS^2 spectrum which is 18 Da lower in mass from the precursor (**Figure S2c**), which is consistent with dehydration that occurs upon imine bond formation. The Schiff base product is confirmed via an additional collisional activation step (MS^3) (**Figure 7.2d**). Collectively, these results demonstrate the ability to covalently modify phosphopeptides in the gas-phase while preserving the integrity of the labile phosphosite.

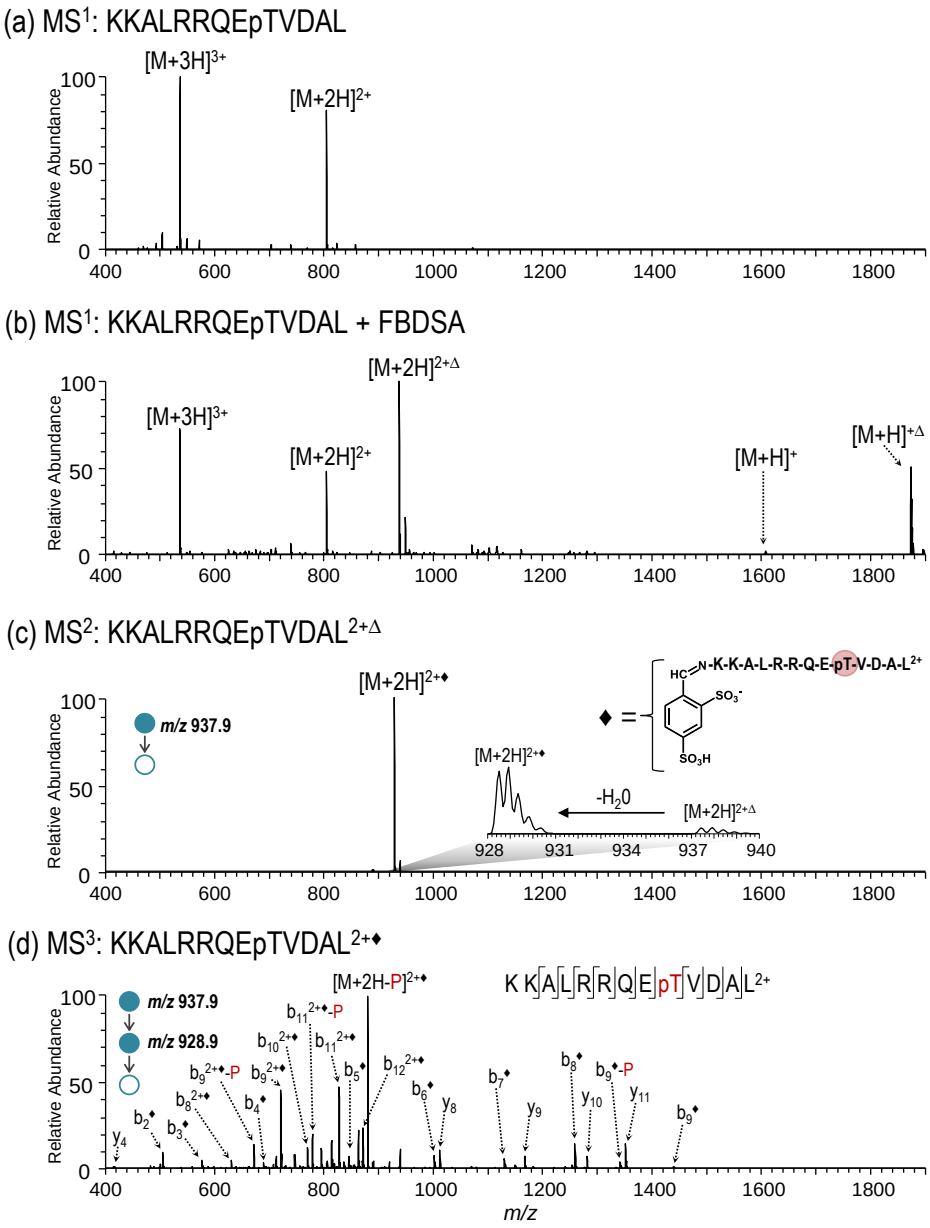


Figure 7.2 Process for online modification of phosphopeptides with FBDSA using a dual spray reactor. ESI spectra for (a) unreacted and (b) dual source reacted KKALRRQEpTVDAL. Charge-reduced electrostatic complexes are formed at atmospheric pressure between multiply charges phosphopeptide cations and FBDSA reagent anions (denoted by the "Δ" subscript), and transferred and mass analyzed in the linear ion trap. (c) Low-energy collisional activation of these ion/ion intermediates promotes concomitant imine bond formation and dehydration to form a covalent Schiff base product (♦). (d) CID of the resulting Schiff base phosphopeptide results in sequence-informative fragmentation.

7.4.2 CID of Unmodified versus FBDSA-Modified Phosphopeptides

As described above and shown in **Figure 7.1**, the product ion spectrum of doubly charged KKALRRQEpTVDAL illustrates the poor performance of CID under proton immobilized conditions; however, a marked improvement in fragmentation is observed following derivatization with FBDSA (**Figure 7.2d**). This is reflected in part by a 65% decrease in product ion signal comprised of non-sequence phosphate neutral loss from the precursor. Suppression of this preferential cleavage is accompanied by a concomitant gain in both the number and relative abundance of diagnostic fragment ions, resulting in an increase in sequence coverage from 42% to 92%. A small subset of product ions containing the phosphothreonine residue exhibit phosphate loss; however, in each case the relative abundance of these ions is lower than that of their corresponding phosphate-retaining forms. Furthermore, the emergence of *y*-ions and singly charged $b^{\blacklozenge}$-ions is consistent with greater proton mobility across the peptide backbone.

This dramatic change in fragmentation behavior upon incorporation of FBDSA is proposed to arise from the disruption and displacement of hydrogen bonding interactions between the phosphate group and basic sites of the peptide that lead to selective phosphate cleavage by CID. This hypothesis is supported by the lower pKa of sulfonate moieties in FBDSA relative to the phosphate group,[44] as well as previous reports describing the gas-phase stability of acid-base interactions between sulfonate moieties and basic side-chains of peptides.[43,45] Consequently, the strength of these interactions, and by effect the degree of change in subsequent fragmentation, should exhibit a dependence on the gas-phase basicity of the interacting side-chains. To explore this further, the relative change in fragmentation following FBDSA derivatization was evaluated for the arginine-containing peptide, RRLIEDAEpYAARG-NH_2, and the lysine-containing peptide, GGGPApTPKKAKKL.

A comparison of the CID product ion spectra for the 2+ charge state of RRLIEDAEpYAARG-NH_2 before and after gas-phase derivatization is shown in **Figure 7.3**. Collisional activation of the underivatized peptide promotes dominant neutral loss of phosphate and ammonia with limited product ion signal arising from cleavage at six of the twelve amide bonds of the peptide backbone (**Figure 7.3a**). Resulting low abundance sequence ions are derived from both termini of the peptide and in all cases are singly charged, which indicates protonation of both the N- and C-terminally located arginine residues. This distribution is likely more energetically favorable than protonation of the two adjacent N-terminal arginines due to coulombic repulsion of side-chains. Despite low overall abundance, the b_6 and y_7 ions originating from cleavage C-terminal to the aspartic acid residue are more abundant than the other *b/y* fragment ions, an outcome consistent with the aspartic acid effect commonly observed under conditions of low proton mobility.[46,47] As demonstrated in **Figure 7.3b**, conversion of peptide RRLIEDAEpYAARG-NH_2 to its Schiff base analogue profoundly alters its fragmentation. The resulting CID spectrum shows complete suppression of selective phosphate cleavage and instead displays extensive pairwise fragmentation across the peptide backbone, resulting in 92% coverage of the *b*-ion series and 75% coverage of the *y*-ion series. Moreover, the collective lack of evidence for charge-directed loss of phosphate or enhanced cleavage C-terminal to aspartic acid is highly indicative of greater proton mobility following FBDSA derivatization.

The effect of FBDSA incorporation on the fragmentation of GGGPApTPKKAKKL is expected to be less pronounced than that of RRLIEDAEpYAARG-NH_2 owing to the lower gas-phase basicity of lysine side-chains relative to arginine and thus resultant weaker acid-basic interactions with FBDSA. The CID spectra of the underivatized and FBDSA-modified peptide are shown in **Figure 7.4**.

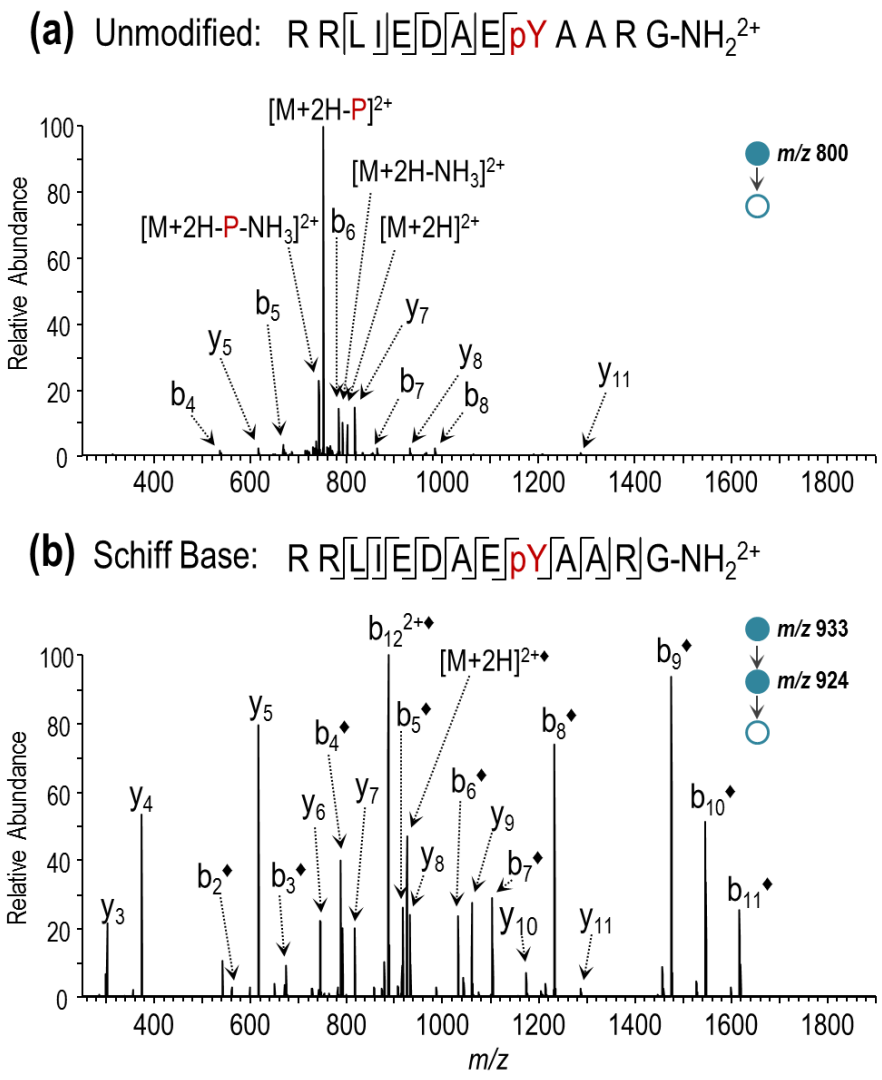


Figure 7.3 CID product ion mass spectra of RRLIEDAEpYAARG-NH_2 (2+) before and after Schiff base modification: (a) MS^2 CID mass spectrum of unlabeled peptide and (b) MS^3 CID mass spectrum following online dual spray reactor-initiated derivatization. The addition of "♦" to the label indicates covalent FBDSA Schiff base modification and "-P" indicates loss of phosphate.

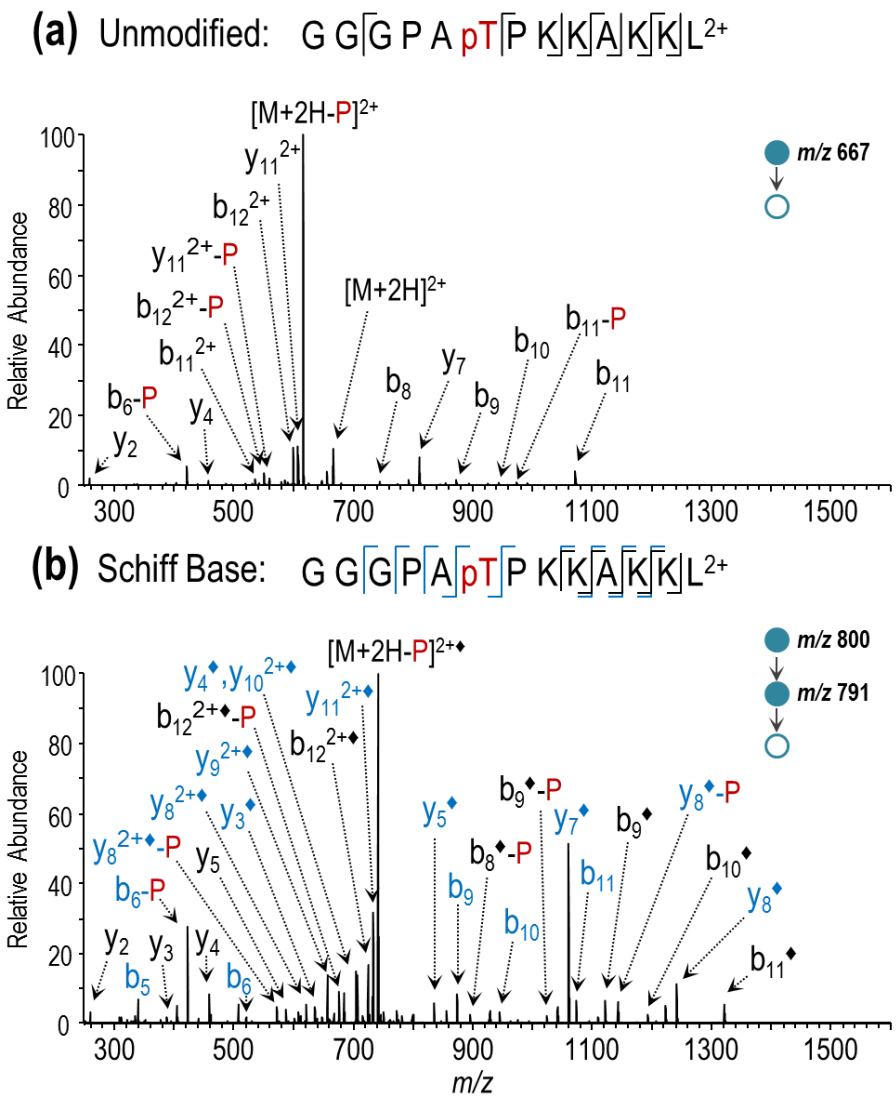


Figure 7.4 CID product ion mass spectra of GGGPApTPKKAKKL (2+) before and after Schiff base modification: (a) MS^2 CID mass spectrum of unlabeled peptide and (b) MS^3 CID mass spectrum following online dual spray reactor-initiated derivatization. The addition of "♦" to the label indicates covalent FBDSA Schiff base modification and "-P" indicates loss of phosphate. Ions shown in blue arise from lysine modification, while those shown in black may arise from either lysine or N-terminally labeled species.

The magnitude of the overall change in MS/MS patterns before and after derivatization generally agrees with this expectation about the impact of the basic side-chains. Notably, the dominance of phosphate neutral loss from the precursor suggests the unimpeded formation of hydrogen bonding interactions that facilitate charge-directed phosphate cleavage pathways. Differences also arise due to the fact that the ε-amino group of the lysine side-chain serves as a substrate for covalent FBDSA attachment. As a result, the location of the addition of the benzene disulfonic acid group is distributed across multiple reactive sites. This is reflected in the greater spectral complexity of the CID fragmentation spectrum of FBDSA-labeled GGGPApTPKKAKKL 2+, which exhibits contributions of product ions arising from various modified forms of the peptide (isomers) (**Figure 7.4b**). Despite these differences, enhanced fragmentation is still observed for the FBDSA modified peptide as demonstrated by more complete coverage of the peptide backbone. Additionally, the enhanced formation of the y_7 ion corresponding to N-terminal to proline cleavage is consistent with greater proton mobility following incorporation of FBDSA.[47,48]

The percent reduction in phosphate neutral loss from the precursor before and after derivatization provides a useful metric by which to evaluate the successful suppression of charge-directed mechanisms that facilitate selective phosphate cleavage. This data is summarized in **Figure 7.5** for four representative singly and doubly charged phosphopeptide/FBDSA complexes generated by front-end ion/ion reactions and subsequently converted to their Schiff base modified forms. Substantial improvements in phosphate retention were observed in all cases following FBDSA incorporation; however, this effect was most pronounced for doubly protonated arginine-containing peptides for which the uninformative phosphate-loss pathways plummeted. This result further supports the hypothesis that the strength of the interaction between the sulfonate

groups of FBDSA and the basic-side chains of the peptide has a direct impact on the observed change in fragmentation of FBDSA-labeled phosphopeptides.

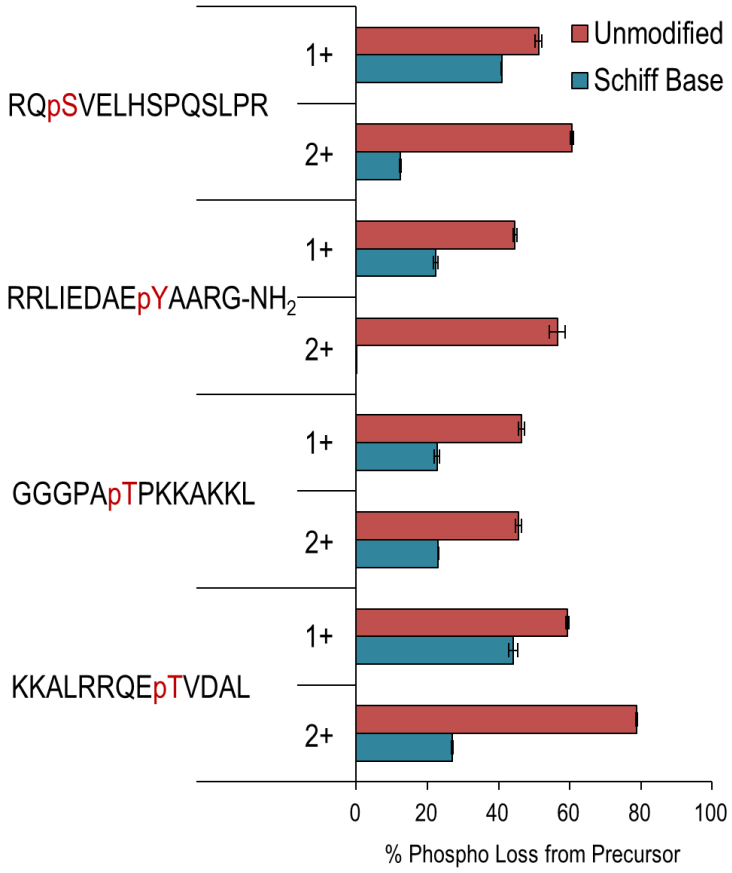


Figure 7.5 Percent of total product ion abundance arising from neutral loss of phosphate from the precursor ion of unmodified and Schiff based-labeled phosphopeptides subjected to CID.

7.4.3 Exploring the Role of the Sulfonate Moiety

In an effort to gain greater insight into the mechanistic role of the sulfonate moiety during collision-induced dissociation, the fragmentation behavior of unmodified and FBDSA-derivatized RRLIEDAEpYAARG-NH_2 and RQpSVELHSPQSLPR was

compared to that of their carbamylated and SPITC-modified analogues prepared in solution. Carbamylation converts primary amines of peptides (in this case the N-terminal amine) to less basic carbamate functionalities that are not expected to interact strongly with the basic side-chains of the peptide, thus serving as an experimental control for N-terminal modification (**Figure 7.6b**). Alternatively, the SPITC reagent introduces a mono-sulfonated phenyl group at the N-terminus of the peptide that more closely resembles FBDSA (**Figure 7.6a**), albeit appended via a rather different pathway (**Figure 7.6c**). As demonstrated in **Figure 7.7**, both reactions are very efficient, and the charge state distributions of the resulting peptides are shifted to lower values (i.e. enhancement of 2+, diminishment of 3+) relative to the unmodified peptide, consistent with removal of the N-terminal protonation site. As expected, this shift is greatest for the SPITC-modified peptide due to the fixed negative charge of the sulfonate moiety.

Energy variable collisional activated dissociation was carried out on the doubly charged precursors for all forms of two representative phosphopeptides, RRLIEDAEpYAARG-NH_2 (**Figure 7.8a**) and RQpSVELHSPQSLPR (**Figure 7.8b**), to evaluate their dissociation thresholds, one measure of stability in the gas phase. In both cases, the survival curves for precursors arising from sulfonate-bearing peptides exhibit a shift toward lower collisional energies relative to unmodified precursors, whereas the carbamylated species show nearly perfect overlap with the dissociation curves of the unmodified peptides. Such variations in fragmentation efficiency curves as a function of collision energy have been purported to reflect the degree of charge sequestration, essentially indicating the energy required for proton mobilization.[47] As such, the fragmentation efficiency curves in **Figure 7.8** suggest greater stabilization of the unmodified and carbamylated peptides (lower proton mobility) relative to the sulfonate-bearing peptides. The trends observed from the energy-variable CID curves correlate

closely with the resulting fragmentation patterns; the unmodified (**Figure 7.8c**) and carbamylated (**Figure 7.8d**) peptides exhibit extensive phosphate loss, in contrast to enhanced backbone cleavages and production of ample diagnostic sequence ions with phosphate retention for the FBDSA-modified peptides (**Figure 7.8e**). The relative abundances of informative sequence ion were similarly enhanced relative to preferential phosphate loss ions for the SPITC-derivatized peptides; however, selective cleavage of sulfanilic acid from the SPITC label via cleavage of the labile C-N bond of the thiocarbamoyl group biased the resulting fragmentation pattern (**Figure 7.9**).

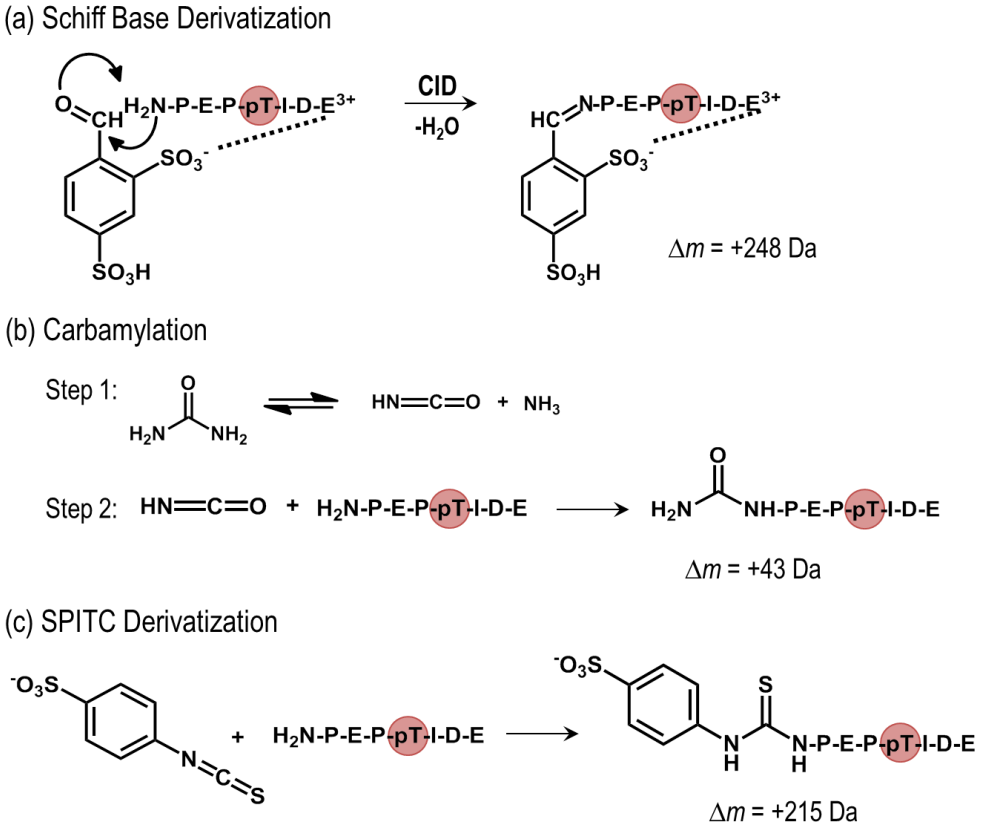


Figure 7.6 Reaction scheme for N-terminal (a) Schiff base derivatization with FBDSA, (b) carbamylation, and (c) 4-sulfophenyl isothiocyanate (SPITC) derivatization of a peptide.

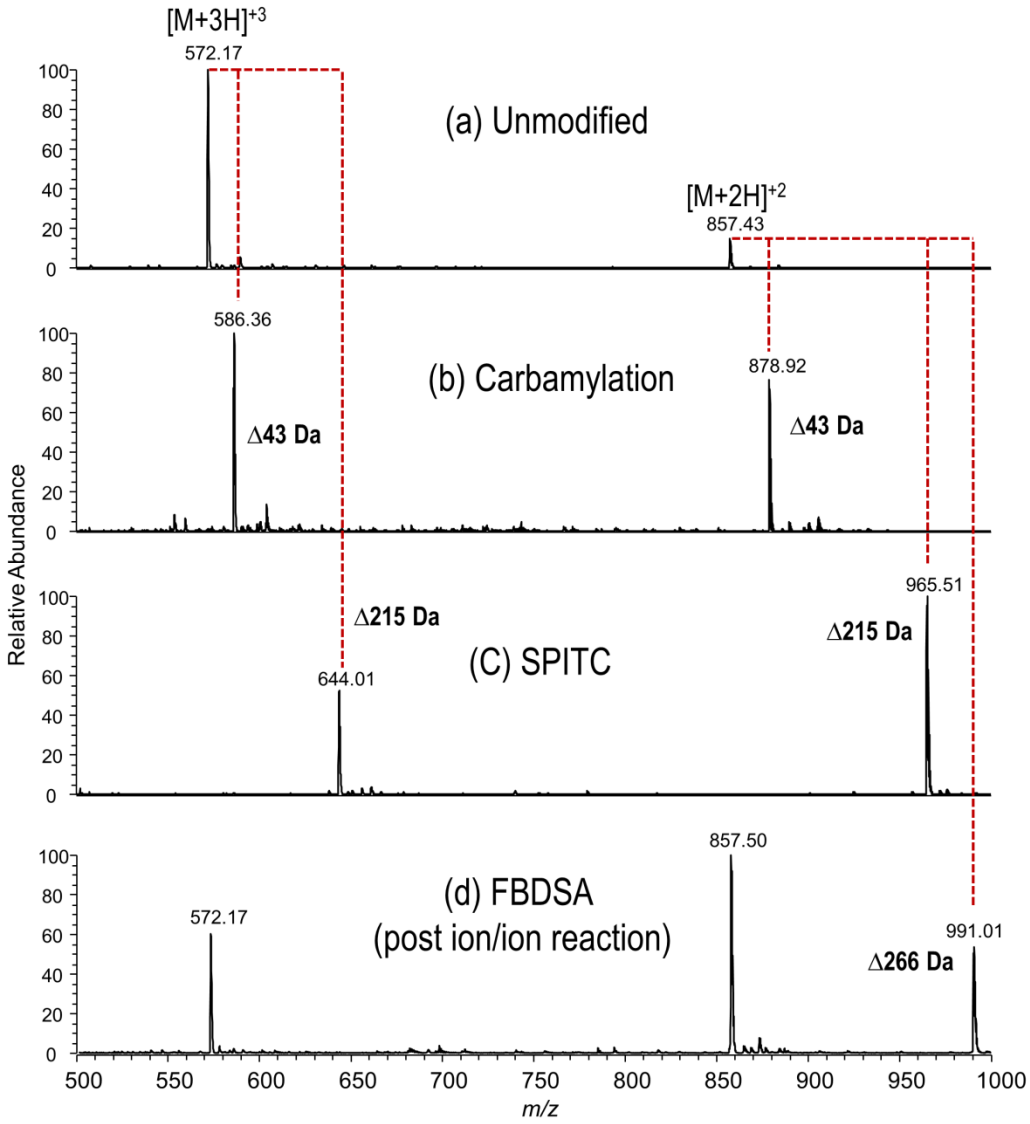


Figure 7.7 MS^1 spectra of RQpSVELHSPQSLPR prior to (a) and after (b) carbamylation, (c) SPITC derivatization, and (d) ion/ion reaction with FBDSA.

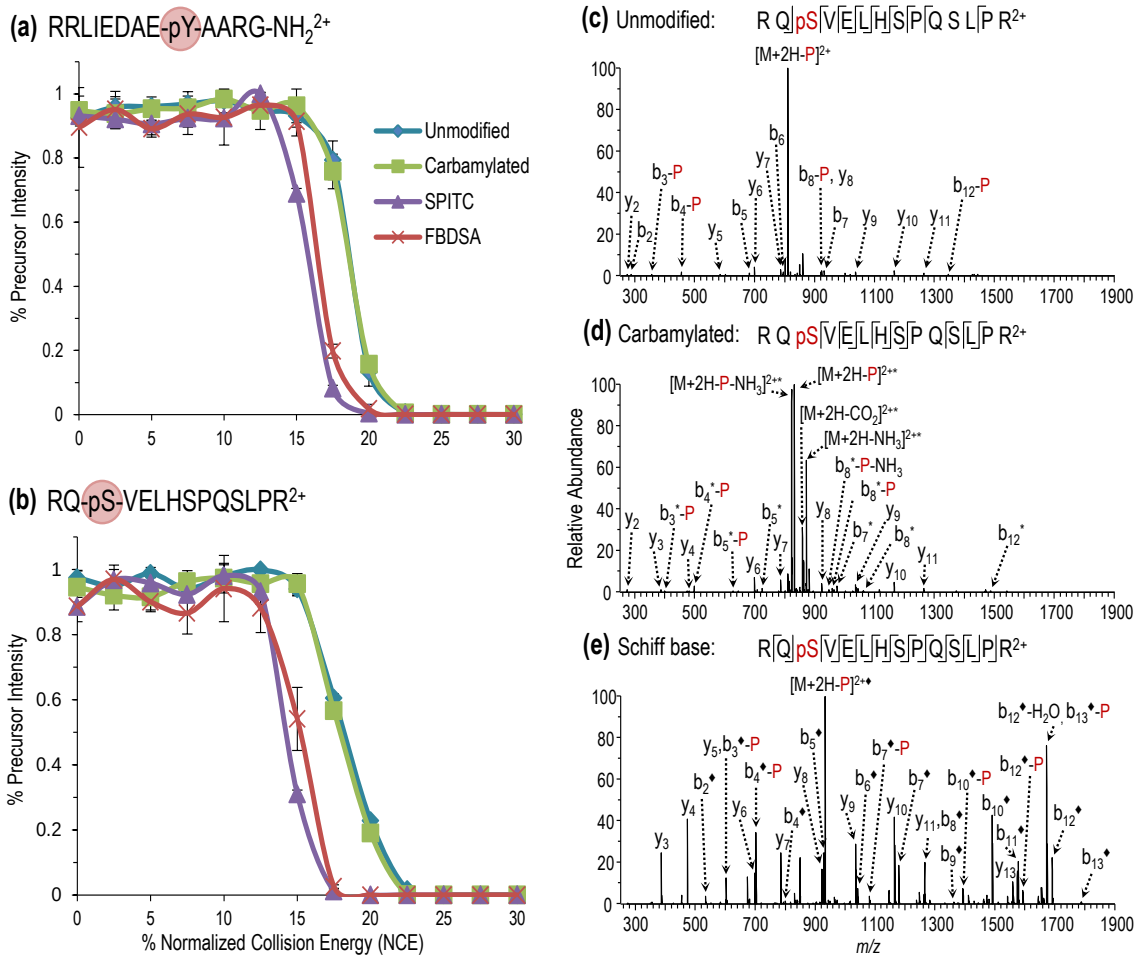


Figure 7.8 Variable energy CID analysis of unmodified and N-terminally carbamylated, SPITC- and FBDSA-modified (a) RRLIEDAEpYAARG-NH_2 (2+) and (b) RQpSVELHSPQSLPR (2+). Normalized precursor abundances are plotted as a function of increasing collision energy. The CID product ion spectra are shown for RQpSVELHSPQSLPR (2+) in the following states: (c) unmodified, (d) carbamylated (*), and (e) FBDSA Schiff base modified (♦). The addition of "-P" to the label indicates loss of phosphate.

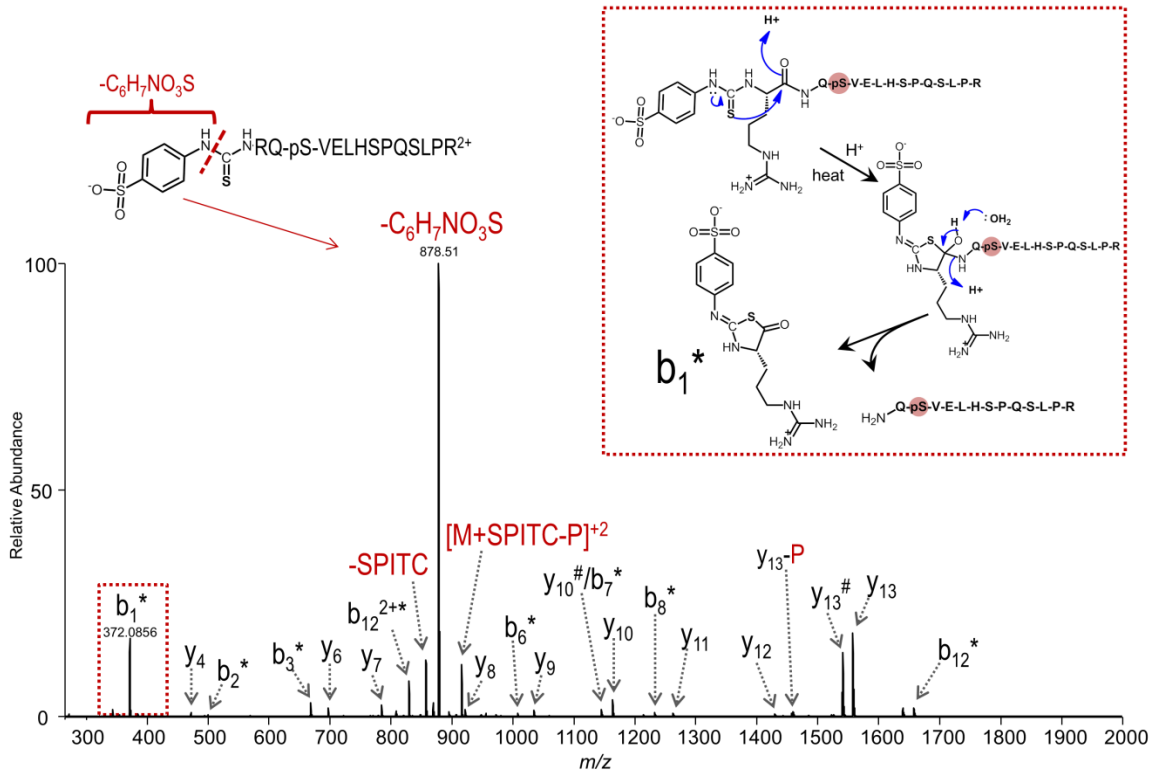


Figure 7.9 CID product ion spectrum of SPITC modified RQpSVELHSPQSLPR 2+. The abundant product ion a *m/z* 878 corresponds to sulfanilic acid cleavage from the SPITC tag. A modified b_1 ion results from the Edmund degradation process shown in the inset.

Based on the empirical observations made herein, two competing routes for the collision-induced dissociation of FBDSA-derivatized phosphopeptides are proposed as variations of the charge-directed S_N2 reaction pathway previously described by Reid and co-workers (**Figure 7.10a**).[16] The first pathway gives rise to preferential acid-base interactions between a sulfonate moiety of FBDSA and a protonated basic side-chain in the peptide allowing enhanced formation of sequence-informative product ions (**Figure 7.10b**). However, these acid-base interactions compete with hydrogen-bonding between the phosphate group and basic side-chains that facilitate selective phosphate neutral loss

(**Figure 7.10c**). Our preliminary results point to the gas-phase basicity of the side-chains as a primary determinant of the dominant dissociation pathway following FBDSA incorporation.

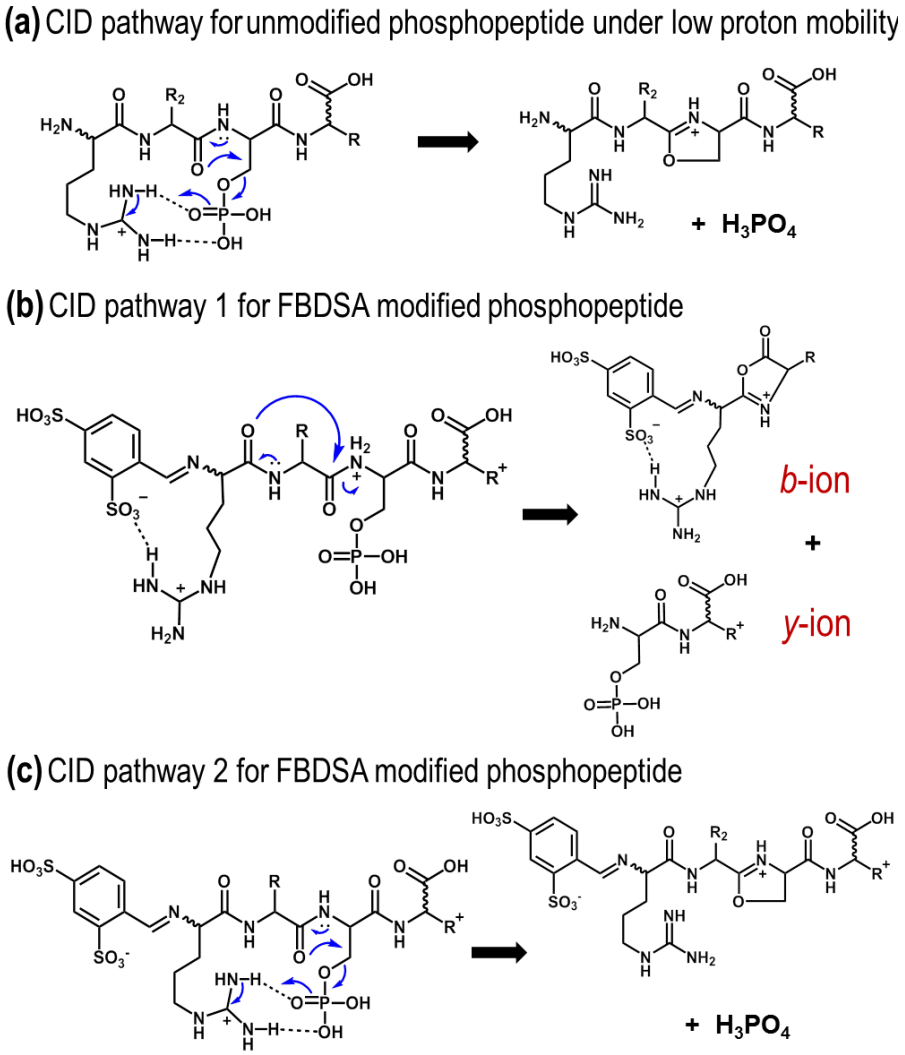


Figure 7.10 (a) Proposed mechanism for charge-directed neutral loss of phosphate for an unmodified phosphopeptide. Proposed competing dissociation pathways for FBDSA-labeled phosphopeptides with (b) and without (c) sulfonate-modulated suppression of charge-directed neutral loss of phosphate.

7.5 CONCLUSIONS

This work represents the first demonstration of ion/ion-mediated bioconjugation of peptides bearing one of the most common types of post-translational modifications, phosphorylation, in the gas-phase. Electrostatic phosphopeptide/FBDSA complexes were shown to undergo facile collision-induced conversion to covalent products with complete preservation of phosphosite integrity. The resulting FBDSA-derivatized phosphopeptides exhibited suppressed phosphate elimination that is the dominant process of conventional protonated phosphopeptides along with concomitant enhancement in the formation of sequence-informative product ions. The sulfonate moieties introduced upon FBDSA-incorporation are critical to this observed change in fragmentation behavior, as validated through comparative fragmentation analysis with other charge-site mediating modifications (i.e. carbamylation and thiocarbamoylation). Our results provide additional experimental evidence for charge-directed neighboring group participation reactions involved in phosphate neutral loss,[16] which appear to be disrupted by competing hydrogen-bonding interactions between sulfonate groups and the side-chains of basic residues. This was further supported by the relative change in phosphate neutral loss observed depending on the gas-phase basicity of the side-chains present, with the greatest changes observed for FBDSA-labeled phosphopeptides containing arginine residues. This suggests that the relative strengths of the hydrogen-bonding interactions play an important role in subsequent fragmentation of the peptides upon CID.

The use of a front-end dual spray reactor to facilitate the ion/ion reaction step allows this method to be widely adaptable to nearly all mass spectrometer platforms, making it a viable option for phosphopeptide analysis when conventional CID-based characterization yields ambiguous sequence and phosphosite information. Moreover, previous demonstrations of reactor compatibility with high-throughput chromatographic

workflows[39] hint at the potential utility of dual source integration into phosphoproteomic analysis.

7.6 REFERENCES

(1) Cohen, P. *Nat. Cell Biol.* **2002**, *4* (5), E127–E130.
(2) Ubersax, J. A.; Ferrell Jr, J. E. *Nat. Rev. Mol. Cell Biol.* **2007**, *8* (7), 530–541.
(3) Johnson, L. N. *Biochem. Soc. Trans.* **2009**, *37* (4), 627–641.
(4) Hunter, T. *Cell* **1995**, *80* (2), 225–236.
(5) Lahiry, P.; Torkamani, A.; Schork, N. J.; Hegele, R. A. *Nat. Rev. Genet.* **2010**, *11* (1), 60–74.
(6) Hendriks, W. J. A. J.; Pulido, R. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **2013**, *1832* (10), 1673–1696.
(7) Humphrey, S. J.; James, D. E.; Mann, M. *Trends Endocrinol. Metab.* **2015**, *26* (12), 676–687.
(8) Mann, M.; Ong, S.-E.; Grønborg, M.; Steen, H.; Jensen, O. N.; Pandey, A. *Trends Biotechnol.* **2002**, *20* (6), 261–268.
(9) Riley, N. M.; Coon, J. J. *Anal. Chem.* **2016**, *88* (1), 74–94.
(10) Michalski, A.; Cox, J.; Mann, M. *J. Proteome Res.* **2011**, *10* (4), 1785–1793.
(11) Bauer, M.; Ahrné, E.; Baron, A. P.; Glatter, T.; Fava, L. L.; Santamaria, A.; Nigg, E. A.; Schmidt, A. *J. Proteome Res.* **2014**, *13* (12), 5973–5988.
(12) Palumbo, A. M.; Smith, S. A.; Kalcic, C. L.; Dantus, M.; Stemmer, P. M.; Reid, G. E. *Mass Spectrom. Rev.* **2011**, *30* (4), 600–625.
(13) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.
(14) Boersema, P. J.; Mohammed, S.; Heck, A. J. R. *J. Mass Spectrom.* **2009**, *44* (6), 861–878.
(15) Jones, A. W.; Cooper, H. J. *Analyst* **2011**, *136* (17), 3419–3429.
(16) Palumbo, A. M.; Tepe, J. J.; Reid, G. E. *J. Proteome Res.* **2008**, *7* (2), 771–779.
(17) Chi, A.; Huttenhower, C.; Geer, L. Y.; Coon, J. J.; Syka, J. E. P.; Bai, D. L.; Shabanowitz, J.; Burke, D. J.; Troyanskaya, O. G.; Hunt, D. F. *Proc. Natl. Acad. Sci.* **2007**, *104* (7), 2193–2198.
(18) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14* (19), 1793–1800.
(19) Nagaraj, N.; D'Souza, R. C. J.; Cox, J.; Olsen, J. V.; Mann, M. *J. Proteome Res.* **2010**, *9* (12), 6786–6794.
(20) Kim, T.-Y.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (12), 2334–2341.
(21) Shin, Y. S.; Moon, J. H.; Kim, M. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (1), 53–59.
(22) Madsen, J. A.; Kaoud, T. S.; Dalby, K. N.; Brodbelt, J. S. *PROTEOMICS* **2011**, *11* (7), 1329–1334.
(23) Fort, K. L.; Dyachenko, A.; Potel, C. M.; Corradini, E.; Marino, F.; Barendregt, A.; Makarov, A. A.; Scheltema, R. A.; Heck, A. J. R. *Anal. Chem.* **2016**, *88* (4), 2303–2310.
(24) Swaney, D. L.; McAlister, G. C.; Wirtala, M.; Schwartz, J. C.; Syka, J. E. P.; Coon, J. J. *Anal. Chem.* **2007**, *79* (2), 477–485.

(25) Frese, C. K.; Zhou, H.; Taus, T.; Altelaar, A. F. M.; Mechtler, K.; Heck, A. J. R.; Mohammed, S. *J. Proteome Res.* **2013**, *12* (3), 1520–1525.
(26) Shaffer, C. J.; Slováková, K.; Tureček, F. *Int. J. Mass Spectrom.* **2015**, *390*, 71–80.
(27) Meyer, H. E.; Hoffmann-Posorske, E.; Korte, H.; Heilmeyer, L. M. G. *FEBS Lett.* **1986**, *204* (1), 61–66.
(28) Oda, Y.; Nagasu, T.; Chait, B. T. *Nat. Biotechnol.* **2001**, *19* (4), 379–382.
(29) Molloy, M. P.; Andrews, P. C. *Anal. Chem.* **2001**, *73* (22), 5387–5394.
(30) Thompson, A. J.; Hart, S. R.; Franz, C.; Barnouin, K.; Ridley, A.; Cramer, R. *Anal. Chem.* **2003**, *75* (13), 3232–3243.
(31) Knight, Z. A.; Schilling, B.; Row, R. H.; Kenski, D. M.; Gibson, B. W.; Shokat, K. M. *Nat. Biotechnol.* **2003**, *21* (9), 1047–1054.
(32) Diedrich, J. K.; Julian, R. R. *Anal. Chem.* **2011**, *83* (17), 6818–6826.
(33) Diedrich, J. K.; Julian, R. R. *J. Am. Chem. Soc.* **2008**, *130* (37), 12212–12213.
(34) McLachlin, D. T.; Chait, B. T. *Anal. Chem.* **2003**, *75* (24), 6826–6836.
(35) Gronert, S.; Huang, R.; Li, K. H. *Int. J. Mass Spectrom.* **2004**, *231* (2–3), 179–187.
(36) Gronert, S.; Li, K. H.; Horiuchi, M. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (12), 1905–1914.
(37) Leitner, A.; Foettinger, A.; Lindner, W. *J. Mass Spectrom.* **2007**, *42* (7), 950–959.
(38) Lanucara, F.; Lee, D. C. H.; Eyers, C. E. *J. Am. Soc. Mass Spectrom.* **2013**, *25* (2), 214–225.
(39) Cotham, V. C.; Shaw, J. B.; Brodbelt, J. S. *Anal. Chem.* **2015**, *87* (18), 9396–9402.
(40) Han, H.; McLuckey, S. A. *J. Am. Chem. Soc.* **2009**, *131* (36), 12884–12885.
(41) Hassell, K. M.; Stutzman, J. R.; McLuckey, S. A. *Anal. Chem.* **2010**, *82* (5), 1594–1597.
(42) Stutzman, J. R.; McLuckey, S. A. *Anal. Chem.* **2012**, *84* (24), 10679–10685.
(43) Stutzman, J. R.; Luongo, C. A.; McLuckey, S. A. *J. Mass Spectrom.* **2012**, *47* (6), 669–675.
(44) Schug, K. A.; Lindner, W. *Chem. Rev.* **2005**, *105* (1), 67–114.
(45) Woods, A. S.; Wang, H.-Y. J.; Jackson, S. N. *J. Proteome Res.* **2007**, *6* (3), 1176–1182.
(46) Tsaprailis, G.; Somogyi, A.; Nikolaev, E. N.; Wysocki, V. H. *Int. J. Mass Spectrom.* **2000**, *195-196*, 467–479.
(47) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35* (12), 1399–1406.
(48) Bleiholder, C.; Suhai, S.; Harrison, A. G.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 1032–1039.

Chapter 8

Conclusions

8.1 Summary of Chapters

Tremendous advances in instrumentation and ion activation methods have established mass spectrometry as an essential analytical tool for the structural characterization of therapeutic antibodies, and have led to its increasing role in new frontiers related to vaccine development and antibody-based biomarker and drug discovery. Nevertheless, there remain opportunities for extending the versatility of tandem mass spectrometric approaches that are anticipated to have important biotechnological implications, both with respect to elevating quality control standards and as a means to more effectively evaluated antibody-mediated immunity and leverage its therapeutic potential. The work presented in this dissertation sought to advance mass spectrometry-based antibody analysis using several variations of ultraviolet photodissociation (UVPD) to either enhance selectivity for regions of diagnostic value in the context of IgG mixtures or to facilitate more detailed structural characterization in the interest of improving quality control.

Using site-specific derivatization, strategic proteolysis, and chromophore-mediated 351 nm UVPD, as described in Chapter 3, a method was developed to streamline the identification of unique antibodies in mixtures by enhancing selectivity for their diagnostic antigen-binding CDR-H3 regions. As demonstrate in the amino acid frequency plot shown in **Figure 8.1**, the highly conserve cysteine residue adjacent to the heavy chain CDR-H3 serves as an ideal analytical target for site-directed tagging with thiol-selective AlexaFluor 350 maleimide. Tailoring enzymatic digestion to include the

modified cysteine residue and adjacent CDR-H3 region allowed facile discrimination of these peptides within a high-throughput 351 nm UVPD workflow.

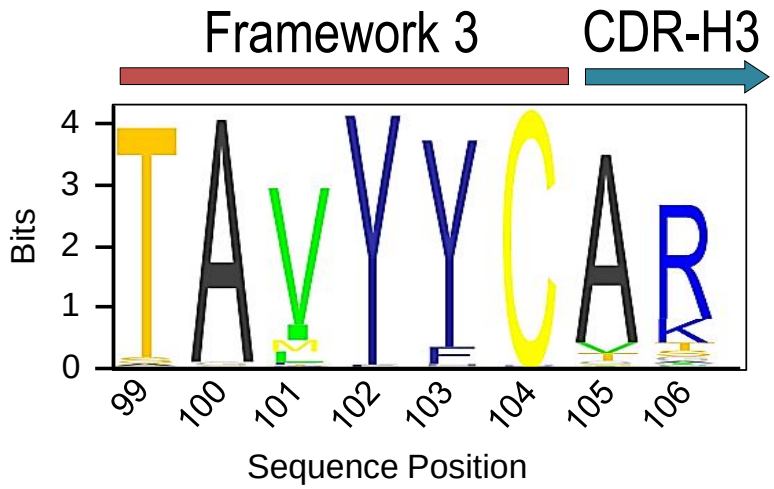


Figure 8.1 Sequence logo demonstrating the amino acid frequencies at the C-terminal portion of framework 3 and the N-terminal region of CDR-H3 of the IgG heavy chain based on IMGT reference genes (www.imgt.org).

Lys-C proteolysis combined with 193 nm UVPD was utilized in Chapter 4 to enhance the characterization of monoclonal antibodies. While restricted digestion with Lys-C effectively reduced sample complexity to facilitate more efficient sampling of the resulting IgG peptide population, the greatest merit of this method resulted from the generation of peptides spanning the entire length of the CDR-H3 region. The sequencing power of 193 nm UVPD was then exploited to comprehensively characterize these diagnostic peptides, often yielding 100% coverage of the CDR-H3 sequence and affording more confident peptide spectral matching compared to collision- and electron-based methods. As a simple, but informative test to evaluate the potential utility of this strategy for proteomic analysis of antibody mixtures, a sample containing several unknown anti-influenza monoclonal antibodies discovered in the serum repertoire of a

post-vaccinated donor was subjected to Lys-C digestion and analyzed by middle-down 193 nm UVPD. The resulting MS/MS data was searched against a database constructed by next-generation sequencing of V-genes from donor B cells. The searches resulted in unambiguous antibody identification.

The utility of 193 nm UVPD for the characterization of antibody subunits was demonstrated in Chapter 5. The tunable nature of energy deposition by UVPD afforded a degree of control over which regions of the subunit sequence could be interrogated in great detail, which is not generally observed for more conventional activation methods. Moreover, UVPD yielded up to 60% coverage of the IgG sequence within a single targeted LC-MS/MS run while also maintaining the integrity of the labile N-linked glycosylation site. Collectively, this allowed confident glycosylation site localization in the Fc/2 heavy chain constant domain and confirmation of the antigen-binding CDR sequences in the variable portions of the Lc and Fd subunits.

A frontend dual spray reactor was developed and implemented in Chapter 6 for the high-throughput bioconjugation of peptides. While not employed directly for the analysis of antibodies, this method affords a rapid and universal means for enhancing primary sequence characterization of peptides when paired with 193 nm UVPD and was shown to be readily implemented into an automated LC-MS^n workflow.

The dual source method was further extended in Chapter 7 for the improved characterization of phosphopeptides. Gas-phase derivatization of phosphopeptides with 4-formyl-1,3-benzenedisulfonic acid (FBDSA) resulted in suppressed phosphate elimination during collisional activation in addition to a concomitant enhancement in sequence-informative product ion formation. Further interrogation of this change in dissociation behavior following derivatization revealed the critical role of sulfonate

interference with charge-directed phosphate neutral loss mechanisms that typically govern the fragmentation of phosphopeptides.

8.2 FUTURE DIRECTIONS

Perhaps the most intriguing implications of the methods developed in this dissertation arise from the potential merits of adapting UVPD-based strategies to advance proteomic analysis of serological antibody repertoires. While the scope of the initial evaluations described in Chapters 3 and 4 were limited to low-complexity samples, they provide compelling evidence for the utility of UVPD for differentiating clonally unique antibodies in complex mixtures.

Owing to the fast and extensive fragmentation afforded by 193 nm UVPD, as well as the ability to generate full length CDR-H3 peptides using restricted Lys-C digestion that are readily adaptable to front-end separations, the middle-down approach described in Chapter 4 likely holds the greatest potential for improving serum immunoproteomics. As a logical next step, this method should be evaluated for complex serum antibody repertoires in combination with Ig-seq V-gene database searching. Although this approach many not achieve the same sensitivity or depth as established bottom-up methods, it is anticipated to provide complementary antibody identifications of much higher confidence.

To further improve this method and increase its sensitivity for CDR-H3 peptides, one can envision coupling restricted Lys-C digestion with thiol-selective enrichment prior to LC-MS/UVPD analysis as illustrated in **Figure 8.2**. This method would take advantage of the highly conserved cysteine residue located N-terminal to the CDR-H3 region (**Figure 8.1**) in a similar manner to the selective tagging strategy described in Chapter 3.

Selective pull-downs on thiol capture resins should dramatically reduce sample complexity while also increasing the relative abundance of Lys-C CDR-H3 peptides for improved detection.

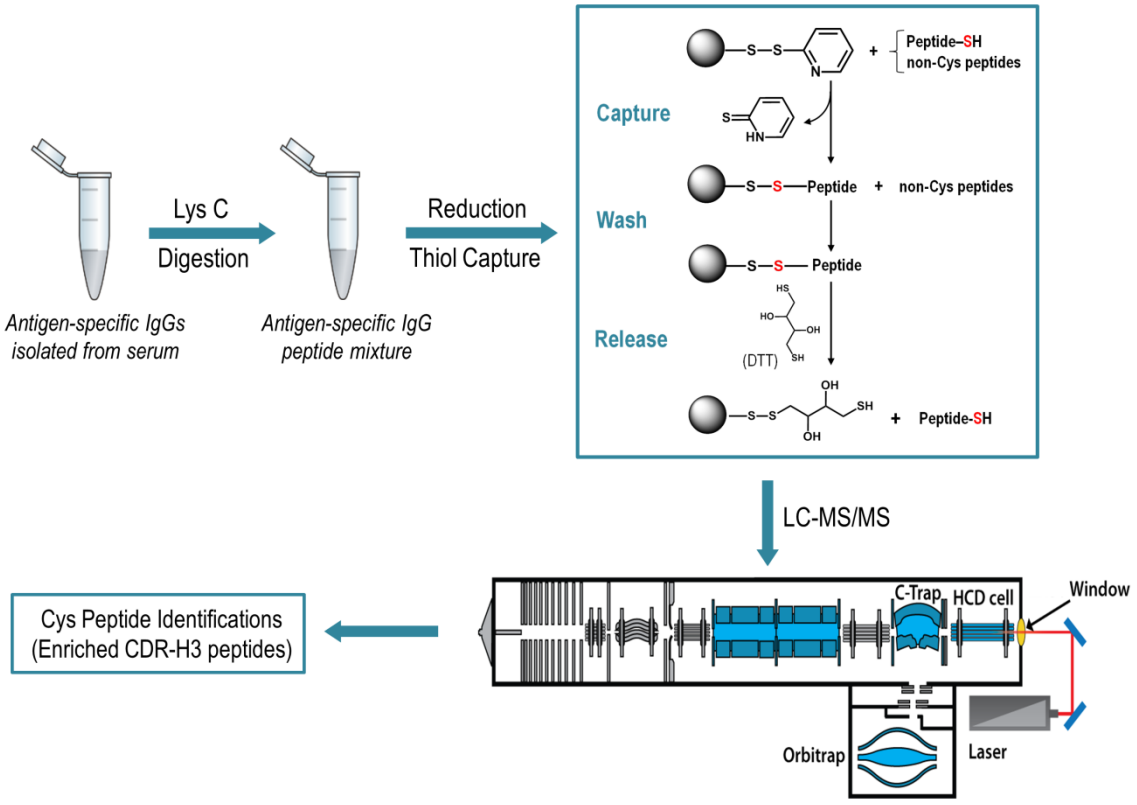


Figure 8.2 Proposed selective enrichment workflow for Lys-C generated CDR-H3 peptides.

References

CHAPTER 1

(1) Weiner, L. M.; Surana, R.; Wang, S. *Nat. Rev. Immunol.* **2010**, *10* (5), 317–327.
(2) Chan, A. C.; Carter, P. J. *Nat. Rev. Immunol.* **2010**, *10* (5), 301–316.
(3) Wang, W.; Singh, S.; Zeng, D. L.; King, K.; Nema, S. *J. Pharm. Sci.* **2007**, *96* (1), 1–26.
(4) Walsh, G.; Jefferis, R. *Nat. Biotechnol.* **2006**, *24* (10), 1241–1252.
(5) De Groot, A. S.; Scott, D. W. *Trends Immunol.* **2007**, *28* (11), 482–490.
(6) Harding, F. A.; Stickler, M. M.; Razo, J.; DuBridge, R. B. *mAbs* **2010**, *2* (3), 256–265.
(7) Tjalsma, H.; Schaeps, R. M. J.; Swinkels, D. W. *PROTEOMICS – Clin. Appl.* **2008**, *2* (2), 167–180.
(8) Georgiou, G.; Ippolito, G. C.; Beausang, J.; Busse, C. E.; Wardemann, H.; Quake, S. R. *Nat. Biotechnol.* **2014**, *32* (2), 158–168.
(9) Lavinder, J. J.; Horton, A. P.; Georgiou, G.; Ippolito, G. C. *Curr. Opin. Chem. Biol.* **2015**, *24*, 112–120.
(10) Boutz, D. R.; Horton, A. P.; Wine, Y.; Lavinder, J. J.; Georgiou, G.; Marcotte, E. M. *Anal. Chem.* **2014**, *86* (10), 4758–4766.
(11) Parkin, J.; Cohen, B. *The Lancet* **2001**, *357* (9270), 1777–1789.
(12) Akira, S.; Uematsu, S.; Takeuchi, O. *Cell* **2006**, *124* (4), 783–801.
(13) Mogensen, T. H. *Clin. Microbiol. Rev.* **2009**, *22* (2), 240–273.
(14) Paul, W. E. *Cell* **2011**, *147* (6), 1212–1215.
(15) Nemazee, D. *Nat. Rev. Immunol.* **2006**, *6* (10), 728–740.
(16) Jung, D.; Giallourakis, C.; Mostoslavsky, R.; Alt, F. W. *Annu. Rev. Immunol.* **2006**, *24* (1), 541–570.
(17) McHeyzer-Williams, M.; Okitsu, S.; Wang, N.; McHeyzer-Williams, L. *Nat. Rev. Immunol.* **2012**, *12* (1), 24–34.
(18) Kurosaki, T.; Kometani, K.; Ise, W. *Nat. Rev. Immunol.* **2015**, *15* (3), 149–159.
(19) Batista, F. D.; Harwood, N. E. *Nat. Rev. Immunol.* **2009**, *9* (1), 15–27.
(20) Murphy, K. *Janeway's Immunobiology, Eighth Edition*; Garland Science, 2011.
(21) Nimmerjahn, F.; Ravetch, J. V. *Nat. Rev. Immunol.* **2008**, *8* (1), 34–47.
(22) Borrok, M. J.; Jung, S. T.; Kang, T. H.; Monzingo, A. F.; Georgiou, G. *ACS Chem. Biol.* **2012**, *7* (9), 1596–1602.
(23) Ricklin, D.; Hajishengallis, G.; Yang, K.; Lambris, J. D. *Nat. Immunol.* **2010**, *11* (9), 785–797.
(24) Diebolder, C. A.; Beurskens, F. J.; Jong, R. N. de; Koning, R. I.; Strumane, K.; Lindorfer, M. A.; Voorhorst, M.; Ugurlar, D.; Rosati, S.; Heck, A. J. R.; Winkel, J. G. J. van de; Wilson, I. A.; Koster, A. J.; Taylor, R. P.; Saphire, E. O.; Burton,

D. R.; Schuurman, J.; Gros, P.; Parren, P. W. H. I. *Science* **2014**, *343* (6176), 1260–1263.
(25) Schroeder, H. W.; Cavacini, L. *J. Allergy Clin. Immunol.* **2010**, *125* (2), S41–S52.
(26) Wine, Y.; Boutz, D. R.; Lavinder, J. J.; Miklos, A. E.; Hughes, R. A.; Hoi, K. H.; Jung, S. T.; Horton, A. P.; Murrin, E. M.; Ellington, A. D.; Marcotte, E. M.; Georgiou, G. *Proc. Natl. Acad. Sci.* **2013**, *110* (8), 2993–2998.
(27) Cheung, W. C.; Beausoleil, S. A.; Zhang, X.; Sato, S.; Schieferl, S. M.; Wieler, J. S.; Beaudet, J. G.; Ramenani, R. K.; Popova, L.; Comb, M. J.; Rush, J.; Polakiewicz, R. D. *Nat Biotech* **2012**, *30* (5), 447–452.
(28) Sato, S.; Beausoleil, S. A.; Popova, L.; Beaudet, J. G.; Ramenani, R. K.; Zhang, X.; Wieler, J. S.; Schieferl, S. M.; Cheung, W. C.; Polakiewicz, R. D. *Nat. Biotechnol.* **2012**, *30* (11), 1039–1043.
(29) Lavinder, J. J.; Wine, Y.; Giesecke, C.; Ippolito, G. C.; Horton, A. P.; Lungu, O. I.; Hoi, K. H.; DeKosky, B. J.; Murrin, E. M.; Wirth, M. M.; Ellington, A. D.; Dörner, T.; Marcotte, E. M.; Boutz, D. R.; Georgiou, G. *Proc. Natl. Acad. Sci.* **2014**, *111* (6), 2259–2264.
(30) Schatz, D. G. *Immunol. Rev.* **2004**, *200* (1), 5–11.
(31) Schroeder Jr, H. W. *Dev. Comp. Immunol.* **2006**, *30* (1–2), 119–135.
(32) Walsh, G. *Nat. Biotechnol.* **2014**, *32* (10), 992–1000.
(33) Ecker, D. M.; Jones, S. D.; Levine, H. L. *mAbs* **2015**, *7* (1), 9–14.
(34) Aggarwal, S. *Nat. Biotechnol.* **2014**, *32* (1), 32–39.
(35) Zeck, A.; Regula, J. T.; Larraillet, V.; Mautz, B.; Popp, O.; Göpfert, U.; Wiegeshoff, F.; Vollertsen, U. E. E.; Gorr, I. H.; Koll, H.; Papadimitriou, A. *PLOS ONE* **2012**, *7* (7), e40328.
(36) Yang, Y.; Strahan, A.; Li, C.; Shen, A.; Liu, H.; Ouyang, J.; Katta, V.; Francissen, K.; Zhang, B. *mAbs* **2010**, *2* (3), 285–298.
(37) Miyazaki, C.; Iba, Y.; Yamada, Y.; Takahashi, H.; Sawada, J.; Kurosawa, Y. *Protein Eng.* **1999**, *12* (5), 407–415.
(38) Winkler, K.; Kramer, A.; Küttner, G.; Seifert, M.; Scholz, C.; Wessner, H.; Schneider-Mergener, J.; Höhne, W. *J. Immunol.* **2000**, *165* (8), 4505–4514.
(39) Liu, H.; Gaza-Bulseco, G.; Faldu, D.; Chumsae, C.; Sun, J. *J. Pharm. Sci.* **2008**, *97* (7), 2426–2447.
(40) Zhang, Z.; Pan, H.; Chen, X. *Mass Spectrom. Rev.* **2009**, *28* (1), 147–176.
(41) Zhang, H.; Cui, W.; Gross, M. L. *FEBS Lett.* **2014**, *588* (2), 308–317.
(42) Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianférani, S. *Anal. Chem.* **2013**, *85* (2), 715–736.
(43) Lundell, N.; Schreitmüller, T. *Anal. Biochem.* **1999**, *266* (1), 31–47.
(44) Bandeira, N.; Pham, V.; Pevzner, P.; Arnott, D.; Lill, J. R. *Nat. Biotechnol.* **2008**, *26* (12), 1336–1338.
(45) de Costa, D.; Broodman, I.; VanDuijn, M. M.; Stingl, C.; Dekker, L. J. M.; Burgers, P. C.; Hoogsteden, H. C.; Sillevis Smitt, P. A. E.; van Klaveren, R. J.; Luider, T. M. *J. Proteome Res.* **2010**, *9* (6), 2937–2945.

(46) Dekker, L. J. M.; Zeneyedpour, L.; Brouwer, E.; Duijn, M. M. van; Smitt, P. A. E. S.; Luider, T. M. *Anal. Bioanal. Chem.* **2011**, *399* (3), 1081–1091.
(47) Srebalus Barnes, C. A.; Lim, A. *Mass Spectrom. Rev.* **2007**, *26* (3), 370–388.
(48) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121* (4), 806–812.
(49) Zhou, H.; Ning, Z.; E. Starr, A.; Abu-Farha, M.; Figeys, D. *Anal. Chem.* **2012**, *84* (2), 720–734.
(50) Kelleher, N. L.; Thomas, P. M.; Ntai, I.; Compton, P. D.; LeDuc, R. D. *Expert Rev. Proteomics* **2014**, *11* (6), 649–651.
(51) Cui, W.; Rohrs, H. W.; Gross, M. L. *Analyst* **2011**, *136* (19), 3854–3864.
(52) Tsybin, Y. O.; Fornelli, L.; Stoermer, C.; Luebeck, M.; Parra, J.; Nallet, S.; Wurm, F. M.; Hartmer, R. *Anal. Chem.* **2011**, *83* (23), 8919–8927.
(53) Fornelli, L.; Damoc, E.; Thomas, P. M.; Kelleher, N. L.; Aizikov, K.; Denisov, E.; Makarov, A.; Tsybin, Y. O. *Mol. Cell. Proteomics* **2012**, *11* (12), 1758–1767.
(54) Mao, Y.; Valeja, S. G.; Rouse, J. C.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2013**, *85* (9), 4239–4246.
(55) Kleemann, G. R.; Beierle, J.; Nichols, A. C.; Dillon, T. M.; Pipes, G. D.; Bondarenko, P. V. *Anal. Chem.* **2008**, *80* (6), 2001–2009.
(56) Srzentić, K.; Fornelli, L.; Laskay, Ü. A.; Monod, M.; Beck, A.; Ayoub, D.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (19), 9945–9953.
(57) Cannon, J.; Lohnes, K.; Wynne, C.; Wang, Y.; Edwards, N.; Fenselau, C. *J. Proteome Res.* **2010**, *9* (8), 3886–3890.
(58) 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. *Nat. Methods* **2012**, *9* (8), 822–824.
(59) Pang, Y.; Wang, W.-H.; Reid, G. E.; Hunt, D. F.; Bruening, M. L. *Anal. Chem.* **2015**, *87* (21), 10942–10949.
(60) Fornelli, L.; Ayoub, D.; Aizikov, K.; Beck, A.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (6), 3005–3012.
(61) Wang, D.; Wynne, C.; Gu, F.; Becker, C.; Zhao, J.; Mueller, H.-M.; Li, H.; Shameem, M.; Liu, Y.-H. *Anal. Chem.* **2015**, *87* (2), 914–921.
(62) Nicolardi, S.; Deelder, A. M.; Palmblad, M.; van der Burgt, Y. E. M. *Anal. Chem.* **2014**, *86* (11), 5376–5382.
(63) Yan, B.; Valliere-Douglass, J.; Brady, L.; Steen, S.; Han, M.; Pace, D.; Elliott, S.; Yates, Z.; Han, Y.; Balland, A.; Wang, W.; Pettit, D. *J. Chromatogr. A* **2007**, *1164* (1–2), 153–161.
(64) An, Y.; Zhang, Y.; Mueller, H.-M.; Shameem, M.; Chen, X. *mAbs* **2014**, *6* (4), 879–893.
(65) Tran, J. C.; Zamdborg, L.; Ahlf, D. R.; Lee, J. E.; Catherman, A. D.; Durbin, K. R.; Tipton, J. D.; Vellaichamy, A.; Kellie, J. F.; Li, M.; Wu, C.; Sweet, S. M. M.; Early, B. P.; Siuti, N.; LeDuc, R. D.; Compton, P. D.; Thomas, P. M.; Kelleher, N. L. *Nature* **2011**, *480* (7376), 254–258.
(66) Garcia, B. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (2), 193–202.

(67) Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L. *Anal. Chem.* **2011**, *83* (17), 6868–6874.
(68) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.
(69) Roepstorff, P.; Fohlman, J. *Biol. Mass Spectrom.* **1984**, *11* (11), 601.
(70) Biemann, K. *Biol. Mass Spectrom.* **1988**, *16* (1–12), 99–111.
(71) Louris, J. N.; Cooks, R. G.; Syka, J. E. P.; Kelley, P. E.; Stafford, G. C.; Todd, J. F. J. *Anal. Chem.* **1987**, *59* (13), 1677–1685.
(72) McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3* (6), 599–614.
(73) Dongré, A. R.; Jones, J. L.; Somogyi, Á.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118* (35), 8365–8374.
(74) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35* (12), 1399–1406.
(75) Vaisar, T.; Urban, J. *J. Mass Spectrom.* **1996**, *31* (10), 1185–1187.
(76) Bleiholder, C.; Suhai, S.; Harrison, A. G.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 1032–1039.
(77) Tsaprailis, G.; Somogyi, A.; Nikolaev, E. N.; Wysocki, V. H. *Int. J. Mass Spectrom.* **2000**, *195–196*, 467–479.
(78) Gu, C.; Tsaprailis, G.; Breci, L.; Wysocki, V. H. *Anal. Chem.* **2000**, *72* (23), 5804–5813.
(79) Huang, Y.; Triscari, J. M.; Tseng, G. C.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *Anal. Chem.* **2005**, *77* (18), 5800–5813.
(80) Fekete, S.; Guillarme, D.; Sandra, P.; Sandra, K. *Anal. Chem.* **2016**, *88* (1), 480–507.
(81) Desaire, H. *Mol. Cell. Proteomics* **2013**, *12* (4), 893–901.
(82) Dunbar, R. C. *J. Am. Chem. Soc.* **1971**, *93* (18), 4354–4358.
(83) Freiser, B. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1974**, *96* (20), 6260–6266.
(84) Ly, T.; Julian, R. R. *Angew. Chem. Int. Ed.* **2009**, *48* (39), 7130–7137.
(85) Brodbelt, J. S. *Chem. Soc. Rev.* **2014**, *43* (8), 2757–2783.
(86) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. *J. Proteome Res.* **2010**, *9* (8), 4205–4214.
(87) Vasicek, L.; Brodbelt, J. S. *Anal. Chem.* **2010**, *82* (22), 9441–9446.
(88) Robinson, M. R.; Madsen, J. A.; Brodbelt, J. S. *Anal. Chem.* **2012**, *84* (5), 2433–2439.
(89) Robinson, M. R.; Moore, K. L.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (8), 1461–1471.
(90) Fort, K. L.; Dyachenko, A.; Potel, C. M.; Corradini, E.; Marino, F.; Barendregt, A.; Makarov, A. A.; Scheltema, R. A.; Heck, A. J. R. *Anal. Chem.* **2016**, *88* (4), 2303–2310.
(91) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.

(92) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(93) Cannon, J. R.; Kluwe, C.; Ellington, A.; Brodbelt, J. S. *PROTEOMICS* **2014**, *14* (10), 1165–1173.
(94) O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2014**, *136* (37), 12920–12928.
(95) Cammarata, M. B.; Thyer, R.; Rosenberg, J.; Ellington, A.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2015**, *137* (28), 9128–9135.
(96) Holden, D. D.; McGee, W. M.; Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 1008–1016.
(97) Shaw, J. B.; Robinson, E. W.; Paša-Tolić, L. *Anal. Chem.* **2016**, *88* (6), 3019–3023.
(98) Woody, R. W.; Koslowski, A. *Biophys. Chem.* **2002**, *101–102*, 535–551.
(99) Cui, W.; Thompson, M. S.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (8), 1384–1398.
(100) Kim, T.-Y.; Thompson, M. S.; Reilly, J. P. *Rapid Commun. Mass Spectrom.* **2005**, *19* (12), 1657–1665.
(101) Yoon, S. H.; Chung, Y. J.; Kim, M. S. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (5), 645–655.
(102) Thompson, M. S.; Cui, W.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (8), 1439–1452.
(103) Madsen, J. A.; Ko, B. J.; Xu, H.; Iwashkiw, J. A.; Robotham, S. A.; Shaw, J. B.; Feldman, M. F.; Brodbelt, J. S. *Anal. Chem.* **2013**, *85* (19), 9253–9261.
(104) O'Brien, J. P.; Pruet, J. M.; Brodbelt, J. S. *Anal. Chem.* **2013**, *85* (15), 7391–7397.
(105) Robotham, S. A.; Kluwe, C.; Cannon, J. R.; Ellington, A.; Brodbelt, J. S. *Anal. Chem.* **2013**, *85* (20), 9832–9838.
(106) Robotham, S. A.; Horton, A. P.; Cannon, J. R.; Cotham, V. C.; Marcotte, E. M.; Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (7), 3990–3997.
(107) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120* (13), 3265–3266.
(108) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (26), 9528–9533.
(109) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14* (19), 1793–1800.
(110) Chi, A.; Huttenhower, C.; Geer, L. Y.; Coon, J. J.; Syka, J. E. P.; Bai, D. L.; Shabanowitz, J.; Burke, D. J.; Troyanskaya, O. G.; Hunt, D. F. *Proc. Natl. Acad. Sci.* **2007**, *104* (7), 2193–2198.
(111) Wiesner, J.; Premsler, T.; Sickmann, A. *PROTEOMICS* **2008**, *8* (21), 4466–4483.
(112) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72* (3), 563–573.

(113) Breuker, K.; Oh, H.; Lin, C.; Carpenter, B. K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (39), 14011–14016.
(114) Olsen, J. V.; Schwartz, J. C.; Griep-Raming, J.; Nielsen, M. L.; Damoc, E.; Denisov, E.; Lange, O.; Remes, P.; Taylor, D.; Splendore, M.; Wouters, E. R.; Senko, M.; Makarov, A.; Mann, M.; Horning, S. *Mol. Cell. Proteomics* **2009**, *8* (12), 2759–2769.
(115) Chi, A.; Bai, D. L.; Geer, L. Y.; Shabanowitz, J.; Hunt, D. F. *Int. J. Mass Spectrom.* **2007**, *259* (1–3), 197–203.
(116) McAlister, G. C.; Berggren, W. T.; Griep-Raming, J.; Horning, S.; Makarov, A.; Phanstiel, D.; Stafford, G.; Swaney, D. L.; Syka, J. E. P.; Zabrouskov, V.; Coon, J. J. *J. Proteome Res.* **2008**, *7* (8), 3127–3136.
(117) Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119* (9), 2240–2248.
(118) Breuker, K.; Oh, H.; Horn, D. M.; Cerda, B. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124* (22), 6407–6420.
(119) Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. *Mol. Cell. Proteomics* **2007**, *6* (11), 1942–1951.
(120) Tran, B. Q.; Barton, C.; Feng, J.; Sandjong, A.; Yoon, S. H.; Awasthi, S.; Liang, T.; Khan, M. M.; Kilgour, D. P. A.; Goodlett, D. R.; Goo, Y. A. *J. Proteomics* **2016**, *134*, 93–101.
(121) Bondarenko, P. V.; Second, T. P.; Zabrouskov, V.; Makarov, A. A.; Zhang, Z. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (8), 1415–1424.

CHAPTER 2

(1) Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. *Science* **1989**, *246* (4926), 64–71.
(2) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Mass Spectrom. Rev.* **1990**, *9* (1), 37–70.
(3) Cech, N. B.; Enke, C. G. *Mass Spectrom. Rev.* **2001**, *20* (6), 362–387.
(4) Douglas, D. J.; Frank, A. J.; Mao, D. *Mass Spectrom. Rev.* **2005**, *24* (1), 1–29.
(5) Makarov, A. *Anal. Chem.* **2000**, *72* (6), 1156–1162.
(6) Michalski, A.; Damoc, E.; Lange, O.; Denisov, E.; Nolting, D.; Müller, M.; Viner, R.; Schwartz, J.; Remes, P.; Belford, M.; Dunyach, J.-J.; Cox, J.; Horning, S.; Mann, M.; Makarov, A. *Mol. Cell. Proteomics* **2012**, *11* (3), O111.013698.
(7) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(8) Vasicek, L. A.; Ledvina, A. R.; Shaw, J.; Griep-Raming, J.; Westphall, M. S.; Coon, J. J.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 1105–1108.

(9) Gardner, M. W.; Smith, S. I.; Ledvina, A. R.; Madsen, J. A.; Coon, J. J.; Schwartz, J. C.; Stafford, G. C.; Brodbelt, J. S. *Anal. Chem.* **2009**, *81* (19), 8109–8118.
(10) Rani, M.; Bolles, M.; Donaldson, E. F.; Van Blarcom, T.; Baric, R.; Iverson, B.; Georgiou, G. *J. Virol.* **2012**, *86* (17), 9113–9121.
(11) Ito, H.; Sagane, Y.; Miyata, K.; Inui, K.; Matsuo, T.; Horiuchi, R.; Ikeda, T.; Suzuki, T.; Hasegawa, K.; Kouguchi, H.; Oguma, K.; Niwa, K.; Ohyama, T.; Watanabe, T. *FEMS Immunol. Med. Microbiol.* **2011**, *61* (3), 323–331.
(12) Krebber, A.; Bornhauser, S.; Burmester, J.; Honegger, A.; Willuda, J.; Bosshard, H. R.; Plückthun, A. *J. Immunol. Methods* **1997**, *201* (1), 35–55.
(13) Hayhurst, A.; Happe, S.; Mabry, R.; Koch, Z.; Iverson, B. L.; Georgiou, G. *J. Immunol. Methods* **2003**, *276* (1–2), 185–196.

CHAPTER 3

(1) Yates, J. R. *J. Mass Spectrom.* 1998, *33* (1), 1–19.
(2) Aebersold, R.; Mann, M. *Nature* 2003, *422* (6928), 198–207.
(3) Patterson, S. D.; Aebersold, R. H. *Nat. Genet.* 2003, *33*, 311–323.
(4) Domon, B.; Aebersold, R. *Science* 2006, *312* (5771), 212–217.
(5) Mann, M.; Kelleher, N. L. *Proc. Natl. Acad. Sci.* 2008, *105* (47), 18132–18138.
(6) Peng, J.; Gygi, S. P. *J. Mass Spectrom.* 2001, *36* (10), 1083–1091.
(7) Walther, T. C.; Mann, M. *J. Cell Biol.* 2010, *190* (4), 491–500.
(8) Hood, B. L.; Malehorn, D. E.; Conrads, T. P.; Bigbee, W. L. *Methods Mol. Biol. Clifton NJ* 2009, *520*, 107–128.
(9) Coombes, K. R.; Morris, J. S.; Hu, J.; Edmonson, S. R.; Baggerly, K. A. *Nat. Biotechnol.* 2005, *23* (3), 291–292.
(10) Nesvizhskii, A. I.; Vitek, O.; Aebersold, R. *Nat. Methods* 2007, *4* (10), 787–797.
(11) Laskin, J.; Futrell, J. H. *Mass Spectrom. Rev.* 2003, *22* (3), 158–181.
(12) McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* 1992, *3* (6), 599–614.
(13) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U. S. A.* 2004, *101* (26), 9528–9533.
(14) Mikesh, L. M.; Ueberheide, B.; Chi, A.; Coon, J. J.; Syka, J. E. P.; Shabanowitz, J.; Hunt, D. F. *Biochim. Biophys. Acta BBA - Proteins Proteomics* 2006, *1764* (12), 1811–1822.
(15) Zubarev, R. A. *Curr. Opin. Biotechnol.* 2004, *15* (1), 12–16.
(16) Molina, H.; Matthiesen, R.; Kandasamy, K.; Pandey, A. *Anal. Chem.* 2008, *80* (13), 4825–4835.
(17) Molina, H.; Horn, D. M.; Tang, N.; Mathivanan, S.; Pandey, A. *Proc. Natl. Acad. Sci.* 2007, *104* (7), 2199–2204.
(18) Vasicek, L.; O'Brien, J. P.; Browning, K. S.; Tao, Z.; Liu, H.-W.; Brodbelt, J. S. *Mol. Cell. Proteomics* 2012, *11* (7).

(19) Reiter, L.; Claassen, M.; Schrimpf, S. P.; Jovanovic, M.; Schmidt, A.; Buhmann, J. M.; Hengartner, M. O.; Aebersold, R. *Mol. Cell. Proteomics* 2009, *8* (11), 2405–2417.
(20) Elias, J. E.; Haas, W.; Faherty, B. K.; Gygi, S. P. *Nat. Methods* 2005, *2* (9), 667–675.
(21) Stahl, D. C.; Swiderek, K. M.; Davis, M. T.; Lee, T. D. *J. Am. Soc. Mass Spectrom.* 1996, *7* (6), 532–540.
(22) Kim, S.; Davis, M.; Sinn, E.; Patten, P.; Hood, L. *Cell* 1981, *27* (3 Pt 2), 573–581.
(23) Honjo, T.; Habu, S. *Annu. Rev. Biochem.* 1985, *54* (1), 803–830.
(24) de Costa, D.; Broodman, I.; VanDuijn, M. M.; Stingl, C.; Dekker, L. J. M.; Burgers, P. C.; Hoogsteden, H. C.; Sillevis Smitt, P. A. E.; van Klaveren, R. J.; Luider, T. M. *J. Proteome Res.* 2010, *9* (6), 2937–2945.
(25) Wine, Y.; Boutz, D. R.; Lavinder, J. J.; Miklos, A. E.; Hughes, R. A.; Hoi, K. H.; Jung, S. T.; Horton, A. P.; Murrin, E. M.; Ellington, A. D.; Marcotte, E. M.; Georgiou, G. *Proc. Natl. Acad. Sci.* 2013.
(26) Cheung, W. C.; Beausoleil, S. A.; Zhang, X.; Sato, S.; Schieferl, S. M.; Wieler, J. S.; Beaudet, J. G.; Ramenani, R. K.; Popova, L.; Comb, M. J.; Rush, J.; Polakiewicz, R. D. *Nat Biotech* 2012, *30* (5), 447–452.
(27) Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* 2011, *22* (2), 197–206.
(28) Ly, T.; Julian, R. R. *Angew. Chem. Int. Ed.* 2009, *48* (39), 7130–7137.
(29) Reilly, J. P. *Mass Spectrom Rev* 2009, *28* (3), 425–447.
(30) Brodbelt, J. S.; Wilson, J. J. *Mass Spectrom Rev* 2009, *28* (3), 390–424.
(31) Crowe, M. C.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* 2004, *15* (11), 1581–1592.
(32) Crowe, M. C.; Brodbelt, J. S. *Anal. Chem.* 2005, *77* (17), 5726–5734.
(33) Diedrich, J. K.; Julian, R. R. *J. Am. Chem. Soc.* 2008, *130* (37), 12212–12213.
(34) Diedrich, J. K.; Julian, R. R. *Anal. Chem.* 2011, *83* (17), 6818–6826.
(35) Diedrich, J. K.; Julian, R. R. *Anal Chem* 2010, *82* (10), 4006–4014.
(36) Diedrich, J. K.; Julian, R. R. *Anal. Bioanal. Chem.* 2012, *403* (8), 2269–2277.
(37) Wilson, J. J.; Brodbelt, J. S. *Anal Chem* 2007, *79* (20), 7883–7892.
(38) Gardner, M. W.; Brodbelt, J. S. *Anal Chem* 2009, *81* (12), 4864–4872.
(39) Enjalbert, Q.; Simon, R.; Salvador, A.; Antoine, R.; Redon, S.; Ayhan, M. M.; Darbour, F.; Chambert, S.; Bretonnière, Y.; Dugourd, P.; Lemoine, J. *Rapid Commun. Mass Spectrom.* 2011, *25* (22), 3375–3381.
(40) Enjalbert, Q.; Girod, M.; Simon, R.; Jeudy, J.; Chirot, F.; Salvador, A.; Antoine, R.; Dugourd, P.; Lemoine, J. *Anal. Bioanal. Chem.* 1–11.
(41) Morea, V.; Tramontano, A.; Rustici, M.; Chothia, C.; Lesk, A. M. *J. Mol. Biol.* 1998, *275* (2), 269–294.
(42) Kiss, C.; Fisher, H.; Pesavento, E.; Dai, M.; Valero, R.; Ovecka, M.; Nolan, R.; Phipps, M. L.; Velappan, N.; Chasteen, L.; Martinez, J. S.; Waldo, G. S.; Pavlik, P.; Bradbury, A. R. M. *Nucleic Acids Res.* 2006, *34* (19), e132.
(43) Rani, M.; Bolles, M.; Donaldson, E. F.; Van Blarcom, T.; Baric, R.; Iverson, B.; Georgiou, G. *J. Virol.* 2012, *86* (17), 9113–9121.

(44) Ito, H.; Sagane, Y.; Miyata, K.; Inui, K.; Matsuo, T.; Horiuchi, R.; Ikeda, T.; Suzuki, T.; Hasegawa, K.; Kouguchi, H.; Oguma, K.; Niwa, K.; Ohyama, T.; Watanabe, T. *FEMS Immunol. Med. Microbiol.* 2011, *61* (3), 323–331.
(45) Krebber, A.; Bornhauser, S.; Burmester, J.; Honegger, A.; Willuda, J.; Bosshard, H. R.; Plückthun, A. *J. Immunol. Methods* 1997, *201* (1), 35–55.
(46) Hayhurst, A.; Happe, S.; Mabry, R.; Koch, Z.; Iverson, B. L.; Georgiou, G. *J. Immunol. Methods* 2003, *276* (1–2), 185–196.
(47) Gardner, M. W.; Vasicek, L. A.; Shabbir, S.; Anslyn, E. V.; Brodbelt, J. S. *Anal. Chem.* 2008, *80* (13), 4807–4819.
(48) Gardner, M. W.; Smith, S. I.; Ledvina, A. R.; Madsen, J. A.; Coon, J. J.; Schwartz, J. C.; Stafford, G. C.; Brodbelt, J. S. *Anal. Chem.* 2009, *81* (19), 8109–8118.
(49) Xu, H.; Freitas, M. A. *PROTEOMICS* 2009, *9* (6), 1548–1555.

CHAPTER 4

(1) Batista, F. D.; Harwood, N. E. *Nat. Rev. Immunol.* **2009**, *9* (1), 15–27.
(2) Nutt, S. L.; Hodgkin, P. D.; Tarlinton, D. M.; Corcoran, L. M. *Nat. Rev. Immunol.* **2015**, *15* (3), 160–171.
(3) Murphy, K. *Janeway's Immunobiology, Eighth Edition*; Garland Science, 2011.
(4) Schroeder, H. W.; Cavacini, L. *J. Allergy Clin. Immunol.* **2010**, *125* (2), S41–S52.
(5) Tonegawa, S. *Nature* **1983**, *302* (5909), 575–581.
(6) Briney, B. S.; Crowe, J. E. J. *B Cell Biol.* **2013**, *4*, 42.
(7) Goodnow, C. C.; Vinuesa, C. G.; Randall, K. L.; Mackay, F.; Brink, R. *Nat. Immunol.* **2010**, *11* (8), 681–688.
(8) Gitlin, A. D.; Shulman, Z.; Nussenzweig, M. C. *Nature* **2014**, *509* (7502), 637–640.
(9) Wilson, P. C.; Andrews, S. F. *Nat. Rev. Immunol.* **2012**, *12* (10), 709–719.
(10) Georgiou, G.; Ippolito, G. C.; Beausang, J.; Busse, C. E.; Wardemann, H.; Quake, S. R. *Nat. Biotechnol.* **2014**, *32* (2), 158–168.
(11) Robinson, W. H. *Nat. Rev. Rheumatol.* **2015**, *11* (3), 171–182.
(12) Lavinder, J. J.; Wine, Y.; Giesecke, C.; Ippolito, G. C.; Horton, A. P.; Lungu, O. I.; Hoi, K. H.; DeKosky, B. J.; Murrin, E. M.; Wirth, M. M.; Ellington, A. D.; Dörner, T.; Marcotte, E. M.; Boutz, D. R.; Georgiou, G. *Proc. Natl. Acad. Sci.* **2014**, *111* (6), 2259–2264.
(13) de Costa, D.; Broodman, I.; VanDuijn, M. M.; Stingl, C.; Dekker, L. J. M.; Burgers, P. C.; Hoogsteden, H. C.; Sillevis Smitt, P. A. E.; van Klaveren, R. J.; Luider, T. M. *J. Proteome Res.* **2010**, *9* (6), 2937–2945.
(14) VanDuijn, M. M.; Dekker, L. J. M.; Zeneyedpour, L.; Smitt, P. A. E. S.; Luider, T. M. *J. Biol. Chem.* **2010**, *285* (38), 29247–29253.

(15) Dekker, L. J. M.; Zeneyedpour, L.; Brouwer, E.; Duijn, M. M. van; Smitt, P. A. E. S.; Luider, T. M. *Anal. Bioanal. Chem.* **2011**, *399* (3), 1081–1091.
(16) Maat, P.; VanDuijn, M.; Brouwer, E.; Dekker, L.; Zeneyedpour, L.; Luider, T.; Smitt, P. S. *J. Autoimmun.* **2012**, *38* (4), 354–360.
(17) Cheung, W. C.; Beausoleil, S. A.; Zhang, X.; Sato, S.; Schieferl, S. M.; Wieler, J. S.; Beaudet, J. G.; Ramenani, R. K.; Popova, L.; Comb, M. J.; Rush, J.; Polakiewicz, R. D. *Nat Biotech* **2012**, *30* (5), 447–452.
(18) Wine, Y.; Boutz, D. R.; Lavinder, J. J.; Miklos, A. E.; Hughes, R. A.; Hoi, K. H.; Jung, S. T.; Horton, A. P.; Murrin, E. M.; Ellington, A. D.; Marcotte, E. M.; Georgiou, G. *Proc. Natl. Acad. Sci.* **2013**, *110* (8), 2993–2998.
(19) Boutz, D. R.; Horton, A. P.; Wine, Y.; Lavinder, J. J.; Georgiou, G.; Marcotte, E. M. *Anal. Chem.* **2014**, *86* (10), 4758–4766.
(20) Ogishi, M.; Yotsuyanagi, H.; Moriya, K.; Koike, K. *Sci. Rep.* **2016**, *6*, 29532.
(21) Lavinder, J. J.; Horton, A. P.; Georgiou, G.; Ippolito, G. C. *Curr. Opin. Chem. Biol.* **2015**, *24*, 112–120.
(22) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R. *Chem. Rev.* **2013**, *113* (4), 2343–2394.
(23) Fang, X.; Zhang, W.-W. *J. Proteomics* **2008**, *71* (3), 284–303.
(24) Cannon, J.; Lohnes, K.; Wynne, C.; Wang, Y.; Edwards, N.; Fenselau, C. *J. Proteome Res.* **2010**, *9* (8), 3886–3890.
(25) 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. *Nat. Methods* **2012**, *9* (8), 822–824.
(26) Tsiatsiani, L.; Heck, A. J. R. *FEBS J.* **2015**, *282* (14), 2612–2626.
(27) Kleemann, G. R.; Beierle, J.; Nichols, A. C.; Dillon, T. M.; Pipes, G. D.; Bondarenko, P. V. *Anal. Chem.* **2008**, *80* (6), 2001–2009.
(28) Fornelli, L.; Ayoub, D.; Aizikov, K.; Beck, A.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (6), 3005–3012.
(29) Cotham, V. C.; Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (7), 4004–4013.
(30) Srzentić, K.; Fornelli, L.; Laskay, Ü. A.; Monod, M.; Beck, A.; Ayoub, D.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (19), 9945–9953.
(31) Pang, Y.; Wang, W.-H.; Reid, G. E.; Hunt, D. F.; Bruening, M. L. *Anal. Chem.* **2015**, *87* (21), 10942–10949.
(32) Reilly, J. P. *Mass Spectrom Rev* **2009**, *28* (3), 425–447.
(33) Ly, T.; Julian, R. R. *Angew. Chem. Int. Ed.* **2009**, *48* (39), 7130–7137.
(34) Brodbelt, J. S. *Chem. Soc. Rev.* **2014**, *43* (8), 2757–2783.
(35) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.
(36) Moon, J. H.; Yoon, S. H.; Kim, M. S. *Rapid Commun. Mass Spectrom.* **2005**, *19* (22), 3248–3252.
(37) Girod, M.; Sanader, Z.; Vojkovic, M.; Antoine, R.; MacAleese, L.; Lemoine, J.; Bonacic-Koutecky, V.; Dugourd, P. *J. Am. Soc. Mass Spectrom.* **2014**, *26* (3), 432–443.

(38) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. *J. Proteome Res.* **2010**, *9* (8), 4205–4214.
(39) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(40) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(41) Shaw, J. B.; Robinson, E. W.; Paša-Tolić, L. *Anal. Chem.* **2016**, *88* (6), 3019–3023.
(42) Lee, J.; Boutz, D. R.; Joyce, M. G.; Vollmers, C.; Chromikova, V.; Leung, K.; Horton, A. P.; DeKosky, B. J.; Lee, C. H.; Lavinder, J. J.; Georgiou, G. *Nat. Med. (publication pending)*.
(43) Vasicek, L. A.; Ledvina, A. R.; Shaw, J.; Griep-Raming, J.; Westphall, M. S.; Coon, J. J.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 1105–1108.
(44) Cannon, J. R.; Edwards, N. J.; Fenselau, C. *J. Mass Spectrom.* **2013**, *48* (3), 340–343.
(45) Safonova, Y.; Lapidus, A.; Lill, J. *Bioinformatics* **2015**, *31* (19), 3213–3215.

CHAPTER 5

(1) Aggarwal, S. *Nat. Biotechnol.* **2014**, *32* (1), 32–39.
(2) Ecker, D. M.; Jones, S. D.; Levine, H. L. *mAbs* **2015**, *7* (1), 9–14.
(3) Nelson, A. L.; Dhimolea, E.; Reichert, J. M. *Nat. Rev. Drug Discov.* **2010**, *9* (10), 767–774.
(4) Weiner, L. M.; Surana, R.; Wang, S. *Nat. Rev. Immunol.* **2010**, *10* (5), 317–327.
(5) Scott, A. M.; Wolchok, J. D.; Old, L. J. *Nat. Rev. Cancer* **2012**, *12* (4), 278–287.
(6) Chan, A. C.; Carter, P. J. *Nat. Rev. Immunol.* **2010**, *10* (5), 301–316.
(7) Beck, A.; Haeuw, J.-F.; Wurch, T.; Goetsch, L.; Bailly, C.; Corvaïa, N. *Discov. Med.* **2010**, *10* (53), 329–339.
(8) Hughes, B. *Nat. Rev. Drug Discov.* **2010**, *9* (9), 665–667.
(9) Wang, W.; Singh, S.; Zeng, D. L.; King, K.; Nema, S. *J. Pharm. Sci.* **2007**, *96* (1), 1–26.
(10) De Groot, A. S.; Scott, D. W. *Trends Immunol.* **2007**, *28* (11), 482–490.
(11) Harding, F. A.; Stickler, M. M.; Razo, J.; DuBridge, R. B. *mAbs* **2010**, *2* (3), 256–265.
(12) Walsh, G.; Jefferis, R. *Nat. Biotechnol.* **2006**, *24* (10), 1241–1252.
(13) Kozlowski, S.; Swann, P. *Adv. Drug Deliv. Rev.* **2006**, *58* (5–6), 707–722.
(14) Janeway, C. A.; Travers, P.; Walport, M.; Shlomchik, M. J. *Immunobiology*, 5th ed.; Garland Science, 2001.

(15) Gabrielli, E.; Pericolini, E.; Cenci, E.; Ortelli, F.; Magliani, W.; Ciociola, T.; Bistoni, F.; Conti, S.; Vecchiarelli, A.; Polonelli, L. *PLoS ONE* **2009**, *4* (12), e8187.
(16) Borrok, M. J.; Jung, S. T.; Kang, T. H.; Monzingo, A. F.; Georgiou, G. *ACS Chem. Biol.* **2012**, *7* (9), 1596–1602.
(17) Zhang, Z.; Pan, H.; Chen, X. *Mass Spectrom. Rev.* **2009**, *28* (1), 147–176.
(18) Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianférani, S. *Anal. Chem.* **2013**, *85* (2), 715–736.
(19) Lundell, N.; Schreitmüller, T. *Anal. Biochem.* **1999**, *266* (1), 31–47.
(20) Srebalus Barnes, C. A.; Lim, A. *Mass Spectrom. Rev.* **2007**, *26* (3), 370–388.
(21) Kleemann, G. R.; Beierle, J.; Nichols, A. C.; Dillon, T. M.; Pipes, G. D.; Bondarenko, P. V. *Anal. Chem.* **2008**, *80* (6), 2001–2009.
(22) Fornelli, L.; Ayoub, D.; Aizikov, K.; Beck, A.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (6), 3005–3012.
(23) Wang, D.; Wynne, C.; Gu, F.; Becker, C.; Zhao, J.; Mueller, H.-M.; Li, H.; Shameem, M.; Liu, Y.-H. *Anal. Chem.* **2015**, *87* (2), 914–921.
(24) Srzentić, K.; Fornelli, L.; Laskay, Ü. A.; Monod, M.; Beck, A.; Ayoub, D.; Tsybin, Y. O. *Anal. Chem.* **2014**, *86* (19), 9945–9953.
(25) Bondarenko, P. V.; Second, T. P.; Zabrouskov, V.; Makarov, A. A.; Zhang, Z. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (8), 1415–1424.
(26) Tsybin, Y. O.; Fornelli, L.; Stoermer, C.; Luebeck, M.; Parra, J.; Nallet, S.; Wurm, F. M.; Hartmer, R. *Anal. Chem.* **2011**, *83* (23), 8919–8927.
(27) Fornelli, L.; Damoc, E.; Thomas, P. M.; Kelleher, N. L.; Aizikov, K.; Denisov, E.; Makarov, A.; Tsybin, Y. O. *Mol. Cell. Proteomics* **2012**, *11* (12), 1758–1767.
(28) Mao, Y.; Valeja, S. G.; Rouse, J. C.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2013**, *85* (9), 4239–4246.
(29) Zhang, J.; Liu, H.; Katta, V. *J. Mass Spectrom.* **2010**, *45* (1), 112–120.
(30) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121* (4), 806–812.
(31) Zhou, H.; Ning, Z.; E. Starr, A.; Abu-Farha, M.; Figeys, D. *Anal. Chem.* **2012**, *84* (2), 720–734.
(32) Zhang, Z.; Shah, B. *Anal. Chem.* **2007**, *79* (15), 5723–5729.
(33) Cannon, J.; Lohnes, K.; Wynne, C.; Wang, Y.; Edwards, N.; Fenselau, C. *J. Proteome Res.* **2010**, *9* (8), 3886–3890.
(34) 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. *Nat Meth* **2012**, *9* (8), 822–824.
(35) Pang, Y.; Wang, W.-H.; Reid, G. E.; Hunt, D. F.; Bruening, M. L. *Anal. Chem.* **2015**, *87* (21), 10942–10949.
(36) Nicolardi, S.; Deelder, A. M.; Palmblad, M.; van der Burgt, Y. E. M. *Anal. Chem.* **2014**, *86* (11), 5376–5382.

(37) Yan, B.; Valliere-Douglass, J.; Brady, L.; Steen, S.; Han, M.; Pace, D.; Elliott, S.; Yates, Z.; Han, Y.; Balland, A.; Wang, W.; Pettit, D. *J. Chromatogr. A* **2007**, *1164* (1–2), 153–161.
(38) An, Y.; Zhang, Y.; Mueller, H.-M.; Shameem, M.; Chen, X. *mAbs* **2014**, *6* (4), 879–893.
(39) Tran, J. C.; Zamdborg, L.; Ahlf, D. R.; Lee, J. E.; Catherman, A. D.; Durbin, K. R.; Tipton, J. D.; Vellaichamy, A.; Kellie, J. F.; Li, M.; Wu, C.; Sweet, S. M. M.; Early, B. P.; Siuti, N.; LeDuc, R. D.; Compton, P. D.; Thomas, P. M.; Kelleher, N. L. *Nature* **2011**, *480* (7376), 254–258.
(40) Garcia, B. A. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (2), 193–202.
(41) Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L. *Anal. Chem.* **2011**, *83* (17), 6868–6874.
(42) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *J. Am. Chem. Soc.* **2013**, *135* (34), 12646–12651.
(43) Cannon, J. R.; Holden, D. D.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (21), 10970–10977.
(44) Cannon, J. R.; Kluwe, C.; Ellington, A.; Brodbelt, J. S. *PROTEOMICS* **2014**, *14* (10), 1165–1173.
(45) Cannon, J. R.; Cammarata, M. B.; Robotham, S. A.; Cotham, V. C.; Shaw, J. B.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. *Anal. Chem.* **2014**, *86* (4), 2185–2192.
(46) Vasicek, L. A.; Ledvina, A. R.; Shaw, J.; Griep-Raming, J.; Westphall, M. S.; Coon, J. J.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 1105–1108.
(47) Bremer, E. T. van den; Beurskens, F. J.; Voorhorst, M.; Engelberts, P. J.; Jong, R. N. de; Boom, B. G. van der; Cook, E. M.; Lindorfer, M. A.; Taylor, R. P.; Berkel, P. H. van; Parren, P. W. *mAbs* **2015**, *7* (4), 672–680.
(48) Michalski, A.; Damoc, E.; Lange, O.; Denisov, E.; Nolting, D.; Müller, M.; Viner, R.; Schwartz, J.; Remes, P.; Belford, M.; Dunyach, J.-J.; Cox, J.; Horning, S.; Mann, M.; Makarov, A. *Mol. Cell. Proteomics* **2012**, *11* (3), O111.013698.
(49) Holden, D. D.; McGee, W. M.; Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 1008–1016.
(50) Roy, J. *Trends Glycosci. Glycotechnol.* **2009**, *21* (118), 105–117.
(51) Solá, D. R. J.; Griebenow, K. *BioDrugs* **2012**, *24* (1), 9–21.
(52) Zhang, L.; Reilly, J. P. *J. Proteome Res.* **2009**, *8* (2), 734–742.
(53) Ko, B. J.; Brodbelt, J. S. *Int. J. Mass Spectrom.* **2015**, *377*, 385–392.
(54) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.
(55) Breuker, K.; Oh, H.; Horn, D. M.; Cerda, B. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124* (22), 6407–6420.
(56) Riley, N. M.; Westphall, M. S.; Coon, J. J. *Anal. Chem.* **2015**, *87* (14), 7109–7116.
(57) Frese, C. K.; Altelaar, A. F. M.; van den Toorn, H.; Nolting, D.; Griep-Raming, J.; Heck, A. J. R.; Mohammed, S. *Anal. Chem.* **2012**, *84* (22), 9668–9673.

(58) Brunner, A. M.; Lössl, P.; Liu, F.; Huguet, R.; Mullen, C.; Yamashita, M.; Zabrouskov, V.; Makarov, A.; Altelaar, A. F. M.; Heck, A. J. R. *Anal. Chem.* **2015**, *87* (8), 4152–4158.

CHAPTER 6

(1) Hermanson, G. T. *Bioconjugate Techniques*; Academic Press, 2013.
(2) Stephanopoulos, N.; Francis, M. B. *Nat. Chem. Biol.* **2011**, *7* (12), 876–884.
(3) Mirzaei, H.; Regnier, F. *Anal. Chem.* **2006**, *78* (12), 4175–4183.
(4) Pashkova, A.; Moskovets, E.; Karger, B. L. *Anal. Chem.* **2004**, *76* (15), 4550–4557.
(5) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17* (10), 994–999.
(6) Wiese, S.; Reidegeld, K. A.; Meyer, H. E.; Warscheid, B. *PROTEOMICS* **2007**, *7* (3), 340–350.
(7) Ulbrich, A.; Merrill, A. E.; Hebert, A. S.; Westphall, M. S.; Keller, M. P.; Attie, A. D.; Coon, J. J. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (1), 6–9.
(8) Keough, T.; Youngquist, R. S.; Lacey, M. P. *Proc. Natl. Acad. Sci.* **1999**, *96* (13), 7131–7136.
(9) Madsen, J. A.; Brodbelt, J. S. *Anal. Chem.* **2009**, *81* (9), 3645–3653.
(10) Robinson, M. R.; Madsen, J. A.; Brodbelt, J. S. *Anal. Chem.* **2012**, *84* (5), 2433–2439.
(11) Cotham, V. C.; Wine, Y.; Brodbelt, J. S. *Anal. Chem.* **2013**, *85* (11), 5577–5585.
(12) Diedrich, J. K.; Julian, R. R. *Anal Chem* **2010**, *82* (10), 4006–4014.
(13) Vasicek, L.; O'Brien, J. P.; Browning, K. S.; Tao, Z.; Liu, H.-W.; Brodbelt, J. S. *Mol. Cell. Proteomics* **2012**, *11* (7), 1–10.
(14) Espy, R. D.; Wleklinski, M.; Yan, X.; Cooks, R. G. *TrAC Trends Anal. Chem.* **2014**, *57*, 135–146.
(15) Prentice, B. M.; McLuckey, S. A. *Chem. Commun.* **2013**, *49* (10), 947–965.
(16) Mortensen, D. N.; Williams, E. R. *Anal. Chem.* **2014**, *86* (18), 9315–9321.
(17) Mortensen, D. N.; Williams, E. R. *Anal. Chem.* **2014**.
(18) Fisher, C. M.; Kharlamova, A.; McLuckey, S. A. *Anal. Chem.* **2014**, *86* (9), 4581–4588.
(19) Miladinović, S. M.; Fornelli, L.; Lu, Y.; Piech, K. M.; Girault, H. H.; Tsybin, Y. O. *Anal. Chem.* **2012**, *84* (11), 4647–4651.
(20) Kharlamova, A.; McLuckey, S. A. *Anal. Chem.* **2011**, *83* (1), 431–439.
(21) Grimm, R. L.; Hodyss, R.; Beauchamp, J. L. *Anal. Chem.* **2006**, *78* (11), 3800–3806.
(22) Kim, H. I.; Kim, H.; Shin, Y. S.; Beegle, L. W.; Jang, S. S.; Neidholdt, E. L.; Goddard, W. A.; Heath, J. R.; Kanik, I.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2010**, *132* (7), 2254–2263.

(23) Müller, T.; Badu-Tawiah, A.; Cooks, R. G. *Angew. Chem. Int. Ed.* **2012**, *51* (47), 11832–11835.
(24) Girod, M.; Moyano, E.; Campbell, D. I.; Cooks, R. G. *Chem. Sci.* **2011**, *2* (3), 501–510.
(25) Badu-Tawiah, A. K.; Li, A.; Jjunju, F. P. M.; Cooks, R. G. *Angew. Chem. Int. Ed.* **2012**, *51* (37), 9417–9421.
(26) Chen, H.; Wortmann, A.; Zhang, W.; Zenobi, R. *Angew. Chem. Int. Ed.* **2007**, *46* (4), 580–583.
(27) Peng, I. X.; Ogorzalek Loo, R. R.; Shiea, J.; Loo, J. A. *Anal. Chem.* **2008**, *80* (18), 6995–7003.
(28) Xia, Y.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **2011**, *19* (2), 173–189.
(29) Han, H.; McLuckey, S. A. *J. Am. Chem. Soc.* **2009**, *131* (36), 12884–12885.
(30) Stutzman, J. R.; Luongo, C. A.; McLuckey, S. A. *J. Mass Spectrom.* **2012**, *47* (6), 669–675.
(31) Mentinova, M.; McLuckey, S. A. *J. Am. Chem. Soc.* **2010**, *132* (51), 18248–18257.
(32) Prentice, B. M.; Gilbert, J. D.; Stutzman, J. R.; Forrest, W. P.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (1), 30–37.
(33) McGee, W. M.; McLuckey, S. A. *Proc. Natl. Acad. Sci.* **2014**, *111* (4), 1288–1292.
(34) Hassell, K. M.; Stutzman, J. R.; McLuckey, S. A. *Anal. Chem.* **2010**, *82* (5), 1594–1597.
(35) Stutzman, J. R.; McLuckey, S. A. *Anal. Chem.* **2012**, *84* (24), 10679–10685.
(36) Vasicek, L.; Brodbelt, J. S. *Anal. Chem.* **2010**, *82* (22), 9441–9446.
(37) Gardner, M. W.; Smith, S. I.; Ledvina, A. R.; Madsen, J. A.; Coon, J. J.; Schwartz, J. C.; Stafford, G. C.; Brodbelt, J. S. *Anal. Chem.* **2009**, *81* (19), 8109–8118.
(38) Villén, J.; Beausoleil, S. A.; Gygi, S. P. *PROTEOMICS* **2008**, *8* (21), 4444–4452.
(39) Stutzman, J. R.; Hassell, K. M.; McLuckey, S. A. *Int. J. Mass Spectrom.* **2012**, *312*, 195–200.
(40) Ogorzalek Loo, R. R.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1992**, *3* (7), 695–705.
(41) Chen, H.; Venter, A.; Cooks, R. G. *Chem. Commun.* **2006**, No. 19, 2042–2044.
(42) Law, W. S.; Wang, R.; Hu, B.; Berchtold, C.; Meier, L.; Chen, H.; Zenobi, R. *Anal. Chem.* **2010**, *82* (11), 4494–4500.
(43) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. *J. Proteome Res.* **2010**, *9* (8), 4205–4214.

CHAPTER 7

(1) Cohen, P. *Nat. Cell Biol.* **2002**, *4* (5), E127–E130.

(2) Ubersax, J. A.; Ferrell Jr, J. E. *Nat. Rev. Mol. Cell Biol.* **2007**, *8* (7), 530–541.
(3) Johnson, L. N. *Biochem. Soc. Trans.* **2009**, *37* (4), 627–641.
(4) Hunter, T. *Cell* **1995**, *80* (2), 225–236.
(5) Lahiry, P.; Torkamani, A.; Schork, N. J.; Hegele, R. A. *Nat. Rev. Genet.* **2010**, *11* (1), 60–74.
(6) Hendriks, W. J. A. J.; Pulido, R. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **2013**, *1832* (10), 1673–1696.
(7) Humphrey, S. J.; James, D. E.; Mann, M. *Trends Endocrinol. Metab.* **2015**, *26* (12), 676–687.
(8) Mann, M.; Ong, S.-E.; Grønborg, M.; Steen, H.; Jensen, O. N.; Pandey, A. *Trends Biotechnol.* **2002**, *20* (6), 261–268.
(9) Riley, N. M.; Coon, J. J. *Anal. Chem.* **2016**, *88* (1), 74–94.
(10) Michalski, A.; Cox, J.; Mann, M. *J. Proteome Res.* **2011**, *10* (4), 1785–1793.
(11) Bauer, M.; Ahrné, E.; Baron, A. P.; Glatter, T.; Fava, L. L.; Santamaria, A.; Nigg, E. A.; Schmidt, A. *J. Proteome Res.* **2014**, *13* (12), 5973–5988.
(12) Palumbo, A. M.; Smith, S. A.; Kalcic, C. L.; Dantus, M.; Stemmer, P. M.; Reid, G. E. *Mass Spectrom. Rev.* **2011**, *30* (4), 600–625.
(13) Brodbelt, J. S. *Anal. Chem.* **2016**, *88* (1), 30–51.
(14) Boersema, P. J.; Mohammed, S.; Heck, A. J. R. *J. Mass Spectrom.* **2009**, *44* (6), 861–878.
(15) Jones, A. W.; Cooper, H. J. *Analyst* **2011**, *136* (17), 3419–3429.
(16) Palumbo, A. M.; Tepe, J. J.; Reid, G. E. *J. Proteome Res.* **2008**, *7* (2), 771–779.
(17) Chi, A.; Huttenhower, C.; Geer, L. Y.; Coon, J. J.; Syka, J. E. P.; Bai, D. L.; Shabanowitz, J.; Burke, D. J.; Troyanskaya, O. G.; Hunt, D. F. *Proc. Natl. Acad. Sci.* **2007**, *104* (7), 2193–2198.
(18) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14* (19), 1793–1800.
(19) Nagaraj, N.; D'Souza, R. C. J.; Cox, J.; Olsen, J. V.; Mann, M. *J. Proteome Res.* **2010**, *9* (12), 6786–6794.
(20) Kim, T.-Y.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (12), 2334–2341.
(21) Shin, Y. S.; Moon, J. H.; Kim, M. S. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (1), 53–59.
(22) Madsen, J. A.; Kaoud, T. S.; Dalby, K. N.; Brodbelt, J. S. *PROTEOMICS* **2011**, *11* (7), 1329–1334.
(23) Fort, K. L.; Dyachenko, A.; Potel, C. M.; Corradini, E.; Marino, F.; Barendregt, A.; Makarov, A. A.; Scheltema, R. A.; Heck, A. J. R. *Anal. Chem.* **2016**, *88* (4), 2303–2310.
(24) Swaney, D. L.; McAlister, G. C.; Wirtala, M.; Schwartz, J. C.; Syka, J. E. P.; Coon, J. J. *Anal. Chem.* **2007**, *79* (2), 477–485.
(25) Frese, C. K.; Zhou, H.; Taus, T.; Altelaar, A. F. M.; Mechtler, K.; Heck, A. J. R.; Mohammed, S. *J. Proteome Res.* **2013**, *12* (3), 1520–1525.
(26) Shaffer, C. J.; Slováková, K.; Tureček, F. *Int. J. Mass Spectrom.* **2015**, *390*, 71–80.

(27) Meyer, H. E.; Hoffmann-Posorske, E.; Korte, H.; Heilmeyer, L. M. G. *FEBS Lett.* **1986**, *204* (1), 61–66.
(28) Oda, Y.; Nagasu, T.; Chait, B. T. *Nat. Biotechnol.* **2001**, *19* (4), 379–382.
(29) Molloy, M. P.; Andrews, P. C. *Anal. Chem.* **2001**, *73* (22), 5387–5394.
(30) Thompson, A. J.; Hart, S. R.; Franz, C.; Barnouin, K.; Ridley, A.; Cramer, R. *Anal. Chem.* **2003**, *75* (13), 3232–3243.
(31) Knight, Z. A.; Schilling, B.; Row, R. H.; Kenski, D. M.; Gibson, B. W.; Shokat, K. M. *Nat. Biotechnol.* **2003**, *21* (9), 1047–1054.
(32) Diedrich, J. K.; Julian, R. R. *Anal. Chem.* **2011**, *83* (17), 6818–6826.
(33) Diedrich, J. K.; Julian, R. R. *J. Am. Chem. Soc.* **2008**, *130* (37), 12212–12213.
(34) McLachlin, D. T.; Chait, B. T. *Anal. Chem.* **2003**, *75* (24), 6826–6836.
(35) Gronert, S.; Huang, R.; Li, K. H. *Int. J. Mass Spectrom.* **2004**, *231* (2–3), 179–187.
(36) Gronert, S.; Li, K. H.; Horiuchi, M. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (12), 1905–1914.
(37) Leitner, A.; Foettinger, A.; Lindner, W. *J. Mass Spectrom.* **2007**, *42* (7), 950–959.
(38) Lanucara, F.; Lee, D. C. H.; Eyers, C. E. *J. Am. Soc. Mass Spectrom.* **2013**, *25* (2), 214–225.
(39) Cotham, V. C.; Shaw, J. B.; Brodbelt, J. S. *Anal. Chem.* **2015**, *87* (18), 9396–9402.
(40) Han, H.; McLuckey, S. A. *J. Am. Chem. Soc.* **2009**, *131* (36), 12884–12885.
(41) Hassell, K. M.; Stutzman, J. R.; McLuckey, S. A. *Anal. Chem.* **2010**, *82* (5), 1594–1597.
(42) Stutzman, J. R.; McLuckey, S. A. *Anal. Chem.* **2012**, *84* (24), 10679–10685.
(43) Stutzman, J. R.; Luongo, C. A.; McLuckey, S. A. *J. Mass Spectrom.* **2012**, *47* (6), 669–675.
(44) Schug, K. A.; Lindner, W. *Chem. Rev.* **2005**, *105* (1), 67–114.
(45) Woods, A. S.; Wang, H.-Y. J.; Jackson, S. N. *J. Proteome Res.* **2007**, *6* (3), 1176–1182.
(46) Tsaprailis, G.; Somogyi, A.; Nikolaev, E. N.; Wysocki, V. H. *Int. J. Mass Spectrom.* **2000**, *195-196*, 467–479.
(47) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35* (12), 1399–1406.
(48) Bleiholder, C.; Suhai, S.; Harrison, A. G.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 1032–1039.

VITA

Victoria Christine Cotham was born in La Quinta, California one not-so-quiet Super Bowl Sunday. After completing an International Baccalaureate education at La Quinta High School in 2006, she pursued higher-level education at the Pennsylvania State University, University Park, in the fall of 2006. As a first generation college graduate, she received her Bachelor of Science degree in forensic chemistry at the top of her class in the fall of 2010. While pursuing her undergraduate degree, she completed four years of research experience under the guidance of Dr. Dan Sykes, during which she designed, constructed and characterized small scale instruments for integration into undergraduate analytical chemistry laboratory curriculum. This experience, in addition to receiving a competitive NSF REU summer fellowship where she conducted research to improve the development of phage-based biosensors in the laboratory of Dr. Reginald Penner at the University of California, Irvine, propelled her into graduate studies at the University of Texas at Austin. While working under the supervision of Jennifer S. Brodbelt, she received a prestigious NSF Graduate Research Fellowship that has funded her graduate work towards the development of new high-throughput mass spectrometric-based strategies to improve the characterization of antibodies and proteomic samples.

Permanent email: v.cotham@utexas.edu

This dissertation was typed by the author.