Methods of Protein Sample Preparation Using Digital Microfluidics and Gas-Phase Techniques with Detection by Mass Spectrometry

by

Brendon Seale

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy – Interdisciplinary Chemistry

Department of Chemistry
University of Toronto

© Copyright by Brendon Seale (2018)
Methods of Protein Sample Clean-Up and Separation Using Digital Microfluidics and Gas-Phase Techniques with Detection by Mass Spectrometry

Brendon Seale

Doctor of Philosophy – Interdisciplinary Chemistry

Department of Chemistry
University of Toronto

2018

Abstract

Biological samples are complicated, having many different components at a wide variety of concentrations. Some means to simplify these samples while leaving the discrete components unaltered is required. Small structural differences can render a protein unusable role. Mass spectrometry is a valuable tool to characterize proteins but the number of components in these samples present a challenge for accurate detection. Sample preparation and separation methods are effective tools to alleviate this challenge.

In this dissertation, I present methods for simplifying protein samples for analyze with mass spectrometry using liquid and gas phase methods. In the first chapter, I review some sample preparation options. In Chapter 2, I present a microscale method for high abundance protein depletion using digital microfluidics and functionalized magnetic particles. This is a fast, effective, and multiplexed method for reducing the interference caused from proteins like serum albumin and immunoglobulins resulting in a 4-fold improvement in signal to noise for a low abundant protein. Chapter 3 extends this methodology by employing the same fluid and particle handling technique to immunoprecipitation. This chapter introduces the use of a pre-
concentration method called pre-concentration using liquid intake by paper (P-CLIP), a promising ‘world-to-chip’ interface for many digital microfluidic applications. In Chapter 4, gas-phase vapours are used to simplify analysis of peptides by electrospray mass spectrometry. Acetonitrile vapour suppresses in-source fragmentation of fragile ions by clustering with fragile ions resulting in improved limits of detection. Chapter 5 presents a union of liquid chromatography and differential mobility spectrometry for characterization of biopharmaceutical proteins. This new combination, called DMS-SWATH, increases MS/MS sequence coverage for proteins and alleviates problems found in traditional analysis by mass spectrometry and differential mobility spectrometry. Overall this work presents important steps forward in miniaturizing and ameliorating the use of mass spectrometry for the characterization of proteins.
Acknowledgments

First, I must acknowledge the work of my supervisor, Professor Aaron Wheeler. Dr. Wheeler has shown a tremendous amount of patience and dedication during my 5 years as a student under his guidance. It is doubtful that I would have been able to be here today without his support.

I would like to thank my supervisory committee members, Professors Ulrich Krull and Rebecca Jockusch. They have been incredibly supportive of my progress and are nothing but helpful. I would especially like to thank Dr. Krull for instilling a love of analytical sciences in me from the first day of his second-year analytical chemistry course at UTM. I would also like to thank Professor Kagan Kerman for serving on my committee for my comprehensive exam and Prof. Hui Peng for serving on the committee for the defense of this thesis. Finally, a very special thanks to Professor Alan Doucette of Dalhousie University for serving as external examiner.

I have been very fortunate to have collaborated with many different scientists and groups throughout my stay here. I have had great support from Dr. Richard Oleschuk and his former student Ningi Mei from Queen’s University for getting off the ground on my journey into protein mass spectrometry. I have been fortunate to be a part of an NSERC CREATE program called Mass Spectrometry Enabled Science and Engineering (MS-ESE) run by Dr. Derek Wilson of York University. The MS-ESE program allowed me to be part of an industrial internship at SCIEX.

The team at SCIEX has been incredibly kind and supportive and has led me to focus my efforts completely on the exploration of mass spectrometry. My supervisor at SCIEX, Dr. Yves LeBlanc has been a terrific help in my work and I can honestly say I could not have done what I have in this short amount of time without him. His knowledge of mass spec is only surpassed by his infectious optimism and good nature. I owe a great deal of gratitude to Dr. Chang Liu for orchestrating my internship at SCIEX and to Dr. Larry Campbell for allowing me to help with his projects as well. Working alongside these individuals has been inspiring and I hope to continue to collaborate with them in the future.

The Wheeler group has been good to me and I have met a great number of people who I hope will stay with me as friends and colleagues throughout life’s journey. I must first acknowledge
the hardest working man in science, Dr. Alphonsus Ng who taught me everything I ever needed to know about microfabrication. His enthusiasm was unparalleled and working alongside of him was a great honor. Then, the two people individuals who I owe perhaps the most for showing me how awesome mass spectrometry is: Dr. Andrea Kirby and Dr. Nelson Lafrenière. Without their support and mentorship this thesis would not include much mass spectrometry at all. You guys are the best!

I have had the pleasure of supervising several undergraduate students in my time but none have been as outstanding as Charis Lam, who has gone on to join the group as a PhD student. Other important individuals in the group to mention specifically include Dr. Darius Rackus for his ingenious method of magnetic particle pre-concentration using digital microfluidics and Dr. Dean Chamberlain for his immensely deep knowledge of all things systems biology. I must also thank Cassandra Lord and Lisa Ngo for their work as administrative assistance for our diverse group. I’m sure dealing with all of us is no picnic but you guys have done a remarkable job keeping us all in check and well stocked with supplies.

I cannot forget the other members of the Wheeler group that I have had the pleasure of working with. From the present: Dr. Christian Fobel, Dr. Ryan Fobel, Dr. Sara Abdulwahab, Calvin Chan, Christopher Dixon, Michael Dryden, Man Ho, Julian Lamanna, Betty Li, Nooman Mufti, Haozhong Situ, Alexandros Sklavounos, Ian Swyer, Jeremy Wong Alex Yu. And from the past: Dr. Sam Au, Dr. Dario Bogojevic, Dr. KiHwan Choi, Dr. Irwin Eydelnant, Dr. Lorenzo Gutierrez, Dr. Jihye Kim, Dr. Jared Mudrik, Dr. Mahesh Sarvothaman, Dr. Mohtashim Shamsi, Dr. Steve Shih, and Dr. Edward Sykes. I am grateful for your comradery and friendship and I hope it continues into the future.

For the support, tips, and good conversation (and for putting up with our loud, messy, and sometimes broken mass spectrometers) I thank Dr. Matt Forbes at the Advanced Instrumentation for Molecular Structure (AIMS) Mass Spectrometry Laboratory at U of T. I would also like to thank the support of the Toronto Nanofabrication Centre at U of T led by Dr. Henry Lee and Yimin Zhou.

None of this could have been possible without the emotional (and sometimes financial) support of my loving family. My mother, Linda, my father, Sheldon and my big brother and sister, David
and Jennifer. Without you none of this would happen. I’m sorry it took so long but I promise I’m not going back to school unless it’s to work in one!

Finally, I must thank the love of my life, Emily Cooper. You have seen me through a great many difficult times and your support means everything to me. Without you I would have given up, I’m sure of it. But with you at my side I am positive that this dissertation is only the start of the many amazing opportunities that await me. I hope to have you with me when I pursue whatever comes after. I love you and always will.
Table of Contents

Acknowledgments .................................................................................................................. iv
Table of Contents .................................................................................................................. vii
Overview of Chapters ........................................................................................................... x
Overview of Author Contributions ....................................................................................... xiii
List of Tables .......................................................................................................................... xv
List of Figures ........................................................................................................................ xvi
List of Abbreviations ............................................................................................................. xviii

Chapter 1 Solution- and Gas-Phase Separation and Clean-Up Methods for Proteinaceous Biological Samples ............................................................................. 1
  1.1 Introduction ..................................................................................................................... 1
  1.2 Digital Microfluidics ...................................................................................................... 3
      1.2.1 Principles of Digital Microfluidics ........................................................................ 3
      1.2.2 Digital Microfluidics and sample preparation ....................................................... 7
  1.3 Protein Sample Preparation Methods ............................................................................. 9
      1.3.1 Extractions .............................................................................................................. 9
      1.3.2 Gel Electrophoresis ............................................................................................... 15
      1.3.3 Chromatography .................................................................................................. 16
  1.4 Mass Spectrometry and Ion Mobility ............................................................................. 21
      1.4.1 Mass Spectrometry and Tandem Mass Spectrometry ............................................ 21
      1.4.2 Ion Mobility .......................................................................................................... 32
  1.5 Perspectives ................................................................................................................... 37

Chapter 2 Digital Microfluidic Platform for Human Plasma Protein Depletion .................... 39
  2.1 Abstract ......................................................................................................................... 39
  2.2 Introduction .................................................................................................................... 39
Overview of Chapters

This thesis describes the development of techniques for biological sample clean-up prior to protein analysis by mass spectrometry (MS). There are two main topics of discussion: liquid-phase sample handling by digital microfluidics (DMF) and gas-phase sample processing by the
mass spectrometer. The DMF work pairs the microscale fluid handling technique with immunoaffinity extraction methods, which was performed during my time in the Wheeler group at the University of Toronto. The second topic examines the ability of techniques such as differential mobility spectrometry (DMS) to simplify protein samples in the gas-phase, as they enter the mass spectrometer. This work was completed at SCIEX in Concord, Ontario as part of an industrial internship Collaborative Research and Training (CREATE) program from the National Science and Research Council (NSERC) called Mass Spectrometry Enabled Science and Engineering (MS-ESE). A summary of each chapter is listed below.

**Chapter 1** is a review of relevant literature concerning processing and clean-up of biological samples for protein analysis by MS. Emphasis is placed on the techniques used in the dissertation including solid phase extraction, immunoaffinity extraction, chromatography, and ion mobility mass spectrometry.

**Chapter 2** describes the development of a magnetic bead based immunoaffinity depletion method for the removal of highly abundant serum proteins: human serum albumin (HSA) and immunoglobulin G. The work used DMF to implement an automated, microscale method for protein depletion from samples. This method greatly improves the signal-to-noise ratio of a protein in low abundance after several cycles of depletion.

**Chapter 3** describes methods which are an extension of those presented in Chapter 2, in which a DMF method relying on antibodies bound to magnetic beads was developed to purify a protein target by immunoprecipitation. The addition of acid to the beads frees the bound target from the antibody allowing analysis of the protein without interference from the matrix.

**Chapter 4** describes means of using the hardware for a DMS system to introduce a chemical modifier to gas-phase peptide ions. These ions are prone to in-source and in-system fragmentation, a phenomenon that can be limited by exposure to polar organic vapours. Thus, the new technique has the potential to reduce the complexity of the resulting mass spectra as well as improving the signal-to-noise ratio for selected ions. The mechanism of action of this phenomenon was verified using DMS.
Chapter 5 introduces a method of analysis combining DMS separation with data independent sequential windowed acquisition of all theoretical fragment ion mass spectra (SWATH) MS/MS triggering to form a 2D separation system using HPLC in conjunction with DMS. DMS separation is orthogonal to reversed phase HPLC and allows segregation of co-eluting peptides by scanning a fixed range of compensation voltages. This facilitates the collection of MS/MS spectra with less noise, which results in improved sequence coverage for monoclonal antibody characterization.

Chapter 6 summarizes the original work described in the dissertation, and discusses the advantages and disadvantages of the methods introduced therein. This chapter concludes with a series of suggestions for future work.
Overview of Author Contributions

I have had the opportunity to work with many talented scientists who have contributed to my success. In this thesis I present four projects which are driven intellectually and experimentally by myself. However, many of my colleagues have aided my endeavours and I outline their contributions to my work below.

Chapter 2 describes a digital microfluidic-based method for the depletion of highly abundant serum proteins to improve the detection of low abundance proteins. This was a collaborative project between the Wheeler group at the University of Toronto and the research group of Dr. Richard Oleschuk at Queen’s University in Kingston, Ontario, Canada. I thank Ningsi Mei, David McLeod and Dr. Jiaxi Wang (Queen’s University) for assistance performing MALDI-MS and LC-MS-MS. I thank Dr. Alphonsus Ng (former graduate student and post-doctoral fellow of the Wheeler group for assistance in device fabrication. The text was compiled by myself with contributions to the discussion and experimental pertaining to MALDI-MS by Ningsi Mei and editing by Dr. Wheeler, Dr. Oleschuk, and Dr. Alphonsus Ng. This chapter was originally published in Analytical Chemistry. I share co-first-authorship on this paper with Ningsi Mei: Mei, N.; Seale, B.; Ng, A.H.C.; Wheeler, A.R.; Oleschuk, R. "Digital Microfluidic Platform for Human Plasma Protein Depletion" Analytical Chemistry, 2014, 86, 8466-8472.

Chapter 3 describes the development of a DMF-based method for immunoprecipitation for human serum albumin proteins. I thank Charis Lam (a current graduate student of the Wheeler Lab) for assistance in preliminary experiments which are not presented in this work. Digital microfluidic magnetic bead pre-concentration (P-CLIP) was first conceived by Dr. Darius Rackus (Post-doc in the Wheeler group) but was developed and applied for immunoprecipitation by myself before any further characterization. Access to a fluorescence spectrophotometer was provided by Dan Mathers (formerly of ANALEST facility at U of T). The text was written by myself with editing by Dr. Wheeler. This chapter was published in Analytical Chemistry: Seale, B.; Lam, C.; Rackus, D.G.; Chamberlain, M.D.; Liu, C.; Wheeler, A.R. "Digital Microfluidics for Immunoprecipitation" Analytical Chemistry, 2016, 88, 10223-10230. I thank Dr. Chang Liu of SCIEX for HPLC-MS and Dr. Dean Chamberlain (a research associate in the Wheeler group) for SDS-PAGE during the peer review of the original publication.
Chapter 4 Project directions and concepts were guided by conversations with Dr. Bradley Schneider and Dr. Yves LeBlanc (SCIEX). Technical assistance with instrumentation was provided by Dr. LeBlanc. This work was presented at the 65th Annual Conference on Mass Spectrometry and Allied Topics 2016 with myself as presenting author titled “Mitigating fragmentation of peptides with controlled clustering”. A manuscript prepared by me with editing by Dr. LeBlanc has been submitted for publication.

Chapter 5 describes a technique for the analysis of proteins using LC-MS in conjunction with DMS separation. Technical assistance for instrumentation and software was provided by Dr. Yves LeBlanc (SCIEX). This work is under review for a patent and future publication with the invention disclosed to SCIEX with Dr. LeBlanc and myself as co-inventors.
List of Tables

Table 2-1: Comparison of S/N Ratios for Ion Intensities in MALDI-MS Spectra for the Control and Following a Single and Double Depletion, with DMF/Magnetic Bead Platform .................. 54

Table 4-1: The peak area increases due to the presence of ACN curtain gas modifier on peptide ions spiked into a BSA digest. .................................................................................................................. 97

Table 4-2: Limit of detection improvements in fragile ions upon the addition of ACN to the curtain gas. ........................................................................................................................................ 98

Table 5-1: Summary of ANOVA with Tukey-Kramer Post-Hoc analysis for auto-validated sequence coverage across different data acquisition methods ........................................ 119
List of Figures

Figure 1-1: Different configurations of digital microfluidic devices........................................... 4

Figure 1-2: Schematic representation of the process of DMS ....................................................... 35

Figure 2-1. Device and processing scheme for protein depletion................................................. 46

Figure 2-2. Frames from Video 2-1 depicting the process of protein depletion............................ 48

Figure 2-3. On-chip depletion kinetics and efficacy...................................................................... 50

Figure 2-4. MALDI spectra of depletion sample......................................................................... 53

Figure 3-1. Immunoprecipitation by digital microfluidics (DMF-IP).............................................. 61

Figure 3-2. Effects of immunoprecipitation-elution buffer on protein conformation. ............... 69

Figure 3-3. Optimization of DMF-IP elution efficiency.............................................................. 71

Figure 3-4. SDS-PAGE gel of human serum albumin (HSA) sample........................................ 72

Figure 3-5. DMF-IP with analysis by mass spectrometry ........................................................... 74

Figure 3-6. Pre-concentration of magnetic particles using liquid intake by paper (P-CLIP) ....... 77

Figure 3-7. Representative tandem mass spectra of an HSA peptide with P-CLIP-DMF-IP ...... 79

Figure 3-8. Representative HPLC-MS/MS chromatogram of HSA after P-CLIP-DMF-IP....... 80

Figure 4-1: Infusion MS CV maps of KGAIL under different curtain gas modifiers .............. 88

Figure 4-2: Precursor spectra of KGAIL 3+ and 4+ under nitrogen, acetonitrile and acetone .... 90

Figure 4-3: Infusion MS DP ramp of XIC of KGAIL 4+ under pure nitrogen, nitrogen with acetonitrile and, nitrogen with acetone ................................................................. 92

Figure 4-4: Fragmentation of KGAIL under different current gas modifiers.......................... 94
Figure 4-5: Selected XIC chromatograms of spiked KGAIL 4+, Angiotensin I 4+ and Angiotensin III 3+ ions from BSA digest matrix under ACN and no modifier conditions........ 96

Figure 4-6: Calibration curves for Angiotensin I 4+ and Angiotensin III 3+ with ACN modifier and without.............................................................................................................................................................................................. 99

Figure 4-7: Protein sequencing comparisons with and without ACN modifier.......................... 101

Figure 5-1: Comparison of untargeted data acquisition methods available............................. 114

Figure 5-2: Chromatograms and MS/MS spectral output of a DMS-SWATH experiment ...... 116

Figure 5-3: Sequence coverage differences observed for different data acquisition methods applied to a model mAb digest ................................................................................................................................................................................. 118

Figure 5-4: MS/MS spectral splitting by DMS-SWATH in a BSA protein digest using SWATH and DMS-SWATH acquisition both with ACN chemical modifier ...................................................... 122

Figure 5-5: Breaking ACN clusters using SWATH and DMS-SWATH on direct infusion of KGAILKGAILR peptide ............................................................................................................................................................................................................. 125
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D-LC</td>
<td>Two-Dimensional Liquid Chromatography</td>
</tr>
<tr>
<td>ACN</td>
<td>Acetonitrile</td>
</tr>
<tr>
<td>AngI</td>
<td>Human Angiotensin I</td>
</tr>
<tr>
<td>AngIII</td>
<td>Human Angiotensin III</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>CE</td>
<td>Collision Energy</td>
</tr>
<tr>
<td>CHIP</td>
<td>Chromatin Immunoprecipitation</td>
</tr>
<tr>
<td>CID</td>
<td>Collision Induced Dissociation</td>
</tr>
<tr>
<td>CPCD</td>
<td>Coupled Physical Chemical Dynamics</td>
</tr>
<tr>
<td>CPS</td>
<td>Counts per second</td>
</tr>
<tr>
<td>CV</td>
<td>Compensation Voltage</td>
</tr>
<tr>
<td>DDA</td>
<td>Data dependent acquisition</td>
</tr>
<tr>
<td>DIA</td>
<td>Data independent acquisition</td>
</tr>
<tr>
<td>DMF</td>
<td>Digital Microfluidics</td>
</tr>
<tr>
<td>DMS</td>
<td>Differential Mobility Spectrometry</td>
</tr>
<tr>
<td>DP</td>
<td>Declustering Potential</td>
</tr>
<tr>
<td>DTIMS</td>
<td>Drift Time Ion Mobility Spectrometry</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>ECD</td>
<td>Electron Capture Dissociation</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray Ionization</td>
</tr>
<tr>
<td>FA</td>
<td>Formic Acid</td>
</tr>
<tr>
<td>FAIMS</td>
<td>Field-Asymmetric Ion Mobility Spectrometry</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal Bovine Serum</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein-isothiocyanate</td>
</tr>
<tr>
<td>GC</td>
<td>Gas Chromatography</td>
</tr>
<tr>
<td>HAP</td>
<td>Highly Abundant Protein</td>
</tr>
<tr>
<td>HILIC</td>
<td>Hydrophilic Interaction Liquid Chromatography</td>
</tr>
<tr>
<td>HPLC</td>
<td>High-Performance Liquid Chromatography</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>HSA</td>
<td>Human Serum Albumin</td>
</tr>
<tr>
<td>IDA</td>
<td>Information Dependent Acquisition</td>
</tr>
<tr>
<td>IgA</td>
<td>Immunoglobulin A</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin G</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilized Metal ion Affinity Chromatography</td>
</tr>
<tr>
<td>IMS</td>
<td>Ion Mobility Spectrometry</td>
</tr>
<tr>
<td>IP</td>
<td>Immunoprecipitation</td>
</tr>
<tr>
<td>I.S.</td>
<td>Internal Standard</td>
</tr>
<tr>
<td>ISF</td>
<td>In-source fragmentation</td>
</tr>
<tr>
<td>ITC</td>
<td>Ion Transfer Coefficient</td>
</tr>
<tr>
<td>ITO</td>
<td>Indium-Tin Oxide</td>
</tr>
<tr>
<td>KGAIL</td>
<td>KGAILKGAILR peptide</td>
</tr>
<tr>
<td>LAP</td>
<td>Low Abundance Protein</td>
</tr>
<tr>
<td>LC</td>
<td>Liquid Chromatography</td>
</tr>
<tr>
<td>mAb</td>
<td>Monoclonal Antibody</td>
</tr>
<tr>
<td>MALDI</td>
<td>Matrix-Assisted Laser Desorption Ionization</td>
</tr>
<tr>
<td>MeOH</td>
<td>Methanol</td>
</tr>
<tr>
<td>MMTS</td>
<td>Methyl Methanethiosulfoante</td>
</tr>
<tr>
<td>MS</td>
<td>Mass Spectrometry</td>
</tr>
<tr>
<td>MS/MS</td>
<td>Tandem Mass Spectrometry</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass-to-charge ratio</td>
</tr>
<tr>
<td>NSI</td>
<td>Nano-electrospray Ionization</td>
</tr>
<tr>
<td>oaToF</td>
<td>Orthogonal Acceleration Time-of-Flight</td>
</tr>
<tr>
<td>OGS</td>
<td>Octyl-β-D-glucopyranoside</td>
</tr>
<tr>
<td>PA</td>
<td>Polyacrylamide</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>P-CLIP</td>
<td>Pre-Concentration using Liquid Intake by Paper</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline</td>
</tr>
<tr>
<td>PID</td>
<td>Photo-Induced Dissociation</td>
</tr>
<tr>
<td>QqQ</td>
<td>Triple Quadrupole Mass Spectrometer</td>
</tr>
<tr>
<td>Q-ToF</td>
<td>Quadrupole Time-of-Flight</td>
</tr>
<tr>
<td>Acronym</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>RSD</td>
<td>Relative Standard Deviation</td>
</tr>
<tr>
<td>SA</td>
<td>Sinapinic Acid</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl Sulfate</td>
</tr>
<tr>
<td>SEC</td>
<td>Size-Exclusion Chromatography</td>
</tr>
<tr>
<td>SPE</td>
<td>Solid Phase Extraction</td>
</tr>
<tr>
<td>S/N</td>
<td>Signal-to-Noise ratio</td>
</tr>
<tr>
<td>SV</td>
<td>Separation Voltage</td>
</tr>
<tr>
<td>SWATH</td>
<td>Sequential Windowed Acquisition of All Theoretical fragment ion mass spectra</td>
</tr>
<tr>
<td>TCEP</td>
<td>Tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>TEAB</td>
<td>Tetraethyl Ammonium Bromide</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic Acid</td>
</tr>
<tr>
<td>TNFC</td>
<td>University of Toronto Nanofabrication Centre</td>
</tr>
<tr>
<td>ToF</td>
<td>Time-of-Flight</td>
</tr>
<tr>
<td>UPLC</td>
<td>Ultra-Performance Liquid Chromatography</td>
</tr>
<tr>
<td>XIC</td>
<td>Extraction Ion Chromatogram</td>
</tr>
</tbody>
</table>
Chapter 1
Solution- and Gas-Phase Separation and Clean-Up Methods for Proteinaceous Biological Samples

This dissertation describes the development of techniques and methods for preparing and processing biological samples containing proteins to render them suitable for analysis by mass spectrometry (MS). This work contains contributions to two main areas of protein sample clean-up: automation and miniaturization of sample handling and protein extraction in the solution phase by immunochemistry, and gas-phase separation and clean-up using differential mobility spectrometry. There is a long history of work in these areas, and each has several competing and companion techniques associated with them. This chapter outlines some of the methods available to the researcher for clean-up of proteinaceous samples before or in-line with MS analysis.

1.1 Introduction

Biologist and chemists have been studying proteins since the 18th century\(^1\) and their importance in the life sciences had been well established by the early 20th century. However, the structure and conformation of proteins has long been difficult to determine because of their size and the nature of the samples in which they are found. Since the advent of electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI)\(^2\), MS has become the gold standard for protein identification and sequence analysis, far surpassing traditional chemical techniques such as Edman degradation\(^3\). This revolution has allowed researchers to examine not just one protein in isolation but many, leading to the development of the field of proteomics, the study of all proteins present in a given sample.\(^4\) The size of a proteome is vast. For example, estimations of the human proteome range from a ~20,000 proteins total (calculated from a 1-to-1 ratio of coding genes to proteins) to ~1.5 million proteins in human serum alone when considering the various isoforms each protein can take.\(^5\) That means the number of isoforms of all expressible proteins in the human proteome can total in the billions.\(^6\) Proteomics can be performed in a ‘top-down’ manner in which proteins are introduced to a mass spectrometer intact where they are subsequently fragmented using techniques such as collision induced dissociation (CID) or electron transfer dissociation (ETD).\(^7\) Conversely, ‘bottom-up’ analysis involves chemically or enzymatically breaking proteins into peptides. These peptides are introduced to the mass
spectrometer for analysis and further fragmentation by the same methods in top-down analysis. Computational methods are then used to identify proteins from the captured spectra.

Proteins are of interest to science as they are excellent indicators of disease. The presence of protein where it is normally not found in large amounts (e.g. in the urine) can be a symptom of diabetes. Abnormalities in protein structure are indicative of genetic conditions and can lead to serious health concerns (e.g. the mutation in CFTR, cystic fibrosis transmembrane conductance regulator, leading to mucosal buildup in the lungs causing poor respiratory function in patients with cystic fibrosis). Proteins are also used to treat diseases, using unmodified natural protein like insulin to treat diabetes or laboratory altered or designed proteins such as therapeutic monoclonal antibodies known as biologics or biosimilars. Biologics are massive sellers in the pharmaceutical market, accounting for billions of dollars revenue each year. However, the safety and efficacy of these products is carefully monitored as small changes in protein post-translation modifications can lead to dangerous unforeseen side effects. For instance, glycosylation of monoclonal antibodies in the non-binding Fc region of the protein is a marker used by the body to regulate protein turnover. Therefore, raising antibodies with certain glycosylations can lengthen or shorten the half-life of the drug. Furthermore, improper glycosylation (for instance, with N-glycolylneuraminic acid, a glycosylation common in non-human mammals) can result in a harsh immune response, called ‘serum sickness’. Proper sequence screening provided by MS and tandem MS analysis is essential for minimizing these and other potentially dangerous sequence deviations.

Proteins are recovered from a wide range of biological media including blood, tissue, or cell lysate. These types of samples are incredibly complex; a major complication for analysis. Blood (for example) contains a myriad salts, sugars, fats, proteins, and other constituents that can interfere with separation and analysis. The proteins in these samples are present at vastly different concentrations, with an estimated concentration difference of nine orders of magnitude between the proteins at the lowest abundance to those of highest abundance. Blood contains about 70 mg/mL protein but the majority (>90% of the total protein by mass) of these proteins are albumin and immunoglobulins. Albumin, a transport protein, and immunoglobulins, proteins of the immune response, can bind reagents added downstream in the sample preparation process, potentially rendering them less effective. Furthermore, the presence of salts and other
constituents (for instance, NaCl which is present at ~160 mM in blood)\textsuperscript{12} has a detrimental effect on the most commonly available MS ionization method, ESI, which is most effective with solutions containing 10 mM or lower salt concentrations.\textsuperscript{13,14}

Mitigation of matrix interferences is paramount in proper preparation practices for the analysis of proteins by mass spectrometry. There are many methods and platforms used to purify proteins prior to or in-line with analysis, and those presented in this thesis fall into three general categories: section 1.2 digital microfluidics for fluidic handling at the microscale, section 1.3 in-solution and solid phase sample preparation techniques and section 1.4 gas-phase separations and mass spectrometry. Within each section, the basics of the technology or routine methods are presented along with the state-of-the-art, giving context to the work of the thesis overall.

1.2 Digital Microfluidics

DMF is a microfluidic technology where discrete microlitre-sized droplets are manipulated on an open array of insulated electrodes.\textsuperscript{15} Applying potentials to the electrodes allows for a variety of liquid handling tasks to be performed including splitting, mixing and merging droplets. Droplets can be controlled via a software interface to simplify microfluidic operations.\textsuperscript{16} In contrast to other channel microfluidic systems, DMF devices are more reconfigurable.\textsuperscript{17} A DMF device does not need to be designed for a specific application owing to the generalized array of driving electrodes in combination with the ability to address each droplet individuality. This section reviews the basic operating principles of DMF devices including the theory of droplet motion, device fabrication and the state-of-the-art in combining DMF devices with sample preparation applications.

1.2.1 Principles of Digital Microfluidics

There are two configurations for DMF devices:\textsuperscript{18} open (or one) plate and two plate devices (Figure 1-1). In the one plate configuration (Figure 1-1A), a substrate (commonly glass) holds the driving and ground electrodes in the same plane and they are both insulated by a dielectric layer. The dielectric layer is then rendered hydrophobic by coating with a fluorinated material like Teflon or Cytop which aids in limiting friction forces in droplet movement. Droplets sit on top of the hydrophobic coating, above the driving electrodes. In the two-plate configuration
(Figure 1-1B), the bottom of the device is reminiscent of the one-plate configuration, containing the same driving electrodes, dielectric and hydrophobic layers with sample droplets on top. However, the two plate configuration places the ground electrode (often composed of the transparent, conductive material indium tin oxide\textsuperscript{15} and always coated with the same kinds of hydrophobic materials as the bottom plate) above the droplets, sandwiching them between the top grounding electrode and the bottom plate driving electrodes. Both configurations can be operated with the droplets surrounded by open air or a specialized filler medium such as a fluorinated oil which can aid in droplet motion or limit droplet losses due to evaporation.\textsuperscript{15}

**Figure 1-1**: Different configurations of digital microfluidic devices. Cross-sections and top view show the different layers necessary for device construction. The thickness of the layers is not to scale. A. One plate DMF devices. B. Two plate DMF devices. This schematic also illustrates the necessary components of the electromechanical model of DMF operation.

Microfabrication by photolithography has been the method of choice for fabrication of DMF devices for many years.\textsuperscript{19} Photolithography allows the selective etching of the conductive material (like chromium metal) to create the various driving electrodes necessary for a DMF
device. Afterwards a dielectric coating (paralyene-C, SU-8, SiO₂, etc.)¹⁹ must be applied commonly through spin-coating or chemical vapour deposition. The thickness and properties of the coating dictate the voltages necessary for droplet movement on the completed device as well as the maximum voltages that can be applied before device failure. Finally, hydrophobic materials like Teflon-AF or Cytop are applied by spin-coating to complete the device.¹⁹ In general, this method of fabrication is time consuming and costly so there is a trend toward development of low-cost methods of device manufacture. Most promising is the concept of printing DMF devices on substrates such as paper.²⁰ Inkjet printing has been shown to be a viable solution to bringing down costs of DMF devices while maintaining the same levels of performance necessary for complex operations.²¹

Droplet motion is sometimes described as an extension of the electrowetting-on-dielectric (EWOD) phenomenon.²² EWOD describes the change in contact angle that a droplet undergoes when an electric potential is applied to it. Upon the application of potential, a droplet will wet a surface more readily i.e. the contact angle of the droplet decreases. However, the EWOD description does not account for several observable effects in DMF devices, namely that contact angle change is not necessary to move droplets, droplets can be manipulated at high frequencies of applied A/C potential and droplets of high dielectric constant are still movable.¹⁵ Instead, the electromechanical model of droplet motion is more accepted. In the electromechanical model, the DMF device can be considered in a circuit-like manner.²³ The droplet between the ground and actuating electrode is a capacitor, capable of storing energy. The amount of energy, \( E \), is proportional to the frequency of the applied potential and the amount of droplet covering the activated electrode denoted by its position, \( x \), along the length of the electrode (assuming a square shaped cross-sectional area of the droplet whose sides are length, \( L \)):¹⁵

\[
E(f, x) = \frac{L}{2} \left( x \sum_i \frac{\varepsilon_{r,\text{liquid}} V_{\text{liquid}}^2 (j2\pi f)}{d_i} + (L - x) \sum_i \frac{\varepsilon_{r,\text{filler}} V_{\text{filler}}^2 (j2\pi f)}{d_i} \right)
\]

where \( \varepsilon_{r,\text{liquid}} \), \( V_{\text{liquid}} \), and \( \varepsilon_{r,\text{filler}} \), \( V_{\text{filler}} \) are the relative permittivity and voltage drop of the liquid (i.e. the droplet) and the filler medium (whether it is air or something else) portions over the actuated electrode respectively, \( d_i \) is the thickness of layer \( i \), and layer \( i \) represents either the dielectric, top and bottom hydrophobic, liquid or filler layers. As \( x \) goes from zero to \( L \) i.e. as the droplet moves to cover the electrode completely, work is being done to move the droplet; the
work done is equivalent to the change in energy as $x$ goes from zero to $L$. Therefore, equation (1) can be differentiated with respect to $x$ to present the driving force as a function of frequency:\textsuperscript{15}

$$F(f) = \frac{L}{2} \left( \sum_l \frac{\varepsilon_0 \varepsilon_{r,\text{liquid}} V_{l,\text{liquid}}^2 (j2\pi f)}{d_l} - \sum_l \frac{\varepsilon_0 \varepsilon_{r,\text{filler}} V_{l,\text{filler}}^2 (j2\pi f)}{d_l} \right)$$  \hspace{1cm} (2)

When operating above a critical frequency (which is almost always how DMF devices are driven), an electric field gradient is developed within the droplet between the droplet and the actuated electrode. This causes liquid-dielectrophoretic force to pull the droplet closer to the electrode. In practice, at typical laboratory DMF operation ($f = 10$ kHz, $V = 100$ V$_{\text{rms}}$) droplets are moved with forces of tens of micronewtons, which is sufficient to overcome the resistive forces of friction between the droplet and the plate and viscosity inherent in the system.

An important consideration in DMF droplet movement for biological samples is the topic of biofouling. Despite the addition of hydrophobic layers, generally composed of materials deemed to be ‘non-stick’, DMF devices are prone to fouling when droplets carry solutes with sufficient hydrophobicity.\textsuperscript{24} For example, proteins have many large regions of high hydrophobicity which will interact and adsorb to hydrophobic surfaces. Over time the buildup of these non-specific adsorptions to the surface renders a DMF device unusable. An elegant solution is the addition of poly(ethylene glycol) and poly(ethylene oxide) copolymer surfactants such as BASF’s Pluronic line. It is hypothesized that the surfactant layer which forms on the outside of the droplet prevents long term interaction between dissolved biomolecules and the hydrophobic surfaces, minimizing the fouling effect.\textsuperscript{25} This allows droplet protein concentrations to be more than 1000-fold higher than without a surfactant additive.\textsuperscript{24} The choice of surfactant is largely dependent on the concentration of proteins and other biomolecules in the desired sample droplets as the different formulations have superior anti-fouling properties depending on the sample. With an optimal surfactant additive in the droplet, DMF operation lifetimes can be effectively tripled versus a non-optimal surfactant additive.\textsuperscript{25} Without such additives, realistic protein concentrations could not be used on a DMF device. However, the presence of these additives can be detrimental to many different sample preparation strategies and removal is often necessary before analysis by MS.
1.2.2 Digital Microfluidics and sample preparation

DMF devices share many of the same benefits that the more common channel microfluidics systems have over traditional macroscale techniques including minimization of reagent consumption through smaller sample volumes and limited dead volumes, efficient heat transfer from the small size of devices and an ease of integration to a variety of detection schemes. In addition, DMF devices also have some important advantageous over other microscale techniques. The open nature of DMF devices allows for efficient means of handling solid materials like magnetic particles or solid tissue samples which may otherwise clog a channel-based device. The independently addressable electrodes also allow a reconfigurability that is not present in channel device; if a failure occurs, droplets can be re-routed around the problem area. These benefits have led to the adoption of DMF for many applications including chemical reaction monitoring and synthesis, cell culture and microbiology, clinical diagnostics, and protein sample preparation. This section summarizes two important areas for the context of this dissertation: the use of solid materials such as magnetic particles on DMF devices, and state-of-the-art protein sample processing methods demonstrated with DMF.

Early investigations of DMF applications involve immobilization of molecules to the DMF device surface or onto immobile polymer supports. This is problematic in nature as it reduces the workable area of the DMF device as a portion is now rendered unusable for routine droplet motion and the surface area of the immobilization limits the amount of contact analytes will have with the immobilized enzymes or antibodies. This surface area limitation diminishes the analytical applicability of DMF. To overcome these challenges, magnetic microparticles with functionalized surfaces suspended in DMF compatible solutions were introduced. Magnetic microparticles have been used extensively since the 1990s. In contrast to non-magnetic particles (like those commonly found in chromatography columns) magnetic particles can quickly and easily be removed from sample by the application of an external magnetic force e.g. a bar magnet. Non-magnetic particles require samples to be passed through a packed bed (often under pressure) which can take significant time or by removal through centrifugation and carefully siphoning off the remaining liquid without disturbing the now packed beads, a tedious endeavor. When operated in a column, preparations with non-magnetic particles produce significant dilution of the sample which is not the case when magnetic particles are employed. This is because the entirety of the sample volume can be removed in discrete steps as opposed to
continuous flow through operation. Magnetic microparticles for biomolecular applications are synthesized by embedding magnetic cores into polystyrene spheres through swelling of the beads with organic solvents. The polystyrene coating allows for further functionalization with biomolecules like enzymes or antibodies, solid-phase extraction media like C18 or strong-cation-exchange resins, or various labelling dyes. Magnetic microparticles were first used in a DMF device for a single-plex immunoassay using magnetic particles with antibody functionalization specific to the desired analyte of insulin. Advances in instrumentation have allowed multiple droplets containing magnetic microparticles to be manipulated and magnetic separated simultaneously, allowing true multiplexed operation of DMF devices with these particles. However, these early investigations relied on simple, proof-of-concept sample compositions with minimal complexity, limiting their utility for real-world sample preparations.

The earliest reports of protein sample handling by DMF involved drying the protein samples directly onto the device for processing by MALDI-MS. Only minimal processing was performed in these studies, relying on simple washes with water to purify the sample. Later developments to DMF systems allowed for longer droplet movement times allowing for typical downstream protein processing steps such as reduction, alkylation and digestion to be performed in homogenous and heterogenous phases. Protein sample clean-up methods on DMF devices have so far been few. Jebrail and Wheeler demonstrated a simple protein precipitation strategy merging droplets containing dissolved proteins with organic solvents achieving 80% protein standard recovery and good MS performance but MS analysis from complex samples (cell lysate, fetal bovine serum) was not performed. Solid-phase extraction has been demonstrated on simple single peptide samples using immobile porous polymer monoliths (solid structures formed on the device with controllable porosity) with C12 chemistry and strong cation exchange chemistry. While effective, the use of immobile monoliths limits the reconfigurability of DMF devices and without the use of automated DMF control systems these studies were limited to handling only a single sample at a time as well. For DMF to progress as a protein preparation platform, multiplexing and demonstrations of non-model system utility are essential. It has been shown that precious biological samples such as dried blood spots and core needle biopsies are easily handled by DMF devices suggesting that extending common protein handling tasks like protein depletion or targeted immunoprecipitation could been of great benefit. Demonstration of multiplexed, complex sample handling for protein samples would pave the way forward for an
all-in-one DMF-based automated proteomics platform. Chapters 2 and 3 of this dissertation demonstrate the progress that has been made towards this goal.

1.3 Protein Sample Preparation Methods

1.3.1 Extractions

Extractions are implemented by presenting a sample with a particular chemistry which will selectively bind some or all of the proteins and peptides dissolved in the sample using various mechanisms such as hydrophobic interaction, ion pairing and dipole-dipole interactions. These methods often involve microparticles or resins functionalized with materials including long hydrocarbon chains, metal oxides, or antibodies, with each specific chemistry targeting a different application. The microparticles can be magnetic to allow for easy removal from the sample and to limit the dilution factor, or particles are bound in resins packed into columns or pipette tips to serve a similar function. This section discusses four separate techniques: the untargeted enrichment or extraction of proteins from a sample, targeted enrichment for a select group of proteins (e.g., glycoproteins), high-abundance protein depletion by immunoaffinity, and extraction of targeted proteins by immunoprecipitation.

1.3.1.1 Untargeted Enrichment

Untargeted enrichment typically relies on a fractionation of compounds from the sample into an extraction phase (liquid or solid) of a different chemistry. For example, an aqueous sample could employ a non-polar extraction phase to transfer the more non-polar solutes from the sample to the extraction phase. Liquid-liquid extraction, a classic analytical technique is not favored for protein preparation as many proteins and peptides can precipitate in organic solvents. This leads to the need to resolubilize the proteins/peptides which is often a non-trivial process. This leads more protein preparations using extractions to focus on solid phase extraction (SPE). SPE relies on the distribution of molecules in a sample between a solid phase and a liquid phase (e.g., the sample matrix) based on each molecule’s chemical affinity for one phase or the other. This distribution is an equilibrium process with an equilibrium constant $K_D$ representing the proportion of analytes in one phase or the other (equation 3),

$$K_D = \frac{x_s}{x_l} \quad (3)$$
with \( x_s \) denoting the concentration of analyte adsorbed to the solid phase and \( x_l \) denoting the concentration of analyte remaining in solution. The \( K_D \) is specific for each compound in the sample and is determined by properties of each phase such as hydrophilicity/hydrophobicity, pH, ionic strength and temperature. Controlling these properties allows improved fractionation for desired compounds. To free retained compounds the solid phase is treated with a solution with \( K_D \) favoring the desired analytes, reversing the partitioning. This is a common theme throughout chemical extraction and is also the basis of chromatography as well (see section 1.3.3).

In SPE, the chemistry is chosen to have a high affinity for the target; e.g., for untargeted protein enrichment, the solid phase is generally n-alkyl-based (reversed phase), as proteins and peptides have high affinity for these hydrophobic phases. \(^{62}\) Other phases will enrich a different set of proteins and peptides. Once the target is adsorbed to the solid phase, the remaining liquid phase can be removed, and the sorbent washed to remove non-specifically bound contaminants, and finally the adsorbed species can be eluted in a new solvent for which the analytes have greater affinity than the solid phase. The hydrophobic solid phases that are typically used for protein capture have low affinity for inorganic salts, meaning that much of the salt can be removed prior to elution, resulting in samples that are appropriate for MS analysis. For example, commercial C4 and C18 SPE cartridges can achieve >98% desalting efficiency while maintaining >70% and >50% peptide and protein recovery efficiency. \(^{63,64}\) While these recovery levels are usable they are far from ideal, especially when attempting to recover very low concentration compounds. More selective methods are needed to recover low concentration targets and more selective techniques also allow a deeper probing of the proteome that untargeted methods cannot reach. \(^{65-67}\)

SPE has typically been integrated into microfluidic devices in one of two ways: polymeric monoliths or microparticles. Monoliths are often constructed from methacrylate precursors directly on-device and can be functionalized for different applications. \(^{68,69}\) Microparticles can be manipulated within devices and are commercially available, reducing the fabrication time for the devices. \(^{70,71}\) DMF devices have recently been used with monoliths \(^{41,50}\) or magnetic beads \(^{72}\) for SPE for a number of different applications.
1.3.1.2 Selective Enrichment

In many applications, it is desirable to isolate particular sub-sets of proteins and peptides. For example, protein phosphorylation is essential in a number of biological processes, and disruption of phosphorylation can lead to a number of disease states. However, since phosphorylated proteins are generally found in low abundance, it is often useful to perform some form of enrichment prior to analysis. A popular method that is used to enrich phosphorylated species is metal oxide affinity chromatography. This is often implemented by means of packed titanium oxide microparticles, which can be used for phosphoprotein recovery of >90%. While more selective than untargeted techniques (like SPE), metal oxide affinity chromatography is not as selective as immunoaffinity; for example, in addition to phosphorylated species, acidic proteins, in general, are often retained. The metal oxide affinity chromatography mechanism involves electron pair donation from the phosphate groups oxygen atoms to the metal centre on the solid phase in a bidentate manner. Other functional groups like carboxylic acids that may be found on side chains of peptides can also interact in a similar manner with the metal centre. This effect can be suppressed by using a chemical modifier, 2,5-dihydroxybenzoic acid to displace non-phosphopeptides or by controlling the pH to protonate the carboxyl groups of acid peptides in the loading solution. These methods have also been miniaturized to operate in microfluidic systems.

1.3.1.3 Protein Depletion by Immunoaffinity

As indicated above, the serum proteome encompasses a wide range of concentrations. Low-abundance proteins in serum are numerous and are important as biomarkers for disease, but they are often masked in MS analysis by six highly abundant proteins: serum albumin, immunoglobulins A and G (IgA and IgG), haptoglobin, α-1-antitrypsin and transferrin, which make up 85% of the serum proteome by mass. Given this, there has long been interest in removing these highly abundant proteins prior to analysis, which can result in an increased likelihood to detect and identify low abundant proteins. Historically, several chemistries were used to deplete the high abundance proteins, typically by adsorbing the proteins to solid phases, leaving the solution-phase depleted of these high abundance species for further analysis. For example, cibacron blue is a dye molecule which preferentially binds serum albumin but has limited affinity for other proteins. When albumin is passed through a porous gel bearing
covalently bonded cibacron blue, the albumin becomes immobilized by binding the dye, while other proteins remain in solution. Likewise, IgA and IgG can be selectively removed using bacterial proteins Protein A and Protein G.\textsuperscript{81,82} While protein A/G are still in use, immunoaffinity techniques (described below) are now widely used for both depletion and enrichment via immunoprecipitation (see section 1.3.1.4).

Immunoaffinity techniques rely on antibodies, proteins that are formed biogenically as part of the vertebrate immune response.\textsuperscript{83} Antibodies structures all consist of two segments: the Fc region (crystallizing fragment) and the Fab (antigen binding fragment). Each class of antibody has a different number of these segments with the most analytically important class, immunoglobulin G (IgG), consisting of a single Fc region with two Fab regions above it in the shape of the letter ‘Y’. IgG is formed from two sets of two protein chains, the heavy chain (~55 kDa) and the light chain (~22 kDa) for a total molecular weight of approximately ~150 kDa. The C terminal end of the heavy chains forms the Fc region, while the light chains and the N-termini of the heavy chains form the Fab region. At the N terminal ends of the Fab region exists a region of high sequence variability. This is the site of antigen binding which arises during antibody development from immune receptor T or B cells. These cells bind foreign bodies and develop the specific interactions which will be expressed in the antibodies they produce. Each cell may bind a different area of a target, termed an epitope.\textsuperscript{84} When a set of antibodies is developed (or ‘raised’ by inoculation of a host such as a rabbit with the antigen of interest)\textsuperscript{85} \textit{in vivo} they are known as polyclonal, as the antibodies bind a specific antigen but each will bind a different epitope. Developments in cell culturing and engineering (specifically the hybridoma method\textsuperscript{86} and through genetic recombinant techniques\textsuperscript{87}) have allowed so-called monoclonal antibodies (mAbs) to be produced, where each antibody binds specifically to a single epitope of the target. This can be of great importance when developing therapeutic antibodies as each product produced will be identical, but it can create challenges using mAbs as an analytical tool.\textsuperscript{83} For instance, changing conformations of antigen (through changes in pH, or temperature etc.) can alter the binding capacity of the antibodies as the epitope of binding may have changed. In these cases, polyclonal antibodies may be of greater value to the researcher. Regardless of the method of production, developing an effective antibody is a time consuming and expensive process.\textsuperscript{88}
The affinity of antibody-antigen binding is measured as a dissociation constant, $K_d$, the ratio of the rate of reaction of dissociation by the rate of reaction of association.\textsuperscript{89} Thus, low $K_d$ values indicate a strong antigen affinity as the association rate is much faster than the complex can dissociate. Naturally raised antibodies have $K_d$s around the low nM level.\textsuperscript{90} Directed evolution can push $K_d$s much further, down to fM levels. The selectivity of the antibody is also of importance. A highly selective antibody will have limited cross-reactivity with other targets, despite similarity in sequence and structure. Thus, antibodies with high affinity and selectivity for high-abundance proteins make effective tools for depletion preparations. Steel \textit{et al.}\textsuperscript{91} were the first to use a monoclonal antibody specific for human serum albumin to selectively remove it from solution. Antibodies were couple to a microparticle resin using an N-hydroxysuccinimide coupling reaction to produce an affinity depletion column far more effective and selective at removing albumin than dye-based resins. Importantly, the immunodepletion method developed did not have appreciable non-specific effects. When the column was treated with acid to elute the capture components all the detectable proteins and peptides were from albumin or were fragments of albumin. Solid phases with tethered antibodies specific for each of the high-abundance serum proteins like described are now widely commercially available showing very good depletion efficiency (>80).\textsuperscript{92} When operated outside of a column, microparticles with antibody coatings can have even higher depletion efficiency, owning to the increased surface area of operation.\textsuperscript{93,94} Coupled with magnetic microparticles, this is an effective and fast method of depletion.

Microfluidics has been used to develop rapid, small scale protein depletion platforms as well, suitable for working with very limited quantities of precious samples. For example, McKenzie \textit{et al.}\textsuperscript{95} reported a microfluidic device that used Protein G functionalized microparticles to capture IgG from samples while leaving immunoglobulin M (the desired target) in solution for analysis. As previously discussed, channel-based microfluidic systems have a greater tendency for failure when operating with solid particles. For these reasons, a DMF-based platform can be considered as a viable and robust option for depletion type experiments. Chapter 2 of this dissertation discusses the development of a multi-parameter magnetic microparticle based depletion system for human serum albumin (HSA) and IgG using digital microfluidics.
1.3.1.4 Immunoprecipitation

The methods of purification described above are used to recover many proteins from a sample. But there are some applications in which it is desirable to isolate only one protein (and occasionally its associated isoforms and/or its binding-partners) for analysis. In such cases immunoprecipitation is an excellent method for purification, used throughout biology and biochemistry to selectively isolate a single protein from a complex matrix like cell lysate or blood serum. The principle of immunoprecipitation is similar as that used in immunodepletion. That is, capture antibodies (polyclonal or monoclonal) are raised to bind selectively to a given antigen. The sample is then treated with either (a) soluble capture-antibodies which bind the antigen that can subsequently be precipitated (via a second antibody known as a counter antibody selective for the capture antibody or Protein A\(^97\)), or (b) capture-antibodies bound to the surface of a resin or magnetic particle. In either variant, the sample solution is removed, the insoluble phase is washed to remove non-specific contaminants, and finally the antigen is released from the capture antibody by disrupting the protein-protein interactions. Since the antibodies are permanently bound to a solid phase the only component that becomes re-solubilized is the desired antigen and whatever reagents are used to elute that antigen from the solid phase. Antigens captured using this technique can include (but not limited to) particular proteins, proteins with a particular post-translational modification (such as phosphorylation\(^98\)), proteins modified with epitope tags like FLAG\(^99\) or protein-protein complexes (in a slightly modified procedure called co-immunoprecipitation\(^100\)). However, there are many serious drawbacks to immunoprecipitation preparations. The fact that protein-protein complexes can be capture means that whenever protein-protein complexes are present there is a likelihood that they will be capture even if the intention is to recover a single protein.\(^88\) In such cases, analysis by MS methods can help to not only characterize the protein networks which are recovered, but to determine if there is any biological significance to the interactions discovered.\(^88\) These interactions do not even need to be specific interactions (as in antibody-antigen binding events) but are commonly non-specific interactions, leading to an even more complex set of proteins (and potentially even small molecules and metal ions carried by transport protein like albumins)\(^101\) recovered. Alleviation of these issues requires effective preparation steps often consisting of numerous and careful washes which benefit from expensive automated equipment to achieve automation.\(^102\) This has led many researchers to explore microfluidics as a low-cost
alternative while also enabling smaller sample sizes. Channel based microfluidic devices have been used for immunoprecipitation for various applications ranging from broad enrichment of a particular post-translational modification\textsuperscript{103} to selective isolation of particular targets.\textsuperscript{104} Chapter 3 of this dissertation reports the development of an immunoprecipitation method for HSA using digital microfluidics and magnetic particles which can overcome the challenges of using channel systems with solid particles.

1.3.2 Gel Electrophoresis

Gel electrophoresis (GE) is a technique in which charged biomolecules are asymmetrically transported through a medium upon application of an electric field. For the separation of proteins, the classic method is SDS-PAGE, sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) where a cross-linked polyacrylamide gel is employed as the separation medium using an SDS pre-treatment on the proteins.\textsuperscript{105–107} SDS molecules bind to each protein proportionally to the size of those proteins (approximately 1 SDS per 2 amino acids in the protein) which produces a constant (negative) charge to size ratio as larger proteins hold more SDS. In addition, the proteins are denatured and left in a linear, rod-like conformation.\textsuperscript{108} This allow each protein to move through the gel by way of electromotive force from the applied electric field with a velocity determined by its molecular weight and no other factors like its shape or native charge, leading to separation of the proteins.\textsuperscript{109} Overall speed through the gel can be controlled by the porosity of the gel which is determined by the concentration of polyacrylamide and the degree of cross linking provided by reagents like bisacrylamide.

SDS-PAGE is a simple technique that produces extremely good results, especially when multiple dimensions of separation are integrated. Two dimensional SDS-PAGE resolves 5000+ proteins without difficulty.\textsuperscript{110} The major downside of SDS-PAGE is its limitations in coupling with various MS techniques, which are required for the highest degree of protein characterization.\textsuperscript{111} One major challenge is recovering proteins or peptides from the gel in a manner compatible with MS. Electroelution methods from SDS PAGE are effective at recovering peptides and small proteins but struggle at larger molecular weights rendering the techniques inapplicable to intact protein analysis for top-down proteomics.\textsuperscript{112–114} Construction of dissolvable gels has improved such that these systems can be used with much of the same resolution and efficiency as
traditional PAGE gels.\textsuperscript{112} Furthermore, the presence of SDS on the proteins must be greatly depleted to perform most forms of mass spectrometry (see section 1.4.1.1.2).\textsuperscript{112} Several methods have successfully been employed (100+ fold reduction in SDS) both off-line such as through protein precipitation\textsuperscript{115–117} or molecular weight cut-off filters\textsuperscript{118–120} and in-line with MS detection\textsuperscript{121}. Protein recovery from gels has pushed researchers to investigate separation methods which avoid gel entrapment all together. Separation using 1- and 2-dimensional chromatography without prior gel separations are becoming more popular (section 1.3.3). Microfluidic methods coupling separations and extractions are also becoming more available, especially owing to the ease of integrating directly to MS\textsuperscript{122,123} or additional separation methods like capillary electrophoresis\textsuperscript{124}. These methods have the added benefit of being applicable to extremely small samples, an area that traditional SDS PAGE has not been especially effective at addressing.\textsuperscript{125}

1.3.3 Chromatography

Chromatography is the separation of molecular species as a function of asymmetric partitioning between two phases: a stationary phase and a mobile phase (similar to the SPE techniques described in section 1.3.1.1). As analytes enter the column they are distributed onto the solid phase of the column (equation 3). However, unlike SPE, in chromatography, the mobile phase is flowed continuously through the column, providing a constant supply of fresh solvent. Analytes with high affinity for the solid phase move slowly through the column, while analytes with low affinity move rapidly. Specifically, the time required for a given analyte to elute form the column is the retention time, \( t_R \), given by equation 4,

\[
 t_R = t_S + t_M
\]

where \( t_S \) is the amount of time the analyte remains on the stationary phase and \( t_M \) the time required for a completely unretained constituent to move from the beginning to the end of the column. The retention time of an analyte is specific to the column geometry, chemistry, and mobile phase flow rate. As such, the unitless retention factor \( k_A \) is defined as equation 5:

\[
 k_A = \frac{t_R - t_M}{t_M}
\]
As indicated above, analytes elute from (or exit) a chromatographic separation as a function of their $k_d$. Chromatography is well suited to separate analytes from each other, and from other species in the sample matrix, enabling significant levels of purification. There are two main types of chromatography that are widely used: gas chromatography (GC) and liquid chromatography (LC). Each has many different variants which can involve a change in geometry (e.g., large bore column vs. capillary, high-performance liquid chromatography vs. ultra-high performance liquid chromatography) or chemistry (e.g., reverse phase vs. normal phase vs. strong cation exchange, etc.). GC has been used sparingly for protein and peptide analysis; LC has been far and away the more important technique for proteomics.

1.3.3.1 Liquid Phase

Liquid chromatography (in the form of high-performance liquid chromatography, HPLC) is now used abundantly in proteomic analyses. This was not always the case. In the early 1970s, liquid chromatographs were only capable of separating 2-4 small peptides. By the end of that decade, C18 columns were being used to separate much larger numbers of small peptides. For example, O’Hare and Nice developed an HPLC method to separate 13 different polypeptides of varying lengths as well as some proteins. At that time, ESI was essentially non-existent as an ion source, and LC-MS interfaces were rare; thus, UV-absorbance was used for detection. The development of techniques to sequence peptides by MS was contemporaneous with these LC improvements, but HPLC separations of the day could not be performed in-line with the MS detection. The first HPLC-MS analysis of peptides with direct interfacing between the LC and MS was reported in the early 1980s, via thermospray ionization (complete with in-line enzymatic digestion of the proteins) for the analysis and sequencing of peptides up to 13 residues in length. The later part of the 1980s saw the advent of ESI interfaces for LC-MS and since that time, the vast majority of LC-MS has been performed using ESI or a related technique (see section 1.4.1.1.2). In the 1990s, a decrease in particle-sizes and an increase in liquid pressure limits of pumps lead to the development of ultra-performance liquid chromatography (UPLC). UPLC provides better separation with lower reagent consumption and is overtaking traditional HPLC as the dominant LC method for peptide MS. The advantage of UPLC comes mainly from the smaller particle sizes used in analytical columns. Measurement of column separation efficiency is governed by the van Deemter equation:
\[ H = A + \frac{B}{u} + C_m u + C_s u \] (6)

where \( H \) is the height of a separation plate (a measure of separation efficiency, where a smaller value indicates more effective separation per length of the column), \( A \) denotes Eddy diffusion (the dispersion of solutes caused by the flow profile), \( B \) denotes axial diffusion of solutes, \( C_m \) denotes mass transfer of solutes to the mobile phase, \( C_s \) denotes mass transfer of solutes to the solid phase and \( u \) is the flow velocity. Of importance to the particle size are the \( A \) and \( C_m \) terms which can be expanded to:\(^{137}\)

\[ A = 2\lambda d_p \] (7)

\[ C_m = \frac{f'(k)d_p^2}{D_M} \] (8)

where \( d_p \) is the particle diameter, \( \lambda \) is the efficiency of the column packing, \( f'(k) \) is a function derived from the retention factor of the column and \( D_M \) is the diffusion coefficient in the mobile phase. With a reduction in \( d_p \), both \( A \) and \( C_m \) are minimized, leading to a smaller \( H \) and more efficient separation. More efficient separation results in smaller peak widths (as low as 1 s) of eluted analyte which requires faster detection methods.\(^{135}\) The speed of UPLC methods becomes a challenge for slow mass spectrometry instruments and data acquisition methods. This can limit the quality of the resulting data as many species will co-elute in the timeframe of instrument acquisition. The addition of other separation methods such as ion mobility may alleviate this burden and Chapter 5 of this dissertation investigates the use of one such technique with a data independent MS acquisition method.

The most common stationary-phase chemistry used in HPLC is reversed phase (i.e., non-polar groups such as octadecyl residues) owing to its exceptional range of applications and ability to operate with aqueous mobile phases which allow biological samples to be handled with ease but many others exist.\(^{138}\) Reversed phase columns are popular for protein and peptide analysis because most protein and peptide molecules have hydrophobic regions. Reversed phase HPLC is today used routinely in proteomics for top-down,\(^{139,140}\) bottom-up,\(^{141–143}\) and targeted protein analysis.\(^{144,145}\) For additional selectivity and resolution of charged peptides, ion-pairing agents are often used. These surfactant-like molecules associate with both charged analytes and the
chromatographic column, providing a degree of retention for those analytes. The addition of ion-pair reagents often improves the sensitivity and selectivity for the analysis of peptides at different pHs, allowing for orthogonal 2-dimensional separation based on pI.\textsuperscript{146} Ion-pairing agents can also affect the charge state of peptide analytes, allowing for tunability to resolve a particular, desired charge state.\textsuperscript{147,148}

In some cases, rather than reversed-phase separations, samples containing highly charged peptides are separated using ion-exchange chromatography. In this format, the stationary phase is charged (composed of materials like sulfonic acid derivatives for cation exchange\textsuperscript{149} or ammonium derivatives for anion exchange\textsuperscript{150}), resulting in no retention in components of the same charge, limited retention for components of neutral charge and excellent retention for oppositely charged analytes. This is a highly effective technique for proteome fractionation\textsuperscript{151} and charge-state distribution analysis of particular families of proteins (such as monoclonal antibodies\textsuperscript{152}).

Another chromatographic mode other than reversed phase used in proteomic analyses is size-exclusion chromatography (SEC, also called gel filtration). In SEC, analytes are separated by size (or more specifically hydrodynamic radius). An SEC column is composed of small porous particles with a fixed pore size. Proteins are pushed through the column by the mobile phase flow. The retention in SEC is a function of entropic interactions of large proteins that may not be able to fully penetrate into the pores.\textsuperscript{153} Smaller molecules are more favoured to remain in the pores causing them to have a longer retention time. SEC is an excellent technique for analyzing and quantifying protein aggregates\textsuperscript{154} but also has some applications to studying large protein complexes without use of a targeted antibody or bait protein.\textsuperscript{155–157}

A third chromatography mode (other than reversed phase) that is used in proteomic research (in fact, experiencing a surge in popularity now despite a history dating back more than 20 years\textsuperscript{158}) is a variant of normal phase chromatography called hydrophilic interaction liquid chromatography (HILIC). In HILIC a polar solid phase is used in combination with a mobile phase that is primarily organic, with a small amount of water.\textsuperscript{159} The water enriches the stationary phase and the organic content of the mobile phase is responsible for eluting the analytes from the column. This arrangement alleviates the need for careful solvent preparation in
traditional normal phase chromatography (which must be strictly water-free) as well as providing a suitable solvent system for analysis by MS. For proteins and peptides, retention to the column is a combination of polar interactions: dipole-dipole interaction and hydrogen bonding, with hydrogen bonding being the primary mechanism of action. HILIC has found a niche in glycopeptide and phosphopeptide analysis. Finally, there are myriad affinity chromatography systems used in proteomic research, such as immunoaffinity, other biomolecule affinity separations such as lectin or protein G, and immobilized metal ion affinity chromatography (IMAC), each suited for particular applications (see section 1.3.1.2).

As is the case for gel electrophoresis, one dimension of separation in chromatography often lacks the capacity to resolve analytes from complex mixtures. To solve this problem (as in electrophoresis), a second dimension of orthogonal chromatography can be added to the system, comprising two-dimensional liquid chromatography (2D-LC). There are two main techniques for transferring analyte from the first separation phase to the second: heart cutting (in which only the most ‘interesting’ fractions of the eluent from the first dimension are transferred) or comprehensive (in which all of the effluent from the first column is transferred, effectively making the peak capacity of the system the product of the peak capacities of both dimensions). Heart cutting is an excellent method for enrichment of peptides and saves considerable time on an analysis over comprehensive methods. Comprehensive 2D-LC, despite its extended run times (sometimes 24 hours or more), represents the best method in use today for complex proteomic samples. Comprehensive 2D-LC has become popular for biopharmaceutical applications, including the complete analysis of sets of antibody-drug-candidates.

Finally, like many of the techniques described above, LC has been miniaturized in various microfluidic devices and formats. Microfluidic configurations are desirable for LC, as the small dimensions increase the surface area to volume ratio, which results in increased LC resolution combined with a reduction in sample size necessary for a given analysis. These systems are also straightforward to integrate with MS via electrospray ionization creating all-in-one separation and ionization devices for peptide and protein analysis. The work described in this dissertation did not rely on microfluidic LC; however, conventional HPLC was a critical tool used in much of the work described in the following chapters.
1.4 Mass Spectrometry and Ion Mobility

1.4.1 Mass Spectrometry and Tandem Mass Spectrometry

Despite its existence for more than a century,\textsuperscript{174} it was not until the 1980s\textsuperscript{175–177} and moreso the 1990s\textsuperscript{178–181} that MS began to make a significant impact in the world of protein analysis. A mass spectrometer is an analytical instrument that separates ions based on their mass to charge ($m/z$) ratios. The critical parts of a mass spectrometer include an ionization source to convert analytes (often dissolved in solution-phase) into gas-phase ions, a mass analyzer that is responsible for the $m/z$ separation, and a detector for the separated ions. In many ways, the ionization source is the most critical component for proteomic applications, because of the challenge of getting proteins into the gas phase. Here I review two common forms of ionization (section 1.4.1.1): matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) (noting that the reader can evaluate many other ionization methods elsewhere\textsuperscript{182–184}). The inventors of these techniques shared the 2002 Nobel Prize in Chemistry in no small part because ESI and MALDI made proteins accessible to analysis by MS. Mass separation in protein analysis is accomplished by several different configurations including quadrupoles, ion traps, time-of-flight (ToF) and Fourier Transform-based technologies like ion cyclotron resonance or Orbitraps. The work in this dissertation focuses on mass analyses by quadrupole mass filters and ToF analyzers; thus, the background is restricted to these techniques (section 1.4.1.2). Finally, it should be noted that by combining multiple mass analyzers in series, it is possible to perform tandem MS (or MS/MS) analyses. This process is critical in the identification of proteins by peptide mass fingerprinting, sequencing, and quantification (section 1.4.1.3).

1.4.1.1 Ionization techniques

There are many considerations to make in choosing the ionization source for protein analysis by MS such as the nature of the analyte, and the nature of the companion purification techniques employed. For example, most proteins are non-volatile and employ liquid chromatography separation upstream of analysis, meaning the ionization source must be compatible with non-volatile analytes dissolved in HPLC elution buffer. In these cases, it is useful to employ ionization techniques that can be implemented in-line with HPLC, including thermospray,\textsuperscript{185} field desorption,\textsuperscript{186} or ESI. The former techniques are largely obsolete now because of the
simplicity, generality, and sensitivity of ESI.\textsuperscript{187} For applications which focus on intact proteins and do not require HPLC separation (such as tissue imaging) fast atom bombardment (FAB)\textsuperscript{188}, secondary ion mass spectrometry (SIMS) and MALDI\textsuperscript{189,190} are all functional techniques. But because of challenges related to sample handling and limitations in sensitivity for FAB and SIMS, MALDI-MS is the most commonly used in the field.

1.4.1.1.1 Matrix-assisted Laser Desorption Ionization

MALDI is a ‘soft’ ionization technique – that is, it imparts only enough energy to the analytes to ionize them, prevent extensive fragmentation (a problem for the traditional ‘hard’ ionization techniques, such as electron impact).\textsuperscript{191} In MALDI, a solution of analyte molecules with chromophoric matrix molecules is allowed to crystallize.\textsuperscript{192} The crystalline sample is then illuminated by a fast, pulsed laser, which results in the formation of a plume of material (including desorbed and ionized analyte). There is some disagreement in the literature about the mechanism of ionization in MALDI. The first (and oldest) model, Coupled Physical and Chemical Dynamics (CPCD), assumes an initial state with neutral analytes that become charged by interaction with matrix ions formed by a photoionization process.\textsuperscript{193} Thus in CPCD, photoionization of the matrix is a prerequisite for ionization. The alternate ‘lucky survivor’ model assumes that analytes are pre-charged upon inclusion in the matrix,\textsuperscript{194} which is consistent with the acidic environment of typical MALDI matrix and solvents. In this mechanism, after absorption of the light, clusters form containing analytes and their counter-ions. Most such clusters are neutral and thus cannot be detected by MS; however, a few of the clusters have a non-quantitative mix of analyte ions and counter-ions, leading to a net charge on the analytes. These ‘lucky survivors’ are the analytes detected in the mass spectrometer. The latest evidence suggests that both models may play a role in different circumstances.\textsuperscript{195}

MALDI is particularly useful for ionizing large, intact biomolecules, but it can also be used for smaller peptides\textsuperscript{196,197} (with the caveat that some low-mass peaks are obscured by matrix peaks in MALDI-MS). Chapter 2 of this dissertation describes the use of MALDI-MS to evaluate protein-containing samples before and after depletion of high-abundance species.
1.4.1.1.2 Electrospray Ionization

ESI is another ‘soft’ ionization technique. In ESI analytes are driven through the outlet of a narrow capillary in a strong electric (10⁶-10⁷ V/m) field, all under atmospheric pressure.¹⁹⁸

ESI has a long history before its development (and subsequent Nobel Prize) as an MS process by John Fenn.² In conventional implementation, the sample is dissolved in a polar solvent (such as water, acetonitrile, methanol or a mix of these and similar solvents) and then pumped into an capillary which is placed in the field with the field gradient and capillary outlet directed toward the inlet of the mass spectrometer.²⁰¹ As solvent exits the capillary, a Taylor cone is formed that degenerates into a spray of droplets moving towards the mass spectrometer inlet. As the droplets travel down-field they evaporate and shrink. As the droplet radius decreases but the charge stays constant Coulombic repulsion increases, forcing the droplet to undergo an uneven fission to smaller droplets. These smaller satellite droplets undergo subsequent fissions into smaller droplets still. Due to the loss of solvent from evaporation the concentration of analytes and electrolytes is enriched i.e. charge is concentrated with each evaporation with only 1-2% of the mass but 10-18% of the charge transferred to each of the resulting droplets.¹⁹⁸ As is the case for MALDI, there is some disagreement in the literature about the mechanism of this process. In the charged residue model (CRM), proposed by Dole et al.²⁰⁰ in 1968, the gas-phase analyte ion is the final component of the constantly shrinking and charge-enriched droplets after all excess solvent and electrolytes are lost to splitting and evaporation. In the ion evaporation model (IEM), the ionization of the analyte happens much earlier in the process (while the droplet still exists), and the electric field is sufficient to pull the ions out of the droplets before they have shrunk down to their smallest sizes.²⁰² Much like MALDI, there has been some consensus that CRM and IEM are both observed under various circumstances, perhaps with larger compounds like proteins ionizing by CRM, while small molecules may ionize by IEM. Some have proposed a third mechanism for large, highly disordered molecules like polymers called the Charge Ejection Model.²⁰³

ESI is the gold standard ionization method for most protein and peptide applications. It is useful for intact as well as digested proteins because of its soft ionization and its high sensitivity. But it is not without its limitations; for example, the standard ESI process is significantly altered by the presence of high concentrations (greater than 10⁻⁵ M) of non-volatile compounds such as salts
frequently found in biological matrices.\textsuperscript{204} These non-volatile interferants complicate the ESI process by destabilizing the electrospray conditions and by competing with the desired analyte for available charge. Beaudry and Vachon\textsuperscript{205} found that the presence of charge species in an electrospray alter the conductivity, altering the surface tension and disrupting the forces responsible for forming the electrospray Taylor cone as well as those responsible for droplet fission. Matrix compounds of basicities greater than that of the analyte of interest can compete for available charge in the electrospray droplets, with the result being a loss of ionization efficiency for the weaker bases.\textsuperscript{206} Suppression does not need to come from the sample matrix necessarily, with chromatographic ion-pairing agents like trifluoroacetic acid also being likely suppressors.\textsuperscript{207} In addition, detergents used in protein separation processes like SDS PAGE have been shown to have detrimental results at SDS concentrations above 100 ppm and perhaps even as lower\textsuperscript{121,208}. The ability for matrix interferants to complicate the ionization process of ESI demonstrates the value in sample preparation methods which eliminate as much of the original matrix as possible. Modification of the electrospray process is also an effective method of mitigating some ionization suppression. So-called nanoelectrospray ionization (NSI) is an associated electrospray method which is much tolerant of non-volatile interferants (on the order of 100s of mM vs. tens of μM for conventional ESI).\textsuperscript{209} In NSI, the capillary orifice is much smaller (only 1-2 μm in diameter) which necessitates a significantly lower flow rate (less than 1 μL/min, usually less than tens of nL/min).\textsuperscript{210} The much smaller orifice in NSI leads to much smaller initial droplet sizes in electrospray,\textsuperscript{14} meaning that there are fewer fission events required to produce bare analyte ions. Thus, the electrolyte concentration increase that results from many fission events in conventional ESI is limited in NSI. These and other properties of NSI including its particular geometry in relation to the mass spectrometer result in a higher ion transfer efficiency when compared with ESI, with 1 in 390 ions being successfully detected using NSI vs. 1 in ~200 000 for ESI.\textsuperscript{210}

The dimensions of ESI and NSI capillaries also lend them well to interface microfluidic devices with MS. This has led to a wide variety of microdevices which incorporate many different protein sample handling techniques in an automated fashion directly in-line with MS. ESI capillaries can be constructed directly into microfluidic chips (from glass or plastic)\textsuperscript{211–213} or external emitters can be interfaced to the chips.\textsuperscript{51,214} Devices have been designed for operations
such as sample preparation, chromatography or capillary electrophoresis, and cell culture monitoring, all with direct access to MS detection on the same chip. Chapters 2 through 5 of this dissertation all employ ESI-MS, with Chapters 3 through 5 using this ionization method exclusively.

An important consideration for most ion sources, especially ESI and its related techniques is the concept of in-source fragmentation (ISF). As ions enter the mass spectrometer they are accelerated by a potential difference between the cone inlet and the ion skimmer in a region of low pressure (10⁻² torr). While the mean free path of the ion in this region is quite large, collisions will still occur with the lingering gas molecules in this zone. The collisions can be used to impart a declustering effect on incoming ions which have clustered with polar gas molecules from the electrospray process. Generally, the energy gained is sufficient to only cause the ion-neutral declustering effect. However, the energy of acceleration can be increased such that these collisions are sufficient to cause fragmentation of the incoming ions or the ions themselves are fragile enough that fragmentation will occur regardless. Often, ISF is induced on purpose. This allows researchers to uncover structural information about ions without performing a complete MS/MS spectral acquisition or to study non-covalent clustering between molecules. Other times ISF is an undesirable outcome as it reduces ion intensity for the fragmenting species and can complicate mass spectra by introducing unexpected peaks leading to poor biological or chemical interpretation of the spectra. The presence of additional fragment peaks can have a negative impact on data-dependent acquisition methods (see below). Controlling ISF can be accomplished by reducing the internal energy of incoming ion i.e. effective temperature of these ion. One strategy is reducing the accelerating potential of the incoming ions at the likely expense of allowing more clustered ions to enter the mass spectrometer. To limit clustering, placing a heated transfer capillary after ions pass through the skimmer allows a further increase in internal energy, sufficient to cause final solvent evaporation. And Collette and De Pauw have shown that altering the spray solvent conditions can result in a change in the effective temperature of ions. The addition of glycerol to the solvent resulted in hotter ions than without glycerol, leading to an increase in collision rate downstream. Chapter 4 of this dissertation examines an alternative method for mitigating ISF by controlling the cluster formation.
1.4.1.2 Mass Analysis

Mass analyzers are the component of mass spectrometers responsible for separating ions with different m/z ratios. There are many kinds of mass analyzers, but they all rely on the ability of electric and/or magnetic fields to move charged particles. Mass analyzers used for proteomics include magnetic sectors, quadrupole filters, time-of-flight analyzers, and ion traps. The following sections introduce quadrupole mass filters and time-of-flight mass spectrometers, which were used in the work described in this dissertation.

1.4.1.2.1 Quadrupole Mass Filters

The quadrupole mass filter consists of 2 pairs of matched rods (4 rods total), either cylindrical or hyperbolic in shape on the order of 10-20 cm in length. Each of the pairs has a different electric potential applied to it, a DC potential \( U \), and a radio-frequency (RF) oscillating AC potential \( V \). The two potentials work together to form a mass filter. Large ions respond more strongly to \( U \), as their size and momentum limits their response to the quickly changing \( V \). But some level of response to \( V \) is necessary for ions to traverse the path such that they do not simply collide with the oppositely charged pole. Therefore, a strong \( V \) creates a high pass mass filter. Small ions behave oppositely, their small size allowing rapid changes in direction driven by \( V \). Thus, a weak \( V \) lowers the mass cutoff of the filter. To create a good low pass filter, you generally need the DC as the RF signal is insufficient to alter the trajectory of a large ion to prevent its collision with a pole but it will not be strong enough to shunt a small ion into a pole either. By controlling \( U \) and \( V \), the analyst can use a quadrupole to create a narrow window of mass selection, limiting the ions which maintain a stable, helical trajectory between the poles. By scanning this narrow band from low m/z to high m/z, an entire mass spectrum can be acquired. This is accomplished by holding the U/V ratio constant but changing the magnitude of those potentials. The scanning speed of a quadrupole system is fast (1000s of m/z per second), making the quadrupole an effective, inexpensive, and compact means of mass separation for many applications. For some applications, it is useful to shut off \( U \), operating in what is known as “RF-only mode” to allow a large window of m/z to navigate between the poles. In this mode quadrupoles (and related systems consisting of six or eight poles, known as hexapoles and octapoles respectively) are effective ion focusing systems and are often used as collision cells in tandem mass spectrometer setups. Note that even in this configuration, the system has both an
upper and lower mass limit where ions above or below those masses will not be transmitted. Alterations in quadrupole geometry allow the quadrupole to function as an ion trap, further enhancing its usefulness as a mass analyzer.\textsuperscript{234}

1.4.1.2.2 Time-of-Flight Mass Analyzers

Time-of-Flight (ToF) mass analyzers were very popular in the early days of MS.\textsuperscript{235} Their popularity waned during the 1980’s with the rise of the quadrupole but picked up again after the advent of the MALDI source.\textsuperscript{236} ToF operation is simple, relying on the different durations that ions of equal kinetic energy but different mass require to transverse a fixed distance.\textsuperscript{231} The potential energy of an ion $E_p$ moving through an electric field is given by equation 9,

$$E_p = qV = ezV$$ (9)

where $q$ is the charge of the ion, $e$ is the charge of electron, $z$ is the integer number of electron charges of the ion, and $V$ is the electric potential. In the accelerator region of a ToF mass analyzer, the potential energy of the ion is made to convert to kinetic energy $E_k$ according to equation 10,

$$E_p = ezV = \frac{1}{2}mv^2 = E_k$$ (10)

where $m$ is the mass of the ion and $v$ is its velocity. We can rearrange equation 5 to solve for velocity (equation 11).

$$v = \sqrt{\frac{2ezV}{m}}$$ (11)

ToF mass spectrometers contain a drift tube that is field-free between the accelerator and the detector. If the drift tube has length $L$, then the time $(t)$ required for an ion to traverse $L$ is given by equation 12.

$$t = \frac{L}{v}$$ (12)

Substituting equation (11) into equation (12) and rearranging gives equation 13:
\[ t = \frac{L}{\sqrt{2eV}} \sqrt{\frac{m}{z}} \quad (13) \]

ToF instruments have some major limitations which limited their initial adoption as MS instruments, two of which are the necessity for a pulsed ion source and the variability in each ion’s starting kinetic energy. A pulsed ion source is readily available from MALDI, and it is one of the reasons for the recent revival in ToF instrument use (the other being the virtually unlimited \( m/z \) range of ToF instruments). The more common ESI source presents a challenge for ToF which can be overcome using orthogonal acceleration-ToF (oaToF). In these instruments, the drift tube is rotated 90° from the trajectory of the incoming ion beam. Under normal operation, the ion beam passes by the drift tube without entering; however, at a particular frequency and duration, an additional potential is applied to direct ions into the drift tube in pulses. Despite the unlimited maximum \( m/z \) limit and the mass resolution of the instrument is limited by the initial spread of the ions as they enter the drift tube, both in space and in velocity i.e. their energy spread. This variability in starting kinetic energy of the ions can be addressed using an ion reflector. In a reflector, the kinetic energy of the ions is unaffected, but instead their total flight path is adjusted to compensate for differences in kinetic energy. That is, the reflector is placed at the end of the drift tube, bearing a zone of increasing and opposing electric potential, up to a magnitude slightly larger than the initial accelerating potential. Ions with high energy penetrate farther into the reflector, while ions with lower energy penetrate less far; this difference in travel distance has the effect of focusing all ions with the same \( m/z \) to the same position when exiting the reflector (negating the initial \( E_k \) differences among them).

1.4.1.3 Tandem Mass Spectrometry

Tandem MS (often called MS/MS) is the operation of mass analyzers twice (or more) in series in a given mass spectrometry experiment, coupled with a means of fragmentation between them. Fragmenting ions by techniques like collision induced dissociation (CID), electron capture dissociation (ECD) or photo-induced dissociation (PID) allows for structural information to be obtained from the ions, as the process of fragmentation is guided by the strength of the intramolecular bonds of the ion (noting that each fragmentation technique has its own suite of bonds it is suited to access). The work described in this dissertation relies on CID, which is discussed here.
Tandem MS is accomplished in one of two ways: in space or in time. Tandem MS in time is restricted to ion trap instruments which can alternately serve as mass analyzers and fragmentation chambers by tuning the applied potentials. These instruments were not used in the work of this dissertation. Tandem MS in space is implemented on instruments that include two mass analyzers separated by a collision cell. These mass analyzers are typically quadrupole mass filters as is the case in triple quadrupole instruments (QqQ, used in Chapter 3), but hybrid quadrupole-ToF (Q-ToF) instruments are also very popular for protein analysis due to the increase in mass resolution (used in Chapters 4 and 5). In either case, the first mass analyzer selects precursor ions, which are then fragmented in the collision cell; the second mass analyzer scans the product ions that result from the fragmentation. In CID, the collision cell, a quadrupole (or hexapole, etc.) operated in RF-only mode for both QqQ and Q-ToF instruments, is filled with a neutral gas like He or N$_2$ called the collision gas. As precursor ions enter the cell they collide with the collision gas, converting kinetic energy to internal energy. At sufficient internal energy, bonds will break. The kinetic energy of the precursor ion is dictated by the kinetic energy of the collision gas it collides with and also its own kinetic energy derived from accelerating in the instrument. Thus, the charge of the precursor ion is important since the kinetic energy gained from electric potentials is greater for an increased charge. In sum, precursors with higher charge are more likely to fragment, which is the reason that many proteins and peptides are selected for CID at high charge state. Of critical importance is the predictability of the resulting tandem mass spectrum when peptides are fragmented. Peptides molecules generally break along the peptide backbone and under CID conditions the most common site of fragmentation is at the amide bond. This results in the formation of ‘b’ and ‘y’ ions with ‘b’ ions showing a C terminal aldehyde and ‘y’ ions showing the N terminal amide functionality. The series of ‘b’ or ‘y’ ions in the resulting spectrum can be used to determine the sequence of the peptide as each ion will differ by the predictable mass of the next adjacent residue.

Tandem MS is useful for proteomics using a mass spectrometer, as it permits protein identification by peptide sequencing. For example, in bottom-up processing, proteins are broken into peptides, typically using enzymes like trypsin. These peptides are then analyzed by tandem MS as it is possible that two or more peptides will have the same $m/z$ (and the same retention time when coupled to LC). But, the resulting MS/MS product ion spectra for the
peptides will be different. These MS/MS spectra can then be compared to databases of known or predicted peptide precursors and fragmentation spectra to determine their identity. This requires reducing the MS/MS spectrum to a peptide sequence as described above and then matching the resulting string to other peptide sequences in the database. An early example of this technique was implemented by Eng, McCormack and Yates[^249] which was later developed into the SEQUEST data analysis suite. Using statistical methods and extensive databases, the protein composition of a complex mix of peptides from many proteins can be determined. For instance, SEQUEST employs cross correlation to match a theoretical tandem MS/MS spectrum to the experimental spectrum[^249]. Cross correlation compares two signals point by point and determines the amount of deviation at those points, known as the displacement value, \( \tau \). Two signals with minimal deviation i.e. many \( \tau \) values close to 0 are shown to be the same. These techniques can also be applied to simpler protein digest samples such as those found in a biopharmaceutical context. This forms basis of bottom-up proteomics using a mass spectrometer.

### 1.4.1.3.1 Data Acquisition methods

Collecting MS/MS data for a complex protein sample such as a bottom-up protein preparation is a non-trivial task as even with efficient separation a single MS spectrum might contain dozens of coeluting peptide fragments. Two major strategies have emerged for collecting MS/MS data: data dependent acquisition (DDA) methods where selected peaks from the MS spectrum are chosen for subsequent fragmentation and data independent acquisition (DIA) methods where an attempt is made to collect MS/MS spectra for all species in the initial MS survey. In DDA, potential precursor ions (up to a user defined number for a given MS survey spectrum) for fragmentation are selected in order of decreasing intensity based on a series of user selected criteria.[^250] These selection criteria include minimum intensity, number of allowable repeat detections, and the amount of time a precursor is eligible for fragmentation.[^251] Additional selectivity is provided by user defined exclusion lists or dynamic exclusion algorithms which automatically exclude ions that appear most frequently.[^252] However, two major drawbacks are present for DDA. First, the act of choosing precursors based on minimum intensity and in decreasing order from highest intensity limits sampling to a restricted dynamic range. Peptides of the lowest intensity (those below the minimal intensity threshold) will not be samples and in the cases of many species co-eluting or eluting very quickly medium intensity peptides (those whose
intensity is above the minimum threshold but the maximum number of precursor ions has already been reached) will also be missed.\textsuperscript{253} Secondly, it is unlikely that each precursor ion will be sampled for MS/MS data at its maximum intensity owing to dynamic exclusions and the serial nature of precursor ion selection.\textsuperscript{254} DDA methods fail to produce sufficient MS/MS data for many of the peptide ions discovered at the MS level, only accounting for 16\% of an estimated >100 000 peptide features in a human cell lysate experiment.\textsuperscript{250} DIA methods attempt to address the shortcomings of DDA by capturing MS/MS spectra for all precursor ions detected.

There are many different implementations of DIA but for the work covered in this dissertation we will focus on Sequential Windowed Acquisition of All Theoretical Fragment Ion Mass Spectra (SWATH). SWATH captures MS/MS data from preset isolation windows which transmit all available precursor ions of a given $m/z$ spread.\textsuperscript{255} A SWATH analysis collects data on a number of these windows one after another. These isolation windows, termed ‘swaths’ are used in sequence to cover the entirety of the detectable $m/z$ range of the instrument. Each swath employs a small 1 Da overlap with the previous window in the series to maximize coverage of ions at the edge of a swath. So, with swath widths of 25 Da, a $m/z$ range of 400-1200 can be covered by 32 separate windows. With sufficient time for each acquisition a sweep of the whole $m/z$ range can be accomplished in 3.2 s.\textsuperscript{253} The novelty of SWATH does not come from the wide isolation windows (which had previously been used in other DIA methods)\textsuperscript{256} but from the data interpretation using spectral library searching. First, peak finding is employed to identify the peaks associated with peptide precursor ions at the MS level by matching the expected $m/z$ with those found in the analysis.\textsuperscript{257} Knowing the $m/z$ of the precursor ion directs you to the swath which will contain the fragment ions for the precursor. The $m/z$ for the expected fragment ions can then be searched in that particular swath as well. The resulting extracted ion peaks for the fragments are compared to that of the precursor and those found in the spectral library and a match is determined based on retention time, peak shape, and $m/z$ accuracy of the matches. SWATH allows identification and quantitation at an increased dynamic range (over 4 orders of magnitude) over other data acquisition methods.

However, SWATH is still reliant on the presence of a spectral library for best results and these libraries are built using DDA methodology employing significant sample fractionation methods to increase the depth of the proteomic coverage.\textsuperscript{255} This means that SWATH’s best merit is in
simultaneous quantitation of peptides and proteins in samples. Recently, efforts have been made to allow SWATH to produce spectral libraries directly from the acquired data allowing true untargeted searching and quantitation to be performed simultaneously but robust comparisons and characterization of that method are still ongoing.\textsuperscript{258} Among the difficulties in achieving untargeted identification in SWATH is the complexities inherent in the multiplexed MS/MS spectra. But simplification of these spectra would require additional sample preparation or separation strategies to be employed. Chapter 5 of this dissertation reports on steps toward supplying an additional means of separation to a SWATH analysis.

1.4.2 Ion Mobility

Ion mobility spectrometry (IMS) is the name given to a suite of techniques that are used to sort ions in the gas phase, often applied to the detection of compounds such as drugs or explosives.\textsuperscript{259} IMS has been paired with MS to provide a means of gas phase chemical separation which aids in the reduction of chemical noise, allows differentiation of some isomers and isobars, and provides an estimate of the size of ions.\textsuperscript{260,261} Here I introduce two forms of ion mobility spectrometry: drift time ion mobility spectrometry (DTIMS) and differential mobility spectrometry (DMS, also known as field-asymmetric ion mobility spectrometry, FAIMS).

1.4.2.1 Drift Time Ion Mobility Spectrometry

DTIMS is a simple form of gas-phase ion separation. When used in conjunction with a mass spectrometer (DTIMS-MS), after ionization by ESI (see section 1.4.1.1.2), MALDI (see section 1.4.1.1.1) or other sources, ions enter a drift tube of length $L$ (often tens of centimetres, with longer tubes providing better separation but with a reduction in sensitivity) which contains a chemically inert buffer gas (like nitrogen)\textsuperscript{262} at pressure of 2-10 torr\textsuperscript{263} if placed inside the vacuum of the mass spectrometer (so called reduced pressure IMS, RPIMS)\textsuperscript{264} or at high flow rates (1 L/min)\textsuperscript{263} in atmospheric pressure conditions (so called atmospheric pressure IMS, APIMS)\textsuperscript{264}. The buffer gas flows opposite to the flow of incoming ions.\textsuperscript{265} Upon application of a constant, near-uniform electric potential $E$ parallel to the flow of the ions, ions are propelled at a constant velocity through the tube toward the mass spectrometer on the other side, experiencing collisions (greater than $10^6$ per ion) with the gas as they traverse the tube.\textsuperscript{266} The collisions between the buffer gas and the incoming ions compensates for the energy gained by the ions due
to the electric field. The tendency to collide with the gas medium is known as the collisional cross section of the ion, $\Omega$; a larger $\Omega$ results in slower progress through the drift tube. Further, since ions are being pulled through the tube by way of a weak (1-500 V/cm, depending on buffer gas pressure)\textsuperscript{267} electric field ($E$), ions with higher charge $z$ move more rapidly through the tube. Thus, DTIMS allows separation of ions based on their $\Omega/z$ ratios. The collisional cross section of an ion is proportional to the time required to move through the drift tube, the drift time ($t_D$) (equation 14),

$$\Omega = \frac{(18\pi)^{1/2}Ze}{16(k_BT)^{1/2}} \left( \frac{1}{m_I} + \frac{1}{m_N} \right)^{1/2} \frac{760ETD}{273.2\cdot LPN}$$

where $k_B$ is the Boltzmann constant, $T$ is the temperature of the buffer gas (in Kelvin), $P$ is the pressure of the buffer gas (in torr), $e$ is the elementary charge, $m_I$ is the mass of ion, $m_N$ is the mass of the buffer gas (in kg), and $N$ is the density of the buffer gas (in m\textsuperscript{-3}).

DTIMS-MS is a useful tool in protein analysis and can be used for the study of protein structure.\textsuperscript{262,268,269} Since DTIMS-MS can separate ions based on collisional cross section, it is possible to separate isobaric species such as proteins or peptides with different glycosylation states. For example, Hinneburg \textit{et al.}\textsuperscript{270} have shown that it is possible to separate peptides of identical sequence (in this case YGNPNETQNNSFK, a sequence derived from human butyrylcholinesterase) but different glycosylation positions (the second and the third asparagine residues) and isomerization patterns (e.g., $\alpha$-2,3 vs. $\alpha$-2,6 N-acetylneuraminic acids) of the glycan derivatives. However, validation of this method was only performed on samples containing a single pure protein and were analyzed in an off-line method. This limits the applicability of this technique to routine quality control in a biopharmaceutical setting.

\textbf{1.4.2.2 Differential Mobility Spectrometry}

Differential mobility spectrometry (DMS or field-asymmetric mobility spectrometry, FAIMS) is a second common mode of ion mobility spectrometry. In DMS (Figure 1-2), the driving force for ion-migration toward the detector is provided by a flowing, inert drift gas (as opposed to an electric field like in DTIMS\textsuperscript{271}) between two parallel-plate electrodes with a gap distance of only a few millimetres and a total plate length of no more than 50 millimetres.\textsuperscript{265} The separation is accomplished by applying two asymmetric electric fields (i.e of very different magnitudes)
orthogonal to the direction of travel of the drift gas. The first of the electric fields has high
magnitude and is driven by a pulsed AC separation voltage, SV. The SV pulse consists of two
parts: a brief high field pulse (1000s of V) followed by a low field section (magnitude of less
than a few hundred V) of opposing charge such that the net field strength during the waveform is
0.272 The SV waveform oscillates at high frequency, typically at 1.25 MHz.273 This results in an
electrostatic attraction of analytes toward one of the plates, which depends on the polarity of the
ions relative to the field as well nonlinearly on strength of the field.274 The second, a DC
compensation voltage CV, is also applied, with the opposite polarity to SV at a much lower
magnitude (selectable between typically -100 and 100 V, highly dependent on the analyte of
interest and DMS cell geometry)273. As ions are moved by SV and CV, they also experience
collisions with drift-gas molecules. Thus, each ion’s motion is also a function of its collisional
cross section, Ω (as in DTIMS). Each ion requires a different CV to reach the detector for a
given SV, resulting in separation of ions based on CV.
Figure 1-2: Schematic representation of the process of DMS. a. Incoming ions are pushed between the DMS electrodes by the incoming gas flow. b. The pulsed separation voltage pushes ions toward one plate. c. The compensation voltage brings a single ion back toward the inlet of the mass spectrometer. d. Each ion needs a unique SV and CV in order to enter through the inlet of the instrument.

To enhance the separation of ions in DMS, gas-phase modifiers can be added to the drift gas. This changes the relative molecular clustering around each ion as each ion is now effectively being solvated by the newly added modifier, and the extent of the clustering can be detected in DMS.\textsuperscript{275} The alternation between high and low fields in DMS serves to break up and reform clusters during the separation.\textsuperscript{276} This contrasts with DTIMS, in which the cluster state of ions does not change.\textsuperscript{276} The cluster-pattern formed around an ion is dependent on chemical parameters such as dipole moment and hydrophobicity, meaning that the addition of clustering modifiers enables selective separations in DMS. This results in increased tunability and separation power in DMS as well as potentially increasing the peak capacity of the technique.\textsuperscript{277} For example, Schneider et al.\textsuperscript{276} examine the effect of varying the cluster state on a set of 70 small molecule analytes. They found that with optimal modifier conditions (an environment consisting of nitrogen and 1.5% isopropanol that the peak capacity could be increased by
approximately 3 times (from 13 for a pure nitrogen environment to 44 when isopropanol is added). They observed that the addition of polar modifiers like isopropanol, water or alcohols reduced (from a nitrogen gas control) the CV for the analytes and spread those CVs across a wider range, leading to the peak capacity increase. However, they note that the use of isopropanol as a modifier lead to ion count depletions (of up to 20-fold) for analytes where the gas-phase acidity of the modifier is greater than that of the analyte. This leads to proton transfer away from the analyte, limiting the ability to detect it. For best results, this limitation must be overcome by investigating which modifiers are best suited for the detection of certain classes of molecules i.e. to determine which modifiers to use to give the best separation for peptides or proteins.

The application of FAIMS/DMS to protein and peptide analysis is quite new. For example, little DMS research has been performed on intact proteins, because their strong molecular dipoles result in non-unimodal CV distributions. Some very recent DMS experiments have been used to evaluate changes in conformation of intact proteins, and hydrogen-deuterium exchange for structural determination.278,280 Peptides have also not been popular targets for investigation by DMS until very recently. Synthetic peptides were recently separated using DMS by a number of researchers, and it has been shown that phosphopeptides can be separated from non-phosphopeptides using DMS.286 The use of mixed-mode modifiers (e.g., drift gas modifiers of two or more constituents with different chemistries) has been shown to improve the peak capacity in protein digest samples by up to 20%.287 Most DMS studies to date have only been applied to model systems. One counter-example is a recent study in which DMS was used to quantify the therapeutic peptide pasireotide and the disulfide-bond containing sunflower trypsin inhibitor, a 14 residue peptide, from blood plasma samples. As most peptide analysis by MS involves using LC as a primary separation method there is an interest in coupling DMS to the system for increased separation and specificity. There are LC-DMS-MS studies which rely on targeting specific analytes with known SV and CV parameters, but there is an increasing effort at developing methodology to operate in an untargeted manner (see the early examples of this in Venne et al. and Kapron et al.). Despite these successes, the adoption of LC-DMS has been slow, at least partly because of challenges related to duty cycle where the required dwell time (~50 ms) to acquire a high quality mass spectrum and clear the DMS cell of ions
limits the number of CVs which can be scanned in a typical chromatographic peak. Chapter 5 of this dissertation represent my contribution to answering this challenge.

1.5 Perspectives

In its approximately 25-year history, scientists working in the area of MS proteomics have produced a number of important discoveries, including a draft map of the human proteome (albeit limited to only to proteins which are directly coded for), supporting the development of the therapeutic antibody and biosimilar industry, finding surprising interactions in biological processes, and uncovering biomarkers and elucidating disease processes by monitoring changes in the proteome. But these discoveries have also come with an increasing awareness that complex mixtures must be simplified sufficiently to allow the mass spectrometer to produce high-quality spectra. The work described in this dissertation adds to the ever-growing set of methods and technologies that allow for high quality protein mass spectrometry to proceed, by introducing two liquid phase microfluidic sample handling methods (Chapters 2 and 3) and two gas-phase separation and clean-up methods (Chapters 4 and 5) for protein and peptide samples.

In this dissertation, I present new techniques for simplifying and analyzing proteins samples using MS. In Chapters 2 and 3 I introduce solution-phase microfluidic techniques (immunodepletion and immunoprecipitation respectively) for processing complex biological samples in an automated fashion using DMF. When I began my work, DMF had been used in a variety of “proof of concept” studies, but had not yet been applied to solving many real-world problems. The techniques that I developed expand the DMF toolbox to incorporate cutting-edge sample preparation techniques in cell biology and proteomics that had not yet been effectively miniaturized. Likewise, in Chapter 4, I present a method for eliminating background noise and MS instrument induced fragmentation in protein MS. This was motivated by previous research in the field of ion mobility modifiers which had shown an ability to improve signal-to-noise ratios of various ions, even without the use of DMS as a separation technique. In my work, I have expanded this idea to peptide samples, to evaluate the potential to reduce their complexity. In Chapter 5, I combined HPLC-MS with DMS ion mobility for protein digest samples. This was motivated by a lack of research in using DMS for protein and peptide analysis despite the orthogonal separation capability of DMS. This technique allows a simple method of combining
the two separation techniques and I propose it will increase the adoption of DMS ion separation in the field. I propose that this work represents an important expansion of the fields of DMF and DMS for the analysis of proteins as both techniques can greatly reduce the challenges associated with complex proteomic samples.
Chapter 2
Digital Microfluidic Platform for Human Plasma Protein Depletion

2.1 Abstract

Many important biomarkers for disease diagnosis are present at low concentrations in human serum. These biomarkers are masked in proteomic analysis by highly abundant proteins such as human serum albumin (HSA) and immunoglobulins (IgGs) which account for up to 80% of the total protein content of serum. Traditional depletion methods using macro-scale LC-columns for highly abundant proteins involve slow separations which impart considerable dilution to the samples. Furthermore, most techniques lack the ability to process multiple samples simultaneously. We present a method of protein depletion using superparamagnetic beads coated in anti-HSA, Protein A, and Protein G, manipulated by digital microfluidics (DMF). The depletion process was capable of up to 95% protein depletion efficiency for IgG and HSA in 10 min for four samples simultaneously, which resulted in an approximately 4-fold increase in signal-to-noise ratio in MALDI-MS analysis for a low abundance protein, hemopexin. This rapid and automated method has the potential to greatly improve the process of biomarker identification.

2.2 Introduction

Recently, clinical proteomics has emerged as an important field for the discovery of disease biomarkers. In particular, researchers are now systematically searching the human plasma proteome for biomarkers that can be used to predict the risk of cancer or monitor the progression of disease. However, these efforts are hindered by the complexity of plasma, which has a proteome that spans 10 orders of magnitude in concentration. As such, biomarkers at low concentrations can be masked by highly abundant proteins (HAPs) such as immunoglobulins (Igs) and human serum albumin (HSA). To reduce the complexity of plasma, many proteomic workflows include a pretreatment procedure that depletes HAPs from the sample. These depletion procedures typically use affinity chromatography spin columns, which contain affinity ligands that bind to specific HAPs to remove them from the sample. Although affinity chromatography is a useful pretreatment strategy, there are drawbacks that
limit its effectiveness. First, chromatography is a labor-intensive process, requiring many sample preparation steps (e.g., multiple fluid handling steps followed by centrifugation). Additionally, the depletion process requires at least a 10-fold dilution of the sample in an appropriate loading buffer.\textsuperscript{78} Furthermore, there is a risk of sample loss arising from protein degradation during the long pretreatment procedure (30 min to 2 h), post-extraction concentration steps to counteract the sample dilution, and sample handling during transfer and aspiration. These limitations represent both a major source of variability and a bottleneck for clinical proteomics. To address these limitations, some groups have explored the concept of miniaturizing affinity chromatography using microfluidics.\textsuperscript{310–312} Microfluidic affinity chromatography has the potential to speed up protein depletion, minimize sample dilution, and eliminate the need for centrifugation and trained personnel. In a recent example, McKenzie et al.\textsuperscript{95} demonstrated a pneumatically driven microfluidic device that depletes 66–77\% of immunoglobulins (IgGs) from a complex sample using Protein G functionalized beads dried on the device surface. The work was focused on preventing false positives in IgM assays and did not examine proteomic sample preparation. In parallel, many groups have developed microfluidic systems to conduct immunoassays.\textsuperscript{313} Analogous analyte capture concepts developed for microfluidic immunoassays can be similarly applied to HAP depletion, the difference being that, in immunoassays, the unbound constituents are discarded, while in HAP depletion, the unbound constituents are preserved. Several liquid actuation schemes have been explored for microfluidics;\textsuperscript{314} however, digital microfluidics (DMF) has a number of potential advantages for HAP depletion. In DMF, discrete droplets are manipulated by electrostatic forces on an array of electrodes coated with a hydrophobic insulator.\textsuperscript{15} When a sequence of voltages is applied to the electrodes, droplets can be addressed individually and made to move, merge, mix, and dispense from sample/reagent reservoirs.\textsuperscript{315} Since droplet operations can be conveniently controlled, experimental conditions can be modified to alter the protein depletion time or implement multistage depletion using the same device design.\textsuperscript{301,316,317} DMF has been used in several sample preparation/ extraction strategies, including protein precipitation,\textsuperscript{49} reversed phase\textsuperscript{41} and strong cation exchange\textsuperscript{50} solid-phase extraction, and liquid–liquid extraction.\textsuperscript{318,319} In addition, DMF has been implemented for magnetic bead-based immunoassays,\textsuperscript{20,26,34,320} in which an external magnet facilitates the separation of droplets from antibody-coated beads. To our knowledge, DMF has never been used as a proteomic preparation tool for HAP depletion. We report here the development of a new
protein depletion platform that relies on DMF for liquid handling and superparamagnetic beads (coated with Protein A, Protein G, and Anti-HSA antibodies) for removal of abundant proteins. This new device brings about enhancements to traditional chromatography spin columns or flow-based microfluidic platforms. First, this method is fully automated and does not require external agitation for mixing or centrifugation; after placing the sample in the device, no further manual intervention is required. Second, the device depletes proteins rapidly because of efficient bead/sample mixing during incubation\(^3\)\(^4\) (e.g., \(\sim 9\) min is required to remove 95% of a 0.5 mg/mL protein from solution). Third, the device can be programmed to carry out various permutations of protein depletion, involving the simultaneous or sequential removal of HSA and IgG. Finally, we propose that this has the potential for facile integration with other microfluidic proteomic processing techniques including reduction, alkylation, and digestion\(^3\(^8\),\(^4\(^7\) and separations.\(^3\(^2\)\(^1\)

2.3 Experimental Section

2.3.1 Reagents and Materials

Unless otherwise noted, reagents were purchased from Sigma-Aldrich (Oakville, ON). Deionized water (DI H\(_2\)O) was utilized for all solution preparation and had a resistivity of \(>18\) M\(\Omega\) cm at 25 °C. All protein and processing reagent solutions were prepared in working buffer (aqueous phosphate buffered saline, PBS, containing 1.5 mM KH\(_2\)PO\(_4\), 155 mM NaCl, and 2.7 mM Na\(_2\)HPO\(_4\) at pH 7.2, supplemented with 0.05% w/v Pluronic F-68) prior to use.

2.3.2 On-chip Protein Depletion Reagents

Reagents used on-chip were prepared in-house. Protein solutions of human serum albumin (HSA, molecular weight based on amino acid composition of 66 437 Da) and hemopexin (molecular weight approximately 57 000 Da), were formed from lyophilized solid in PBS buffer, purchased from Life Technologies (Carlsbad, CA). Solutions of human IgG (molecular weight approximately 150 000 Da) were diluted from a stock solution of 4.7 mg/mL in PBS buffer. Superparamagnetic beads with specific functional coatings were obtained from Millipore (Billerica, MA). For the depletion of IgG and HSA, PureProteome Protein A/G Mix Magnetic Beads (LSKMAGAG02) and PureProteome Albumin Magnetic Beads (LSKMAGL02) were used, respectively. The beads are 10 μm in diameter and are coated with a mix of Proteins A and G and anti-HSA, respectively. Bead dilutions were performed by immobilizing the beads in a
magnetic separation rack, removing the supernatant, and adding the desired volume of PBS. Two types of suspensions were used here: Protein A/G beads alone (at a dilution of 1:4 from stock) or a mixture of Protein A/G beads (at a dilution of 1:4 from stock) and anti-HSA beads (at a dilution of 1:2 from stock). Magnetic bead concentrations were determined using capacity values from the supplier. Fluorescein-isothiocyanate (FITC) labeled human IgG at 0.5 mg/mL was purchased from GenScript USA Inc. (Piscataway, NJ). FITC labeled HSA at 1 mg/mL was purchased from Abcam (Cambridge, MA).

2.3.3 Off-Chip MALDI and LC-MS/MS Protein Depletion Analysis Reagents

ZipTip C4 pipet tips and Milli-Q water were purchased from Millipore (Etobicoke, ON). MALDI matrix solution was prepared by dissolving 10 mg of sinapinic acid (SA) in 1 mL of 50:50 DI H$_2$O/acetonitrile (ACN) obtained from Caledon (Georgetown, ON) containing 0.1% trifluoroacetic acid (TFA) from Thermo Scientific Pierce (Nepean, Ontario). Bovine serum albumin (BSA) obtained from Mann Research Laboratories (Port Saint Lucie, FL, USA) was used as the calibration standard for the analysis of all samples.

2.3.4 Device Fabrication and Operation

Device fabrication reagents included photoresist developer MF-321 from Rohm and Haas (Marlborough, MA), chromium etchant CR-4 from Cyantek (Fremont, CA), photoresist stripper AZ-300T from AZ Electronic Materials (Somerville, NJ), Teflon-AF from DuPont (Wilmington, DE) and Parylene-C dimer from Specialty Coating Systems (Indianapolis, IN). Digital microfluidic devices were fabricated in the University of Toronto Nanofabrication Centre (TNFC) clean room facility and were assembled as described previously (Figure 2.1A), and summarized here. Bottom plates of DMF devices consisted of an array of 80 interdigitated working electrodes (2.2 mm x 2.2 mm) connected to 8 larger reservoir electrodes (16.4 mm x 6.7 mm) and 4 waste reservoirs (16.4 x 6.4 mm). Bottom plates were constructed with standard photolithography methods. Briefly, glass substrates (50.8 mm x 76.2 mm x 1.1 mm) coated with chromium (200 nm) and photoresist from Telic Co. (Santa Clarita, CA) were exposed to UV from a Suss MicroTec mask aligner (29.8 mW/cm$^2$, 10 seconds) under an acetate photomask printed at 20,000 dpi (Pacific Arts and Designs, Inc, Markham, ON). The exposed substrates were then developed in MF-321 (3-5 min.) and baked on a hot plate (125°C, 1 min.). Developed
substrates were then etched in CR-4 chromium etchant for 3 minutes before being stripped of remaining photoresist in AZ-300T (5 min.). Substrates were rinsed in isopropanol, contact pads were covered in dicing tape and the substrates were coated with ~7 μm of Paralyene-C by vapor deposition. Subsequently, ~200 nm of Teflon-AF was applied to the devices by spin coating (1% w/w Teflon-AF in Fluorinert FC-40 at 2000 rpm for 45 seconds). Finally, the devices were baked on a hot plate (165⁰C, 10 min.) and the dicing tape removed from the contact pads. Top plates of DMF devices were fabricated by spin coating ~200 nm of Teflon-AF (as above) on unpatterned indium-tin oxide (ITO) coated glass substrate (Delta Technologies Ltd, Stillwater, MN) and then baking on a hot plate (225⁰C, 10 minutes). Devices were assembled by attaching the ITO top plate to the electrode array bottom plate with a spacer formed from two layers of Scotch double-sided tape (3M, St. Paul MN). The total spacer thickness was ~180 μm. An automated two plate actuation system (described in detail elsewhere\textsuperscript{34}) was used to control droplet movement and magnet position for the immobilization of magnetic particles. Droplet movement is controlled via custom Microdrop\textsuperscript{16} software which was interfaced to the control system to engage a magnetic lens assembly. This magnet provides approximately 600 μN of force which exceeds the minimal threshold for magnetic separation (470 μN).\textsuperscript{34} Droplet driving voltages were between 100 and 120 VRMS at 10 kHz (sine wave). During incubation, the droplets were moved in a modified “figure 8” pattern (Video 2-1) to ensure proper dispersal of particles. Waste and unused fluids were removed from devices by wicking using KimWipes (Kimberly-Clark, Irving, TX).

2.3.5 On-Chip Protein Depletion Protocol

On-chip protein depletion was performed in eight steps: (1) One droplet each of superparamagnetic magnetic beads and protein sample (4 μL each) were loaded on the device, (2) beads were actuated into the device array, (3) the magnet was engaged to immobilize the beads onto the device surface, (4) the supernatant was removed from the beads, (5) the protein sample droplet was merged with the beads and the beads were dispersed, (6) the suspension was actively incubated (moved in a figure 8 pattern, as above) for 10 min, (7) the beads were immobilized and the supernatant was separated from the beads, and (8) the depleted protein sample droplet was collected in the reservoir for removal and subsequent analysis (2 min). In practice, the eight-step procedure was typically performed on four protein samples in parallel.
2.3.6 Fluorescent Characterization of On-Chip Depletion

The kinetics of on-chip depletion was probed using FITC-IgG and a suspension of Protein A/G beads, using a variation of the protocol described above. Briefly, steps (1−5) were applied to a sample of FITC-IgG (0.5 mg/mL), which was then incubated for only 30 s in step (6). After steps (7−8), the supernatant droplet was driven to an unused portion of the device, and the fluorescence intensity was probed by loading the device into a plate reader (PHERAstar microplate reader, BMG Labtech, Ortenberg, Germany). Measurements were performed in “well scanning mode” (COSTAR 96 well plate geometry) with three flashes per scan point and a gain of 100 using 485 nm excitation and 520 nm emission. After measurement, the depletion was continued by resuspending the beads in the FITC-IgG droplet and mixing for an additional 30 s, followed by extracting the supernatant (steps 5−8) for another measurement of fluorescence. This procedure (deplete for 30 s, extract supernatant, measure fluorescence, and resuspend beads) was repeated until 10 min of total incubation time was completed. The efficiency of on-chip depletion of both IgG and HSA was tested using a suspension of Protein A/G and anti-HSA beads. The two analytes were evaluated separately. For IgG, a droplet of 0.5 mg/mL FITC-IgG was extracted (steps 1−8) with active incubation for 10 min in step (6). In step (8), the supernatant droplet was driven to an unused portion of the device, and the fluorescence intensity was probed as above. The supernatant droplet was then extracted again (steps 1−8) using a fresh aliquot of beads, followed by a second fluorescence intensity measurement. For HSA, an identical process was used to extract 0.5 mg/mL FITC-HSA, except with a gain of 125 in the fluorescence intensity measurements. For all fluorescence experiments, the intensity data were normalized to the control intensity (before depletion) to obtain relative values. Fluorescence measurements were carried out on three to four samples with fluorescence determined for each sample following 10 min depletions. Blank measurements were taken from onchip regions with no droplets.

2.3.7 MALDI-MS Characterization of On-Chip Depletion

Protein mixture solutions containing two high-abundance proteins (2 mg/mL human IgG and 0.5 mg/mL HSA) and one low-abundance protein (0.1 mg/mL hemopexin) were depleted by digital microfluidics as described above. Samples were collected for MALDI-MS analysis before depletion, after a single round of depletion with one aliquot of mixed beads, and after two rounds
of depletion with two aliquots of mixed beads. Depleted samples were collected for analysis by removing the top plate and transferring the sample droplet by pipet. Four replicates were collected and evaluated for each single and double depletion experiment. After processing by DMF, each sample was purified using a ZipTip C4 according to the manufacturer’s instructions. Briefly, ZipTip C4 tips were wetted in 50% ACN containing 0.1% TFA (5×) and then equilibrated in 0.1% TFA (5×). After equilibrating, fluid in the ZipTip C4 pipet tips was drawn in and out of the tip for 20 cycles for a maximum binding of complex mixtures and then washed with 0.1% TFA (3×). Finally, samples were eluted directly in SA matrix solution onto a stainless steel MALDI target plate. After drying, spots were analyzed using a PerSeptive Biosystems Voyager DE Pro MALDI-TOF Mass Spectrometer (AB Sciex, Framingham, MA, USA) operating over a m/z range of 20 000–200 000. A total of 250 shots were collected per spectrum, with laser power adjusted to optimize the signal-to-noise ratio (S/N). Data were processed by baseline correction, resolution (set to 100), and noise smoothed (default settings) using Voyager Data Explorer software. Signals were extracted from prominent peak heights, and root-mean square noise (NRMS) values were estimated from a spectra region with no prominent peaks (m/z 88 000–120 000).

2.4 Results and Discussion

2.4.1 DMF Device and Method
Figure 2-1. Device and processing scheme. (A) Schematic representation of the digital microfluidic device used for protein depletion in the automated magnetic separation system. Inset shows a cross-section of the device layers when the magnet is in position for magnetic separation. (B) Schematic representation of protein depletion using magnetic beads and digital microfluidics. First, functionalized magnetic beads are isolated from their supernatant by magnetic separation. Second, protein samples are added to the magnetic beads and mixed. Third, application of a magnetic field immobilizes the beads again. Fourth, the immobilized beads are separated from the depleted protein solution by DMF actuation, and the depleted sample is ready for analysis.

Digital microfluidics enables the manipulation of discrete droplets on an array of electrodes and thus offers a number of advantages for sample preparation prior to biochemical analysis. We hypothesized that digital microfluidics would be a convenient platform to use for depletion of HAPs from complex proteomic samples. Current commercially available depletion methods for depletion of a single protein on columns such as the ProtoPrep 20 (Sigma-Aldrich) require between 20 and 60 min for completion for a single sample. The device used here is shown in
Figure 2-1A. Droplets of samples and reagents are loaded into reservoir electrodes, where they can be aliquoted/dispensed, mixed, and separated using a defined voltage program. In addition to droplet manipulation, samples can be further manipulated using antibody-functionalized superparamagnetic particles. The particles can be controlled with magnetic fields, allowing for separation of specific molecules bound to the particles from the remainder of the droplet (the “supernatant”). The interplay between magnetic forces and interfacial forces arising from droplet manipulation can be tuned by moving a magnet vertically under the device (either close to the device, “engaged”, or away from the device, “disengaged”). A general scheme for HAP depletion by DMF and magnetic particle immobilization is depicted in Figure 2-1B. As shown, a droplet containing superparamagnetic particles is positioned over an engaged magnet, the initial supernatant is driven away, and a second droplet containing proteomic analytes is delivered to the immobilized beads. The magnet is then disengaged allowing for resuspension of the particles and active mixing, followed by engaging the magnet a second time to allow for the particles to be separated again. The resulting supernatant droplet should (ideally) contain substantially depleted concentrations of constituents that are bound to the beads. The full process is shown in Figure 2-2, which depicts a series of images from a movie (available as Supporting Video 1). With the device format used here, it was feasible to implement this process in a multiplexed fashion, up to four samples at a time. With the recent report of DMF devices bearing thousands of independently addressable electrodes, we propose that, in the future, much higher levels of multiplexing might be achieved.
2.4.2 On-Chip Depletion Kinetics and Efficacy

Two fluorescent assays were developed to determine appropriate conditions for protein depletion. First, an assay to determine the kinetics of depletion was developed, using fluorescently labeled IgG (FITC-IgG) and Protein A/G-labeled particles (Proteins A and G bind IgG with high specificity). Supernatant droplets containing FITC-IgG were probed repeatedly after successive 30 s incubations with particles using techniques reported previously for

Figure 2-2. Frames from Video 2-1 depicting the process of protein depletion. The dark areas on the array are the magnetic beads.
on-chip fluorescence analysis. Figure 2-3A shows the trend observed for the relative fluorescence intensity as a function of increased contact time with the Protein A/G beads. The fluorescence intensity initially decreases rapidly but gradually stabilizes after 5 min of exposure to the superparamagnetic particles. Following 9 min of exposure, the fluorescence intensity is reduced by >95%, and further depletion time did not result in a substantial fluorescence intensity reduction. As a result, we established a conservative mixing/contact time of 10 min for subsequent depletion experiments. The second assay was developed to evaluate the depletion efficacy for the two most prevalent HAPs in human serum: IgG (again monitored as FITC-IgG) and HSA (monitored as FITC-HSA), with a mixture of particles bearing Protein A/G (for IgG) and anti-HSA (for HSA). In practice (as described below), the two proteins can be depleted simultaneously, but for this assay, because the same fluorophore was used for both analytes, they were probed sequentially. As shown in Figure 2-3B, the relative fluorescence intensity decreased dramatically after a single depletion, with >95% reduction following 10 min of contact. A second depletion step was then studied to explore whether the fluorescent intensity could be reduced further. A second aliquot of functionalized superparamagnetic particles was actuated to the center of the chip and mixed with the sample for an additional 10 min. The second depletion resulted in a further reduction of fluorescence intensity, 98% relative to the initial fluorescence intensities. In the future, if different concentrations of proteins or densities of particles are used, it may be useful to evaluate additional (sequential) depletions. Regardless, the depletion efficiencies shown in Figure 2-3 are similar to those of commercially available extraction methods which remove ≥98% of the high abundance proteins.324
Figure 2-3. On-chip depletion kinetics and efficacy. (A) Graph of mean relative FITC-IgG fluorescence intensity (normalized to $t = 0$) as a function of mixing time using Protein A/G magnetic beads ($n = 3$). After approximately 9 min, the beads have depleted the IgG level by 95%. Error bars represent 1 standard deviation about the mean. (B) Graph of mean relative fluorescence intensity (normalized to control) of FITC-IgG (solid, $n = 4$) and FITC-HSA (hatched, $n = 3$) as a function of one or two 10 min depletion step(s). The magnitude of the blank measurements was multiplied by 100 to illustrate the low background signal of on-chip fluorescent measurements. Error bars represent 1 standard deviation about the mean.
2.4.3 MALDI-MS Analysis of DMF-Based Protein Depletion

Fluorescence measurements provide a quantitative assessment of the amount of a high abundance protein that is removed following a single and double depletion steps, but little information is obtained regarding the specificity of the superparamagnetic particle-based depletion process. Off-chip MALDI-MS analysis was used to evaluate the specificity and detection enhancement afforded through superparamagnetic particle/DMF-based protein depletion. MALDI-MS has been used as a semiquantitative profiling tool for proteomic samples.\textsuperscript{325} A mixture of HSA (0.5 mg/mL), IgG (2 mg/mL), and hemopexin (0.1 mg/mL) was used to represent a protein mixture composed of HAPs (HSA and IgG) and low-abundance proteins (hemopexin). Note that because of reagent availability and solubility issues, this mixture is not an identical match to the concentrations of these proteins in serum, but it does reflect the correct ratio of IgG to hemopexin (in serum, IgG is typically $\sim10$--$20\times$ more concentrated than hemopexin). Figure 2-4 shows three representative MALDI-MS spectra for the protein mixture treated with (A) control (no depletion), (B) a single sample depletion step, and (C) two depletion steps. Four spectra were collected for each condition, and root-mean square noise (NRMS) values were estimated from spectral regions without prominent peaks to determine signal-to-noise (S/N) ratios (Table 2-1: Comparison of S/N Ratios for Ion Intensities in MALDI-MS Spectra for the Control and Following a Single and Double Depletion, with DMF/Magnetic Bead Platform). Initially, the low-abundant protein, hemopexin, produces very low relative signal intensity compared to the highly abundant species (IgG and HSA) in the protein mixture, prior to depletion (Figure 2-4A). The HSA is the most intense signal peak at approximately 152 (S/N = 92.0) while the singly and doubly charged intensities for IgG are 22 (S/N = 23.9) and 35 (S/N = 22.0), respectively. Conversely, the MALDI signal for hemopexin is quite low at approximately 13 (S/N = 14.7). The low signal in the protein sample prior to extraction presumably results from charge competition/ion suppression due to the presence of the HAPs.\textsuperscript{326} The mixed protein sample was depleted using the DMF bead-based protocol (vide supra), and following a single depletion step, the hemopexin MALDI signal is increased 6.7 times to 87 (S/N = 38.9). Conversely, the HSA signal is decreased by 5.2 times to 34 (S/N = 17.6) and now ranks as the fourth most intense peak behind IgG with intensities of 95 (S/N = 40.9) and 67 (S/N = 28.7), for +1 and +2 charge states, respectively (Figure 2-4B). Interestingly, the first depletion produces an increase in both the S/N ratio for the IgG and hemopexin, resulting from reduced ion suppression.
from the simplified matrix. Following a double depletion (Figure 2-4C), the MALDI signals for HSA and IgG are diminished to 5.5 (S/N = 11.1) for HSA, which corresponds to a 6.2 times reduction compared to the first depletion and a total reduction of 27.6 compared to the original sample, and IgG where the signal is reduced to 16 (S/N = 17.5) and 8.1 (S/N = 11.1) for the singly and doubly charged ions, respectively, with an overall average reduction of 2.5 for both IgG ions. Conversely, the hemopexin signal is now the most intense signal at 76 (S/N = 62.3) in the MALDI-MS spectrum with a signal enhancement of 5.6 and an improvement in signal-to-noise ratio of 4.2. Detailed results of S/N for each replicate are tabulated in Table 2-1. A significant enhancement for hemopexin is observed when comparing the signal-to-noise (S/N) ratios of the peaks in the protein mixture before depletion and following a single and double depletion. Furthermore, the signal-to-noise ratios also point to the necessity of conducting a second depletion step as a significant protein concentration remains after the first step to limit the signal-to-noise enhancement. Similarly following protein depletion, a sample can be subjected to ESI-MS/MS.
Figure 2-4. MALDI spectra of sample comprising HSA (0.5 mg/mL), IgG (2 mg/mL), and hemopexin (0.1 mg/mL). (A) Before depletion, (B) after single depletion, and (C) after double depletion.
Table 2-1: Comparison of S/N Ratios for Ion Intensities in MALDI-MS Spectra for the Control and Following a Single and Double Depletion, with DMF/Magnetic Bead Platform

<table>
<thead>
<tr>
<th>analyte and S/N ratio</th>
<th>Hemopexin (M+H)+</th>
<th>HSA (M+H)+</th>
<th>IgG (M + 2H)2+</th>
<th>IgG (M+H)+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (no depletion)</td>
<td>14.65</td>
<td>92.04</td>
<td>21.98</td>
<td>23.35</td>
</tr>
<tr>
<td>Single depletion 1</td>
<td>16.71</td>
<td>16.32</td>
<td>29.85</td>
<td>42.58</td>
</tr>
<tr>
<td>Single depletion 2</td>
<td>38.81</td>
<td>17.61</td>
<td>28.66</td>
<td>40.90</td>
</tr>
<tr>
<td>Single depletion 3</td>
<td>13.53</td>
<td>12.73</td>
<td>25.07</td>
<td>25.87</td>
</tr>
<tr>
<td>Single depletion 4</td>
<td>22.21</td>
<td>15.78</td>
<td>23.81</td>
<td>28.63</td>
</tr>
<tr>
<td>Mean single depletion</td>
<td>22.81</td>
<td>15.61</td>
<td>26.85</td>
<td>34.49</td>
</tr>
<tr>
<td>σ single depletion</td>
<td>11.25</td>
<td>2.07</td>
<td>2.87</td>
<td>8.47</td>
</tr>
<tr>
<td>Double depletion 1</td>
<td>21.58</td>
<td>7.37</td>
<td>11.05</td>
<td>12.10</td>
</tr>
<tr>
<td>Double depletion 2</td>
<td>40.72</td>
<td>10.66</td>
<td>14.54</td>
<td>17.94</td>
</tr>
<tr>
<td>Double depletion 3</td>
<td>31.41</td>
<td>8.12</td>
<td>11.91</td>
<td>14.08</td>
</tr>
<tr>
<td>Double depletion 4</td>
<td>62.25</td>
<td>8.14</td>
<td>11.05</td>
<td>17.45</td>
</tr>
<tr>
<td>Mean double depletion</td>
<td>38.99</td>
<td>8.57</td>
<td>12.14</td>
<td>15.39</td>
</tr>
<tr>
<td>σ double depletion</td>
<td>17.37</td>
<td>1.44</td>
<td>1.65</td>
<td>2.78</td>
</tr>
</tbody>
</table>

2.5 Conclusion

We have demonstrated DMF separation of high-abundance and low-abundance proteins with antibody and bioaffinity protein immobilized magnetic beads for a human plasma protein depletion application. Using this new method, protein depletion was successfully developed on DMF and is a powerful tool for rapid, efficient, and automated sample processing by achieving >95% depletion efficiency in as little as 10 min for multiple samples simultaneously (up to four on the current device). Current commercially available depletion methods for depletion of a single protein on columns require between 20 and 60 min for completion for a single sample. We
eliminate the need for lengthy depletion protocols, high levels of sample dilution, or both. We propose that the new technique has great potential for biomarker identification.
Chapter 3
Digital Microfluidics for Immunoprecipitation

3.1 Abstract

Immunoprecipitation (IP) is a common method for isolating a targeted protein from a complex sample such as blood, serum, or cell lysate. In particular, IP is often used as the primary means of target purification for the analysis by mass spectrometry of novel biologically derived pharmaceuticals, with particular utility for the identification of molecules bound to a protein target. Unfortunately, IP is a labor-intensive technique, is difficult to perform in parallel, and has limited options for automation. Furthermore, the technique is typically limited to large sample-volumes, making the application of IP-cleanup to precious samples nearly impossible. In recognition of these challenges, we introduce a method for performing microscale IP using magnetic particles and digital microfluidics (DMF-IP). The new method allows for 80% recovery of model proteins from ~μL volumes of serum in a sample-to-answer run-time of approximately 25 minutes. Uniquely, analytes are eluted from these small samples in a format compatible with direct analysis by mass spectrometry. To extend the technique to be useful for large samples, we also developed a macro-to-microscale interface called pre-concentration using liquid intake by paper (P-CLIP). This technique allows for efficient analysis of samples >100x larger than are typically processed on microfluidic devices. As described herein, DMF-IP and P-CLIP-DMF-IP are rapid, automated, and multiplexed methods that have the potential to reduce the time and effort required for IP sample preparations with applications in the fields of pharmacy, biomarker discovery and protein biology.

3.2 Introduction

Immunoprecipitation (IP) is a widely used technique in which target molecules in complex samples are captured by antibodies (or other capture-molecules like Protein A or G) bound to insoluble supports. The target molecules are then eluted from the supports without background contaminants, greatly simplifying subsequent quantitation and analysis for applications including medical diagnostics and forensic sciences. Importantly, in contrast to other kinds of pre-concentration/sample cleanup techniques, IP elution conditions are designed to be gentle, such that target molecules maintain their structure and functionality throughout the process. This
allows for capture, cleanup, and recovery while maintaining associated target-binding partners – e.g. an enzyme bound to an inhibitor can be captured on the insoluble support and the complex can be eluted intact.\(^{88,329,330}\) IP has become increasingly prevalent as the pharmaceutical industry has pivoted to focus on the development of biologically derived drugs.\(^{331–333}\) These applications typically involve complex matrices like blood and serum\(^{334}\) or cell lysate,\(^{335}\) and IP can effectively remove the background matrix, preparing the analytes for the analytical method of choice for this application, mass spectrometry (MS).

In its most common format, immunoprecipitation is implemented using magnetic particles with covalently attached antibodies. The process is typically tedious, requiring that a magnet be held adjacent to a vial containing the sample, such that the particles become immobilized, allowing supernatant to be removed and exchanged. Magnetic separation racks (e.g., the MagnaRack from ThermoFisher\(^{336}\)) alleviate some of the tedium, and robotic systems (e.g., the Abnova Precipitor\(^{337}\)) can automate some of the steps in settings for which cost is not a concern. But all of these options are designed to operate with large volumes – typically 100 \(\mu\)L of sample or more. This requirement precludes the use of precious samples such as pin-pricks of blood or core-needle biopsies.

The challenge described above has driven interest in miniaturizing IP using microfluidic devices relying on enclosed microchannels. This approach has been used in the analysis of peptides,\(^{104}\) proteins,\(^{338}\) and DNA\(^{339–342}\) (the latter via chromatin immunoprecipitation; CHIP). While these studies represent important steps forward, the methods do not lend themselves to straightforward integration for analysis by mass spectrometry, as they are limited by the complexities of integrating membranes and sonicators within devices,\(^{104,340}\) MS-incompatible elution media,\(^{338}\) or the necessity of protein elution off-chip into microcentrifuge tubes.\(^{341,342}\) Furthermore, most of these studies also rely on external mechanical pumps which can be cumbersome to operate and automate for multiplexed analysis.

In an effort to address the challenges described above, we introduce a technique for performing IP using digital microfluidics (DMF). DMF is a microscale fluid handling technique that operates by manipulating nL-\(\mu\)L fluid droplets on an open array of insulated electrodes, by applying a programmed series of electrical potentials.\(^{15}\) All of the operations are automated and computer-
controlled, which greatly increases the reproducibility and allows parallel handling of multiple samples simultaneously. Many different types of fluids are compatible with DMF, including protein-rich biological samples and polar organic solvents, meaning that DMF is suitable for the reagents used in IP. Most importantly, DMF is well suited for handling suspended magnetic particles for applications such as heterogeneous immunoassays, oligonucleotide binding assays, and high abundant protein depletion (Chapter 2). Finally, because of the open nature of the DMF array, there is no risk of clogging (a common problem for applications involving precipitation in microchannels).

Here we describe a method for immunoprecipitation by digital microfluidics (DMF-IP) for sample clean-up and targeted protein pre-concentration. We demonstrate that DMF-IP is effective for purifying model protein analytes found in undiluted serum such that the analytes can be quantified by direct-elution mass spectrometry. Furthermore, to extend the technique to be useful for large samples, we developed a macro-to-microscale interface called pre-concentration using liquid intake by paper (P-CLIP). To our knowledge, this report is the first to describe the integration of DMF with immunoprecipitation, and we propose that the new hyphenated technique will be useful for a wide range of applications.

### 3.3 Experimental Section

#### 3.3.1 Reagents and Materials

Unless otherwise specified, reagents were purchased from Sigma-Aldrich (Oakville, ON), and deionized water (DI water) with a resistivity >18 MΩ-cm was used to prepare all aqueous solutions. An aqueous stock suspension of superparamagnetic particles with an anti-human serum albumin antibody coating (LSKMAGL02) was purchased from Millipore (Billerica, MA). Human serum albumin (HSA, molecular weight based on amino acid composition of 66,437 Da) solutions were prepared from lyophilized solids. Fluorescein-isothiocyanate labeled human serum albumin (FITC-HSA) was purchased from GenScript USA Inc. (Piscataway, NJ). All solutions used in DMF devices contained 0.1% Tetronic 90R4 (BASF Corp., Germany) surfactant additive.
3.3.2 Device Operation and Fabrication

Digital microfluidic devices comprised two plates and were assembled (as detailed in Chapter 2) at the University of Toronto Nanofabrication Centre (TNFC). The bottom plates comprised a 15 x 4 array of square driving electrodes (2.2 x 2.2 mm each) and 12 large (16.4 x 6.7 mm) reservoir electrodes. These bottom plates were coated with ~7 μm Parylene-C (Specialty Coating Systems, Indianapolis, IN) by vapor deposition before being spin coated with ~200 nm Teflon-AF (DuPont, Wilmington, DE) and baked at 160°C for 10 minutes. Top plates were 0.7 mm thick indium tin oxide coated glass substrates (Delta Technologies, Ltd., Loveland, CO). The top plates were also coated with ~200 nm of Teflon-AF but the 10-minute baking was performed at 235°C. Top plates were attached to bottom plates using double sided tape spacers with an inter-plate distance of ~180 μm. This results in a unit droplet volume on a single actuation electrode of approximately 1 μL; larger volumes were handled by simultaneously actuating multiple electrodes.

DMF experiments were implemented using an automation system (described in detail elsewhere\textsuperscript{34}) that features a magnet mounted on a step-motor which can be moved to enable separation of magnetic particles on the device. The automation system (including the magnet) was programmed and managed by Microdrop, an open-source application for the manipulation of droplets on DMF devices.\textsuperscript{16} Droplets were driven by applying 100-120 V\textsubscript{RMS} sine waves at 10 kHz between the top plate counter-electrode and successive driving electrodes on the bottom plate. All reagents were loaded onto the device from the edge of the top plate by actuating reservoir electrodes underneath the reagents. Extended droplet mixing was performed as described previously\textsuperscript{34} to ensure rapid dispersion of magnetic particles.

3.3.3 DMF-IP General Procedure

Immunoprecipitation was performed in eight steps as depicted in Figure 3-1 with minor variations for specific applications (outlined in the following sections). (1) A suspension of magnetic particles was loaded to a reservoir and then the entire volume was driven onto the device (volume and density dependent on application). (2) The magnet was engaged and the particles were immobilized to the device surface as the droplet was driven to a waste reservoir. (3) A 4 μL droplet of sample was loaded into a reservoir and driven onto the immobilized
particles, the magnet was disengaged and the droplet was mixed for 10 min to re-suspend the particles. (4) Step (2) was repeated. (5) A 4 µL droplet of tris-HCl buffer (pH 7.6, tris-buffer) was loaded into a reservoir and driven onto the immobilized particles, the magnet was disengaged, and the droplet was mixed for ~3 minutes, followed by a repeat of step (2). (6) Step (5) was repeated. (7) A 4 µL droplet of 100 mM citric acid was loaded onto a reservoir and driven onto the particles, the magnet was disengaged, and the droplet was mixed for 5 minutes (unless otherwise noted). (8) The magnet was engaged to separate the particles for a final time, and the resulting droplet was collected for analysis by disassembling the device.
Figure 3-1. Immunoprecipitation by digital microfluidics (DMF-IP). Left: graphical representation of the eight-step DMF-IP procedure, focused on the surface of a single particle. Centre: graphical representation of the eight-step procedure, focused on the droplets and particles on the DMF device. Right: photographs of DMF device surface during DMF-IP at steps 3) and 6). Food colouring dyes were added to improve droplet visibility. Steps include: 1) Antibody-modified particles are loaded onto the device. 2) Particles are isolated. 3) Particles are exposed to sample. 4) Particles are isolated. 5) Particles are washed. 6) Particles are isolated. 7) Analytes are eluted. 8) Eluted analytes are recovered for analysis.
3.3.4 DMF-IP Elution Efficiency

Elution efficiencies were evaluated using a sample comprising 200 µg/mL FITC-HSA in 80% fetal bovine serum (FBS) and 20% DI water. The protocol followed the eight steps outlined above, but with the following modifications. The stock suspension of magnetic particles was diluted 1/4 in DI water and was loaded as a 4 µL aliquot in step (1). Step (7) was performed with mixing durations of: 1 minute, 2 minutes, 5 minutes and 10 minutes, and all of the mixing steps were performed in the dark. After completing step (8), 1 µL of the eluted droplet was collected and diluted to 100 µL in tris-buffer and vortexed briefly. Each diluted sample was loaded into a 96 well plate and the fluorescence intensity was measured with a plate reader (PHERAstar microplate reader, BMG Labtech, Ortenberg, Germany). Measurements were performed in “endpoint” mode (COSTAR 96 well plate geometry) with 10 flashes per well and a gain of 701 using 485 nm excitation and 520 nm emission filters. The FITC-HSA concentration in each eluted droplet was determined by comparing the measured fluorescence intensity to a calibration curve of standards (0, 10, 50, 100 and 200 µg/mL FITC-HSA in tris-buffer) evaluated using the same conditions. Each calculated eluted concentration was referenced to the initial sample concentration to determine the elution efficiency for each elution time. Each condition was evaluated in triplicate using the same DMF device.

The elution efficiency for multiple elution steps was evaluated as described above but for a single mixing duration: 5 minutes. In these experiments, after step (8), the device was reassembled and steps (7) and (8) were repeated using a fresh aliquot of citric acid. This process was repeated twice more to collect droplets from one, two, three and four elution steps. The collected samples were then measured and the concentration determined and compared to that of the initial droplet as above. Each condition was evaluated in triplicate using the same device.

3.3.5 Sample Clean-Up by DMF-IP

Immunoprecipitation was performed on FBS containing 1 mg/mL HSA using the eight-step protocol described above, with the following modifications. The stock suspension of magnetic particles was diluted 1/2 with DI water and then was loaded as a 5 µL aliquot in step (1). After step (8), 3 µL of the eluted droplet was collected and diluted to 5 µL with DI water and then
processed for MS analysis as described below. Each condition was evaluated in triplicate using the same device.

### 3.3.6 DMF-IP with Pre-Concentration using Liquid Intake by Paper (P-CLIP)

DMF devices were formed as above, except triangular pieces cut from a KimWipe were placed onto some reservoirs such that a point penetrated into the array of actuation electrodes. Another piece of KimWipe was positioned adjacent to each triangle, such that they presented a contiguous network of paper. Samples of FBS containing 30 µg/mL HSA were prepared and treated by two methods. In the first method (with no pre-concentration), the eight-step DMF-IP procedure described above (with a 4 µL aliquot of sample) was performed, using a 1/2 dilution of the stock suspension of magnetic particles in DI water (loaded as a 5 µL aliquot) in step (1). In the second method (powered by P-CLIP), a 100 µL aliquot of sample was combined off-chip with 2.5 µL of the stock suspension of magnetic particles in a microcentrifuge tube, which was then mixed for 1 hour using a rotating vial rack. The 102.5 µL sample was then loaded onto a reservoir on a DMF device (adjacent to the edge of the top plate) where a modified DMF-IP procedure was performed in eight steps. (1A) A “tongue” of sample was pulled from the sample-reservoir to the center of the device and the magnet was engaged. (2A) The tongue was then driven to the opposite side of the device, such that it encountered the KimWipe. (3A) The wicking action of the KimWipe pulled the volume of sample across the device. (4A) The remaining sample was passively drawn into the wipe, leaving the magnetic particles immobilized on the center of the device. During this step the electrodes remained actuated, defining the path for the sample to flow across the device. The electrodes were sequentially turned off (starting with the electrode adjacent to the source-reservoir) as the final volume of solvent was pulled into the waste-reservoir. (5A) – (8A) were identical to steps (5) – (8) above. After step (8A), 3 µL of the eluted droplet was collected and diluted to 5 µL with DI water and then processed for MS analysis as described below. Each condition was evaluated in triplicate using the same device.

### 3.3.7 HSA Conformation Experiments

Human serum albumin (HSA) conformation was monitored using UV fluorescence using techniques similar to those described previously. Briefly, three different buffers were
prepared: (1) 25 mM tris-HCl with 750 mM NaCl (2) 100 mM citric acid with 700 mM NaCl, and (3) 100 mM citric acid. The pH of the buffers were measured to be 7.5, 1.7, and 1.7, respectively, using an AR 50 pH meter (Fisher Scientific, Waltham, MA, USA) equipped with a Beckman Coulter (Brea, CA, USA) pH electrode. HSA was then added to each buffer to a final concentration of 9 μM, and the solutions were allowed to incubate for 1 h at room temperature. Solution (3) was subsequently neutralized by adding Na3PO4 to 200 mM and then allowed to incubate for an additional hour at room temperature (the pH of the neutralized solution was measured to be 7.1). Blank solutions not containing HSA were also prepared for each condition.

Fluorescence intensities were monitored on a Perkin Elmer LS50B Luminescence spectrometer using an excitation wavelength of $\lambda_{ex} = 295$ nm (5.0 nm slit width) and an emission scan range of $\lambda_{em} = 320-360$ nm (7.5 nm slit width, 250 nm/min scan rate) in a quartz cuvette with path length of 1 cm (Precision Cells, Inc., New York). Each HSA-containing solution was evaluated in triplicate with 10 scans averaged for each replicate; blank measurements (not containing HSA) were subtracted from each scan. The data were plotted as a function of emission wavelength to determine the wavelength of maximum intensity $\lambda_{em,max}$.

### 3.3.8 Gel Electrophoresis for sample purity

A 1 mg/mL solution of HSA in FBS was prepared and subjected to the eight-step DMF-IP as described in the main text. Three fractions were collected (by pipette) for analysis: fraction 1 is the product-droplet generated after step 8, fraction 2 is the pooled supernatant droplets collected after step 5, and fraction 3 is the supernatant droplet generated after step 3. (Note that fractions 2 and 3 are typically driven to a waste reservoir and discarded, but in these experiments, they were collected for analysis.) The fractions were analyzed by SDS-PAGE using a 10% polyacrylamide gel with beta-mercaptoethanol as a reducing agent. The gel was stained with Coomassie Blue G-250 (Bio-Rad, Hercules CA., USA) as per manufacturer’s instructions. The stained gel was visualized using a Gel Doc EZ Gel Documentation System (Bio-Rad).

### 3.3.9 Mass Spectrometry

Samples collected from DMF devices (as well as 5 μL control samples including (a) 1 mg/mL HSA in tris-buffer containing 0.1% Tetronic 90R4, or (b) 1 mg/mL HSA in FBS containing 0.1%
Tetronic 90R4) were diluted to 45 μL in a centrifuge tube with an aqueous solution of 9 M urea, the mixture was vortexed briefly, and then allowed to incubate for 5 minutes at room temperature. Then 2.25 μL of aqueous 100 mM tris(2-carboxyethyl)phospine (TCEP) was added, the mixture was vortexed briefly, and then incubated for 30 minutes at room temperature. 4.7 μL of aqueous 100 mM chloroacetamide was added, the sample was vortexed briefly and incubated for 30 minutes at room temperature in the dark. The sample was diluted to 500 μL in aqueous 25 mM ammonium bicarbonate and then an aliquot of aqueous trypsin was added (containing 1.25 μg trypsin for samples that originally contained 1 mg/mL HSA, or 200 ng trypsin for samples that originally contained 30 μg/mL HSA). The mixture was incubated at 37°C for 4 hours, and then quenched by adding 7 μL of 5% formic acid. All samples were frozen at -80°C until analysis.

Before analysis, each sample was thawed and subjected to solid-phase extraction using C18 solid phase extraction pipette tips (Pierce™ C18 Tips, 100 μL bed, #87784) as per the manufacturer’s instructions. In brief, the tips were wetted in 50% acetonitrile (ACN) in DI water and then equilibrated in 0.1% trifluoroacetic acid (TFA) in DI water. The entire sample (above) was then iteratively aspirated into and dispensed from the tip for 10 cycles. The tips were then rinsed twice with 100 μL 0.1% TFA, 5% ACN in DI water. Samples were then eluted in 100 μL 0.1% acetic acid, 60% ACN in DI water. Finally, an internal standard of angiotensin I was added (final concentration 5 μg/mL) and the sample was evaluated by direct injection mass spectrometry at a flow rate of 5 μL/min.

Samples were analyzed using a SCIEX API 4000 (SCIEX, Concord, ON) triple quadrupole mass spectrometer operating in tandem (MS/MS) mode. Samples were introduced via a small-volume, 25 μm internal diameter electrospray ionization interface (SCIEX). The source temperature was 150°C with an electrospray voltage of 5000 V. The collision gas, curtain gas and sheath gas (GS1, GS2) pressures were 4, 10, 15 and 15 psi respectively. Quadrupoles Q1 and Q3 were set at unit resolution. For HSA analysis the precursor tryptic peptide fragment LVNEVTEFAK^2+ was isolated at m/z 575.3 using a declustering potential of 101 V and fragmented at a collision energy of 23 eV and a collision cell exit potential of 8 V. The HSA peptide product ion (NEVTEFAK^+) was isolated at m/z 937.6. The internal standard (I.S.) precursor ion (human angiotensin I, DRVYIHPFHL^3+) was isolated at m/z 433 using a declustering potential of 91 V and fragmented.
with a collision energy of 29 eV and a collision cell exit potential of 10 V. The I.S. product ion (histidine immonium ion fragment) was monitored at m/z 110. The dwell time for both species was 150 ms and 12 scans were collected for each ion in both MS2 scanning (MS2 m/z range of 900-980 for the HSA tryptic fragment and 100-200 for the I.S.) and multiple reaction monitoring (MRM) modes. Peaks were viewed and measured using SCIEX PeakView® v 1.2.2.0 and SCIEX Analyst® v 1.6.2 software. For quantitative analysis of replicate samples, results were expressed as the ratio of peak heights of the analyte product ion to that of the I.S.

3.3.10 P-CLIP-DMF-IP with HPLC-MS

50 ng/mL HSA in FBS samples were prepared and subjected to the P-CLIP-DMF-IP procedure, and then reduced, alkylated, digested and extracted as described in the main text. The samples were then diluted 1:1 in water containing 0.1% formic acid and analyzed by a Shimadzu (Kyoto, Japan) 20AC HPLC operating in-line (via electrospray ionization) with a SCIEX 6500+ QTRAP mass spectrometer (Concord, ON). The HPLC column was a Phenomenex (Torrance, CA) Kinetex (2.1 mm x 100 mm, packed with 2.6 μm dia. particles with 100 Å-thick C18 coating), operated in gradient elution mode at 40°C. The injection volume was 5 μL and the mobile phase was a mixture of solution A (98% water, 2% acetonitrile each with 0.1% formic acid) and B (2% water, 98% acetonitrile each with 0.1% formic acid). The HPLC gradient was a total of 30 minutes with an initial concentration of solution B at 5%, with a linear ramp up to 27% after 15 minutes followed by a second linear ramp up to 97% at 18 minutes where it was held for 5 minutes before linearly ramping down to 5% at 25 minutes where it was held until the end of the run, all at 300 μL/min. The mass spectrometer was operated in MRM mode monitoring the same transitions listed in the main text but with different operating parameters: the source temperature was 400°C and the electrospray voltage was 5500 V. The collision, curtain and sheath gases (GS1 and GS2) were 8, 20, 30 and 40 psi respectively. The collision energy, cell exit potential and declustering potential for the HSA peptide were 25 eV, 14 V and 90 V respectively and for angiotensin I (internal standard) they were 30 eV, 15 V and 90 V. Chromatograms were treated with Gaussian smoothing (0.5-point width) using PeakView 1.2.2 (SCIEX).
3.4 Results and Discussion

3.4.1 Digital Microfluidic Immunoprecipitation (DMF-IP)

IP is a powerful technique for working with biological samples, particularly for purifying a specific target while maintaining the target’s conformation and inter-molecular associations, but it can be time-consuming and tedious and uses large volumes of samples. The overall goal of this work was to address the challenges of automating and miniaturizing IP using digital microfluidics and antibody-functionalized magnetic particles.

As shown in Figure 3-1, an eight-step procedure was developed to facilitate DMF-IP on sample volumes ranging from 3-5 µL. The left column shows a graphical representation of the surface of a magnetic particle during each step, and the central column shows a graphical representation of what is occurring on the DMF device. In step (1), magnetic particles are loaded into reservoirs and then driven onto the electrodes. In step (2), the magnet underneath the device is engaged to immobilize the particles on the device surface, while the suspending liquid is removed. In step (3) (depicted in the upper inset photograph), a sample (red) containing the target analyte (black pentagons) and other solutes (purple triangles and orange hexagons) is moved to the immobilized particles, the magnet is disengaged and the particles are actively re-suspended. In step (4), the particles are isolated from the sample via magnetic separation, and in steps (5-6) (depicted in the lower inset photograph), the particles are washed in washing buffer (green) twice. Then, in step (7) the particles are exposed to elution buffer (blue), and finally, in step (8) the particles are removed and the purified target is recovered. In practice, the device was capable of processing four samples simultaneously. The only manual steps required were the loading of sample and reagents and the removal of waste.

Human serum albumin (HSA) was chosen as a model analyte, in recognition of its importance as a target for monitoring the metabolism of drugs or environmental toxins in pharmaceutical or clinical settings.\(^{351}\) (For example, the presence of a mixed disulfide HSA adduct associated with environmental toxicants is often used to diagnose retardation of uterine growth in pregnant mothers.\(^{352}\)) For all results described here, the sample comprised HSA dissolved (at various concentrations) in bovine serum.
A number of different elution conditions were evaluated, eventually settling on an elution buffer of 100 mM citric acid as this buffer was found to not (permanently) alter the conformation of the analyte (a key trait of good IP methodology). HSA has a “normal” conformation $N$ at neutral pH, but adopts “extended” conformation $E$ at pH $<3$.\textsuperscript{353} The native fluorescence of HSA (originating from a single tyrosine residue) reflects this change – the $\lambda_{em,max}$ is blue-shifted for $E$ relative to $N$.\textsuperscript{350} This phenomenon was used here to evaluate the effects of 100 mM citric acid on HSA conformation. Three HSA solutions were prepared (with approximately equivalent ionic strength): (1) HSA in neutral buffer (pH 7.5), (2) HSA in citric acid (pH 1.7), and (3) HSA in citric acid (pH 1.7) which was incubated for an hour and then subsequently brought to neutral pH (7.1). Fluorescence spectra of each sample are shown in Figure 3-2. As expected, reducing the pH from neutral (1) to acidic (2) resulted in a reduction in emission-maximum from $\lambda_{em,max}(1) = 346 \pm 2$ nm to $\lambda_{em,max}(2) = 338 \pm 1$ nm (Figure 3-2). This suggests that the immunoprecipitation-elution buffer (100 mM citric acid) described in the main text causes a change in conformation from $S$ to $E$. But the spectrum of solution (3), which has an emission max of $\lambda_{em,max}(3) = 345 \pm 1$ nm, suggests that the conformation change is not permanent. When evaluated in replicates, these data were statistically indistinguishable from those of solution (1) (Student’s $t$-test, $a = 0.05$, $p = 0.840$).
Figure 3-2. Effects of immunoprecipitation-elution buffer (100 mM citric acid) on protein conformation. Fluorescence spectra of 9 μM HSA in (1) neutral buffer (pH 7.5, blue), in (2) 100 mM citric acid (pH 1.7, orange), and in (3) 100 mM citric acid that was subsequently neutralized (pH 7.1, gray). The wavelengths of maximum emission are $\lambda_{em,\text{max}(1)} = 346 \pm 2$ nm, $\lambda_{em,\text{max}(2)} = 338 \pm 1$ nm, and $\lambda_{em,\text{max}(3)} = 345 \pm 1$ nm, respectively.

An assay was developed relying on fluorescently labeled HSA (FITC-HSA) to determine (a) the optimum elution step duration, and (b) the optimum number of elution steps required to maximize the recovered protein. For the former, it was determined that a one-minute elution step facilitates recovery of $72 \pm 1$ % (average ± 1 std. dev.) of the analyte (Figure 3-3). Increasing the elution step duration to 2 min, 5 min and 10 min yielded recoveries of $80 \pm 8$ %, $80 \pm 7$ % and $72 \pm 2$ % with an overall mean recovery of $76 \pm 5$% across all time points. The differences between these recoveries were not significant (Single factor ANOVA, $\alpha = 0.05$, $p = 0.170$); thus, we
arbitrarily chose to use 5-minute elution steps for the remainder of the work described here. For the latter experiment, as shown in Figure 3-3, a single 5-minute elution step recovered $80 \pm 7\%$ of analyte, a second step recovered an additional $3 \pm 1\%$. Third and fourth treatments did not yield detectable amounts of analyte.
Figure 3-3. Optimization of DMF-IP elution efficiency. Samples comprised 4 µL aliquots of 200 µg/mL FITC-HSA in 80% fetal bovine serum. (A) Bar graph of elution efficiencies as a function of elution step duration. (B) Bar graph of elution efficiency as a function of the number of 5-minute elution steps. Error bars represent ± 1 standard deviation.

In practice, it was decided that a single 5-minute elution step (with 80% analyte recovery) was sufficient for the work described here, which resulted in a run-time (from sample loading to
analyte recovery) of 25 min. The purity of the analyte eluted under these conditions was also assessed by protein gel electrophoresis relative to the wash-droplets and depleted serum after IP. As shown in Figure 3-4, the serum sample contains a large number of bands representing different proteins, while the DMF-IP sample shows a single band at ~66 kD, suggesting that the sample has been substantially purified. Note that this test does not distinguish between HSA and bovine serum albumin (in the matrix), which co-elute. Note that 80% recovery is consistent with or higher than the recoveries reported for conventional\textsuperscript{354,355} and microfluidic\textsuperscript{104,338,341,342} IP protocols, with the caveat that recoveries depend on antibody-antigen binding kinetics (making comparisons between different systems problematic). Most importantly, the ability of the new method reported here to operate on 3-5 μL volumes in a format convenient for mass spectrometry should potentially allow seamless integration with analysis of precious samples such as dried blood spots\textsuperscript{51,52} or core-needle biopsies.\textsuperscript{28,319}

\textbf{Figure 3-4.} Image of a Coomasie Blue-stained SDS-PAGE gel of human serum albumin (HSA) under reducing conditions. Each lane contains a sample collected from a DMF-IP device at different stages in the eight-step procedure. Lane 1 (left): fully processed sample collected after completing step (8). Lane 2: combined wash buffer droplets collected after step (4). Lane 3: serum sample supernatant droplet collected after being treated with magnetic particles in step (3). Lane 4 (right): protein mass ladder.
The gold standard analytical technique used for biologically derived drug-analysis is mass spectrometry (MS). To evaluate the compatibility of this technique with MS, fetal bovine serum samples containing 1 mg/mL HSA were analyzed before and after performing DMF-IP. Figure 3-5 shows representative ESI-MS-MS spectra of three types of samples (A) HSA in DI water, (B) HSA in bovine serum (not subjected to DMF-IP) and (C) HSA in bovine serum which was subjected to DMF-IP. In each analysis, samples were processed and MS-MS transitions were monitored from a precursor ion (the tryptic peptide of HSA, LVNEVTEFAK²⁺, m/z 575.3) to all product ions in the m/z range of 900-980. As shown in a representative spectrum of the standard in water (Figure 3-5A), a strong signal from the y8 product-ion (NEVTEFAK⁺) is observed at m/z 937.6 with very little background signal. In contrast, in a representative spectrum of the standard in serum with no IP (Figure 3-5B), there is low signal at the designated transition. Finally, in a representative spectrum of the standard in serum processed by DMF-IP (Figure 3-5C), the y8 product ion signal is high. This highlights the need for IP; even when using an exquisitely selective technique like MS-MS (targeting a known analyte’s transition), IP-pulldown is required for robust detection. Specifically, in this case, when using an internal standard (to compensate for variations in ionization efficiency), the DMF-IP method rescues a signal that is nearly undetectable in the control, improving the signal by a factor of 3.7 (the ratio of HSA product ion intensity/Internal standard product ion intensity for the DMF-IP sample divided by the same ratio for the untreated sample) while maintaining an acceptable level of variance on the DMF-IP recovered samples of 7% relative standard deviation (RSD, n = 4).
Figure 3-5. DMF-IP with analysis by mass spectrometry. Representative tandem mass spectra of a tryptic peptide of HSA ($m/z$ 575.3 → $m/z$ 937.6) prepared from 4 μL samples of 1 mg/mL HSA dissolved in (A) DI water (with no serum), (B) fetal bovine serum without DMF-IP, and (C) fetal bovine serum processed by DMF-IP.
For the data shown in Figure 3-5, the samples were processed by IP on-chip (as per Figure 3-1), but the processing steps of reduction, alkylation, and digestion, as well as solvent-exchange by solid-phase extraction (SPE) were carried out off-chip. In the future, it should be straightforward to integrate the latter steps on-chip as well (as reported previously for DMF-proteomic processing\textsuperscript{36,38,356} and DMF-SPE\textsuperscript{41,50,72}), such that the DMF device generates fully processed analyte (i.e., tryptic peptides in MS-compatible solvent from proteins captured by IP) directly from undiluted serum. Note that the detection method in this work was direct-infusion MS/MS; in the future it should also be possible to integrate the method described here with an automated DMF-HPLC-MS interface (as reported previously\textsuperscript{357}). This would improve the sensitivity of the technique, and would provide a complete sample-to-answer solution appropriate for a wide range of applications in biologic pharmaceutical analysis.\textsuperscript{331–333}

### 3.4.2 DMF-IP with Pre-Concentration using Liquid Intake by Paper

As described above, the capacity to work with small-volume samples is a unique advantage of the technique described here. But this advantage can also be a limitation – when not working with a precious sample, it often is desirable to be able to analyze a large volume, allowing the observation of analytes present at very low concentration. This limitation is a critical drawback for DMF and most other microscale analysis techniques. To overcome this limitation, we developed a new macro-to-microscale interface that integrates pre-concentration with DMF-IP, allowing for the evaluation of much larger sample-sizes.

In developing the new method, we considered a variety of formats. For example, one mechanism to evaluate a 100 \(\mu\)L sample-volume would be to dispense (for example) 25 4-\(\mu\)L droplets of sample onto the electrode array, process them in parallel or in series, and then pool the final eluents together for analysis. This idea was discarded as being unwieldy and slow; we wanted to develop a much simpler and faster method for straightforward operation. Our solution\textsuperscript{358} relies on (a) pre-mixing a large-volume sample (in this case,100 \(\mu\)L) with magnetic particles (off-chip), followed by (b) isolation and resuspension of the particles on-chip in a small-volume droplet (in this case, 4 \(\mu\)L) for further processing. We call the new technique “pre-concentration using liquid intake by paper” (P-CLIP).
The new P-CLIP method is depicted in Figure 3-6. A key feature is the incorporation of lateral-flow pumping$^{359,360}$ to facilitate rapid and passive handling of the large fluid volume. As shown, a large volume of sample (pre-mixed with magnetic particles) is loaded onto the device, and then a portion of the sample is driven across the array while the magnet is engaged (Figure 3-6A). Upon reaching the other side of the device, the sample contacts an immobilized paper pad, which continuously wicks the sample across the device (Figure 3-6B). Because the magnet is engaged, the magnetic particles become immobilized over the top of the magnet on the device surface (Figure 3-6C). Finally, a combination of wicking and DMF actuation removes the remainder of the sample from the magnetic particles (Figure 3-6D), and the standard DMF-IP protocol can then be implemented as normal in small droplet volumes (Figure 3-6E-F). We have demonstrated the compatibility of P-CLIP with immunoassays under various conditions$^{361}$; here we report its application to DMF-IP.
**Figure 3-6.** Pre-concentration of magnetic particles using liquid intake by paper (P-CLIP).

Frames from a movie showing A) initial device setup featuring two 100 μL serum samples (red, bottom) and absorbant pads (top); B) “tongues” of fluid from the samples driven by DMF actuation over the engaged magnet; C) particles collecting onto the device surface as the fluid wicks into the absorbant pad; D) the sample nearly completely absorbed, leaving a small volume (and bolus of particles) behind; E) wash droplets moving toward the particles; F) resuspension of pre-concentrated particles into wash droplets. Food colouring dyes added to improve visibility.

To test the effectiveness P-CLIP for the analysis of dilute analytes, FBS samples were spiked with 30 μg/mL HSA and analyzed using DMF-IP (Figure 3-1) and P-CLIP-DMF-IP (Figure 3-6).
using direct injection mass spectrometry. Representative spectra are shown in Figure 3-7. As shown, a small volume of sample (i.e., with no pre-concentration by P-CLIP) produces a product ion peak that has low signal-to-noise (Figure 3-7A). While this type of sample might be easily detected if evaluated with HPLC-MS/MS and/or with a newer spectrometer, there simply is not enough analyte to produce a usable signal in the system used here. In contrast, when 100 μL of the same sample was concentrated to 4 μL using P-CLIP, the HSA product ion peak has high intensity (Figure 3-7B), representing a signal increase of 12 times (with 14% RSD on P-CLIP samples, n = 4) relative to the measurement without P-CLIP. The increased variance for this technique is likely a function of variations in bead-retrieval efficiencies during the concentration step (a topic of on-going study\textsuperscript{361}). A preliminary investigation combining P-CLIP-DMF-IP with HPLC-MS/MS shows effective detection of HSA at concentrations as low as 50 ng/mL in serum. The retention time for the HSA peptide was found to be 8.8 minutes and, with peak area 1247 (Figure 3-8) relative to peak area in a blank sample of 13. This result suggests that P-CLIP-DMF-IP in combination with HPLC-MS is suitable for quantitation of low-abundance protein targets. Finally, in this work, 100 μL samples were processed two-at-a-time; in the future, slight device modifications should allow for many more pre-concentration operations to be conducted in parallel on a single device. More work is needed, but we propose that P-CLIP is an attractive solution to the long-standing problem of limited sample sizes in DMF.
Figure 3-7. Representative tandem mass spectra of a tryptic peptide of HSA ($m/z$ 575.3 → $m/z$ 937.6) prepared from 30 µg/mL HSA in serum, analyzed from (A) a 4 µL sample processed by DMF-IP with no pre-concentration (as in Figure 3-1), or (B) a 100 µL sample pre-concentrated and processed by P-CLIP-DMF-IP (as in Figure 3-6).
Figure 3-8. Representative HPLC-MS/MS chromatogram of HSA after P-CLIP-DMF-IP. A sample of FBS containing 50 ng/mL HSA was processed using P-CLIP-DMF-IP and then analyzed, monitoring the same peptide transition (575.3 → 937.6) described in the main text. The retention time of HSA was found to be 8.8 minutes (highlighted in blue). Blank samples showed negligible signal intensity at that retention time.

3.5 Conclusion

We have developed the first digital microfluidic method for performing immunoprecipitation. The method was applied to analyzing human serum albumin in serum samples using digital microfluidic manipulation of antibody-functionalized magnetic particles with and without pre-concentration. We believe this technique has great potential for the analysis of proteins in the fields of pharmacy and medicine including elucidation of protein structure and associated molecules.
Chapter 4
Enhancing Signal and Mitigating In-Source Peptide Fragmentation using Controlled Clustering by Gas-Phase Modifiers

4.1 Abstract

In-source fragmentation is a phenomenon where molecular ions are activated and fragment in the as they enter the vacuum region of the instrument, and consequently can complicate the spectrum and the analysis. This phenomenon can be minimized by controlling the environment where ions are sampled from the atmospheric region, but this approach can also have a negative effect on overall ion sensitivity. In this study, we introduce gas-phase modifiers (acetonitrile, acetone, cyclohexane, water, and methanol) to the curtain gas to mitigate in-source fragmentation. These modifiers cluster with incoming ions, increasing the energy barrier to fragmentation and consequently reducing the complexity of mass spectra. The clustering is monitored via differential mobility spectrometry and precursor mass spectrum-scanning. The ion-modifier clusters routinely survive into the mass spectrometer showing that highly charged peptides cluster most strongly with acetonitrile and acetone, and that when they cluster with acetonitrile, they produce a large increase in signal intensity for the most highly charged and fragile ions. This results in a significant reduction, up to 90% with some modifiers, in in-source fragmentation for these fragile highly charged peptides, increasing the overall analytical sensitivity and decreasing the limits of detection by up to 82% depending on the analyte. The new technique has no significant detrimental effect on the peptide mass fingerprinting of a BSA or mAb protein digest, but it does reduce the amount of redundant and data-deficient spectra needed to produce adequate sequence coverage using independent data acquisition methods by ~40%. We propose that this technique could have a great benefit in the fields of proteomics and peptidomics where in-source fragmentation and chemical noise routinely mask targets of biological importance.
4.2 Introduction

The analysis of compounds by mass spectrometry (MS) is a ubiquitous technique across many different fields from the detection of pollutants in environmental samples\textsuperscript{362–364} to the screening of biological samples for disease conditions\textsuperscript{365–367}. Amongst the many methods for MS ionization is analyte introduction using atmospheric pressure electrospray ionization (ESI). ESI is known to be a low-energy, gentle form of ionization; analytes are ionized but not fragmented in contrast to traditional techniques like electron impact, in which incoming compounds are fragmented from their initial state.\textsuperscript{368} But even if ions are formed gently (as in ESI), downstream of ionization, the acceleration of newly formed ions as they move into the vacuum of the mass spectrometer can cause fragmentation of the ions by increasing their internal energy.\textsuperscript{223,369} When this occurs, the phenomenon is frequently referred to as in-source fragmentation or in-source collision induced dissociation (in-source CID). While this phenomenon can be exploited for many applications (including identification of selected molecule classes for further analysis in-line\textsuperscript{370}, production of MS/MS data on single-stage instruments\textsuperscript{371}, generation of molecular structural information\textsuperscript{372}, or toxicology screening\textsuperscript{373}), this process is often undesirable, resulting in compounds not being detected, or in the resulting fragments complicating mass spectra. In especially complex samples like those routinely surveyed for peptidomic or proteomic analyses it has been suggested that in-source fragmentation could account for up to 60\% of all non-tryptic peptides\textsuperscript{374} in standard protein digest samples and affect more than 15\% of metabolites\textsuperscript{222} in a standard yeast cell lysate metabolomics analyses.

Today, reduction of in-source fragmentation is possible by reducing the energetics at the source of fragmentation.\textsuperscript{223} For example, since the majority of fragmentation occurs as ions enter the vacuum region of the instrument as directed by the orifice voltage, reduction of that voltage can limit the fragmentation.\textsuperscript{375} The highly energetic environment is caused by the dramatic changes in pressure as well as various ion accelerators\textsuperscript{376} which aid in focusing and declustering incoming ions. These adjustments may also negatively impact the overall signal intensity, resulting in a balance between preventing fragmentation and achieving the highest possible signal. Careful tuning has been shown to improve signal intensity by more than two times.\textsuperscript{377}
Our approach was inspired by the work of DeMuth et al.\textsuperscript{378} who use organic modifiers to shelter non-covalently bound protein complexes. They show that the introduction of polar organic modifiers to the curtain gas of a nano-ESI source prevents the loss of the heme group of myoglobin and hemoglobin. They propose that clustering of polar species around the analyte creates a delayed desolvation which prevents fragmentation caused by ion acceleration at the entrance of the mass spectrometer. Addition of stabilization reagents in solution also appear to increase ion stability in the field-free region.\textsuperscript{379} We postulate that the same mechanism could be employed on a traditional ESI source to prevent the fragmentation of fragile compounds like highly charged peptides. The clustering of the organic modifier around the analytes necessitates an increase of energy to remove the cluster before fragmentation of the analyte could occur.

In this preliminary study, we use gas-phase polar compounds to shelter fragile peptides in model systems from fragmentation. We monitor the clustering behavior of five different polar modifiers using differential mobility mass spectrometry (DMS) before focusing on the use of acetonitrile as a modifier to a) prevent fragmentation of a number of peptides and b) increase the signal intensity of the highest detectable charge state of these peptides. This results in lower limits of detection for these ions and a simplification of the resulting mass spectra, which we propose could aid in peptide searches in the future, as the removal of in-source fragmented peptides will leave only those peptides of biological or proteomic relevance behind.

4.3 Experimental Section

4.3.1 Materials

Peptides human angiotensin I (DRVYIHPFHL, AngI), human angiotensin III (RVYIHPF, AngIII) bradykinin (RPPGFSPFR), and the proteins bovine serum albumin and trypsin were purchased from Sigma Aldrich (Oakville, ON, Canada) and used without further purification. Peptide KGAILKGAILR (referred to as “KGAIL”) was purchased from SynPep (Dublin, CA, USA) and used without addition purification. Chemical reagents dithiothreitol (DDT), iodoacetamide, ammonium hydroxide, and ammonium bicarbonate were purchased from Sigma Aldrich (Oakville, ON, Canada). An intact monoclonal antibody (mAb) standard (Waters Intact mAb Mass Check, 720004420EN) was purchased from Waters (Milford, MA, USA). HPLC
grade acetonitrile (ACN) and methanol (MeOH) were purchased from Caledon Laboratory Chemicals (Georgetown, ON, Canada). HPLC grade acetone, HPLC grade cyclohexane, and formic acid (FA) were purchased from Sigma Aldrich (Oakville, ON, Canada). Distilled, deionized water (resistance ≥18 MΩ) was produced in-house by a Millipore Integral 10 water purification system (Billerica, MA, USA).

4.3.2 Modifier clustering

Direct infusion of 1 pmol/μL KGAIL peptide in 50% MeOH and 50% water at flow rate of 10 μL/min on a SelexION technology equipped SCIEX 5600+ TripleToF LC-MS system (Concord, ON, Canada) was used to evaluate the different curtain gas modifiers. The ion spray voltage was 5500 V with drying gases (GS1, GS2) flowing at 30- and 15-psi respectively. ToF-MS scans were performed at a declustering potential (DP) of 100 V, collision energy (CE) of 10 eV and an ion transfer coefficient (ITC) of 40% with 250 ms of accumulation time across a m/z range of 150-2000 m/z. ACN, acetone, MeOH, and cyclohexane were supplied at 1.5% of the 20-psi nitrogen curtain gas using the built-in solvent pump SelexION DMS controller. An external Perkin Elmer series 200 micropump (Waltham, MA, USA) supplied water directly into a 10-psi nitrogen curtain gas at a flow rate of 37.6 μL/min, resulting in an approximate 1.5% water enriched gas. The DMS internal temperature was kept at 150°C.

Compensation voltage (CV) mapping was performed at a fixed separation voltage (SV) of 3500 V, ramping the CV at 0.2V increments across an appropriate range. The resulting data was processed by producing extracted ion chromatograms (XIC) of the m/z of three separate ions of KGAIL, 4+: 285.6, 3+: 380.5, and 2+: 570.3 with a m/z width of 0.3. Gaussian smoothing (1-point width) was applied to each XIC individually. Precursor ion scans were performed of KGAIL under nitrogen without modifier and with ACN and acetone modifiers. Precursor ions for 4+ and 3+ KGAIL (285.6 and 380.5 m/z, respectively, width of 2 m/z) were acquired with a q1 m/z width of 280-360 and 370-450 m/z respectively and a scan time of 2 seconds. A total of 30 scans was acquired and the scans were averaged using Research PeakView v.1.2.2.0 software (SCIEX). DP ramps were performed of KGAIL infusions with ACN, acetone and without any curtain gas modifier using similar MS conditions except the DP was ramped from 0 to 225V with SV, CV set to 0 V. XICs of 4+ KGAIL were obtained from these scans and plotted. Finally,
ToF-MS spectra were obtained of the same infusion of KGAIL under all curtain gas configurations with SV, CV set to 0 V and a fixed DP of 100V. Spectra were collected for 1 minute (119 in total) and the peak area of KGAIL ions and fragments was determined from the average of the collected spectra. Each condition was repeated in triplicate for an assessment of reproducibility and all data was collected using Analyst TF v.1.7.1 (SCIEX).

4.3.3 Protein digestion and LC-MS instrumentation

A standard trypsin digest of bovine serum albumin (BSA) was performed. Briefly, 100 μL of 20 mM DTT was added to 2 mg of BSA in a polypropylene vial and boiled for 15 minutes. Then 100 μL of iodoacetamide was added to the vial and it was placed in the dark for 30 minutes. A 50-mM ammonium bicarbonate/ammonium hydroxide buffer was prepared at pH 8.5 and 600 μL was added to the vial. A total of 39 μg of trypsin was added and the vial was incubated at 37°C overnight. The resulting digest was diluted to 1 pmol/μL with water containing 0.1% FA for analysis by LC-MS. KGAIL, AngI, AngIII and bradykinin were added to the diluted BSA digest in some samples to evaluate known fragile peptides (concentrations ranging from 3.4 - 273, 1.7 – 133, 4.1 – 333 and 1.7 – 133 ng/mL respectively). The digestion procedure was also performed on a monoclonal antibody sample as well, with a final concentration of 1 pmol/μL.

A Shimadzu Prominence LC system equipped with a Phenomenex Aeris 2.6 μm PEPTIDE XB-C18 100 mm x 2.1 mm LC column was coupled to the same LC-MS system used to monitor clustering. Two solvents were employed in a 30-minute gradient elution, solvent A: 97.9% Water, 2% ACN, 0.1% FA and solvent B: 97.9% ACN, 2% Water, 0.1% FA. The gradient elution was as follows: 5% B for 1 min., ramp to 35% B for 14 min, ramp to 60% B for 5 min., ramp to 95% B for 2.5 min., hold B at 95% for 2.5 min., ramp to 5% B in 0.1 min., hold at 5% B for 4.9 min. The column temperature was held at 45°C and the flow rate was conserved throughout at 0.5 mL/min. Sample injections of 10 μL performed. The ion source conditions were as follows: curtain gas flow of 30 psi, sheath gases (GS1, GS2) of 50 psi each, source temperature of 450°C and an ion spray voltage of 5000 V. The MS system was operated in independent data acquisition (IDA) mode. IDA conditions involved a ToF-MS survey scan (200 ms) followed by up 10 MS/MS scans (30 ms each) on the most intense ions using an intensity
threshold of 50 cps. Ions (and isotopes up to 4 Da) were excluded for 2 sec. after a single occurrence.

Data collected from samples containing the spiked peptides were analyzed using PeakView to produce XICs for the 4+, 4+, 3+ and 3+ ions of KGAIL, AngI, AngIII and bradykinin respectively. From the XICs, the peak area of each ion was determined and the mean of three replicates (blank corrected) of four concentrations was subjected to linear regression to produce calibration curves. Limits of detection (LODs) were calculated as the concentration (from the regression) that correlated with the signal of the mean of the blank plus three times the standard deviation of the mean. Data from samples which did not contain additional peptides were analyzed using BioPharmaView v.2.0r3682 (SCIEX) for sequence coverage. Protein sequences were obtained from the UniProt database (P02769) for BSA and from Waters for the mAb sample and sequenced using the Peptide Mapping command (with a maximum charge state of 10+, a minimum peptide length of 3, a deconvolution tolerance of ±10 ppm, an XIC m/z width of 0.075 Da, and automatic recalibration). The percentage of sequence coverage and the percentage of auto-validated sequence coverage (mean of three replicates) were both evaluated by Student’s t-test between samples analyzed in the presence of ACN and controls. The number of IDA counted spectra was determined for each sample using PeakView’s IDA Explorer and compared between ACN treatment and control using Student’s t-test.

4.4 Results and Discussion

4.4.1 Cluster formation and fragmentation reduction

DMS is an effective means of probing clustering behavior of ions.274,276 The formation of clusters amplifies the mobility differences between ions based on their ability to cluster. Incoming ions which cluster strongly will have very different mobility versus those that cluster poorly. We employed DMS to observe the clustering potential of the KGAIL peptide with ACN, acetone, cyclohexane, water and MeOH. Figure 4-1 shows plots of ion intensity as a function of CV for the KGAIL ions (4+ red, 3+ black, and 2+ blue) under the different modifiers and a pure nitrogen control at a fixed SV of 3500 V. All the modifiers showed some change in the CV response of KGAIL, with cyclohexane showing the smallest change overall (approximately 3V
less than the control). Polar, aprotic modifiers (ACN, acetone) produced strong negative CV shifts, consistent with other research.\textsuperscript{380,381} It is hypothesized that the strong dipole moment of these modifier compounds (3.92 D and 2.88 D for ACN and acetone, respectively\textsuperscript{382}) enables extensive cluster formation on the densely charged KGAIL ions. Using ACN also showed a pronounced increase in 4+ ion intensity. Polar, protic modifiers (MeOH, water) produced weaker negative shifts but also exhibited a proton stripping effect in which the highest charge state ions were suppressed in the resulting ionograms, an effect which has been seen in other protic modifiers.\textsuperscript{282} Specifically, using water as a modifier eliminated the 4+ KGAIL ion, while MeOH resulted in the loss of both the 4+ and 3+ ions. Charge stripping can also be seen under other modifiers. For example, at the CV maximum of the 4+ ion under ACN, a small amount of 3+ ion is detected, suggesting a transfer of charge from the 4+ ion to the chemical modifier. Only ACN and acetone showed appreciable amounts of all three charge states (+2 to +4)
Figure 4-1: Infusion MS CV maps of KGAIL under different curtain gas modifiers. A solution of KGAIL was infused into the instrument and the CV was ramped at 0.2 V with a fixed SV of 3500 V. The change in the CV of maximum intensity for a given ion between modifier conditions indicates different clustering environments of that ion. A. The control condition under pure nitrogen shows all ions separated and with positive CVs. B. In the presence of ACN all CVs shift far negative and the relative intensity of the 4+ ion is increased over the control. C. The CV range of all ions under acetone is greatly compressed. D. Cyclohexane exhibits a small CV shift for all ions of ~3V (relative to the control), indicating limited clustering. E. Water demonstrates a charge stripping effect; the 4+ ion has been lost from the ionogram. F. Methanol shows an extreme case of charge stripping where only the 2+ ion remains in the ionogram.
Precursor scans provide a rudimentary method to observe incoming clustered ions if those clusters survive the entrance to the mass spectrometer. Figure 4-2 shows representative precursor mass spectra of 4+ and 3+ KGAIL ions under a nitrogen control, ACN modifier and acetone modifier. The precursor scans were set up to select for the expected $m/z$ of the 3+ and 4+ ions (380.5 and 285.6 $m/z$ respectively) in q3 while scanning an 80 $m/z$ window above those $m/z$s in q1. Control spectra show no clustering (as expected) but with ACN or acetone present, additional peaks beyond the bare ion $m/z$ are visible. With ACN, the KGAIL 3+ ion shows a strong peak at the expected ion $m/z$ with an additional peak +14 $m/z$ greater than the expected, consistent with the addition of one ACN molecule to the ion. Acetone shows a similar pattern with additional peaks at +19 and +30 $m/z$ greater than the expected peak, suggesting clusters of one and two additional acetone molecules on the KGAIL ion. Increasing to the 4+ charge state creates a complex set of clusters for ACN and acetone. The 4+ KGAIL precursor scan with ACN shows some bare ion but the predominant signal intensity originates from clusters with 2 or 3 additional ACN molecules, +20 and +30 $m/z$ respectively. Weak signals from clusters with 1 and 4 additional ACN molecules are also present. Acetone shows similarly low amounts of bare ion with strong signals for clusters of 2, 3 and 4 additional acetone molecules. This evidence shows that ion + modifier clusters survive into the vacuum of the mass spectrometer, only to be dissociated in the collision cell. This has implications for performing tandem MS experiments on highly charged ions such as KGAIL 4+, as most of the ion intensity is not present (prior to the collision cell) at the expected $m/z$ due to significant clustering.
Figure 4-2: Precursor spectra of KGAIL 3+ (380.5 m/z, left) and 4+ (285.6 m/z, right) under nitrogen (red), acetonitrile (blue) and acetone (black). Positive numbers on mass spectra indicate the number of modifier molecules clustered to the selected m/z ion. The spectra indicate that modifier/analyte clusters survive to q1 of the mass spectrometer and are cleared with 10 eV of collision energy in the collision cell before detection.

The clustering illustrated in Figures 4-1 and 4-2 can be used to reduce the potential for in-source fragmentation of ions. To evaluate this phenomenon, in-source fragmentation can be induced on our instrument by increasing the DP. Figure 4-3 shows the effect of increased DP on the XIC of
4+ KGAIL. The control (red, nitrogen only), shows a significant drop in ion intensity beginning at 80 V DP, dropping to less than 40% of the initial intensity at 100 V DP, the default and typical DP used on these instruments. By 130 V DP, all the 4+ ion signal is essentially eliminated. Evidence of b and y ion fragments is also seen at these voltages (see below). In contrast when ACN or Acetone are present the loss in signal intensity is much reduced. The intensity of the signal at the highest measured potential, 225 V, is still greater than 80% for the peptide under both ACN and acetone. We interpret this result as being evidence that the clustering of the ion shelters it from the energetic process of entering the mass spectrometer. The instability seen under the presence of acetone is likely a function of normal fluctuations in ion signal at relatively low overall intensity as acetone does not exhibit an improvement in 4+ ion intensity.
Figure 4-3: Infusion MS DP ramp of XIC of KGAIL 4+ under pure nitrogen (red), nitrogen with acetonitrile (blue) and, nitrogen with acetone (black). Ramping DP creates artificial in-source fragmentation of KGAIL 4+. At standard DP operating conditions (80-100 V) ion intensity is < 50%. The addition of ACN or acetone eliminates the fragmentation allowing >85% signal intensity at 225V DP. Acetone likely exhibits increased variability due to the limited signal intensity of the ion.

The mass spectrum of 4+ KGAIL ion without any gas modifier shows the presence of b- and y-ions, with b_8, y_4 and y_7 showing very strong intensities (Figure 4-4A). Upon the addition of ACN to the curtain gas the intensities of the fragment ions drops and the intensity of the 4+ ion itself is increased (Figure 4-4B). The sheltering of the ion (by clustering) seems to have prevented its fragmentation in the ion source, resulting in a much strong signal. The peak area of all ions and the sum of all peak areas for identifiable fragments were compared to the total peak area derived from KGAIL to evaluate the effect of each modifier (Figure 4-4C). Without any modifier, the majority of the KGAIL intensity is from the 3+ charge state but approximately 10% is due to fragment ions. Under ACN, the peak area intensity of the KGAIL ions is more evenly
distributed, with much stronger 4+ ion than any of the other modifiers. Furthermore, the percent of the total area derived from fragments is reduced by 66%. Acetone and cyclohexane produce even fewer fragments, with reductions of 84% and 90% respectively. Unfortunately, acetone does not produce as strong a 4+ ion intensity as ACN, and cyclohexane appears to suppress the 4+ ion as well. The charge stripping modifiers water and MeOH produce spectra with enriched 2+ ion intensities at the expense of the other charge states. Under MeOH, there is no detectable 4+ or 3+ ion, with 97% of the total peak area from the 2+ ion state. Both the charge stripping modifiers reduce the fragmentation of KGAIL, as observed by the limited amount (2.3 ± 0.7 % of the total KGAIL area) of fragmentation peak area observed in the ToF-MS spectra (Figure 4-4C). Overall, ACN is the most effective modifier in terms of a) reducing the fragmentation of the fragile ion and b) enhancing the intensity of the ions with highest charge states (in this case the 4+ charge state).
Figure 4-4: Fragmentation of KGAIL under different current gas modifiers. A. Under pure nitrogen control (left) KGAIL shows significant fragmentation of the 4+ ion which results in several b- and y-ions as well as a limited signal for the ion itself. But, with ACN present (right) the intensity of the 4+ ion is dramatically increased while the fragment ions are strongly decreased. B. The % total peak area from each ion or the sum of all fragments for KGAIL under different modifiers. Both ACN and acetone show an increase in 4+ ion percentage. All organic modifiers limit the total peak area from fragments. MeOH shows most of its peak area from a single state (2+ ion). Error bars represent 1 standard deviation about the mean (n = 3).
4.4.2 Application to fragile peptides in protein digests

To assess the application of an ACN curtain gas modifier to a protein digest, four fragile peptides (AngI, AngIII, bradykinin and KGAIL) were spiked into a BSA digest. This simulates a ‘real-world’ matrix while still providing a set of analytes with known behavior. It was hypothesized that the presence of the ACN modifier would improve the ability to detect the fragile ions and that the improvement would be associated to the overall fragility of the ion. Specifically, very fragile ions such as KGAIL 4+ may show improvements while more stable ions may show limited differences. The samples were analyzed in triplicate by HPLC-MS with and without the presence of ACN in the curtain gas, and the ions of interest (4+ KGAIL, 3+ bradykinin, 4+ AngI, and 3+ AngIII) were extracted from the chromatograms. Three example-spectra are shown in Figure 4-5, KGAIL 4+ (13.7 ng/mL), AngI (6.7 ng/mL) and AngIII (16.7 ng/mL). KGAIL shows an improvement in peak area under ACN, with a 5.7x increase versus the control, while maintaining an acceptable relative standard deviation (RSD) of the peak area under ACN of 8% (Figure 4-5A). While detectable in the control, AngI (Figure 4-5B) shows a modest improvement in peak area with ACN, approximately 2x over the control condition. When measuring AngIII 3+, the presence of ACN enhances the ion’s peak by 2.8 x versus the control under only nitrogen (Figure 4-5C) with an acceptable RSD of 8% for the ACN treated sample (n = 3). Table 4-1 shows the peak area improvement of each spiked ion with an ACN modifier. The scale of each improvement suggests the overall fragility of the ions, with KGAIL 4+ being the most fragile and bradykinin 3+ the least.
Figure 4-5: Selected XIC chromatograms of spiked KGAIL 4+ (285.7 ± 0.3 m/z), Angiotensin I 4+ (324.9 ± 0.3 m/z) and Angiotensin III 3+ (311.18 ± 0.03) ions from BSA digest matrix under ACN (blue) and no modifier (red) conditions. All ions show a marked increase in chromatogram peak area with ACN in the curtain gas. Insets show a zoomed x-axis around the ion peaks. A. KGAIL 4+ has a peak area increase of 5.7 x in the presence of the control with a peak area RSD of 8%. B. Angiotensin I 4+ has a peak area increase of 1.7 times in the presence of ACN with an RSD of the peak area of 6% (n = 3). C. Ang III 3+ has a peak area increase of 2.8 times in the presence of ACN with an RSD of the peak area of 8% (n = 3).
Table 4-1: The peak area increases due to the presence of ACN curtain gas modifier on peptide ions spiked into a BSA digest.

<table>
<thead>
<tr>
<th>Ion</th>
<th>Conc. (ng/mL)</th>
<th>Increase w/ ACN</th>
<th>%RSD of peak area under ACN</th>
</tr>
</thead>
<tbody>
<tr>
<td>KGAIL 4+</td>
<td>13.6</td>
<td>5.6 x</td>
<td>8%</td>
</tr>
<tr>
<td>AngI 4+</td>
<td>6.7</td>
<td>2.2 x</td>
<td>6%</td>
</tr>
<tr>
<td>AngIII 3+</td>
<td>16.7</td>
<td>2.8 x</td>
<td>8%</td>
</tr>
<tr>
<td>Bradykinin 3+</td>
<td>6.7</td>
<td>1.7 x</td>
<td>22%</td>
</tr>
</tbody>
</table>

The increase in peak area upon the addition of ACN suggests an overall improvement in the sensitivity and detection limit to the fragile ions in question. We propose that this could allow for the detection of previously unknown peptides in biological samples that would otherwise be undetectable due to their fragility. To test this explicitly, calibration curves were collected to determine sensitivities and limits of detection (LODs). All the spiked peptides showed an improvement in their LODs and an increase in the slope of their calibration line (Table 4-2).

Peptide ions that were previously identified as being very fragile (e.g., KGAIL 4+), show only marginal improvements to their LOD (i.e., a 16% reduction for KGAIL 4+), while the least fragile peptides show improvements (e.g., AngIII 3+ shows a 94 % reduction in LOD). This could be due to an improvement in reproducibility when ACN is present. Figure 4-6A shows the calibration curves for AngI 4+ ion with ACN (blue) and a control with no modifier (red). The control line has greater uncertainty (with %RSD range across the data points of 26 – 44% versus a %RSD range across the data points under ACN of 4 – 15 %), likely a combined effect of reduction in matrix interference afforded by the presence of ACN, and reducing fragmentation of the peptide. This is a novel observation which has previously only been reported when the DMS system is engaged for separation. The calibration curve of AngIII is show in Figure 4-6B. The control curve had a substantially larger LOD (130 ng/mL vs. 7.7 ng/mL with ACN present). The benefit is likely due to a reduction to the relative standard deviation (RSD) at each measured point. The control data had approximately ~50% RSD at each data point vs. 1 – 8 % RSD when
measured with ACN present. In combination with an increase in peak area, this results in much improved detection limit using ACN for this charge state.

**Table 4-2**: Limit of detection improvements in fragile ions upon the addition of ACN to the curtain gas.

<table>
<thead>
<tr>
<th>Ion</th>
<th>Limit of detection with no modifier (ng/mL)</th>
<th>Limit of detection with ACN (ng/mL)</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>KGAIL 4+*</td>
<td>4.4</td>
<td>3.7</td>
<td>16%</td>
</tr>
<tr>
<td>AngI 4+</td>
<td>5.6</td>
<td>3.6</td>
<td>36%</td>
</tr>
<tr>
<td>AngIII 3+</td>
<td>130</td>
<td>7.7</td>
<td>94%</td>
</tr>
<tr>
<td>Bradykinin 3+</td>
<td>3.9</td>
<td>1.2</td>
<td>70%</td>
</tr>
</tbody>
</table>

*KGAIL 4+ ion LOD is calculated from a calibration curve covering only three concentrations (3.4 ng/mL, 13.7 ng/mL and 68.3 ng/mL due to non-linear behavior of the ion.*
Figure 4-6: Calibration curves for Angiotensin I 4+ and Angiotensin III 3+ with ACN modifier and without. The presence of ACN (blue) increases the slope of the calibration line in relation to the control (red). In addition, the uncertainty of the data is greatly reduced with ACN modifier by up to 3 times for AngI 4+ and 6 times for AngIII 3+. A. Calibration curves for Angiotensin I 4+ ion. B. Calibration curves for Angiotensin III 3+ ion. The lowest point on the curve is omitted from the plot due to its immense uncertainty. Error bars represent 1 standard deviation about the mean (n = 3).

Most importantly, the benefits of using ACN to increase sensitivity do not negatively impact the results of peptide mass fingerprinting of the protein matrix. Figure 4-7A shows the % sequence coverage determined by BioPharmaView’s algorithm, which matches m/z of potential peptides in MS1, for a digest of a mixture of a mAb (chosen as a stand-in for common protein therapeutics) and BSA (chosen to represent a generic protein), analyzed with and without an ACN modifier. Statistically there is no difference between the sequence coverage with the modifier for the mAb (p = 0.363, α = 0.05) or BSA (p = 0.224, α = 0.05). A similar result is seen with BioPharmaView’s auto-validation algorithm which employs MS/MS spectra to verify its matches in MS1. Specifically, there is no significant impact on the auto-validated sequence coverage for the mAb (p = 0.410, α = 0.05) or BSA (p = 0.172, α = 0.05) when using ACN as a modifier. This suggests that the clustering seen in Figure 4-2 does not have a detrimental impact on most of the peptides present in the sample. However, it remains possible that there are some
peptides that cannot be detected in MS/MS scans using IDA due to their m/z being altered by clustering. This is because the IDA method selects for the exact m/z of the precursor, which would only be detected after declustering in the collision cell. Stated another way, in this case, q1 would be effectively at set at different, erroneous mass. Thus, data independent acquisition methods such as Sequential Window Acquisition of All THeoretical mass spectra (SWATH)\textsuperscript{257} with sufficiently wide isolation windows may be necessary to achieve complete MS/MS information from samples analyzed with strongly clustering modifiers like ACN or acetone.

Another major benefit of using ACN as a modifier for the IDA method is a significant reduction in the number of spectra captured in the analysis (Figure 4-7B). Both the mAb and BSA digest had large reductions in the total number of IDA events using ACN when compared to the control (reductions of 42% and 40% respectively). This indicates that a good portion of the spectra collected by IDA do not add further support for the identification of the protein. We speculate that these extraneous spectra (that are not identified when using ACN) may be be derived from contaminants and noise which is eliminated by the gas-phase modifier (Figures 4-5 and 4-6) or peptide fragments created by in-source fragmentation (Figure 4-4). The eliminated spectra provide either no support for the protein identification or redundant support. The removal of these artifact spectra could be of great benefit to more complex analyses where elimination of noise is crucial to proper interpretation of the biological significance of proteins and peptides.
Figure 4-7: Protein sequencing comparisons with and without ACN modifier. The sequence coverage of both mAb and BSA were statistically unchanged in the presence of ACN despite the dramatic reduction in the number of spectra captured by IDA. A: Total MS1 based sequence coverage and % of computer auto-validated sequence coverage was no different with ACN. B. The total number of spectra captured by the IDA method. The presence of ACN reduces the total number of IDA ‘hits’ necessary to achieve maximal coverage. Error bars represent 1 standard deviation about the mean (n = 3).
4.5 Conclusion

We have demonstrated the use of polar gas-phase modifiers in the curtain gas to shelter fragile peptide ions from in-source fragmentation. The modifiers cluster with the ions which results in a larger energy barrier to fragmentation which we can monitor via DMS and precursor scans. The net result is significant reduction in fragmentation of fragile ions which leads to a decrease in limits of detection for these ions in a protein digest matrix. Furthermore, this method does not interfere with routine peptide mass fingerprinting of the protein digest but does eliminate much of the redundant data produced by data independent means of MS/MS acquisition.
Chapter 5
DMS-SWATH: A New Method Combining HPLC and DMS for the Analysis of Protein Digests

5.1 Abstract

Full characterization of a therapeutic protein requires the use of untargeted data acquisition approaches. But many of these approaches requires strong chemical separation that can be challenging to achieve with liquid chromatography (LC) alone. In response to this challenge, we introduce DMS-SWATH, a technique which combines differential mobility separation (DMS) with LC by leveraging Sequential Windowed Acquisition of All Theoretical fragment ion mass spectra (SWATH) for the analysis of single-protein digests. SWATH is a data acquisition method which relies on simultaneous fragmentation of wide \( m/z \) windows to capture MS/MS data for all possible ions. DMS-SWATH employs a DMS-transparent survey scan in conjunction with a series of large \( m/z \) SWATH windows to produce MS/MS spectra with each window separated using DMS at different compensation voltages. In this technique, DMS separates the components prior to analysis, simplifying the resulting MS/MS spectra, and removing many chemical interferences. This allows a significant improvement in MS/MS spectral quality and recovery, resulting in a 12-25% increase in auto-validated (where recovered peptide fragments are matched to expected theoretical fragments) sequence coverage compared to other common methods. Overall, by applying DMS-SWATH, more peptides were recovered and matched with higher scores using BioPharmaView software resulting in higher confidence in routine analysis. In addition, DMS-SWATH resolved peptides which co-eluted and whose masses placed them within the same SWATH window. Finally, DMS-SWATH can be combined with chemical modifiers to improve the separation capability of DMS and provides a means to deal with
persistent clustering to precursor ions which would otherwise prevent data dependent acquisition methods from producing high quality MS/MS spectra. DMS-SWATH is a powerful new means of untargeted data acquisition which we propose will have a large impact on the analysis of biopharmaceuticals like monoclonal antibody characterization.

5.2 Introduction

Identification and characterization of components from biological samples is necessary in many fields including medicine\textsuperscript{384} and the -omics fields such as proteomics\textsuperscript{385,386} and metabolomics\textsuperscript{387,388}. For example, the pharmaceutical industry is particularly interested in the characterization of therapeutic proteins like monoclonal antibodies (mAbs),\textsuperscript{389–391} which can be produced with a wide range of different post-translational modifications\textsuperscript{392}, disulfide linkages\textsuperscript{393} and contaminants.\textsuperscript{394} The rigours of biopharmaceutical development and testing demands that these variations be evaluated in detail at many stages of a drug’s development and production. Many mAbs are also digested prior to analysis to gain additional insight into their particular structure.\textsuperscript{395–397} Owing to these complexities, separation methods are often essential for ensuring useful analytical results.\textsuperscript{398} These samples are typically evaluated by liquid chromatography in conjunction with mass spectrometry (MS) and/or tandem mass spectrometry (MS/MS).\textsuperscript{399,400} But, even the most advanced 1-dimensional LC separation is not always sufficient to provide full resolution of all components, or to separate the analyte peaks from chemical noise inherent to the system or sample matrix.\textsuperscript{401} To solve the problem, many researchers turn to multidimensional LC methods, which are far more complex and time-consuming to perform.\textsuperscript{401–403} To avoid these complexities, gas-phase separation methods in the ion mobility family are emerging as a useful alternative.\textsuperscript{404,405}
Ion mobility is a family of methods which separate ions based on their differing mobility as they pass through gaseous media or fields. The simplest member of the family is drift tube ion mobility spectroscopy, in which ions move through a gas-filled tube under a low electric field. Ions with larger collisional cross-sections move more slowly through the field than smaller ions, allowing for separations, even of isobaric species. Alternate approaches include differential mobility spectrometry (DMS) and field asymmetric ion mobility spectrometry (FAIMS), both of which rely on similar separation principles but with different geometries.

In DMS, a radio-frequency alternating separation voltage (SV) is applied to two parallel plates. Ions are pushed through the plates by a buffer gas, where they move perpendicularly to the gas flow in response to the SV and as a function of their mobility, toward one of the plates. A DC compensation voltage (CV) is applied to pull ions back toward the region between the plates, in line with the inlet of the mass spectrometer. When CVs are scanned across a range of values, ions are sorted through the DMS system and sequentially detected by the mass spectrometer. Many applications of DMS involve the use of a chemical modifier in the curtain-gas. These modifiers alter the mobilities of incoming ions by clustering around them. In periods of high field, the clusters break up, and then reform in periods of low field. This allows further separation of ions based on their ability to cluster (or not) with the modifier gas which can lead to an increase in overall peak capacity (a measure of separation power). DMS has been used for a wide range of targeted and untargeted applications including drug analysis, proteomics, food analysis, and metabolomics.

There have been many approaches to combining LC separation with DMS in an untargeted manner. Despite these successes, the adoption of LC-DMS has been slow at least partly because of challenges related to duty cycle. Specifically, it is difficult to scan all the
CVs available to a DMS system during the limited duration of a typical chromatogram peak (1-10 seconds in a standard protein digest preparation). This limitation is driven by the required dwell time (30-50 ms) for the mass spectrometer to acquire a high-quality MS spectrum, combined with an additional dwell time to clear the DMS cell (5-7 ms in planar geometry) between runs. In addition, most bottom-up protein characterization methods rely on data acquisition methods such as Information Dependent Acquisition (IDA) or Sequential Windowed Acquisition of All Theoretical fragment ion mass spectra (SWATH). These methods employ an MS-level survey scan to orient and manage the data acquisition and interpretation as it is collected. Adding DMS to such systems and continually scanning CVs would likely complicate the procedure, as it is would be necessary to perform each scan without DMS separation. To perform an effective analysis of a protein digest sample it would be necessary to employ a) a survey scan to acquire precursor ions for peptide mass fingerprinting, b) fast cycling of CV to acquire all essential sample components, and c) MS/MS capability.

Here, we introduce DMS-SWATH, a new method for combining LC with DMS for the analysis of protein digests. DMS-SWATH employs a novel approach, combining a ‘DMS transparent’ scan at the MS level with a traditional DMS CV sweep performed at the MS/MS level. Briefly, DMS-SWATH first performs a survey scan (ToF-MS) with SV and CV set to 0 V. This is used to identify peptides from protein digests by their precursor ion \( m/z \). Then, SWATH is employed in conjunction with DMS separation to capture MS/MS spectra of product-ion species from the survey scan. In SWATH, analytes in windows of fixed \( m/z \) widths are transferred to the collision cell and fragmented all at once, ensuring that all ions (not just those with high intensities in the survey scan) are fragmented. In DMS-SWATH we use a large SWATH window with DMS separation at a fixed SV while scanning CVs. In so doing, we trade the mass separation of the
SWATH window for the LC-orthogonal separation provided by DMS. This technique allows for an increase in MS/MS auto-validation and spectral quality in a typical mAb digest over traditional IDA or SWATH analysis and it can effectively sort out co-eluting components of similar m/z.

5.3 Materials and Methods

5.3.1 Materials

Bovine serum albumin (BSA), dithiothreitol (DDT), iodoacetamide, ammonium hydroxide, ammonium bicarbonate, tetraethyl ammonium bromide (TEAB), octyl-β-D-glucopyranoside (OGS), tris(2-carboxyethyl)phosphine (TCEP), and methyl methanethiosulfonate (MMTS) were purchased from Sigma Aldrich (Oakville, ON, Canada). Trypsin was acquired from Promega (Madison, WI, USA). An intact monoclonal antibody (mAb) standard (Waters Intact mAb Mass Check, 72000420EN) was purchased from Waters (Milford, MA, USA). A peptide, KGAILKGAILR (Synpep, Dublin, CA, USA), was used to evaluate cluster breaking efficiency. It was prepared at 1 pmol/μL in 50% methanol (MeOH) and water. HPLC grade acetonitrile (ACN) and MeOH were purchased from Caledon Laboratory Chemicals (Georgetown, ON, Canada). HPLC grade formic acid (FA) was purchased from Sigma Aldrich (Oakville, ON, Canada). Distilled, deionized water (resistance ≥18 MΩ) was produced in-house by a Millipore Integral 10 water purification system (Billerica, MA, USA).

5.3.2 Sample preparation

Both mAb and BSA samples were subjected to trypsin digestion under slightly different conditions. mAb digestion was performed using TEAB, OGS, TCEP and MMTS. Briefly, the mAb protein was prepared in 500 mM TEAB at 5 μg/mL and then denatured by adding 1 mg
OGS and incubating at 95°C for 5 minutes. Afterwards 50 mM TCEP was added and the mix was left to sit at 60°C to reduce disulfide bonds over 60 minutes. Once completed cysteines were blocked using MMTS and the final mix was diluted with 0.67% OGS in water. Trypsin was added in a 1:20 trypsin:protein ratio and incubated at 37°C overnight for digestion. BSA digestion was performed using DTT, iodoacetamide and ammonium bicarbonate. Briefly, 100 μL of 20 mM DTT was added to 2 mg of BSA in a polypropylene vial and boiled for 15 minutes. Then 100 μL of iodoacetamide was added to the vial and it was placed in the dark for 30 minutes. 600 μL of a 50-mM ammonium bicarbonate/ammonium hydroxide buffer at pH 8.5 was added to the vial. 39 μg of trypsin was then added and the vial was incubated at 37°C overnight. Both digestes were diluted to 1 pmol/μL with water containing 0.1% FA for analysis by LC-MS.

5.3.3 Liquid chromatography

A Shimadzu Prominence LC system equipped with a Phenomenex Aeris 2.6 μm PEPTIDE XB-C18 100 mm x 2.1 mm LC column coupled to a Sciex 5600 TripleToF mass spectrometer with a SelexION DMS unit attached. Two solvents were employed in a 30-minute gradient elution, solvent A: 97.9% Water, 2% ACN, 0.1% FA and solvent B: 97.9% ACN, 2% Water, 0.1% FA. The gradient elution was as follows: 5% B for 1 min, ramp to 35% B for 14 min, ramp to 60% B for 5 min, ramp to 95% B for 2.5 min, hold B at 95% for 2.5 min, ramp to 5% B in 0.1 mi., hold at 5% B for 4.9 min. The column temperature was held at 45°C and the flow rate was conserved throughout at 0.5 mL/min. Sample injections of 10 μL were performed. The ion source conditions were as follows: curtain gas flow of 30-psi, sheath gases (GS1, GS2) of 50-psi each, source temperature of 450°C and an ion spray voltage of 5000 V. All acquisitions were performed in triplicate except DMS-SWATH which was performed 6 times.
5.3.4 Mass spectrometry and data analysis

All mass spectrometry was performed with the following parameters, with additional MS/MS parameters as dictated by the data acquisition method (see sections 5.3.5 through 5.3.8). ToF-MS and MS/MS scans were performed at a declustering potential of 80 V with 100% ion transmission coefficient (ITC). ToF-MS survey scans were performed with an accumulation time of 200 ms, 10 eV collision energy (CE), across a 150-2000 m/z window. Peptide mass fingerprinting was performed using BioPharmaView v.2.0r3682 (SCIEX) for sequence coverage. Protein sequences were obtained from the UniProt database (P02769) for BSA and from Waters for the mAb sample with protein modifications manually added to BioPharmaView using those sources. The proteins were sequenced using the Peptide Mapping command (with a maximum charge state of 10+, a minimum peptide length of 3 amino acid residues, a deconvolution tolerance of ±10 ppm, an XIC m/z width of 0.075 Da, and automatic recalibration). Peptide score data was transposed from BioPharmaView to Microsoft Excel, where the highest score for a given peptide (across all CVs if necessary) was determined. Differences in auto-validated sequence coverage were tested using ANOVA with Tukey-Kramer Post-Hoc analysis using Microsoft Excel. A Grubbs test on the auto-validated sequence coverage excluded one set of DMS-SWATH data. Other chromatographic, ionographic and spectral features were evaluated using Research PeakView v.1.2.2.0.

5.3.5 Information Dependent Acquisition (IDA) analysis

IDA uses the ToF-MS survey scan to select for (trigger on) ions on which it will acquire MS/MS data. The following parameters were used to trigger MS/MS acquisition for up to 10 ions, starting with the ion of greatest intensity and working down in intensity. All ions of greater than 50 counts per second (cps) intensity were monitored. Ions and their isotopes within 4 Da were
excluded for 2 s after they were captured once. Identical \( m/z \) peaks within 6 Da of each other were ignored. Dynamic background exclusion and dynamic collision energy were enabled. Total period cycle time was 501 ms.

5.3.6 SWATH Analysis

SWATH is a data independent acquisition method used to capture MS/MS data from a range of \( m/z \) simultaneously, guaranteeing that all ions are captured with MS/MS data.\(^{257}\) In the SWATH analyses used here, 18 windows of 50 \( m/z \) each were employed across a range from 350 to 2000 \( m/z \) with a 1 Da overlap with the previous window: 350-400, 399-450, 449-500, etc. Each window was captured at a CE of 40 eV with 15 eV CE spread and an accumulation time of 30 ms. Total period cycle time was 741 ms.

5.3.7 Combined DMS analysis

The SWATH conditions above were combined with DMS. In this manner three separate injections were performed each at a different CV and the results of each were combined into a single analysis using BioPharmaView. DMS was performed with SV of 3500 V, DMS offset voltage of -3V, and three separate CVs: 8, 12, and 16 V. Another experiment was performed as above except the ToF-MS survey scan was performed with the use of DMS separation i.e. SV and CV were both 0 V, like that seen in DMS-SWATH below.

5.3.8 DMS-SWATH

DMS-SWATH modifies the SWATH protocol to allow each separate step of a SWATH analysis to operate at a different SV and CV. This was accomplished in Analyst TF v.1.7.1 (SCIEX) software with the parameters being modified independently in each step via the ‘Advanced MS’ tab. In this report, the ToF-MS survey was performed with SV and CV = 0 V with 12 additional
SWATH steps operating at $SV = 3500$ V and CV ranging from $5$ V to $16$ V in $1$ V increments. Each SWATH step employed a wide SWATH window (350-1250 m/z) with an accumulation time of 32.4 ms. This window is slightly increased relative to the other analyses to incorporate a lag time between switching CV values. The total cycle time was 592 ms. To address a limitation in BioPharmaView, each resulting data file was split into 12 separate files each containing a copy of the ToF-MS survey scan and one DMS-SWATH chromatogram at a single CV value. These were then analyzed simultaneously using BioPharmaView. This methodology was applied to a mAb digest sample. A variant method was used to confirm the presence of precursor ions at each CV; instead of producing MS/MS spectra via large SWATH windows, in the variant method, those steps were replaced with 100 ms accumulation ToF-MS scans at each CV.

**5.3.8.1 ACN-DMS modification**

The feasibility of DMS-SWATH with a chemical modifier was assessed using a BSA digest sample with ACN as a chemical modifier. Briefly, ACN was delivered to the curtain gas via the SelexION modifier delivery system at a total modifier composition of 1.5%. The DMS-SWATH conditions were modified to reflect the change in CV behavior in the ACN environment with CVs ranging from -13 to 9 V in 2 V increments. Otherwise the DMS-SWATH procedure was unchanged.

SWATH was also evaluated for its ability to deal with $m/z$ shifting caused by ACN clustering on precursor ions. Direct infusion of a model peptide KGAILKGAILR (molecular weight: 1141.78 g/mol) was performed at 10 μL/min under identical chemical modifier conditions as above. The SV and CV were 3500 and -9.6 V respectively; these were found to be the optimal conditions for separation of the 4+ ion of the peptide. MS/MS was performed on the precursor ion $m/z$ (285.6
m/z) at CE = 13 eV in the presence of ACN and without. A SWATH method was employed with windows of 50 m/z units from 280 to 580 m/z (6 windows total) with a CE of 13 eV and CE spread of 15 eV. Data was acquired for 50 cycles with each cycle approximately 600 ms. The spectra were summed and analyzed using PeakView. DMS-SWATH was then performed on a BSA digest containing 68 ng/mL KGAILKGAILR with a SWATH window size of 250-1250 m/z and CE of 13 eV and CE spread of 15 eV. The resulting MS/MS spectra (6 total) from CV = -7V across the eluting peak were averaged and compared to the direct infusion experiments.

5.4 Results and Discussion

5.4.1 Methodology

DMS-SWATH differs from other untargeted methods in that specificity arises from DMS separation instead of from mass selection. This principle is illustrated in Figure 5-1, which compares three strategies (including DMS-SWATH) for untargeted data acquisition. In IDA and other data-dependent methods (Figure 5-1A), a survey scan is an essential first step. Following a set of criteria in terms of charge states and intensity, peaks in the survey scan are selected for MS/MS fragmentation in order of decreasing intensity. Exclusion rules and exceptions lists are essential for reducing redundant data and lower intensity ions are routinely missed. Missing MS/MS data for these ions limits the overall MS/MS sequence coverage of IDA. SWATH and other data independent methods (Figure 5-1B) solve the problem of bias towards high-intensity ions in the survey scan. In SWATH, a window of a defined m/z width is isolated in Q1, and all of the captured ions are fragmented simultaneously. The window is then shifted towards higher mass by the defined width with an additional 1 m/z overlap. For instance, if the m/z window width is 50 m/z units, the first window might be 350-400 m/z. The subsequent window is 399-450, with 449-500, etc., following, making a total of 18 separate windows for a m/z range of 350-
1250 (the range used in this study). While SWATH improves upon IDA by not missing low-intensity ions from the survey scan, in SWATH, the captured ions are all fragmented together (many of which are chemical noise not analytes of interest), which leads to complex MS/MS spectra. These complex spectra confound the ability of software algorithms to automatically match MS/MS spectra with the expected peptides in a protein digest, leading to lower than desired MS/MS sequence coverage. This can be alleviated to some degree by reducing the window size and using more windows to cover the whole mass range. But this comes at the expense of time: a longer time to cycle through the entire mass range can result in fast chromatographic peaks being insufficiently sampled or missed entirely. The complexity of MS/MS spectra (and the resulting reductions in automated protein sequence coverages) in SWATH motivated us to develop DMS-SWATH, which uses DMS-based pre-separations to reduce the complexity of the MS/MS spectra. DMS-SWATH thus produces less complex MS/MS spectra, that allow for higher validated sequence coverage than SWATH alone, which allows for routine protein analysis with higher confidence in identifications. As is the case for IDA and SWATH, DMS-SWATH begins with a MS level survey scan (Figure 5-1C). This is essential for determining the precursor ions in the sample. Then, DMS separation is enabled and a series of MS/MS spectra are obtained. Each spectrum is obtained at an exceptionally wide SWATH isolation window (350-1250 m/z for instance), but each spectrum is less complicated than a typical SWATH spectrum, as it is obtained for analytes collected at different DMS CVs. In preliminary analyses, it was found that all detectable peptides (for the targets evaluated in this work) fall between 5 and 16 V CV. Under ACN chemical modified conditions (as described in Chapter 4), the range can be made broader, with an optimal CV spread of -13 to 9 V. The primary limitation of DMS-SWATH is a lack of tunability for collision energy. As each
spectrum can contain ions of any \( m/z \), a large collision energy is employed to ensure fragmentation of even the most durable ions, which can result in the loss of fragile ions. However, peptides tend to exist at different charge states and it was found that in most cases, at least one charge state survives to be analyzed.

**Figure 5-1:** Comparison of untargeted data acquisition methods available. A. Information dependent acquisition (IDA) involves selecting peaks from an initial survey scan to fragment with MS/MS. The peak selection criteria are user defined and parameters available include dynamic exclusion of peaks which have already been sampled and isotope peak exclusions. The most intense peak which matches the approved criteria is selected first with subsequent lower intensity peaks sampled afterwards. B. Sequential Windowed Acquisition of All Theoretical fragment ion mass spectra (SWATH) is a data independent technique where all the ions present in a survey scan are fragmented in distinct SWATH windows of a selected \( m/z \) range. All ions present in the window are fragmented simultaneously. C. DMS-SWATH eschews the mass selectivity of both IDA and SWATH for selectivity and separation based on DMS. In DMS-SWATH a survey scan is acquired at the MS level without the use of DMS. Then, a single large SWATH window (350-1250 \( m/z \)) is selected and MS/MS data is acquired from that window at a series of CV values.
As described above, DMS-SWATH data is a survey scan combined with a series of MS/MS spectra from large, identical SWATH windows at different CVs (Figure 5-2). The first scan is identical to what one would find in a traditional IDA or SWATH analysis (Figure 5-2A, black trace). Tracing the total ion current for each scan, the differences brought on by separating the ions by DMS are apparent (Figure 5-2A, multi-coloured traces). This indicates that different species are isolated in each CV window. The power of DMS separation is clearer when examining the MS/MS spectra (Figure 5-2B). Each ion captured in each spectrum is derived from different components isolated by the DMS separation. To provide MS/MS validated sequence coverage, the precursor ions captured in the MS survey scan must be matched to MS/MS spectra. In a typical SWATH acquisition, this is trivial; the precursor m/z correlates to a SWATH window which contains that m/z. In a DMS-SWATH acquisition, the MS/MS fragments of a precursor ion are separated by CV. Currently there is no method of accurately predicting the expected CV of a given ion. Thus, in the method introduced here, the precursor-to-MS/MS-fragment correlation was performed by splitting the resulting data file into 12 separate files; each new file consisted of the survey scan and one of the MS/MS scans at a single CV. These 12 files were then processed independently and their results combined into a single MS/MS sequence coverage result. This increases the processing time but is unavoidable until CV prediction is possible.
**Figure 5-2:** Chromatograms and MS/MS spectral output of a DMS-SWATH experiment. A. DMS-SWATH total ion chromatograms. Each DMS-SWATH experiment consists of a ToF-MS survey scan (black) and 12 DMS separated MS/MS spectra (multi-coloured), each at a unique CV. Certain components are emphasized at different CV values. B. MS/MS spectra from a DMS-SWATH experiment from a single peak in the chromatogram (10.9 min retention time). A 350-1250 m/z window from the survey scan mass spectrum (black) is separated by DMS across a CV range of 5-16 V (in 1 V increments) then fragmented by collision induced dissociation resulting in 12 different MS/MS spectra (multi-coloured). Each separate MS/MS spectrum contains fragments from different precursors based on their separation by DMS.
5.4.2 mAb sequence coverage

The new method was tested by applying it to analyze a model mAb digest. In this test, the mAb sequence coverage was evaluated using five different methods: DMS-SWATH, SWATH, combined SWATH at 3 different CVs, combined SWATH at 3 different CVs with a DMS-transparent survey scan, and IDA. All methods returned approximately the same MS-level sequence coverage, with a slight increase in the combined methods (likely a function of the repeated injections necessary to collect data at each CV). But critically, the auto-validated sequence coverage shows differences between each the different techniques, with 94 ± 3, 82 ± 3, 86 ± 4, 84 ± 8 and 68 ± 6 % coverage from DMS-SWATH, SWATH, combined 3 x CV SWATH, combined 3 x CV with transparent DMS ToF-MS, and IDA, respectively (Figure 5-3A). Specifically, there is a significant improvement for the DMS enabled methods (DMS-SWATH and the combined SWATH methods) when evaluated in pairwise comparisons (α = 0.05, Table 5-5). Note that auto-validation in this study is implemented by BioPharmaView for any MS/MS spectrum with a score greater than 3. Scores are the -log of a p-value calculated examining the probability that an observed spectrum is different than a spectrum composed of random peaks; stated another way, a high score denotes the presence of many predicted b- and y-ion fragments at high intensities. The sum of all peptide scores for a method gives an indication of the quality of the MS/MS spectra for those peptides. We can visualize the sequence validation and overall quality of spectra for each method by plotting the sum of all scores greater than 0 (i.e. all peptides with some MS/MS data) relative to the total number of auto-validated peptides for each method (Figure 5-3B). Here we see the main benefit of DMS-SWATH: it produces more auto-validated peptide matches than the other methods (x-axis) and the overall quality of the matches is greater than other SWATH methods and comparable to IDA (y-axis).
Figure 5-3: Sequence coverage differences observed for different data acquisition methods applied to a model mAb digest. The evaluated acquisition methods were DMS-SWATH (red), SWATH (light green), 3 separate SWATH injections at distinct CV values combined (turquoise), 3 separate SWATH injections at distinct CV values with a DMS-transparent ToF-MS combined (dark green) and IDA (purple). Error bars represent 1 standard deviation around the mean. A. Mean sequence coverage (MS) and mean auto-validated sequence coverage (MS/MS) of all methods. B. The mean sum of all scores greater than 0 relative to the mean number of auto-validated peptides matches by each method. The results show that DMS-SWATH validates more peptides by MS/MS than the other methods while also producing high quality MS/MS spectra.
Table 5-1: Summary of ANOVA with Tukey-Kramer Post-Hoc analysis for auto-validated sequence coverage across different data acquisition methods

<table>
<thead>
<tr>
<th></th>
<th>Q-STAT</th>
<th>Q-CRIT</th>
<th>P-VALUE</th>
<th>SIGNIFICANT? (α = 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMS-SWATH vs. SWATH</td>
<td>4.6749</td>
<td>4.508</td>
<td>0.04087</td>
<td>Yes</td>
</tr>
<tr>
<td>DMS-SWATH vs. 3CV SWATH</td>
<td>3.2759</td>
<td>4.508</td>
<td>0.20513</td>
<td>No</td>
</tr>
<tr>
<td>DMS-SWATH vs. 3CV SWATH with ToF-MS</td>
<td>4.0637</td>
<td>4.508</td>
<td>0.08472</td>
<td>No</td>
</tr>
<tr>
<td>DMS-SWATH vs. IDA</td>
<td>10.406</td>
<td>4.508</td>
<td>7.04 x 10⁻⁵</td>
<td>Yes</td>
</tr>
<tr>
<td>SWATH vs. 3CV SWATH</td>
<td>-1.2512</td>
<td>4.508</td>
<td>0.89714</td>
<td>No</td>
</tr>
<tr>
<td>SWATH vs. 3CV SWATH with ToF-MS</td>
<td>-0.5466</td>
<td>4.508</td>
<td>0.99461</td>
<td>No</td>
</tr>
<tr>
<td>SWATH vs. IDA</td>
<td>5.12650</td>
<td>4.508</td>
<td>0.02363</td>
<td>Yes</td>
</tr>
<tr>
<td>3CV SWATH vs. 3CV SWATH with ToF-MS</td>
<td>0.70459</td>
<td>4.508</td>
<td>0.98599</td>
<td>No</td>
</tr>
<tr>
<td>3CV SWATH vs. IDA</td>
<td>6.37776</td>
<td>4.508</td>
<td>0.00523</td>
<td>Yes</td>
</tr>
<tr>
<td>3CV SWATH with ToF-MS vs. IDA</td>
<td>5.67317</td>
<td>4.508</td>
<td>0.01217</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The improvement in MS/MS validated sequence coverage of DMS-SWATH is attributed to two properties. First, DMS-SWATH allows for a significant reduction in meaningless noise found in the MS/MS spectra. For example, in the mAb dataset, DMS-SWATH is more likely than the other techniques to recover a high-quality spectrum for a large peptide, heavy chain tryptic peptide T13

(GYFPEPVTVTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTVPSSTWPSETVTCNV AHPS
STK, monoisotopic mass 6543.1203 with a single methylthio-addition on the cysteine residue, 9% of total sequence). Only high charge states of this peptide (6+ and 7+) can be isolated on the hardware used here, and these charge states tend to produce peaks in the survey-scan that have low intensity. DMS separation improves the relative peak intensities for these analytes by reducing the amount of chemical noise present in the MS/MS spectra. This peptide is particularly important for overall protein sequence coverage, given its longer-than-average length. Different hardware with wider isolation windows may change these results, as the lower charge
state ions of this peptide produce peaks in the survey scan that are more intense, but they exist outside of the isolatable mass range on our system. The second property of DMS-SWATH that facilitates higher auto-validated sequence coverage is identical to the advantage of traditional SWATH over IDA-like acquisition: all ions are collected for fragmentation whereas IDA can miss triggering on ions with low peak intensities in the survey scan. Many of these ions are small, but together they can make a significant difference in sequence coverage. This is also why IDA can miss the large heavy chain T13 peptide as well as the isolatable charge state peaks have low intensity.

5.4.3 DMS separation of SWATH neighbours

An increase in peak capacity is necessary to prevent species from co-eluting from LC columns. Changes in the mechanics of chromatography to develop more efficient columns or systems like UPLC or low flow-rate nanoLC improve the peak capacity but can introduce new problems in terms of robustness. SWATH acquisition is usually effective at handling co-eluting components as they are unlikely to fall within the same SWATH window. But that likelihood is not zero, and in the rare cases in which two co-eluting species are ‘SWATH neighbours’ (i.e. their m/z ratios put them within the same SWATH window), a ‘chimeric’ MS/MS spectrum showing fragment ions from multiple species is the result. Improved separation capability is required to resolve this SWATH neighbor/chimeric spectrum problem. The addition of DMS separation does not require altering the chromatography and can resolve components which co-elute and exist in the same SWATH window. Specifically, by engaging DMS separation using DMS-SWATH, SWATH neighbour peptides can be separated. Figure 5-4 demonstrates this phenomenon using a BSA protein digest under ACN-modified DMS separation conditions. In traditional SWATH, peptides T22$^+$ (GACLPPK, m/z 758.4156 with a single iodoacetamide modification on the cysteine
residue) and T50$^{2+}$ (EYEATLEECCAK, $m/z$ 751.8032, also with iodoacetamide modifications on the cysteine residues) co-elute at 4.9 minutes. With 50 $m/z$ SWATH windows both peptides lie within the 749-800 $m/z$ window, resulting in a chimeric MS/MS spectrum (Figure 5-4A). In contrast, in DMS-SWATH (Figure 5-4B), the difference in charge of the two analytes allows DMS to easily separate the two peptides, with T22$^+$ being recovered at CV = 1 V and T50$^{2+}$ recovered at CV = -5V. The result is two MS/MS spectra which show no apparent contamination by other sources. With many studies showing the ability of DMS to separate even very similar compounds it seems likely that many SWATH neighbour spectra could be resolved by applying DMS-SWATH.$^{411,422,423}$
Figure 5-4: MS/MS spectral splitting by DMS-SWATH in a BSA protein digest using SWATH and DMS-SWATH acquisition both with ACN chemical modifier. DMS-SWATH can separate precursor ions which co-elute and exist in the same SWATH window. A. BSA peptides T22 and T50 co-elute under these chromatographic conditions. This results in a chimeric SWATH MS/MS spectrum showing fragments of both peptides. B. Applying DMS-SWATH separates T22 and T50 by CV. This means that the resulting MS/MS spectra are not mixed.

5.4.4 DMS-SWATH and ACN clusters

The most effective separation by DMS is often achieved through the addition of chemical modifiers like ACN, acetone or isopropanol. As indicated in Chapter 4, however, the clustering effect that supplies the increase in peak capacity of this technique also can lead to changes to the m/z of many species caused by persistent clustering of ions into the mass
spectrometer (Chapter 4, Figure 4-2). These clusters require a modest amount of energy to clear which currently (in our system) can only be provided in the collision cell. This means methods such as IDA will not operate as expected as the survey scans (where ions have been subjected to some collision energy for proper ion transmission to the ToF sector) do not reflect the \( m/z \) condition of ions as they enter the isolating quadrupole. This phenomenon was investigated here using the peptide KGAILKGAILR under a 1.5% ACN curtain gas modifier. Without a chemical modifier, the 4+ precursor ion (285.6 \( m/z \)) can be effectively isolated and fragmented (Figure 5-5A). Upon the addition of ACN, the resulting MS/MS spectrum shows low-intensity peaks (Figure 5-5B). Specifically, the expected fragment ion peaks are present but their intensities are several orders of magnitude lower than without the ACN modifier. SWATH analysis can solve this problem if the isolation window is large enough to encompass the \( m/z \) of the precursor ion plus the \( m/z \) of any precursor + modifier molecule clusters. Using a SWATH window of 50 \( m/z \) from 280 to 330 \( m/z \) (which is large enough to encompass the 4+ ion and clusters consisting of the 4+ ion with up to 4 additional ACN molecules), the resulting MS/MS fragmentation spectrum is like the control, albeit with slightly higher peak intensities (Figure 5-5C). The increase in intensity is likely a function of the reduction in the in-source fragmentation of the 4+ ion (Chapter 4). A disadvantage of this strategy is that the SWATH window must be selected specifically to encompass the clusters and ions; for example, a more common window such as 250-300 \( m/z \) would not have the same results. DMS-SWATH provides the ultimate solution. Since the SWATH window in DMS-SWATH is so large, all the precursor clusters can be routinely recovered. Using a BSA digest matrix spiked with KGAILKGAIL peptide, DMS-SWATH was capable of producing a modest MS/MS spectrum at an LC timescale (Figure 5-5D). The number of fragments recovered in the DMS-SWATH experiment is less than in the
direct infusion experiments seen in Figure 5-5A-C, but note that only six spectra are averaged using DMS-SWATH at this timescale (compared to 50 with direct infusion). Despite the limited scan averaging available the peaks recovered are of higher intensity than those found by isolating and fragmenting the 4+ precursor ion directly. This suggests that DMS-SWATH may be an effective way to handle the possibility of unexpected clustering in DMS separation analyses.
**Figure 5-5:** Breaking ACN clusters using SWATH and DMS-SWATH on direct infusion of KGAILKGAILR peptide. DMS is on with parameters set to optimally transmit the 4+ ion. SWATH-based techniques can alleviate the persistence of modifier clusters preventing proper MS/MS isolation. A. Without ACN, the 4+ precursor of KGAILKGAILR can be isolated and fragmented to produce this MS/MS spectrum. B. Upon the introduction of ACN, isolation of the 4+ precursor at its expected \( m/z \) (285.6) produces a weak fragmentation spectrum. C. SWATH data acquisition was used with a SWATH window of 280-330 \( m/z \). This encompasses the 4+ precursor as well as the precursor plus up to 4 ACN molecules associated with it. The resulting spectrum is comparable to the control without ACN. D. DMS-SWATH can also alleviate this effect to some degree. KGAILKGAILR was spiked into a BSA digest and DMS-SWATH acquisition was performed. While limited in intensity due to the time constraints of LC acquisition, several fragment ions are recoverable at much higher intensity than a comparable selective fragmentation of the 4+ ion under ACN condition.
5.5 Conclusion

DMS-SWATH is a new method of untargeted analysis for protein digests. The combination of LC and DMS separation with SWATH-like data acquisition allows for an improvement in MS/MS auto-validation as demonstrated by the evaluation of a model mAb digest. In addition, the inclusion of DMS separation can resolve co-eluting peptides which are found in identical SWATH windows as well as providing a means of breaking clusters formed by chemical modifiers. The results reported here represent the first step in evaluating this method of untargeted analysis, focusing only on simple, single protein digests. Future work is necessary to explore the applicability of DMS-SWATH for other types of analysis.
Chapter 6
Concluding Remarks and Future Prospects

6.1 Concluding remarks

Analysis of proteins in biological samples by mass spectrometry provides a wealth of information in terms of protein identity, structure, and quantity. But a biological sample cannot simply be analyzed directly by the mass spectrometer. Appropriate sample preparation is essential to purify, concentrate, eliminate interferants, and otherwise render the sample amenable to MS. There are many different techniques to accomplish these tasks: precipitation, immuno-based separations, solid phase extraction, various liquid chromatographic methods, and gas-phase separations such as ion mobility. Any given sample could require one or more of these methods to produce an appropriate sample condition for the sophisticated MS analyses available today. Because of this, technologies and methods which can reduce the labor required in sample preparation are highly desirable. In this dissertation, I present new means to handle complex biological samples for protein analysis in two broad categories a) miniaturization and automation by DMF and b) gas-phase clean-up methods by the introduction of chemical modifiers and DMS. In the first category, I describe the applicability of the DMF platform in performing immuno-based sample preparation methods using DMF, specifically immunodepletion and immunoprecipitation, both indispensable techniques in the field of protein analysis. In the second category, I describe the use of gas-phase chemical modifiers to reduce in-source fragmentation and chemical noise in complex protein samples as well as the use of DMS gas-phase separation in conjunction with and orthogonal to HPLC to improve the sequence coverage in routine protein digest analysis. I summarize these contributions below, and conclude by suggesting avenues of potential future work.

6.1.1 Digital Microfluidic Platform for Human Plasma Protein Depletion
(Chapter 2)

This chapter describes the development of a DMF-based means to deplete samples of highly abundant proteins like albumin and immunoglobulins. We take advantage of DMF’s ability to handle solid materials on-chip and use functionalized microparticles suspended in a medium to bind human serum albumin and human IgG proteins from a sample. The depletion process is fast
and efficient, achieving >95% removal of interfering proteins in only 20 minutes. More importantly, the signal-to-noise of proteins initially present at low abundance improves by 4-fold. In comparison to commercially available systems, the DMF immunodepletion method was as efficient but could be operated with much less sample. Sample sizes were only 4 μL compared to several 100s of μL needed for more traditional systems.

The main drawback to this method is the lack of ‘real-world’ sample analysis. All testing was performed using standard solutions which lack the complications of, for instance, a clinical blood sample. Re-evaluation of the method using a blood or serum sample would be prudent. Also, many commercial depletion methods can process up to five different highly abundant proteins at once or in series. Adding differently functionalized microparticles to the chip would allow for more depletion targets. Finally, this method involved analysis by MALDI-MS. More recent improvements in DMF-MS interfacing have resulted in the possibility of direct analysis from the DMF-chip to LC-MS. Combining those technologies would be an important step toward developing an all-in-one protein sample handling and analysis system.

6.1.2 Digital Microfluidics for Immunoprecipitation (Chapter 3)

This chapter describes the development of a DMF-based immunoprecipitation platform. This chapter is an extension from Chapter 2, once again relying on DMF’s capacity for handling microparticles. In this chapter, after the microparticles bound an albumin protein target, they were acidified which caused the release of the target from the microparticles. The resulting solution was then analyzed using direct infusion ESI-MS. Recovery of the target protein was 80% with a signal improvement of 3.7-fold over a sample not processed by DMF-IP. This chapter also introduced a method of pre-concentration of microparticles on DMF chips called P-CLIP: pre-concentration using liquid intake by paper. P-CLIP facilitates the use of sample sizes >100 times larger than traditional performed by DMF and results in significant pre-concentration of analytes. This allowed for analysis of HSA concentrations down to 50 ng/mL when combined with HPLC-MS/MS.

Further development of DMF-IP would necessitate performing the technique on a more biologically relevant target. The capability to use IP to select for protein complexes should be leveraged, perhaps via a DMF-IP method for a commonly expressed protein tag such as FLAG.
This would enable DMF-IP to be deployed in a wide variety of biological applications. Furthermore, owing to DMF’s ability to perform cell culture, the prospect of a device capable of both producing cells and preparing them for analysis is real. As is the case for the work described in Chapter 2, an integrated means of connecting to HPLC-MS would further aid in maturing the platform. P-CLIP is already seeing some use outside of IP as a diagnostic tool. The greatest potential for P-CLIP is in the ability to concentrate samples down to a scale that is more appropriate for DMF actuation. So far, a ‘real-world-to-chip’ via P-CLIP has not been demonstrated. In such a scenario, magnetic particles are added to a sample such as a standard blood draw for clinical analysis. The entirety of the sample can then be concentrated on-chip by isolating the magnetic particles which have bound the analytes via a magnet. With the wide range of particle functionalization available commercially, it is a ‘buyer’s market’ as to what kinds of diagnostics might be demonstrated using this method.

6.1.3 Enhancing signal and mitigating in-source peptide fragmentation using controlled clustering by gas-phase modifiers (Chapter 4)

This chapter describes the use of DMS hardware to supply a chemical modifier to the curtain gas of a mass spectrometer. This allows the modifier to cluster with the incoming analyte ions thereby preventing their fragmentation as they enter the energetic conditions present at the inlet of the instrument. Using acetonitrile (ACN) as a modifier reduced the intensity of fragment ion peaks by approximately 3-fold. As a result, the limits of detection of fragile ions spiked into a protein digest matrix were reduced by up to 5-fold. There are two additional discoveries that arose from this project. The first is the persistence of the modifier to cluster with the ions. Analyte ions were found to exist clustered to up to 4 molecules of chemical modifier as far as the collision cell of the instrument. This suggests a limitation of chemical modifiers in DMS-based analysis with targeted MS/MS fragmentation such as multiple reaction monitoring (MRM), as the precursor ions may not be found at the correct m/z. The second is the charge preservation effect of ACN on the ions. Only ACN was found to increase the intensity of the most highly charged ion. This means that the incorporation of ACN into the curtain gas can allow recovery of highly charged ions at strong intensities for additional experiments like electron capture dissociation.
Reduction of peptide fragmentation could be of immense benefit in the field of peptidomics. Being able to isolate peptides which are biological in origin from those that are the result of in-source fragmentation is paramount in the understanding of how peptides are used in biological processes like cell signaling or cell growth. The focus of this work was on the reduction of in-source peptide fragmentation but the technique may be useful for other molecules that can fragment, as well. It would be beneficial to explore the use of chemical modifiers to broaden the scope of the technique outside of protein analysis. While the challenge of persistent clustering is somewhat resolved by Sequential Windowed Acquisition of All Theoretical fragment ion mass spectra (SWATH) based acquisition methods (see Chapter 5), it would be prudent to investigate a more robust hardware solution to clearing the clustering modifier molecules from analyte ions. This would greatly benefit many targeted, DMS-based methods which may not be achieving their highest potential due to hidden clusters. Finally, an in-depth investigation if the clustering behavior of ions with modifiers in a DMS system would be essential to further explore the physics of DMS separation. Chemical modelling to identify the probable locations of adduct association is necessary, as well as further investigation into the kinds of molecules which form these persistent clusters.

6.1.4 DMS-SWATH: A new method combining HPLC and DMS for the analysis of protein digests (Chapter 5)

This chapter describes the development of a new methodology combining HPLC with DMS separation for the improvement in MS/MS sequence auto-validation of protein digests. DMS-SWATH adds DMS-based separation in conjunction with HPLC to enhance separation and detection of peptides. DMS-SWATH showed high auto-validated peptide mass-fingerprinting of monoclonal antibody digests relative to other data acquisition methods. Additionally, DMS-SWATH could separate peptides which would ordinarily co-elute and exist in the same SWATH window, which simplified the resulting MS/MS spectra. Finally, DMS-SWATH, as well as other SWATH methods, was shown to be a useful solution to the problem of persistent clusters formed on some compounds when operating DMS with a chemical modifier.

The work described in chapter was a preliminary investigation into a new data acquisition method and much work is still needed to validate it. This will include a thorough comparison between the various data acquisition methods operating on different hardware systems to
understand whether these benefits are universal. DMS-SWATH was designed to aid in therapeutic protein analysis so it would be useful to test the method with unknowns to determine whether full characterization of such samples is possible. Testing the applicability of the technique for other sample types is needed as well. It is possible that DMS separation may be of greater importance in applications like untargeted analysis of environmental samples for contaminants where many compounds are very similar in structure and mass. The additional peak capacity provided by DMS would be necessary to separate these compounds.

6.2 The future

Protein analysis is a critical step in many important applications, both in industry and academia. Many challenges still exist in terms of sample preparation. In this thesis, I describe my work that was focused on addressing these challenges: sample preparation by DMF and MS signal enhancement using DMS and allied methods. While these methods have shown significant improvements to various stages in protein analysis, in the future I would expect a true all-in-one platform to be developed. DMF is a platform that has shown to be amenable to cell culture as well as protein analysis (as demonstrated in Chapters 2 and 3). This suggests to me that the combination of these methods could be attractive for a wide range of applications. Proteomics is one such application; DMF could allow cells to be grown, exposed to stimulants, their proteins extracted, cleaned-up and injected into the mass spectrometer for analysis, all on a single, easy to use platform.

Chapter 2 presented a means of removing interfering proteins from a sample to detect proteins of low abundance. However, many applications involving protein depletion do not focus on the detection of proteins at all. The incorporation of protein depletion tools on DMF may one day enable sample processing for applications like metabolomics or peptidomics. DMF can handle a wide variety of samples including solid tissues and the targeted protein depletion presented in Chapter 2 could be used to remove proteins that are carried over from tissue extraction methods. Since tissue samples are often in limited quantities it is imperative to develop methods which can handle very small amounts of sample. Interfacing to LC-MS would create an all-in-one microscale -omics platform. Use of gas-phase sheltering like presented in Chapter 4 would also increase the stability of these analytes while simultaneously preventing confounding signals.
caused by in-source fragmented ions. Doing so would mean that any ions detected will have been derived from biological processes as opposed to in-source fragmentation.

In Chapter 3, I presented a method of immunoprecipitation using DMF, a technique which is very popular for cell biology applications, especially useful for investigating protein complexes. The next step for DMF-IP is to extend the method to automating the analysis of protein complexes. The multiplexing capability presented by DMF would allow for a sample to be immunoprecipitated for several different targets simultaneously. The method could also be extended beyond antibody-based capture methods. Molecularly imprinted polymers and aptamers have different means of capture and release, and thus might be used to form a system that is more amenable to keeping proteins and associated molecules intact during the precipitation process. Additional sample processing steps can be accomplished using magnetic particles besides clean-up methods like those presented in Chapters 2, 3 and elsewhere. Magnetic beads with protease functionality are commonly employed and their integration with DMF based analysis is inevitable. This would add one further protein processing step to the DMF toolbox.

One challenge for integrating DMF with MS is the lack of means to implement in-line chemical separations. The simplest DMF-MS interfaces reported in the literature have been direct infusion via electrospray interfaces. This limits the samples that can be effectively analyzed as the most complex samples such as blood contain materials rendering them difficult to electrospray and/or manipulate with DMF. Incorporating DMF with LC is possible but this strategy has its own challenges in terms of device design and autosampler compatibility. Incorporating DMS with DMF is an attractive option, eliminating the need for external fluid pumps while still providing chemical separation. DMS could perform complex condensed-phase sample preparation steps while DMS could clean-up the components of the sample that could not be eliminated or were added to the solution for DMF compatibility (e.g., surfactant additives to improve droplet movement). DMS is known to be highly effective in combating chemical contamination, which seems a nice fit for the fact that biomolecule-rich samples often must contain some amount of surfactant additive for proper DMF operation. The interface between DMS and DMF could come in the form of a nanospray emitter with a curtain gas for DMS similar to the Bruker nanoBooster source. Alternatively, another new MS source called the Open Port
Probe\textsuperscript{428,429} could be even more valuable as a DMF-DMS-MS interface. Incorporating the open port on or within a DMF device should be quite facile, and operation would only require moving a DMF droplet to the port. As soon as it contacts the flowing liquid of the probe, the droplet will be carried off for analysis.

One of the drawbacks of current DMS technology for peptide and protein analysis is the limited resolution of the system. Peptides can be recovered at a broad range of CV values (between 3 and 5 V in width, Chapter 5). This means that there is a significant amount of ‘co-elution’ when separating peptides solely by DMS. Chapter 5 addresses this challenge through the incorporation of LC methods in conjunction with DMS. However, the use of LC creates extended run times and a reliance on liquid pumps and columns. Increasing the resolution of the system will limit the amount of co-elution and help to eliminate the need for LC separation in less complex systems such as single protein digests which often have fewer than 100 components. To do so, the SV and CV must be increased with control as well as widening the gap spacing between the electrodes. The difficulty will be in maintaining efficient ion transmission into the mass spectrometer at increased sizes. In conjunction with DMF sample preparation techniques like presented in Chapters 2 and 3, a high resolution DMS separation system could enable an automated, single-chip microscale analysis method for protein therapeutic characterization, greatly reducing the time and cost in identifying, verifying, and characterizing new biologics and biosimilar formulations.

Overall, this dissertation introduces and explores methods for improving biological sample handling and detection, specifically for proteins from liquid samples. These are important, incremental steps toward building robust, automated, and thorough methods for protein and peptide analysis. It is my hope that these methods will eventually find their way toward improving the ‘deep’ analysis like that performed in the proteomics field. However, the techniques described in this dissertation are not exclusive to protein samples. In fact, it would not be surprising to see the principles introduced by my dissertation adopted in fields like metabolomics or clinical diagnostics where the reductions in chemical interferences would be welcomed.
References

(1) Osborne, T. B. The vegetable proteins; Longmans: London, 1909.
(2) Fenn, J. B. Angew. Chemie Int. Ed. 2003, 42 (33), 3871–3894.
(27) Sista, R. S.; Eckhardt, A. E.; Srinivasan, V.; Pollack, M. G.; Palanki, S.; Pamula, V. K. Lab Chip 2008, 8 (12), 2188.
Protoc. 2014, 9 (1), 64–75.


(91) CALBIOCHEM. 2005, p 7.
fundamentals and instrumentation; Elsevier, 2017.


(197) Pappin, D. J. C. In Protein Sequencing Protocols; Humana Press: New Jersey, 1997; pp
165–174.


(337) Abnova. PrecipitorTM; Neihu, Taiwan, 2013.


(353) Peters, T. All About Albumin: Biochemistry, Genetics, and Medical Applications, 1st ed.;


(401) Sorensen, M.; Harmes, D. C.; Stoll, D. R.; Staples, G. O.; Fekete, S.; Guillarme, D.; Beck,


