Fundamental Studies of Protein Ionization for Improved Analysis by Electrospray Ionization Mass Spectrometry and Related Methods

Kevin A. Douglass
Western Michigan University, aartdouglass@hotmail.com

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations
Part of the Organic Chemistry Commons

Recommended Citation
https://scholarworks.wmich.edu/dissertations/240
FUNDAMENTAL STUDIES OF PROTEIN IONIZATION FOR IMPROVED ANALYSIS BY ELECTROSPRAY IONIZATION MASS SPECTROMETRY AND RELATED METHODS

by

Kevin A. Douglass

A dissertation submitted to the Graduate College in partial fulfillment of the requirements for the degree of Doctor of Philosophy Department of Chemistry Western Michigan University April 2014

Doctoral Committee:

Andre R. Venter, Ph.D., Chair
John B. Miller, Ph.D.
Yirong Mo, Ph.D.
Blair R. Szymczyna, Ph.D.
Bruce Bejcek, Ph.D.
Mass spectrometry (MS) is an analytical technique in which a sample is converted to gas-phase ions that are subsequently separated and detected. It offers great speed, selectivity, and sensitivity during analysis, characteristics which have enabled it to become a leading method for the study of proteins. The applications of MS for these biologically significant macromolecules range from accurately determining identity and sequence to shedding light on post-translational modifications and protein-molecule interactions. As a first step towards analysis by MS, gas-phase protein ions must be formed. A common method for ionization is electrospray ionization (ESI), where a liquid sample including the protein is charged, nebulized, and evaporated, resulting in bare protein ions. Although ESI has been used in this way for over two decades, many aspects of the protein charging mechanism remain unclear.

To address this problem, my research has focused on (1) identifying the factors that determine the extent of protein multiple charging during ESI and (2) improving the ionization of proteins by desorption electrospray ionization (DESI). DESI is a method similar to ESI, except that the sample is desorbed from a surface by the spray instead of being present in it from the onset. As a result of (1), a simple model was developed that enables the accurate prediction of protein multiple charging observed during ESI-MS if
the protein sequence is known. Furthermore, the enhancement of multiple charging that is observed upon the addition of certain organic reagents, a phenomenon known as supercharging, was investigated and a novel mechanism of protein supercharging was proposed. For (2), the difficulty in analyzing large proteins by DESI-MS was studied using an innovative approach where DESI was separated into its individual sub-processes and their individual contributions to the DESI process were evaluated. As a result, core limitations to the DESI-MS of large proteins were identified. The results of my cumulative research efforts should lead to the improved MS analysis of proteins by spray ionization methods, including ESI and DESI.
ACKNOWLEDGMENTS

It is hard to believe that I’m already finishing my dissertation, preparing to defend, and anticipating life after death—doctorate. Joking aside, I have enjoyed my time at Western Michigan University immensely, in no small part to the wonderful faculty and friends I have met there. I owe everyone at least a small “thank you” for the support and encouragement they have provided me; most deserve more. I would like to especially thank:

Dr. Andre Venter, for pushing me to achieve beyond what I would have on my own. You’ve helped me to be confident in my abilities and to think outside the confines of popular thought. My success as a student is a reflection of your success as an advisor.

My labmates. Shashank, Afrand, Benedicte, Amy, Kari, Gregg, Elahe, Sepideh, and Wisam, it has been a pleasure working beside you. I wish you all the best.

Pam McCartney and Robin Lenkart. Thank you for patiently helping me to stay afloat amidst the murky sea of administration. I’m sure I’ll only bother you each a dozen more times before I graduate.

My committee members. I would like to thank Dr. John B. Miller, Dr. Yirong Mo, Dr. Blair Szymbczyna, and Dr. Bruce Bejcek for many thoughtful discussions and the times you have each given me advice on my projects.

The Department of Chemistry faculty. Regardless of whether you were part of my committee or not, you were always there to give me advice or simply talk about school,
work, or life. Some of my best memories over the last few years are of conversations I had with faculty members.

Sean Bashaw and Dr. Kevin Blair. It’s your hard work and dedication to the students, more so than the nuts and bolts, that keep the instruments in this department together.

Finally and most of all, I would like to thank my amazing wife, Jessi. Your continuous love and support have helped me get through these last four years with most of my sanity remaining. Marrying you will always be the greatest accomplishment of my life.

Kevin A. Douglass
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................... ii

LIST OF TABLES ....................................................................................................................... x

LIST OF FIGURES ..................................................................................................................... xi

LIST OF ACRONYMS ................................................................................................................ xiv

PREFACE ................................................................................................................................... 1

CHAPTER

1. INTRODUCTION .................................................................................................................... 1
   1.1 A Brief History of Electrospray Ionization ...................................................................... 2
   1.2 The Dissertation Structure ............................................................................................ 5
   1.3 References ....................................................................................................................... 6

2. FUNDAMENTALS OF PROTEIN IONIZATION BY ELECTROSPRAY IONIZATION ............... 9
   2.1 Introduction to Electrospray Ionization .......................................................................... 9
   2.2 The Formation of Charged Droplets ............................................................................. 10
   2.3 The Disintegration of Charged Droplets .................................................................... 11
   2.4 Analyte Ionization ......................................................................................................... 13
   2.5 Protein Multiple Charging ............................................................................................ 13
   2.6 References ...................................................................................................................... 15

3. A NEW ALGORITHM FOR CHARGE STATE PREDICTION ............................................ 18
   3.1 Introduction ...................................................................................................................... 18
# Table of Contents---Continued

**CHAPTER**

3.2 General Background of the Extent of Protein Multiple Charging .......... 19

3.3 Experimental Details................................................................. 23

3.4 Results and Discussion .............................................................. 24

3.4.1 Coulombic Repulsion and Protein Charging .................................. 24

3.4.2 Determining the Optimum Separation between Charged Residues................................................................. 26

3.4.3 A Detailed Example of the Counting Algorithm .............................. 27

3.4.4 Evaluating the Accuracy of the Algorithm ..................................... 32

3.4.5 Implications of the New Model on Protein Charging by ESI .......... 35

3.5 Conclusion ....................................................................................... 36

3.6 References ....................................................................................... 37

4. CHARGE STATE MODIFICATION: SULFOLANE ADDUCTS AND SUPERCHARGING ................................................................. 41

4.1 Introduction....................................................................................... 41

4.2 Modification of the Charge State Distribution................................... 42

4.3 Experimental Details........................................................................ 44

4.4 Investigating Supercharging of Cytochrome c by Sulfolane Addition .... 46

4.4.1 Concentration Dependence ........................................................... 46

4.4.2 Comparison of Supercharging in the Positive and Negative Ion Modes.................................................................................. 48

4.4.3 Sulfolane Adducts in the Positive Ion Mode .................................... 51
# Table of Contents---Continued

## CHAPTER

4.4.4 Further Evidence for the Direct Interaction of Sulfolane with Protein Ions .......................................................................................................................... 53

4.4.5 Multiple Charging and the Amino Acid Sequence ........................................ 57

4.4.6 The Correlation between Supercharging and Reagent Dipole Moment .............................................................................................................................. 59

4.5 Conclusion ............................................................................................................. 62

4.6 References ............................................................................................................. 63

5. EVIDENCE FOR A DIRECT INTERACTION MODEL OF SUPERCHARGING PROTEINS IN THE POSITIVE ION MODE ........................................ 68

5.1 Introduction ............................................................................................................ 68

5.2 Modification of Protein Multiple Charging and Reagent GB ................................ 69

5.3 Experimental Details ............................................................................................ 71

5.4 Results and Discussion .......................................................................................... 75

5.4.1 Supercharging in the Negative Ion Mode ....................................................... 75

5.4.2 Supercharging in the Positive Ion Mode ........................................................ 77

5.4.3 The Mechanism of Supercharging Related to Reagent GB ............................ 80

5.5 Conclusion ............................................................................................................. 85

5.6 References ............................................................................................................. 86

6. AMBIENT IONIATION: DESORPTION ELECTROSPRAY IONIZATION (DESI) AND RELATED METHODS .......................................................... 91

6.1 Introduction to Ambient Ionization ........................................................................ 91

6.2 Mechanistic Aspects of Spray Desorption Ambient Mass Spectrometry ............... 94
Table of Contents---Continued

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2.1 The Primary Spray ...................................................................</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>6.2.2 Dissolution/Extraction ................................................................</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>6.2.3 Desorption .............................................................................</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>6.2.4 Ionization .............................................................................</td>
<td>101</td>
<td></td>
</tr>
<tr>
<td>6.3 Liquid Microjunction Methods ..................................................</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td>6.4 References ...............................................................................</td>
<td>107</td>
<td></td>
</tr>
<tr>
<td>7. DECONSTRUCTING DESORPTION ELECTROSPRAY IONIZATION........................</td>
<td>114</td>
<td></td>
</tr>
<tr>
<td>7.1 Introduction ..............................................................................</td>
<td>114</td>
<td></td>
</tr>
<tr>
<td>7.2 The Optimization of DESI .......................................................</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>7.3 Experimental Details ...................................................................</td>
<td>119</td>
<td></td>
</tr>
<tr>
<td>7.3.1 Instrumental Details ................................................................</td>
<td>119</td>
<td></td>
</tr>
<tr>
<td>7.3.2 Surfaces and Samples ................................................................</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td>7.3.3 The Methods ..........................................................................</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td>7.3.3.1 Desorption Electrospray Ionization (DESI) ............................</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td>7.3.3.2 Spray Desorption Collection (SDC) .......................................</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td>7.3.3.3 Reflective Electrospray Ionization (RESI) .............................</td>
<td>122</td>
<td></td>
</tr>
<tr>
<td>7.4 Results and Discussion ..................................................................</td>
<td>123</td>
<td></td>
</tr>
<tr>
<td>7.4.1 Choosing the Sample Surface ..................................................</td>
<td>123</td>
<td></td>
</tr>
<tr>
<td>7.4.2 Optimization of the Applied Potential .....................................</td>
<td>124</td>
<td></td>
</tr>
<tr>
<td>7.4.3 Optimization of the Nebulizing Gas Pressure ............................</td>
<td>125</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER

7.4.4 Optimization of the Sprayer-to-Impact Zone Distance .................. 128
7.4.5 Optimization of the Impact Zone-to-Inlet Distance .................... 130
7.5 Conclusion ................................................................................ 132
7.6 References ............................................................................... 133

8. INVESTIGATING PROTEIN ANALYSIS BY DESORPTION ELECTROSPRAY IONIZATION MASS SPECTROMETRY .............. 138
8.1 Introduction .............................................................................. 138
8.2 History of Protein Analysis by Desorption Electrospray Ionization .... 139
8.3 Experimental Details .................................................................. 145
  8.3.1 Sample Preparation .............................................................. 145
  8.3.2 Experimental Sequence ....................................................... 145
  8.3.3 Conditions for DESI, RESI, and SDC .................................. 146
  8.3.4 Conditions for ESI ................................................................. 146
8.4 The Mass Dependent Instrumental Response and Desorption ........... 148
8.5 Investigating Other Sources ....................................................... 151
8.6 The Effect of Ineffective Protein Dissolution ................................ 155
8.7 The Addition of Solution Additives for Improved Protein DESI ......... 159
8.8 Signal Distribution and Sensitivity ............................................. 162
8.9 Conclusion ............................................................................... 163
8.10 References ............................................................................. 164

9. CONCLUSION AND FUTURE PERSPECTIVES .......................... 169
CHAPTER

9.1 Protein Charge State Prediction and Modification ........................................... 170

9.1.1 Charge State Prediction .................................................................................. 170

9.1.2 Charge State Modification ............................................................................. 171

9.2 The Improved DESI-MS Analysis of Proteins .................................................. 174

9.3 References ........................................................................................................... 175

APPENDICES .............................................................................................................. 178

A. MICROSOFT EXCEL TOOLS FOR ANALYZING PROTEIN
   MASS SPECTROMETRY DATA .............................................................................. 178

B. CONSTRUCTION AND CALIBRATION OF DESI SPRAY SOURCES .... 187
LIST OF TABLES

3.1 Sequences of proteins expressed in-house. Directionality is from the N-terminus to the C-terminus ................................................................. 23

3.2 Information and experimental results for the proteins investigated in the current study. The data for human apo-transferrin was acquired from the literature where the spectra had been obtained from denaturing solutions. "Sequences for proteins with "---" can be found in Table 3.1 ......................... 32

4.1 Instrumental parameters used during ESI-MS. Parameters promoting adduct retention are presented in column A. Parameters in column B were used to compare the multiple charging of cyt c in the positive and negative modes. ... 45

4.2 Chemical properties of several reagents and the ACS and HOCS for cyt c obtained in ESI-MS upon their addition to a denaturing solution of cyt c. $T_b$ = boiling point, $\mu$ = experimental dipole, $\mu^*$ = calculated dipole. A “-” indicates that no data could be found.................................................. 59

5.1 Gas phase basicities and boiling points (if available) for all supercharging reagents investigated in this chapter. Values with a ‘*’ indicate that the value was calculated using the methods described in the experimental section and a ‘-‘ indicates that no data was found. For m-nitrobenzyl alcohol, the values obtained from ref were calculated using B3LYP and (full MP2) using the 6-311++G(2p,2d) basis set.................................................. 72

5.2 Proton affinities for solvent species...................................................... 84

6.1 Descriptive statistics summarizing DESI parameters found in the literature ... 96

7.1 Standard operating conditions for SDC, RESI, and DESI used in the present chapter........................................................................................................... 119

8.1 Unless otherwise specified in the text, the experimental parameters used for DESI-MS, SDC-MS, RESI-MS, and ESI-MS........................................ 148
LIST OF FIGURES

2.1 A schematic of the electrospray process in the positive ion mode. Ionic species within the solvent are enriched at the solvent meniscus due to the penetration of the electric field into the protruding solvent. This results in the formation of a Taylor cone at the capillary end and the ejection of a charged solvent jet. The jet breaks apart into small, charged droplets. Reproduced with permission from Kebarle, P.; Verkerk, U. H., Electrospray: From ions in solution to ions in the gas phase, what we know now. Mass Spectrom. Rev. 2009, 28 (6), 898-917. Copyright © 2009 Wiley Periodicals, Inc. ................................................................. 10

2.2 An example of the evaporation/fission pathway of charged droplets formed by electrospray. Reproduced with permission from Kebarle, P.; Verkerk, U. H., Electrospray: From ions in solution to ions in the gas phase, what we know now. Mass Spectrom. Rev. 2009, 28 (6), 898-917. Copyright © 2009 Wiley Periodicals, Inc. ................................................................. 12

2.3 A mass spectrum of the protein cytochrome c. ................................................. 14

3.1 The average difference between the predicted HICS and experimental HICS for 8 proteins between 6 and 27 kDa as a function of the number of residues separating like charges. Error bars represent one standard deviation. ............... 26

3.2 a) The amino acid sequence of equine myoglobin with the basic residues lysine (K), arginine (R), and histidine (H), as well as the n-terminus, highlighted in red. The numbers and brackets indicate the charge assignments to the charge groupings, which are defined in the text. b) Possible charge assignments for the residues K67-H72 (charge assignments 10 and 11). ................................................................. 28

3.3 Spectra for several proteins acquired under denaturing conditions (50/50/0.1 MeOH/H₂O/formic acid). Below each spectrum is a representation of the respective amino acid sequence, where basic residues are highlighted in red and charge groupings are underlined (representations for all other proteins are provided in Figure 3.4). The values in parenthesis indicate the ratio of the number of groupings to the total number of basic amino acids. For each protein shown, with the exception of BSA, the HICS corresponds to within one or two charges to the number of groupings of charge sites................. 29
List of Figures---Continued

3.4 Representations of the respective amino acid sequences of each the remaining proteins not presented in Figure 3.3, as well as the negative ion mode predictions, where basic (acidic) residues are highlighted in red (blue) and predicted charge groupings are underlined. The values in parenthesis indicate the ratio of the number of groupings to the total number of basic amino acids. ................................................................. 30

3.5 A correlation plot of the predicted and experimental HICS for denatured proteins for both the positive and negative (inset) modes. The dashed line represents a perfect agreement between the predicted and experimental data. Green triangles indicate observed charging for BSA (66kDa) and human apo-transferrin (79kDa) after treatment with DTT. .................................................. 33

3.6 Correlation plots for the experimental highest intensity charge state (HICS) and the predicted HICS or molecular mass for the proteins in Table 3.2 with molecular masses between 6 and 27 kDa. .............................................................. 33

3.7 (a) Ball and stick model of bovine serum albumin (BSA). (b) Ball and stick model showing only the amino acids involved in disulfide bridges. Eight of the groupings, numbered in the panel, are pairs of adjacent disulfide bridges. (c) Close-up of an adjacent pair of disulfide bridges................................. 35

4.1 (A) 10 µM cyt c in 50:50:0.1 MeOH:H₂O:FA. (B) 10 µM cyt c in 50:50:0.1 MeOH:H₂O:FA with 300 mM sulfolane added. Charged sulfolane (M) species were observed when sulfolane was added to cyt c samples (inset). Peak assignment was verified by MS/MS. ........................................... 47

4.2 The average charge state of three cyt c concentrations in 50:50:0.1 MeOH:H₂O:FA for a wide range of added sulfolane (log[S]). Irrespective of protein concentration, supercharging only became apparent at sulfolane concentrations above 10⁻³ M. Error bars represent 1 standard deviation of 3 replicate analyses. ........................................................................ 48

4.3 A comparison of positive and negative ESI-MS of 10 µM cyt c in 96:4 MeOH:H₂O with and without 100 mM sulfolane added. (A) and (B) are positive and negative mode spectra, respectively, with no sulfolane added. (C) and (D) are positive and negative mode spectra, respectively, with sulfolane added. Supercharging upon sulfolane addition was only observed in the positive mode................................................................. 49

4.4 Adducts observed with cyt c in negative ion mode. The -8 charge state is shown. Peak labels are described in the text................................................................. 50
4.5 Spectra of cyt c in in 50:50:0.1 MeOH:H_2O:FA with various additives. (A) Cyt c with no supercharging. Hydrogen sulfate adducts were observed at low intensity. (B) Cyt c with 30 nM sulfuric acid added. No supercharging was observed although hydrogen sulfate adduction increased dramatically. (C) Cyt c with 300 mM sulfolane. Hydrogen sulfate adducts were observed with much greater intensity than in the blank and sulfolane adducts were observed for the higher charge states. (D) The ratio of the sulfolane adducted signal to the total signal for cyt c plotted against the charge state. The error bars represent 1 standard deviation of 5 replicate analyses. Sulfolane adduction was seen to dramatically increase with increasing charge state.

4.6 A simplified schematic of the front ion optics for the LTQ mass spectrometer. Ions transfer from atmospheric pressure to the intermediate pressure region through the heated inlet, where a voltage applied to the tube lens focuses the ions through the skimmer cone into the high vacuum region. The dotted line shows the path of a small molecule while the dashed line shows the path of a large one.

4.7 Heat maps of denatured (a) and supercharged (b) cytochrome c. The color scale spans the range of intensities within each data set, red for low and green for high. (c) A plot of the optimum tube lens values for each charge state.

4.8 Amino acid sequence for bovine cyt c. Basic amino acids are highlighted in red. Brackets and numbers indicate groups expected to have a single charge under denaturing conditions in the absence of supercharging.

4.9 A plot of the HOCS and ACS of cyt c in 50:50:0.1 MeOH:H_2O:FA obtained with several supercharging reagents against their respective dipole moments. Refer to Table 4.2 for the reagent number key. There is a clear trend of increasing supercharging ability with increasing dipole moment.

4.10 Spectra of 10 μM cytochrome c in 50:50:0.1 MeOH:H_2O:FA supercharged with several reagents at 100 mM concentration.

5.1 Chemical structures of the supercharging reagents investigated for both the positive (1-9) and negative (A-H) ion modes.
5.2 Correlation of calculated (MP2(full)/6-311++G(2d,2p) and experimentally determined\(^{39}\) GBs for (in order from left to right) 1,3,2-dioxathiolane 2,2-dioxide, 1,3-propanesultone, ethylene sulfite, dimethyl sulfone, sulfolane, dimethyl sulfoxide, and tetrahydrothiophene 1-oxide. ................................................................. 74

5.3 Plots of the average charge state of (a) ubiquitin and (b) myoglobin supercharged in the negative ion mode with the reagents A-H in Table 1 as a function of the reagents’ GBs. Unfilled diamonds represent compounds with boiling points lower than that of water and the dashed lines represent the average charge states of the proteins without the addition of a supercharging reagent.................................................................................. 76

5.4 Plots of the average charge state of (a) ubiquitin, (b) myoglobin, and (c) cytochrome c supercharged in the positive ion mode with the reagents 1-5 in Table 1 as a function of the reagents’ GBs. (d) Cytochrome c was reanalyzed with the addition of reagents with GBs between 799 and 840 kJ mol\(^{-1}\). Unfilled diamonds represent reagents possessing a nitro group instead of sulfuryl groups. ........................................................................................................ 78

5.5 Combined extracted ion chromatogram (EIC) of the +19, +18, and +17 charge states of cyt c, which are typically enhanced upon supercharging by m-nitrobenzyl alcohol (m-NBA). (b) Averaged spectrum of cyt c with no m-NBA. (c) Averaged spectrum showing the supercharging observed at the tail of the reagent plug, where diffusion has reduced the concentration of reagent. An identical spectrum was obtained for the front. (d) Averaged spectrum showing the heavy adduction and loss of signal observed for the region of constant m-NBA concentration................................................................. 80

5.6 GB\(_{\text{app}}\) of denatured cytochrome c as function of protein charge state. Values represented by filled diamonds were calculated by Schnier et al.\(^{48}\) while values represented by unfilled diamonds were extrapolated from this data. .... 83

6.1 Popular ambient ionization methods (pink) divided by ionization process (green) and the sample processing methods (black). The ionization mechanism of each ambient method is indicated by its font color: black represents electrospray ionization (ESI), blue represents atmospheric pressure chemical ionization (APCI), and white indicates that sample processing occurs without subsequent ionization................................................................. 92
6.2 A desorption electrospray ionization (DESI) source. A pneumatically-assisted electrospray is directed toward a sample present on a surface. Sample is extracted into the charged spray and carried to the mass spectrometer inlet in reflected droplets. Ionization proceeds via the typical electrospray mechanisms.

6.3 Schematics of the (a) liquid microjunction surface sampling probe (LMJ-SSP) and (b) nano-desorption electrospray (nano-DESI) ambient ionization sources.

7.1 Schematics of (a) desorption electrospray ionization (DESI), (b) reflective electrospray ionization (RESI), and (c) spray desorption collection (SDC).

7.2 A diagram showing several parameters that are important to the DESI experiment. The spray voltage is excluded, which is typically applied between the solvent needle driving the sample and the MS system.

7.3 Results for the optimization of the applied potential for SDC (squares), RESI (triangles), and DESI (diamonds). Error bars represent one standard deviation of triplicate analyses. Lines connecting individual data points are presented only to guide the reader’s eye.

7.4 Results for the optimization of the nebulizing gas pressure for SDC (squares), RESI (triangles), and DESI (diamonds). Error bars represent one standard deviation of triplicate analyses. Lines connecting individual data points are presented only to guide the reader’s eye.

7.5 Results for the optimization of the sprayer-to-impact zone distance for SDC (squares), RESI (triangles), and DESI (diamonds). Error bars represent one standard deviation of triplicate analyses. Lines connecting individual data points are presented only to guide the reader’s eye.

7.6 Results for the optimization of the impact zone–to-inlet distance for SDC (squares), RESI (triangles), and DESI (diamonds). Error bars represent one standard deviation of triplicate analyses. Lines connecting individual data points are presented only to guide the reader’s eye.

8.1 Logarithmic plot of protein DESI-MS detection limit vs. molecular mass collected on Thermo Fischer Scientific LTQ (♢) and LCQ (⦁).
List of Figures---Continued

8.2 Schematic of the various methods used in the current study. (a) Protein mixture infused and pointed directly at MS for ESI experiments, or directed at clean surface for RESI; (b) Solvent only is infused and pointed at sample containing surface. Secondary droplets from sample surface analyzed directly for DESI analysis or collected for SDC. SDC collected material subsequently infused exactly as in (a). ................................................................. 144

8.3 (a) An Eppendorf tube was used for collection following SDC. (b) The Eppendorf tube was modified by adding two small holes to the bottom for rinsing and squeezing it into an elongated shape to improve collection efficiency. (c) The collection tube was rinsed several times with the same aliquot of solvent to reconstitute the collected sample. ............................................................... 147

8.4 Plot of the relative HICS intensities for the direct ESI-MS, SDC-ESI-MS, and DESI-MS of a protein mixture with the tube lens optimized for m/z 765. 149

8.5 Direct ESI-MS (a,e), SDC-ESI-MS (b,f), DESI-MS (c,g), and RESI-MS (d,h) of a protein mixture at low (a-d) and high (e-h) tube lens settings. Triangles mark a peak attributed to one of the four proteins. ................................. 152

8.6 Cytochrome c (red triangles) and bovine serum albumin (unlabeled peaks under the brackets) analyzed simultaneously at the tube lens values of +160 V and + 250 V for direct ESI-MS, SDC-ESI-MS, and DESI-MS. ................................. 153

8.7 (a) β-lactoglobulin spectrum acquired by direct ESI-MS. (b) β-lactoglobulin spectrum acquired by DESI-MS. (c) Cytochrome c spectrum acquired by DESI-MS ................................................................. 158

8.8 (a) Protein mixture deposited in 50/47/3 MeOH/H$_2$O/AcOH (b) The same protein mixture with the addition of 200 mM ammonium acetate. (c) The same protein mixture with the addition of 200 mM ammonium bicarbonate... 161

8.9 DESI-MS spectra for cytochrome c (a) and β-lactoglobulin (b) analyzed individually. The maximum signal intensity of the β-lactoglobulin is ~50x less than that of cytochrome c ................................................................. 162
### LIST OF ACRONYMS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACS</td>
<td>average charge state</td>
</tr>
<tr>
<td>APCI</td>
<td>atmospheric pressure chemical ionization</td>
</tr>
<tr>
<td>API</td>
<td>atmospheric pressure interface</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>CEM</td>
<td>charge ejection model</td>
</tr>
<tr>
<td>CNT</td>
<td>charge neutralization theory</td>
</tr>
<tr>
<td>CRM</td>
<td>charge residue model</td>
</tr>
<tr>
<td>CS</td>
<td>charge state</td>
</tr>
<tr>
<td>CSD</td>
<td>charge state distribution</td>
</tr>
<tr>
<td>cyt c</td>
<td>cytochrome c</td>
</tr>
<tr>
<td>DEMI</td>
<td>desorption electrospray/metastable-induced ionization</td>
</tr>
<tr>
<td>DESI</td>
<td>desorption electrospray ionization</td>
</tr>
<tr>
<td>DICE</td>
<td>desorption ionization my charge exchange</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EASI</td>
<td>easy ambient sonic-spray ionization</td>
</tr>
<tr>
<td>ECD</td>
<td>electron capture dissociation</td>
</tr>
<tr>
<td>EESI</td>
<td>extractive electrospray ionization</td>
</tr>
<tr>
<td>ESI</td>
<td>electrospray ionization</td>
</tr>
<tr>
<td>ETD</td>
<td>electron transfer dissociation</td>
</tr>
<tr>
<td>FT-ICR</td>
<td>Fourier transform ion cyclotron resonance</td>
</tr>
</tbody>
</table>
List of Acronyms---Continued

GB .................................................................................... gas-phase basicity
GB_{app} ................................................................................ apparent gas-phase basicity
GB_{int} ................................................................................ intrinsic gas-phase basicity
GFP ..................................................................................... green fluorescent protein
HICS .................................................................................... highest intensity charge state
HOCS .................................................................................. highest observed charge state
IEM ........................................................................................ ion evaporation model
MAII .................................................................................. matrix-assisted inlet ionization
MALDI .................................................................................. matrix-assisted laser desorption
MeOH ...................................................................................... methanol
m-NBA .................................................................................. m-nitrobenzyl alcohol
MS ........................................................................................ mass spectrometry
MW ......................................................................................... molecular weight
m/z ......................................................................................... mass-to-charge ratio
LESA .................................................................................. liquid extraction surface analysis
LMJ-SSP .............................................................................. liquid microjunction surface sampling probe
NCBI ..................................................................................... national center for biotechnology information
PDB ...................................................................................... protein data bank
PMMA .................................................................................. poly(methyl methacrylate)
PTFE .................................................................................... polytetrafluoroethylene
List of Acronyms---Continued

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESI</td>
<td>reflective electrospray ionization</td>
</tr>
<tr>
<td>RSD</td>
<td>relative standard deviation</td>
</tr>
<tr>
<td>SDC</td>
<td>spray desorption collection</td>
</tr>
<tr>
<td>TLC</td>
<td>thin layer chromatography</td>
</tr>
<tr>
<td>TM-DESI</td>
<td>transmission mode desorption electrospray ionization</td>
</tr>
<tr>
<td>TNT</td>
<td>trinitrotoluene</td>
</tr>
<tr>
<td>TOF</td>
<td>time-of-flight</td>
</tr>
</tbody>
</table>
PREFACE

The present work describes my efforts to progress the science of protein ionization for MS analysis, particularly by investigating the electrospray ionization (ESI) and desorption electrospray ionization (DESI) mechanisms. The results presented herein were obtained during the last four years under the guidance of Dr. Andre Venter at Western Michigan University and my doctoral committee, and most of it has been published in leading peer-reviewed journals, including The Journal of the American Society for Mass Spectrometry, Analytical Chemistry, and The Journal of Mass spectrometry. Chapters that are comprised primarily of one of my previously published works are preceded by an acknowledgment referencing the source material. My sincere thanks go out to the original copyright holders for their kind permission in letting me use this material in my dissertation.
“How now will the future reckon with these beams? Is biology the next arena in which they will exercise their powers? Wielded by sufficiently alert and able advocates will they coax the marvelously complex molecules of living systems into surrendering vital secrets?”

Meng, Mann, and Fenn

CHAPTER 1

INTRODUCTION

1.1 A Brief History of Electrospray Ionization

Mass spectrometry (MS) is an analytical technique in which a sample is converted to gas-phase ions that are subsequently separated and detected in a high-vacuum environment using electric and magnetic fields. It offers great speed, selectivity, and sensitivity, characteristics which have enabled it to become a leading method for chemical analysis, in particular for biologically significant macromolecules such as proteins. Early MS analysis was mostly limited to small, volatile compounds which could be thermally vaporized without degrading. In the late 1960s, it was realized by Malcolm Dole at Northwestern University, IL, that non-volatile macromolecules such as polystyrene could be easily vaporized and ionized by a process known as electrospray ionization (ESI), where a dilute liquid sample is charged, nebulized, and evaporated, resulting in intact ions that can be transferred into the vacuum region of the MS. Nearly twenty years later, John Fenn, a professor in the Department of Mechanical Engineering at Yale, demonstrated that ESI could be used to ionize protein molecules for MS analysis, important work that was recognized by the 2002 Nobel Prize in Chemistry. Since that
time, ESI has become one of the most popular methods of protein ionization for MS analysis.

ESI-MS enables a researcher to easily and quickly introduce a protein sample to a mass spectrometer, whether directly or after chromatographic separation. Depending on the capabilities of the mass spectrometer, a number of important protein analyses can be performed. It was soon realized that a major benefit of protein ESI is the high level of multiple charging that is typically observed, which opens up the field of protein analysis to instruments with limited mass ranges such as those with ion trap or quadrupole mass analyzers, since increased charge reduces the m/z value of an ion. With accurate mass instruments, such as those using time-of-flight (TOF), orbitrap, or Fourier transform ion cyclotron resonance (FT-ICR) mass analyzers, protein parent masses can be determined with accuracies as low as 1 ppm. High accuracy enables protein identification with high confidence as well as accurate sequence information when fragmentation strategies are employed. These types of analyses usually require highly charged ions, which is a signature ability of ESI that is not typically achieved through other ionization methods such as matrix-assisted laser desorption ionization (MALDI). High accuracy mass determination also enables the study of protein modifications such as mutations or post-translational modifications, as well as covalent or non-covalent interactions including protein-protein or protein-molecule interactions.

By the time Dole had adopted the electrospray phenomenon for use in analytical chemistry, it had already been long studied by scientists primarily interested in aerosol research. Hence, a good deal of the physics surrounding the generation and evolution of an electrospray was already known. For instance, the electrospray phenomenon was first...
described in 1917, while the systematic disintegration of charged droplets, such as those generated by an electrospray, was described as early as 1882. The concept of charging an analyte molecule using electrospray was novel, though, and nothing was yet known about this analyte charging process. In an effort to explain how molecules become ionized by ESI, Dole originally postulated the Charge Residue Model (CRM), which proposes that if an electrospray solution is sufficiently dilute, the droplet breakup process that occurs during ESI will eventually result in small, charged droplets each containing a single analyte molecule. Upon final solvent evaporation, charges present on the droplet can be transferred to the analyte.

Interestingly, although the CRM was the first model introduced for analyte charging by ESI and was developed before any experimental evidence was collected, it remains perhaps the most widely-adopted theory for the charging of macromolecules by ESI to this day. Other ionization models have been proposed and continue to be developed, including Iribarne and Thomson’s Ion Evaporation Model (IEM) and Kaltashov’s emission model. There are many important details of protein ionization by ESI that are not yet understood. Thus, it continues to be an area of research which is actively debated within the mass spectrometry community. Topics of particular interest pertain to where, when, and how charges are acquired by protein molecules and which factors determine the extent of multiple charging that is characteristic of proteins ionized by ESI. In fact, it may appear that we are sure of very little regarding protein ionization! This has made for an exciting period of research on protein ionization for mass spectrometry analysis, which I am grateful to be a part of. The following chapters, divided into two main sections, follow my work from fundamental studies of protein
charge state determination and charge state modification during ESI (Part I) to protein analysis using spray-based ambient ionization methods (Part II), which share many characteristics of traditional ESI.

1.2 The Dissertation Structure

Part I begins in Chapter 2 with a review of the fundamentals of protein ionization by ESI. This chapter will lay the foundation and provide context for Part I and, to a lesser extent, Part II. Chapter 3 describes a simple charge state prediction algorithm which can be used to accurately determine the extent of multiple charging observed for protein ions formed by ESI. Unlike previous methods, it does not require the use of computers and only requires that a protein’s amino acid sequence be known. In Chapter 4, the modification of protein charge state distributions is discussed, comprised mainly of original work investigating the supercharging of cytochrome c with the reagent sulfolane. Supercharging is the phenomenon whereby the extent of multiple charging can be increased for protein ions formed during ESI, usually upon the addition of small amounts of certain reagents to the electrospray solution. Highly charged protein ions are desirable for ETD or ECD fragmentation experiments since the information content of such analyses are improved over those using lower charge state ions. Additionally, some mass analyzers perform better at lower m/z ratios, for example the orbitrap. Part I finishes with Chapter 5, which builds upon the work presented in Chapter 4 in the pursuit of a direct interaction model of supercharging, where we propose that supercharging efficacy for any particular reagent is related to a direct interaction between the supercharging reagent and the protein.
Part II begins with Chapter 6, which covers the history of ambient ionization with particular attention paid to desorption electrospray ionization (DESI). The goal of ambient ionization is to introduce samples for mass spectrometry analysis with as little sample preparation as possible, ideally none. DESI, introduced in 2004,\textsuperscript{18} was one of the first such methods and has been widely adopted by the mass spectrometry community. However, protein analysis by DESI remains a challenge due to a mass-dependent response which decreases with increasing protein mass. To study this problem, novel methods of analysis were developed which enabled us to study the sub-processes of DESI individually in order to determine the effect that each has on the overall process.\textsuperscript{19} This work is presented in Chapter 7. The application of these novel methods of analysis towards the investigation of the mass-dependent response observed for proteins is presented in Chapter 8.\textsuperscript{20} Finally, the dissertation concludes with my final remarks and future perspectives in Chapter 9.

1.3 References


CHAPTER 2

FUNDAMENTALS OF PROTEIN IONIZATION BY ELECTROSPRAY IONIZATION

2.1 Introduction to Electrospray Ionization

There are many reasons for electrospray ionization (ESI) becoming one of the most widely used ionization methods for mass spectrometry. Firstly, ionization occurs at atmospheric pressure, so samples do not need to be transferred to a high vacuum environment prior to ionization. Secondly, it requires that the sample is in the form of a conducting liquid, making ESI a natural choice for coupling liquid chromatographic separations with mass spectrometric detection, an incredibly powerful combination. Thirdly, ESI is amenable to the ionization of a large range of compounds. Since ionization typically occurs by protonation/deprotonation during ESI, any chemical species with even slight Brønsted base or acid properties is likely able to be ionized. And yet, even other compounds without Brønsted base or acid properties can become ionic species by forming stable interactions with cations or other positively-charged species such as ammonium, or negatively-charged species such as chloride. This leads to an important fourth characteristic: ESI is a soft ionization method, meaning that ionization is not very energetic and no fragmentation of the parent molecule occurs as a result. In fact, it is gentle enough that noncovalent intermolecular interactions, such as those holding protein-ligand, protein-protein, and other multimeric complexes together, can be preserved into the gas phase. Finally, ESI often leads to multiple charging for compounds with multiple potential charging sites, such as proteins, nucleic acids, and other polymers.
In particular, these final two traits have had an important impact on protein analysis by mass spectrometry.

In general, electrospray ionization is comprised of three main steps. In the first step, the liquid sample is converted to charged, airborne droplets, or nebulized, at the electrospray emitter. Secondly, these droplets undergo a succession of evaporation and disintegration events to form many smaller charged droplets, some of which contain the analyte of interest. Finally, analyte molecules are ionized and released to the gas-phase, where they are subsequently transported into and through the mass spectrometer to be analyzed.

2.2 The Formation of Charged Droplets

![Schematic of the electrospray process](https://example.com/schematic.png)

**Figure 2.1.** A schematic of the electrospray process in the positive ion mode. Ionic species within the solvent are enriched at the solvent meniscus due to the penetration of the electric field into the protruding solvent. This results in the formation of a Taylor cone at the capillary end and the ejection of a charged solvent jet. The jet breaks apart into small, charged droplets. Reproduced with permission from Kebarle, P.; Verkerk, U. H., Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **2009**, *28* (6), 898-917. Copyright © 2009 Wiley Periodicals, Inc.

Figure 2.1 presents a schematic of the electrospray process, the first step of ESI. An electrospray is initiated by applying a high electric potential between a conducting
liquid sample eluting from a capillary and a ground electrode, usually the mass spectrometer inlet. The electric field is very high at the tip of the small capillary opening and penetrates into the sample extruding from the capillary exit. Due to the high electric field, ions within the solvent are separated. These ions could be charged analyte molecules, components of solvent additives such as salts or organic acids, or ionized species resulting from oxidation or reduction of the solvent occurring at the interface between the metal emitter and the sample. Electrolyte species of the polarity opposite that of the ground electrode are pulled to the solvent meniscus, deforming it into a cone called a ‘Taylor cone’.  

If the field strength is high enough, the pull on the electrolytes becomes strong enough to overcome the solvent surface tension and a thin solvent jet forms at the tip of the Taylor cone which expels solvent carrying an excess of charge. The Coulombic repulsion between like charges then causes the jet to break apart, forming a plume of many, small, charged droplets.

2.3 The Disintegration of Charged Droplets

The charged droplets formed by the electrospray, each carrying a set amount of excess charge of a single polarity, drift through the air towards the counter electrode, usually the mass spectrometer inlet. Collisions of the droplets with the atmospheric gas lead to solvent evaporation and droplet shrinkage. As the droplets shrink and increase in curvature, the electric field normal to the droplet surface increases as the droplet charge remains constant. Eventually, the field strength at the droplet surface will reach a critical value at which point repulsion between like charges present on the droplet becomes high enough to overcome the attractive surface tension force. This value is determined by the Rayleigh limit criterion for charged droplet instability.
\[ q^2 = 64\pi^2\varepsilon_0\gamma a^3 \]  

(2.1)

where \( q \) is the droplet charge, \( a \) is the droplet radius, \( \varepsilon_0 \) is the permittivity of free space, and \( \gamma \) is the surface tension of the droplet. The resulting instability leads to a droplet fission event, where many small progeny droplets are expelled from the droplet in a jet carrying a disproportionate amount of charge (~15%) from the parent droplet relative to their mass (~2%).\(^5\) This evaporation/fission process is illustrated in Figure 2.2. If a parent or progeny droplet is small enough in radius, such as droplets formed after several successive fission events, then charge evaporation begins to occur instead of droplet fission. In this case, the charge of the droplet is maintained below the Rayleigh limit by charges lifting from the droplet until the droplet is completely vaporated.\(^6\)

![Figure 2.2](image-url)  

**Figure 2.2** An example of the evaporation/fission pathway of charged droplets formed by electrospray. Reproduced with permission from Kebarle, P.; Verkerk, U. H., Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **2009**, *28* (6), 898-917. Copyright © 2009 Wiley Periodicals, Inc.
2.4 Analyte Ionization

At some point along the repetitious droplet evaporation/fission pathway, analyte molecules originally present in the electrosprayed solution become ionized. Historically, the mechanism of ion formation during ESI has been believed to occur along one of two pathways, the charge residue model (CRM) introduced by Dole in 1968\textsuperscript{7} or the ion evaporation model (IEM) introduced by Iribarne and Thomson in 1976.\textsuperscript{6,8-9} According to the CRM, the solvent evaporation and successive fission events result in the formation of progeny droplets containing a single analyte molecule, which becomes charged through a charge transfer process between the charge carriers on the surface of the droplet and the analyte molecule upon final droplet evaporation.\textsuperscript{7} According to the IEM, charge-bearing analyte molecules present on a charged droplet’s surface can be lifted into the gas phase by a strong electric field normal to the droplet surface.\textsuperscript{8} Although other theories have been proposed,\textsuperscript{10-11} at the current time it is believed that large globular molecules such as proteins are predominantly charged by a CRM-like process\textsuperscript{10, 12-13} whereas smaller molecules are primarily charged by the IEM process.\textsuperscript{14-15}

2.5 Protein Multiple Charging

During ESI, molecules can acquire more than a single charge. This is especially true if the particles are large, like proteins. The number of charges on a molecule is called its ‘charge state’ (\(z\)). There may be several different charge states that are possible for a molecule, and the collection of all the charge states observed in the spectrum is called the ‘charge state distribution’ (CSD). For a given set of experimental conditions, the CSD will appear approximately normally distributed\textsuperscript{16} with a characteristic highest observed charge state (HOCS) and highest intensity charge state (HICS) at positions in the mass
Accordingly, the intensity of each charge state is indicative of its probability. The most probable charge state will have the highest intensity peak in the mass spectrum (HICS), but other peaks will still be observed corresponding to charge states with lower probabilities. For example, Figure 2.3 shows a mass spectrum of cytochrome c, a 12 kDa protein. Several peaks are observed symmetrically distributed around the HICS, or the +15 charge state. In reality, even at low concentrations billions of molecules or particles are being analyzed every second, so the mass spectrum gives a very accurate depiction of the probability of each charge state. It should be noted that the charge states observed in the gas phase do not necessarily reflect the ions’ net charge in solution.\textsuperscript{17–18} While it has been previously common in the literature to refer to the HOCS as the ‘maximum charge state’, HOCS removes any ambiguity between the charge state with the highest intensity (HICS) and the charge state

![Figure 2.3 A mass spectrum of the protein cytochrome c.](image)
with the highest number of charges (HOCs). In the following chapter, it will be shown that these charge states are able to be predicted based on our current understanding of protein ionization by ESI-MS.

2.6 References


CHAPTER 3

A NEW ALGORITHM FOR CHARGE STATE PREDICTION

Reprinted (adapted) with permission from

K. A. Douglass, A. R. Venter.
Predicting the Highest Intensity Ion in Multiple Charging Envelopes Observed for Denatured Proteins during Electrospray Ionization Mass Spectrometry by Inspection of the Amino Acid Sequence.
Analytical Chemistry 2013, 85, 8212.

Copyright 2013 American Chemical Society

3.1 Introduction

In this chapter, a simple, manual method for predicting the highest intensity charge states (HICS) of denatured protein ions generated by electrospray ionization (ESI) based on an inspection of the proteins’ amino acid sequence is proposed. The HICS is accurately predicted by identifying groupings of nearby basic amino acids in the positive ion mode, or acidic amino acid residues in the negative ion mode. The method assumes that the likelihood of having more than one charge per group of proximal potential charge sites becomes less likely due to Coulombic repulsion of like charges. It is shown empirically that a spacing of at least three non-charged residues is required between charged amino acids for the charge state with the highest intensity. Verification of this method is presented and its limitations are identified. It is fast, inexpensive, and provides similar, although less detailed, information as state of the art methods that rely on computational calculations. With a few exceptions, the HICS of the investigated proteins were predicted to within an average of one charge state of the experimental data. For those proteins whose HICS were not accurately estimated, the experimental values below
those of the predictions. Upon reduction of the disulfide bonds in these proteins, the experimental HICS became closer to the predicted values, suggesting that charging lower than the prediction can be attributed to conformational inflexibility of those proteins.

3.2 General Background of the Extent of Protein Multiple Charging

Electrospray ionization mass spectrometry (ESI-MS) has become an attractive tool for the study of proteins and other biological macromolecules in part because it is a soft ionization technique that can preserve solution-phase analyte characteristics and interactions into the gas phase\(^1\) as well as produce extensive multiple charging so that proteins can be analyzed even on instruments with limited mass ranges. Protein conformation prior to ionization has a profound effect on the degree of multiple charging observed during ESI-MS analysis,\(^2\) which enables ESI-MS to be used for the investigation of protein conformation in solution. Protein molecules in their native or folded state typically have charge state distributions (CSD) present at lower m/z values relative to their denatured, or unfolded, state. The cause of this shift from higher to lower m/z values upon unfolding is attributed to several factors, including an increase in the steric accessibility of possible charge sites,\(^3\)-\(^5\) a reduction in the electrostatic repulsion upon increased distance between charge sites,\(^6\)-\(^7\) and the disruption of charge-neutralizing interactions between oppositely-charged residues.\(^8\) Extensive work relating to the ionization mechanisms for both folded and unfolded states has been done, and understanding the mechanisms of protein charging during ESI-MS has enabled the highest observed charge state (HOCS), or the charge state with the highest number of charges, to be predicted with confidence.\(^5\), \(^8\)-\(^13\)
For folded proteins, ionization by ESI is usually described in terms of the charge residue model (CRM). According to this model, the maximum number of charges on a protein ion is determined by the number of charges able to be supported on a solvent droplet having a similar surface area as the molecule, which in turn is determined by the Rayleigh equation (Equation 2.1) and therefore dependent on the solvent surface tension. Additional support for this model can be inferred from the good log-log correlation between the solvent accessible surface area and the average charge state determined for a large set of proteins, since it is supposed that the charged droplet evaporates to the size and shape of the protein molecule. However, some studies have shown that the multiple-charging of proteins cannot always be explained by surface tension alone, indicating that other mechanisms must play a part. One such mechanism is described by the conformation-dependent neutralization theory (CNT), where the extent of multiple charging relies on the neutralization of oppositely-charged residues, which, in folded proteins, can be found within the protein interior as part of ion-ion pairs. Recently, a theory compatible with both the CRM and CNT which relies only on the apparent gas-phase basicity ($\text{GB}_{\text{app}}$) of the protein relative to that of the solvent was proposed.

For unfolded proteins the HOCS usually appears limited by the number of basic residues (lysine, histidine, and arginine) present on the protein, although charging beyond this value can be observed under certain experimental conditions. It is also generally agreed that the extent of multiple charging is determined by the GB$_{\text{app}}$ of the protein ions relative to the GB of the solvent. Using the positive ionization mode as an example, each amino acid residue of a protein molecule possesses an intrinsic GB,
which has been found to be higher by 50 to 70 kJ/mol than that of the free amino acid due to stabilizing intramolecular interactions in the polymer chain.22-23 With each additional charge gained by the protein, the GBs of the remaining residues are decreased by the resulting Coulombic repulsion. The GB\textsubscript{app} of a particular charge state is assumed to be that of the residue with the highest GB after accounting for the additional Coulombic energy added by the presence of all the existing charges. The HOCS is determined by the last charge state that has a GB\textsubscript{app} higher than that of the solvent. For several proteins, both folded and unfolded, the change in GB\textsubscript{app} with increasing charge has been measured\textsuperscript{13, 23} and/or calculated\textsuperscript{10-11, 23} to decrease linearly with an increasing number of charges. Good agreement between the HOCS of protein ions and the GB of the solvent has been demonstrated.\textsuperscript{10-11, 23}

Alternatively, it has been proposed that the multiple charging of denatured proteins might occur during ejection from charged droplets in a manner similar to the ion evaporation model (IEM) proposed by Iribarne and Thomson,\textsuperscript{24} where the extent of multiple charging (HOCS) would be determined by the size, or more specifically the charge density, of the originating droplet.\textsuperscript{25-26} According to this charge ejection model (CEM), protein molecules migrate immediately to the surface of the droplet upon unfolding to minimize unfavorable interactions between the polar solvent and the newly-exposed nonpolar residues. One of the terminal ends then emerges from the droplet, followed by the sequential ejection of the rest of the chain. During ejection, charges equilibrate between the droplet and the protein on a timescale much faster than the ejection process.\textsuperscript{26} Molecular dynamics studies validating the CEM were able to
accurately reproduce experimental data for cytochrome c and myoglobin without considering the gas-phase basicities of potential charge sites.\textsuperscript{27}

Less work on the origin of the highest intensity charge state (HICS) has been done than for the HOCS. The HICS is of particular interest for quantitative protein analysis since it is the most intense peak in the CSD. Inferences can be made based on the explanations proposed for the width of the CSDs for denatured proteins. Typically, CSDs for folded proteins are much narrower (encompass fewer charge states) than those for unfolded proteins.\textsuperscript{3} One possible explanation suggests that the heterogeneity of conformations available to unfolded proteins leads to an increase in the width of the charge state distribution.\textsuperscript{2-3,27} While folded proteins possess a few unique conformations, unfolded proteins, or proteins with various degrees of unfolding, possess numerous conformations able to encompass a larger range of protonation.\textsuperscript{8,28} The most prominent conformation or subset of conformations resulting in the same extent of multiple charging would determine the HICS. Another hypothesis suggests that since ESI produces a range of droplet sizes which can support a larger or smaller number of protons depending on how large or small the droplets are, protein ions which originate from these droplets would therefore have a larger or smaller degree of protonation, respectively.\textsuperscript{27} In this case, the HICS would be indicative of the mean droplet size. In this chapter, a novel method of predicting the HICS for ions of denatured proteins formed by ESI is described which proposes that the HICS depends on the Coulombic repulsion between nearby like charges.
Table 3.1 Sequences of proteins expressed in-house. Directionality is from the N-terminus to the C-terminus.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilson’s disease protein domains 5 &amp; 6</td>
<td>APQKCFLQIKGMTSASSVSNIERNLQKEAGVLSVLYALMAGKAEIKYDPEVIQPLEIAQFIQDGLGFEAAAVMEDYAGSDGNIETITGMTSASSVHNIESKLTRTNGITYASVALATSKALKVFDPEIIIPDDIHKIEEIGHASLAQ</td>
</tr>
<tr>
<td>Green fluorescent protein</td>
<td>GTCSTTLIAIAGMTASCAGHSIEGMSQLEGVQQISVSLAEGTATVLYNPAVISPEELRAIAEDMGFEASVVSCESTNPLGNHSAGNSMVQTDDGTPTSLQEVAPHTGRLPANHAPDLAKSPQSTRAVMPKHEFSVDMTCGFGCAEAVSRVLKGGVVKYDIDLPNKVCIESEHSMDTLATLKKTGTKTVSYLGLE</td>
</tr>
<tr>
<td>Xanthine alkaloid methyltransferase</td>
<td>GSSHHHHHHHSSGLVPRGSHMVDVKVLCMNTGEGEYLLNSKFTNTIAKSIPTLKRAIESLPFEHELHNLVADLGCASTSTSNITMPITIVQTVVNRCELHNIPEFQYFLNLPSDFNLFLKGLNGLGVSIEEFENTSLVMGAPGSHLRFLPNLTIHVYSVHLSKVPDLKDEKINGPNKGTFTYIKSTSPGVREAYLAQFKDDTLFLKSSRAEMYGNRVVLYLHGLRSQDFSEKEKLQPLWILLSKIASRLVSKGWDEEKLDFEVPYYTPSEQEVKLVEGESYAVELMFTFTKDGARNEGIWSDARGFNYLRSFTETMISHHHFPGQILDELYDEIHNLPQDFATQCSFVGLKRN</td>
</tr>
</tbody>
</table>

3.3 Experimental Details

Bovine aprotinin, bovine ubiquitin, bovine cytochrome c, bovine hemoglobin, equine myoglobin, bovine serum albumin, and dithiothreitol (DTT) were purchased from Sigma-Aldrich (St. Louis, MO, USA). β-lactoglobulin and α-chymotrypsinogen were purchased from MP Biomedical (Solon, OH, USA). Green fluorescence protein (GFP), a mutant form of the N-terminal domains 5 and 6 of the Wilson disease protein (WDP-5,6), and a xanthine alkaloid methyltransferase were expressed and purified in-house. Sequence information for these proteins is provided in Table 3.1. Formic acid and methanol were purchased from Fluka Analytical (St. Louis, MO, USA). Water was provided in-house from a Purelab Ultra water purification system (Elga LLC, Woodridge, IL, USA). All reagents were used as received without any additional purification. With the exception of aprotinin and the proteins expressed in-house, all other protein
sequence data was obtained from the National Center for Biotechnology Information or the RCSB Protein Data Bank (PDB).

A linear ion-trap mass spectrometer (LTQ, Thermo Scientific, Waltham, MA, USA) equipped with an Ionmax pneumatically-assisted electrospray source was used for all ESI-MS analysis. Protein samples were prepared to 10 μM in either 50/50/0.1 MeOH/H₂O/formic acid for positive mode analysis or 80/19/1 MeOH/H₂O/NH₄OH for negative mode analysis. The instrumental parameters included a spray voltage of +5 or -4 kV for the positive or negative ion mode, respectively, a capillary voltage of +/-20 V, a capillary temperature of 250°C, a sheath gas flow rate of 10 arbitrary units, and a sample delivery rate of 5 μL/min. The tube lens value is an important mass-dependent ion transmission parameter for the LTQ and is further investigated in Chapter 4. Because of the large range of protein masses investigated in the present study, this value was set to the optimum for each protein with a typical value of around +/-150 V. The spectra were generated by averaging approximately 50 scans for each protein.

For proteins treated with DTT prior to analysis, DTT was added to an aqueous solution of the protein so that the concentrations of the protein and DTT were 20 μM and 10 mM, respectively. The solutions were analyzed after 12-24 hours at room temperature. MeOH and formic acid were added so that the final solvent composition was 50/50/0.1 MeOH/H₂O/formic acid with a protein concentration of 10 μM.

3.4 Results and Discussion

3.4.1 Coulombic Repulsion and Protein Charging

The extent of multiple charging (HOCS) for a denatured protein by ESI has been estimated by considering the electrostatic modification of the protein’s GB_{app} upon
multiple charging.\textsuperscript{10-11,23} It was shown that when basic residues are present as a grouping of two or more nearby residues, the intrinsic GBs (\(GB_{int}\)) of the residues decrease with the presence of charges within the grouping because of the additional Coulombic repulsion between like charges. The additional Coulombic energy term modifying the GBs is inversely proportional to the distance \((r)\) between the charges \((q)\) on sites \(i\) and \(j\) according to Coulomb’s law,

\[
GB_{app} = GB_{int} - \sum \frac{1}{4\pi \varepsilon_0 \varepsilon} \frac{q_i^2}{r_{ij}}
\]  

(3.1)

where \(\varepsilon_0\) is the permittivity of free space and \(\varepsilon\) is the relative permittivity of the surrounding medium. The largest Coulombic contributions appear when two charges are supported on adjacent or nearby residues because of the inverse relationship between the energy and distance between charges, which causes the Coulomb potential to vary significantly at short distances. For instance, the spacing between the \(\alpha\)-carbons of consecutive amino acid residues is approximately 4 Å.\textsuperscript{26} Coulomb potentials for charge separations of 8 and 12 Å are 86.8 and 57.9 kJ/mol, respectively, assuming a relative permittivity of two.\textsuperscript{23}

Based on these published results, it seems probable that a grouping of nearby residues on a protein chain would end up with only a single charge because of the closeness of the basic sites, and so we recently proposed that the HICS of an unfolded protein can be predicted without any calculations or modeling simply by inspection of the amino acid sequence followed by an assignment of charges based on charge site proximity.\textsuperscript{32} According to this method, the HICS is predicted in the positive ion mode by assigning the maximum number of charges to the basic residues and the N-terminus, with the provision that at least three residues separate each charge. Similarly for the negative
ion mode, the maximum number of charges is assigned to the acidic residues aspartic acid, glutamic acid, and the C-terminus, again with at least three residues separating each charge. The requirement for a minimum of 3 separating residues was determined empirically as shown in Figure 3.1.

3.4.2 Determining the Optimum Separation between Charged Residues

Charge separations of 0 to 5 residues were evaluated for their predictive accuracies. For each charge separation, the average difference between the predicted and experimental HICS (ΔHICS) was calculated for the 8 proteins in Table 3.2 between 6 and 27 kDa,

\[
\Delta HICS = \frac{\sum_{i=1}^{n}(x_{i,p} - x_{i,e})}{n}
\]  

(3.2)

where \(x_{i,p}\) is the predicted HICS for the \(i\)th protein, \(x_{i,e}\) is the experimentally observed HICS, and \(n\) is the number of proteins included in the calculation. The best results were obtained for minimum separations of 3 residues with an average ΔHICS of -0.25 ± 1.83

Figure 3.1 The average difference between the predicted HICS and experimental HICS for 8 proteins between 6 and 27 kDa as a function of the number of residues separating like charges. Error bars represent one standard deviation.
charge states. At separations less than 3 residues, the HICS becomes increasingly overestimated, for example a separation of 2 results in an average $\Delta$HICS of $+0.88 \pm 1.46$ charge states, while at separations greater than 3 residues the HICS becomes increasingly underestimated, for example a $\Delta$HICS of $-2.00 \pm 2.20$ charge states for a separation of 4. Therefore, a minimum charge separation of 3 residues was used within the current study. The relatively small difference in the accuracy between the minimum charge separations of 2 or 3 residues could indicate that the exact separation depends somewhat on the identity of the separating residues, as this would influence the real distance between nearby potential charge sites and their individual basicities, as discussed below. However, this refinement would make this simple HICS prediction method too complicated to achieve by direct inspection using simple algorithms, and would begin to require extensive computational approaches.

3.4.3 A Detailed Explanation of the Counting Algorithm

Following the example given in Figure 3.2a for myoglobin, and starting at the N-terminus side, basic residues G1, K16, H24, R31, H36 and K42 (charge groups 1-6) are spaced more than 3 positions apart and are likely to be protonated in the HICS. Residue K45 is not assigned a charge because there are less than three residues separating it from K42. When multiple basic or acidic amino acids are spaced closer together than 3 positions on the sequence, it is sometimes possible that more than one charge assignment exists that results in the same maximum number of charges overall. For example, only one of the basic residues K47-K50 (charge group 7) can be charged in order to maintain a minimum three-residue separation from the other charges, but which residue is assigned the charge does not affect the overall count. This method does not contain the necessary
level of detail to determine which specific amino acid in a group will become charged. Some flexibility exists in group assignment as long as the maximum number of charge sites are identified. For example K45 is not shown grouped together with K42 in the example because doing so would change the appearance of charge group 7, but without changing the maximum number of charges. This flexibility in group assignment is further illustrated in Figure 3.2b, which shows different ways to place charges on nearby basic residues for the residues K77-H82 of the myoglobin sequence shown in Figure 3.2a (charge sites 10 and 11). Only the bottom three assignments are considered valid for predicting the HICS by this method, because the aim is to count the maximum number of possible charge sites using the algorithm outlined above. For simplicity sake, only one permutation for charge assignment to each group is shown for each sequence in Figures 3.3 and 3.4. Following this method for the entire sequence, the HICS for myoglobin is predicted to be the +20 charge state (Figure 3.2). This prediction is in good agreement with the experimental
Figure 3.3 Spectra for several proteins acquired under denaturing conditions (50/50/0.1 MeOH/H₂O/formic acid). Below each spectrum is a representation of the respective amino acid sequence, where basic residues are highlighted in red and charge groupings are underlined (representations for all other proteins are provided in Figure 3.4). The values in parenthesis indicate the ratio of the number of groupings to the total number of basic amino acids. For each protein shown, with the exception of BSA, the HICS corresponds to within one or two charges to the number of groupings of charge sites.
**Positive Ion Mode:**

Bovine aprotinin:

Bovine hemoglobin a-chain:

GFP:

Xanthine alkaloid methyltransferase:

Human apo-transferrin

**Negative Ion Mode:**

Bovine aprotinin:

Bovine ubiquitin:

Bovine cytochrome c:

Bovine hemoglobin a-chain:

Equine myoglobin:

**Figure 3.4** Representations of the respective amino acid sequences of each the remaining proteins not presented in Figure 3.3, as well as the negative ion mode predictions, where basic (acidic) residues are highlighted in red (blue) and predicted charge groupings are underlined. The values in parenthesis indicate the ratio of the number of groupings to the total number of basic amino acids.
result shown in Figure 3.3c, which shows an ESI-MS spectrum for bovine myoglobin acquired under denaturing conditions.

The simplest way to implement this predictive method in practice is to start from the N-terminus side of the sequence and assign a charge to every possible charge site (basic residues and the N-terminus for the positive ion mode, acidic residues and C-terminus for the negative ion mode), ensuring that at least three residues separate each new charge from the previous. Using this method the maximum number of charges was always found. The maximum charge does not depend on whether one starts assigning from the N- or C-terminus, although a different charge assignment could result. The maximum assignable number of charges was shown to be the best predictor for the HICS for the proteins investigated in this study.

The remaining panels in Figure 3.3 show the distributions of basic amino acid residues in the sequences of four other proteins and the ESI-MS spectrum of each, also acquired under denaturing conditions. Basic residues are highlighted in red and groupings of nearby basic residues are underlined assuming at least 3 uncharged amino acids between each charged amino acid. The amino acids marked with a caret (^) indicate charged amino acids using the simplified charge assignment strategy described in the previous paragraph. Bear in mind that this method does not provide the means to predict which residue in a group is most likely to be protonated during an actual ESI experiment, only how many charges will be present in the highest intensity charge state. The predicted charge distributions for the other positive ion mode proteins, as well as the negative ion mode examples, are presented in Figure 3.4. Sequence information for each
protein and a comparison of predicted versus experimental HICS is presented in Table 3.2.

Table 3.2 Information and experimental results for the proteins investigated in the current study. The data for human apo-transferrin was acquired from the literature\textsuperscript{22} where the spectra had been obtained from denaturing solutions. *Sequences for proteins with “---” can be found in Table 3.1.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence ID\textsuperscript{a}</th>
<th>Mass (kDa)</th>
<th>Positive Ion Mode HICS</th>
<th>Negative Ion Mode HICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Aprotinin</td>
<td>Ref [29]</td>
<td>6.6</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>2 Ubiquitin</td>
<td>444791</td>
<td>8.6</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>3 Cytochrome c</td>
<td>NP_001039526.1</td>
<td>12.3</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>4 Hemoglobin (α chain)</td>
<td>NP_001070890.2</td>
<td>15.1</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>5 WDP-5.6</td>
<td>---</td>
<td>16.0</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>6 Myoglobin</td>
<td>NP_001157488.1</td>
<td>17.0</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>7 α-Chymotrypsinogen</td>
<td>2CGA_A</td>
<td>25.7</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>8 GFP</td>
<td>---</td>
<td>26.9</td>
<td>24</td>
<td>28</td>
</tr>
<tr>
<td>9 Xanthine alkaloid methyltransferase</td>
<td>---</td>
<td>42.6</td>
<td>38</td>
<td>47</td>
</tr>
<tr>
<td>10 Bovine serum albumin</td>
<td>CAA76847.1</td>
<td>66.4</td>
<td>68</td>
<td>45</td>
</tr>
<tr>
<td>11 Human apo-transferrin</td>
<td>2HAU_A</td>
<td>79</td>
<td>70</td>
<td>36</td>
</tr>
</tbody>
</table>

3.4.4 Evaluating the Accuracy of the Algorithm

The HICS for 11 proteins in the positive mode spanning a mass range of 6 to 79 kDa were estimated by inspection of the amino acid sequence and compared to their experimentally observed HICS. The correlation plot in Figure 3.5 shows good agreement between the predicted and experimental HICS for the 8 proteins with molecular masses between 6 and 27 kDa. The $R^2$ value for the correlation in this region is 0.94 (Figure 3.6a). For comparison, a molecular weight fit yields $R^2 = 0.63$ (Figure 3.6b). Similarly, the HICS was estimated to within an average of 1.6 charge states of the experimental values for several of these proteins in the negative ion mode (Figure 3.5, inset).

The experimental HICS falls short of the predicted HICS for the proteins in the set with the highest masses: BSA (66kDa) and apo-transferrin (79 kDa). A similar
observation was made by Loo and coworkers while investigating the correlation between the HOCS and the number of basic residues of a protein. In their study, the HOCS for proteins with masses 36 kDa or less were accurately predicted, while the HOCS for proteins with masses 36 kDa or less were accurately predicted, while the
proteins 66 kDa or larger were observed below their predicted values. After the larger proteins were treated with 1,4-dithiothreitol (DTT), a reagent used to reduce disulfide linkages, the experimental HOCs were increased significantly, although they remained below the predicted values. Their experiment demonstrates the importance of a protein’s complete unfolding in both minimizing the Coulombic repulsion between like charges supported on the same protein molecule and providing solvent access to previously inaccessible areas of the protein. Similarly, we believe that the extent of multiple charging for the large proteins investigated during our experiment was less than predicted because of incomplete denaturing, due either to disulfide linkages creating conformational inflexibility or residual tertiary structure. Figures 3.3e-f show the mass spectrum of BSA before and after treatment with DTT. After treatment, the HICS of BSA shifted from +45 to +58, a 39% increase in charging. Although the HICS is still less than the predicted HICS of +77, this represents a significant increase in multiple charging. BSA has a unique network of disulfide linkages: of the 17 disulfide linkages in BSA, 16 form 8 pairs of adjacent linkages (Figure 3.7). It is possible that treatment with DTT did not result in complete reduction of all the disulfide bonds for BSA. For human apo-transferrin, Thevis et al. observed an increase of the HICS from +36 to +89 upon treatment with DTT, bringing the HICS to within 11% of the predicted value. This observation may make it possible to use this method to estimate the degree of conformational flexibility of proteins based on the difference between their predicted and observed HICS and could potentially serve as a rapid quality control measure.
3.4.5 Implications of the New Model on Protein Charging by ESI

Our results might also have implications for the origin of the charge state distribution. It is expected that most conformations would not resemble the linear chain implied by our model. Charge states lower than the HICS might arise from protein conformations where fewer potential protonation sites (basic amino acids) are solvent accessible, or if the distance between some of the exposed potential protonation sites is reduced due to preserved protein secondary or tertiary structure, which would decrease the $\text{GB}_{\text{app}}$ of the protein ion. Ion mobility studies have shown that protein ions with lower charge states exhibit more compact formations in the gas-phase than those with higher
Charge states higher than that of the HICS occur when the number of charges exceeds the number of the charging groups. These charge states represent ions where one or more groupings of nearby basic residues on the amino acid chain contain more than one charged residue. Recall that the HOCS is usually several charge states higher than the HICS and is determined by the highest available charge state with a GB\textsubscript{app} higher than that of the solvent. Charge states higher than the HICS are still accessible based on differences between the GB\textsubscript{app} of the protein and the GB of the solvent up to the HOCS. Their intensity is reduced relative to the HICS due to the diminishing difference between the GB\textsubscript{app} of the charge states and that of the solvent as well as the additional Coulombic repulsion between nearby like-charges and thus decreased interaction potential\textsuperscript{7} between charges and potential charge sites. Williams et al. determined that the effect of existing charges on the interaction potential of possible charge sites occurs for distances $\leq 10$ Å.\textsuperscript{7} There is evidence that for higher charge states, residues other than lysine, histidine, or arginine can also become charged,\textsuperscript{6, 23, 36} which could also explain charge states higher than the HICS, or even the HOCS.

### 3.5 Conclusion

The simple count of groupings of nearby basic or acidic amino acid residues in protein sequences accurately predicts the highest intensity charge state (HICS) for the majority of denatured proteins used in the present study. For those proteins whose HICS were not accurately predicted by this method, we have shown that this is likely due to some conformational inflexibility of these proteins.

This method does not account for all energy contributions arising from intramolecular interactions or from residual secondary and tertiary structure, which could
decrease the distance between charges as well as the GB_{app} of the protein. The accuracy of the model may be improved by considering the actual amino acid residue sequence in each group because of the differences in the lengths of the basic residue side groups as well as their respective gas-phase basicities, topics for further investigation. However, these considerations would increase the complexity of the model and begin to resemble some of the sophisticated computational techniques already developed,^{10-11, 26} detracting from the aim of the simple algorithm to provide a fast and easy estimation of the HICS.

We foresee several uses for the method presented here. The prediction of the HICS allows one to determine how much of the charge envelope would fall within the mass range of the analyzer and might prove useful in designing quantitative protein analyses. This method also provides information regarding a protein’s conformation, or degree of conformational flexibility, based on the agreement between the expected and experimental HICS. While the method presented in the current study can only be applied to denatured proteins with known sequences, an increasingly large set of proteins have sequence data that are publicly available through protein sequence databases such as the NCBI or PDB.^{30-31}

3.6 References


4.1 Introduction

In this chapter, work is presented\textsuperscript{1} which studied the shift in the charge state distribution of the protein cytochrome c (cyt c) upon the addition of small amounts of reagent, a phenomenon known as ‘supercharging’. Both the highest intensity charge state (HICS) and highest observed charge state (HOCS) are shifted to values higher than those predicted by the methods described in Chapter 3. In particular, the popular supercharging reagent sulfolane\textsuperscript{1-8} was investigated. It was discovered that the addition of sulfolane on the order of 1 mM or greater to denaturing solutions of cyt c results in supercharging independent of protein concentration over the range of 0.1 to 10 µM. Also, while supercharging was observed in the positive ion mode, no change in the charge state distribution was observed in the negative ion mode for this reagent, which ruled out polarity-independent factors such as protein conformational changes or solvent surface tension effects. Instead, a series of sulfolane adducts observed with increasing intensity concurrent with increasing protein ion charge state suggested that a direct interaction
between sulfolane and the charged sites of cyt c plays an important role in supercharging. We investigated the possibility of charge delocalization occurring through large-scale dipole reordering of the highly polar supercharging reagent, which could reduce the electrostatic barrier for proximal charging along the cytochrome c amino acid chain. To support this theory, supercharging was shown to increase with increasing dipole moment for several supercharging reagents structurally related to sulfolane.

4.2 Modification of the Charge State Distribution

In practice, the extent of multiple charging for proteins in ESI-MS is dependent on several experimental parameters. Solvent composition determines protein conformation and charge availability according to the surface tension and Rayleigh limit, and can impose an upper limit on charging based on gas-phase basicity. Denaturing solutions, such as those with high acid, base, or alcohol content, result in CSDs shifted to higher charge states due to protein unfolding. Interactions with other molecules in the gas phase, whether as part of the solvent system or introduced, can shift CSDs to higher or lower charge states due to charge transfer reactions. Some instrumental settings, such as source voltages, sprayer tip diameter, and gas pressures, can also shift the CSD.

There is a special case of CSD modification called “supercharging”, where small amounts of additives such as m-nitrobenzyl alcohol, sulfolane, or other reagents are added to electrosprayed solutions, usually containing proteins, which results in a dramatic enhancement of multiple charging. Increased multiple charging of proteins has many benefits, such as reducing the mass range required for protein analysis as well as
improved top-down protein sequencing.\textsuperscript{4, 26} First reported in 2000,\textsuperscript{27} supercharging has been the subject of considerable debate.

Initial results from the Williams’ group, obtained with denatured proteins in acidified solution, suggested a rough relationship between the surface tension of the supercharging reagent and the average charge state.\textsuperscript{11-12} As supercharging reagents enrich in the droplet during solvent evaporation owing to their low volatility, there is a concomitant increase of the surface tension of the droplet. This increased surface tension would increase charge availability according to the Rayleigh limit theory (Equation 2.1).

However, subsequent work from Grandori and coworkers indicated that in the absence of conformational changes, multiple charging is largely independent of surface tension,\textsuperscript{25, 28-29} a result that was later supported by Loo et al.\textsuperscript{30}

Recently, and possibly due to an increasing interest in using ESI-MS as a tool for top-down proteomics and for studying protein interactions, the majority of reported supercharging research has investigated the charge enhancement of proteins in the native state. In 2009, Loo et al. determined that the supercharging of noncovalent protein complexes from aqueous solutions does not appear to be caused by conformational changes in solution.\textsuperscript{30} However, soon afterwards Williams et al. showed similarities between results obtained from thermal denaturation and from supercharging of aqueous solutions, suggesting that supercharging from aqueous solutions is due to conformational changes in the electrospray droplets resulting from droplet heating.\textsuperscript{31} It was proposed that more extended protein conformations result in a final droplet size with larger charge availability according to the Rayleigh limit theory (Equation 2.1). While several studies have since appeared that support the role of conformational changes occurring late during
ESI as a potential source of supercharging in aqueous solutions\textsuperscript{6-7,32}, other studies have questioned this model.\textsuperscript{8,30} It is clear that there is still a good deal of uncertainty that surrounds the supercharging phenomenon.

As presented in this chapter, work was carried out to investigate the supercharging of cytochrome c with the reagent sulfolane and other compounds.\textsuperscript{1} By ensuring that the protein was originally denatured in the electrospray solution, properties other than protein conformation could be investigated as potential sources of supercharging by removing any uncertainty about conformational changes. First, the dependence of supercharging on reagent concentration is presented. Second, differences between supercharging in the positive and negative ion modes are discussed. Third, reagent adduction is correlated to protein charge state, suggesting a direct interaction between the reagent and protein charge sites. Fourth, additional evidence for a direct interaction is presented as the tube lens voltage is compared to charge state intensity for supercharged and unsupercharged samples. Finally, the dipole moment of several supercharging reagents are compared to their supercharging efficacies.

### 4.3 Experimental Details

Bovine cyt c, sulfolane, sulfolene, and dimethyl sulfoxide were purchased from Sigma-Aldrich, LC-MS grade methanol and formic acid were purchased from EMD Chemicals, and water was obtained from Fluka. 3-chlorothietane-1,1-dioxide (CAS 15953-83-0) was purchased from Synthonix (Wake Forest, NC). All chemicals were used as received without further purification.

Stock solutions of cyt c were prepared in water followed by dilution to the specified concentration and solvent composition of 50:50:0.1 MeOH:H\textsubscript{2}O:formic acid
Table 4.1. Instrumental parameters used during ESI-MS. Parameters promoting adduct retention are presented in column A. Parameters in column B were used to compare the multiple charging of cyt c in the positive and negative modes.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>A</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capillary temperature (°C)</td>
<td>180</td>
<td>270</td>
<td>270</td>
</tr>
<tr>
<td>Capillary voltage (V)</td>
<td>+43</td>
<td>+43</td>
<td>-43</td>
</tr>
<tr>
<td>Tube lens (V)</td>
<td>+110</td>
<td>+110</td>
<td>-250</td>
</tr>
<tr>
<td>Spray voltage (kV)</td>
<td>+5</td>
<td>+3.5</td>
<td>-3.5</td>
</tr>
<tr>
<td>Solvent flow rate (µL/min)</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Sheath gas (arb. units)</td>
<td>5</td>
<td>25</td>
<td>25</td>
</tr>
</tbody>
</table>

(FA) unless otherwise noted. Samples were prepared daily at room temperature and were analyzed immediately following preparation. For samples containing supercharging reagents, aqueous solutions containing the supercharging reagent were added as part of the water fraction of the cyt c samples.

For electrospray ionization experiments, a linear ion trap mass spectrometer (LTQ, Thermo Scientific, CA) equipped with the Ionmax ESI source was used. The MS was tuned by automated routine on the m/z 941 (+13 charge state) peak of the cyt c envelope in the positive mode and m/z 1746 (-7 charge state) in the negative mode. Instrumental parameters are given in Table 4.1. The position of the sprayer tip relative to the mass inlet was optimized manually for the optimum signal stability and intensity. Initially, relatively mild ionization conditions were chosen to retain adducts during ion formation (parameter set A in Table 4.1). For comparing positive and negative mode supercharging, as well as comparing several supercharging reagents, the sheath gas and capillary temperature were increased to reduce adduct formation and give cleaner spectra (parameter set B in Table 4.1). Samples were introduced by infusion and spectra were obtained by spectral averaging with 1 minute acquisition times using automatic gain.
control. Trap fill times were typically 1 to 10 ms. Protein charge states were determined using the formula\(^\text{33}\)

\[
z_i = \frac{m_{i-1}+1}{m_{i-1} - m_i}
\]  

(4.1)

where \(z_i\) is the charge state of the \(i^{\text{th}}\) charge state, \(m_i\) is the mass-to-charge ratio of the \(i^{\text{th}}\) charge state, and \(m_{i-1}\) is the mass-to-charge ratio of the \((i-1)^{\text{th}}\) charge state. The average charge state \(z_{\text{avg}}\) was determined using the formula\(^\text{34}\)

\[
z_{\text{avg}} = \frac{\sum_{i=1}^{N} z_i W_i}{\sum_{i=1}^{N} W_i}
\]  

(4.2)

where \(N\) is the total number of protein charge states observed in the mass spectrum and \(W_i\) is the signal intensity of the \(i^{\text{th}}\) charge state.

4.4 Investigating Supercharging of Cytochrome c by Sulfolane Addition

4.4.1 Concentration Dependence

The alcohol and pH induced denaturing of cytochrome c were previously studied.\(^\text{14, 35-36}\) In the present solvent system of 50:50:0.1 MeOH:H\(_2\)O:FA, cyt c is expected to obtain an expanded, highly helical conformation.\(^\text{37}\) The ESI-MS of 10 \(\mu\)M cyt c in Figure 4.1A shows a bimodal distribution mostly centered on the +15 charge state, confirming that the protein is primarily in the denatured state. Figure 4.1B shows the shift in the CSD from an average of +15 to +21 with the addition of 300 mM sulfolane to 10 \(\mu\)M cyt c. Of the several known supercharging reagents, sulfolane was chosen for this
study because it showed the highest supercharging ability in a recent publication by Lomeli et al.\textsuperscript{5} and is miscible with water. Several peaks corresponding to charged sulfolane products were observed. The inset in Figure 4.1B shows the low m/z range of the spectrum of a typical 10 µM cyt c sample to which 300 µM sulfolane was added. A base peak at m/z 121 is observed, corresponding to [M+H]\textsuperscript{+}, where M is sulfolane, as well as peaks at m/z 153 and m/z 241, corresponding to [M+CH\textsubscript{3}OH\textsubscript{2}]\textsuperscript{+} and [2M+H]\textsuperscript{+}, respectively.

Figure 4.2 shows the change in average charge state for several concentrations of cyt c with the addition of increasing amounts of sulfolane. High concentrations of
supercharging reagents are required to observe any significant supercharging effect. For known reagents, it has been demonstrated that maximum supercharging is achieved when the concentration of the reagent is in the millimolar range,\textsuperscript{5,13} well above the micromolar analyte concentration typical in ESI-MS. In the present study, no supercharging was observed for cyt c until the sulfolane concentration of the sample was on the order of 1 mM or higher. It appears that the onset of supercharging depends on the concentration of sulfolane largely independently of the concentration of the protein, at least over the range of 0.1 to 10 µM cyt c. However, the concentration of sulfolane is 10^3 to 10^5 times higher than that of the protein and thus is present in large molar excess.

4.4.2 Comparison of Supercharging in the Positive and Negative Ion Modes

It has been proposed that an increase in surface tension is the origin of supercharging for denatured proteins.\textsuperscript{11-12} However, several groups have also reported that supercharging appears to be independent of surface tension in the absence of other effects such as denaturing.\textsuperscript{25, 28-30} One way to test the effect of surface tension on

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Figure4.png}
\caption{The average charge state of three cyt c concentrations in 50:50:0.1 MeOH:H\textsubscript{2}O:FA for a wide range of added sulfolane (log[S]). Irrespective of protein concentration, supercharging only became apparent at sulfolane concentrations above 10^{-3} M. Error bars represent 1 standard deviation of 3 replicate analyses.}
\end{figure}
supercharging is to compare supercharging in the positive and negative modes. Very few studies of supercharging in the negative mode have been reported.\textsuperscript{6, 12} Some supercharging reagents, such as glycerol, that increase charging in the positive mode have little to no effect in the negative mode\textsuperscript{6} or even reduce charging\textsuperscript{12}. Here we also studied the effect of sulfolane on supercharging in the positive and negative modes for cyt c (Figure 4.3). 10 μM cyt c with and without 10 mM sulfolane was prepared in 96:4 MeOH:H\textsubscript{2}O, a solvent system suitable for positive and negative modes. Cyt c has 24 amino acid side-chains including the N-terminus and 13 acidic amino acid side-chains including the C-terminus. Using the methodology described in Chapter 3, cyt c is predicted to have a HICS of $+14$ in the positive ion mode and $-7$ in the negative ion.

\textbf{Figure 4.3} A comparison of positive and negative ESI-MS of 10 μM cyt c in 96:4 MeOH:H\textsubscript{2}O with and without 100 mM sulfolane added. (A) and (B) are positive and negative mode spectra, respectively, with no sulfolane added. (C) and (D) are positive and negative mode spectra, respectively, with sulfolane added. Supercharging upon sulfolane addition was only observed in the positive mode.
mode. Without sulfolane added, the spectra of cyt c in the positive (Figure 4.3A) and negative (Figure 4.3B) modes reflect this difference in available charge sites with a highest intensity charge state (HICS) of +12 in the positive mode and a HICS of -7 in the negative mode. Upon addition of sulfolane, the HICS in the positive mode shifts to +15 (Figure 4.3C) while no change in the HICS in the negative mode was observed (Figure 4.3D). Several adducts were observed in the negative ion mode which correspond to losses of 44 Da as well as water and methanol adducts (Figure 4.4). The subscripts on the labels, corresponding to the numbers below, denote i or j:

1. \([\text{M-i44-8H}]^{8-}\) Possible loss(es) of i CO\(_2\) molecule(s)
2. \([\text{M-i44-8H+H}_2\text{O}]^{8-}\) with water adduct
3. \([\text{M-8H+jMeOH}]^{8-}\) j methanol adduct(s)

This disparity for positive and negative ion modes is not consistent with surface tension being responsible for supercharging of proteins from denaturing solutions. If an increase in surface tension is the cause of supercharging, then it should occur in both the positive
and negative modes since surface tension’s effect on the Rayleigh limit is independent of
the polarity of the charge.\textsuperscript{38}

4.4.3 Sulfolane Adducts in the Positive Ion Mode

It was noted that high concentrations of sulfolane result in protein spectra abundant with adducts, especially under typical instrumental conditions such as parameter set A presented in Table 4.1. To study these adducts, the instrumental conditions were selected to minimize adduct loss. Figure 4.5 contains several plots of intensity against charge state for samples of cyt c with various additives. Without the addition of supercharging reagent, a series of low intensity peaks corresponding to cyt c with one or more adducts was observed (Figure 4.5A). Upon deconvolution these peaks were assigned a mass of 97 Da. It is known that protein solutions are prone to contain sulfate, a remnant of the protein purification process.\textsuperscript{39} When the intensity of the protein with sulfate adducts is plotted for each charge state, it can be seen that the CSD of the adducted series follows very closely to that of the CSD of the protein without adducts, but shifted one charge state lower. These adducts are thus most likely charge-neutralizing hydrogen sulfate. No hydrogen sulfate adducts were observed on the HOCS. During supercharging, these adducts were observed to increase in intensity concurrently with increasing sulfolane concentration. Additional sulfate may be produced by the oxidation of sulfolane at the positive electrode (the electrospray capillary in positive mode) to sulfur dioxide and subsequently to the sulfate ion.\textsuperscript{40} Supercharging and hydrogen sulfate adduct formation were observed to increase simultaneously upon increasing concentration of sulfolane (Figure 4.1B). A sample of cyt c was analyzed containing 30 nM sulfuric acid directly as an additive (Figure 4.5B). No supercharging was observed
upon the addition of sulfuric acid, though hydrogen sulfate adduct formation increased remarkably. The result confirms that hydrogen sulfate presence and adduction is not significant to the supercharging process.

While closely inspecting the spectra for samples containing high concentrations of sulfolane (e.g. Figure 4.1B), another series of adducts was observed. Unlike hydrogen sulfate, these adducts were only observed on higher charge states. A plot of the intensity of the adducted (triangles) and nonadducted peaks (circles) is shown in Figure 4.5C. Deconvolution of these adducts revealed a mass difference of 120, corresponding to the

Figure 4.5 Spectra of cyt c in 50:50:0.1 MeOH:H₂O:FA with various additives. (A) Cyt c with no supercharging. Hydrogen sulfate adducts were observed at low intensity. (B) Cyt c with 30 nM sulfuric acid added. No supercharging was observed although hydrogen sulfate adduction increased dramatically. (C) Cyt c with 300 mM sulfolane. Hydrogen sulfate adducts were observed with much greater intensity than in the blank and sulfolane adducts were observed for the higher charge states. (D) The ratio of the sulfolane adducted signal to the total signal for cyt c plotted against the charge state. The error bars represent 1 standard deviation of 5 replicate analyses. Sulfolane adduction was seen to dramatically increase with increasing charge state.
mass of sulfolane. Adducts of supercharging reagents were previously reported for high charge states of supercharged proteins.\textsuperscript{5,30} Sulfolane adducts were present even up to the +24 peak of the protein CSD, which is the HOCS for cyt c during the present study. Bovine cyt c has 24 basic residues; 2 arginines, 3 histidines, 18 lysines, and the n-terminus. Since no charging beyond +24 was observed in the present system, even under highly-acidic conditions, it seems plausible that the amount of attainable supercharging may be limited by the number of basic residues present on the denatured protein available for protonation. Thus, sulfolane adducts are observed for the highest possible charge state for cyt c in the present system. When the ratio of the sum of the intensities of these sulfolane adducts to the total signal of the corresponding charge state is plotted against the charge state, it can be seen that the most adduct formation occurs for the higher charge states and decreases as the charge state is decreased (Figure 4.5D). Sulfolane adducts are likely weakly bound by noncovalent interactions and are believed to mostly be lost before detection.

\textbf{4.4.4 Further Evidence for the Direct Interaction of Sulfolane with Protein Ions}

It cannot be confirmed using normal methods of mass spectrometric analysis whether protein ions are heavily adducted prior to mass selection and detection if these adducts do not survive these processes. Fortunately, the interface of the LTQ mass spectrometer used in the current study has a component of the ion optics system called the tube lens, which guides ion exiting the heated inlet capillary to the skimmer nozzle orifice separating the intermediate and high vacuum regions of the mass spectrometer (Figure 4.6). The tube lens voltage is an adjustable parameter that aids in desolvation and ion transfer and is mass and charge dependent. Since heavily adducted protein ions would
have a larger mass than their un-adducted counterparts, the optimum tube lens voltage should be different for both. This property was exploited to determine whether protein ions formed from supercharging solutions were more heavily adducted than non-supercharging proteins before transfer to the high vacuum region of the mass spectrometer, where these adducts could be lost before detection.

Protein ions formed by ESI encompass a wide range of charge states. Since each protein ion has roughly the same mass but a different charge, each charge state should have its own optimum tube lens value. For this reason, the optimum tube lens value for each charge state of cytochrome c was determined. The heat maps in Figure 4.7 show the relative intensities of each individual charge state of cytochrome c, supercharged (b) and not (a), as the tube lens voltage was varied from +30 V to +250 V. Details for the generation of the heat maps is described in Appendix A. The optimum tube lens value for each charge state, defined as the tube lens value which gives the highest intensity, is plotted in Figure 4.7c. A clear shift in the maximum is observed between the two heat

Figure 4.6 A simplified schematic of the front ion optics for the LTQ mass spectrometer. Ions transfer from atmospheric pressure to the intermediate pressure region through the heated inlet, where a voltage applied to the tube lens focuses the ions through the skimmer cone into the high vacuum region. The dotted line shows the path of a small molecule while the dashed line shows the path of a large one.
Figure 4.7 Heat maps of denatured (a) and supercharged (b) cytochrome c. The color scale spans the range of intensities within each data set, red for low and green for high. (c) A plot of the optimum tube lens values for each charge state.
maps. The maximum intensities for the supercharged sample are shifted to higher charge states (lower m/z) from +15 for the non-supercharged to +19. The optimum tube lens voltage which gives the maximum intensity within each data set remains unchanged between the two samples at around +160 V.

Interestingly, when the optimum tube lens voltage for each charge state is compared between both samples (Figure 4.7c), a shift of the optimum tube lens voltage is observed. For every charge state, the supercharged sample required a higher tube lens value to achieve its maximum intensity. In the mass spectrum, the unadducted and protonated-only peak is the base peak for each charge state of both the supercharged and non-supercharged samples. If the unadducted peak for each charge state was the most common species in the intermediate pressure region when guided by the tube lens to the skimmer orifice, it would be expected that each charge state would have the same optimum tube lens voltage whether originating from the supercharged sample or not. The large shift in optimum tube lens voltage indicates that either the ions are either charge reduced while passing through this region or more massive. Charge reduction could occur if charge-neutralizing adducts such as hydrogen sulfate are attached to the protein. If these adducts are lost after transmission through the intermediate pressure region, the protein would have a higher charge at the time of detection. Alternatively, if many neutral adducts are attached to the protein then the additional mass would require a higher tube lens value for effective transmission. Because charge-reducing adducts are typically higher for lower charge states than for higher charge states\[^{39}\] and the difference between optimum tube lens voltages appears to increase with increasing charge states, the latter
explanation seems the most likely. This explanation supports an interaction between supercharging reagent and highly charged protein ions.

### 4.4.5 Multiple Charging and the Amino Acid Sequence

As described in Chapter 3, molecular modeling experiments have successfully predicted the charge states of proteins, including cyt c, observed in ESI-MS by calculating the apparent basicity of the protein in the gas phase.\(^{18, 41-42}\) Charges were assigned one at a time to the residues with the highest basicities, where the basicities of the residues were recalculated following each charge assignment. In the charge assignments reported by Peschke et al., only two charges were assigned within 5 residues of another charged site within the first 15 charges.\(^{18}\) The first two charges were assigned to the 2 arginine residues while 12 of the remaining 13 were assigned to lysine residues. These results support that charge assignment is heavily directed by electrostatic interactions, since the 18 lysine residues are differentiated only by their location along the peptide chain and their proximity to other charged sites. While cyt c was considered in the native state for their investigation, it is expected that this trend should also apply to the denatured state.

Our experimental results for denatured cyt c appear to be consistent with electrostatics-driven charging. Figure 4.8 shows the amino acid sequence for bovine cyt c with the basic residues highlighted. There are 14 distinct groups of single or proximally located basic sites where a charged basic site is at least 3 residues from another charged site. If a basic site within a group is charged, the electrostatic barrier for the charging of proximal basic residues within the same group will be elevated. The average charge state observed during ESI-MS with no supercharging reagent added was between +14 and +15.
Thus, it appears that a single charge per group is the most probable state in the typical denaturing solvent system, consistent with the results presented in Chapter 3.

When a supercharging reagent is added to an electrosprayed protein solution, the HICS increases to a level where multiple basic amino acids within a grouping have become charged. Concurrently, sulfolane adducts are observed to increase dramatically with increasing charge state. One possible explanation is that supercharging occurs through direct interactions between the supercharging reagent and charged basic sites. It has been demonstrated computationally that polar solvents are capable of diffusing charges through large-scale dipole ordering. Solvent molecules align around a charge such that the ends with the partial charges opposite that of the charged analyte are directed towards the charge. In this way, the partial opposing charges of the solvent
molecules effectively neutralize some portion of the charge on the analyte, where that portion of charge is now distributed along the periphery of this ordered solvation shell. It is supposed that several of these ordered solvation shells can form around a single charge, spreading the charge across a considerable distance. If sulfolane behaves in a similar fashion by delocalizing the charges on basic sites, the electrostatic effect on proximal basic sites could likely be reduced. Like many other identified supercharging reagents, sulfolane is highly polar with a dipole moment of 4.35 D, much higher than water’s 1.85 D or methanol’s 1.70 D.

**4.4.6 The Correlation between Supercharging and Reagent Dipole Moment**

To test the effect of reagent polarity on supercharging ability, the dipole moments of several reagents, structurally related to sulfolane, were compared to the average charge state obtained by cyt c upon addition of each reagent. This information is summarized in Table 4.2. Each sample consisted of 10 µM cyt c with 100 mM supercharging reagent in

<table>
<thead>
<tr>
<th>Reagent</th>
<th>(T_b) (°C)</th>
<th>(\mu) (D)</th>
<th>(\mu^*) (D)</th>
<th>ACS</th>
<th>HOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Water</td>
<td></td>
<td>1.85</td>
<td>2.16</td>
<td>14.3</td>
<td>19</td>
</tr>
<tr>
<td>2. Benzyl alcohol</td>
<td>205</td>
<td>1.71</td>
<td>1.79</td>
<td>14.7</td>
<td>20</td>
</tr>
<tr>
<td>3. Dimethyl sulfoxide</td>
<td>189</td>
<td>3.96</td>
<td>4.44</td>
<td>16.1</td>
<td>21</td>
</tr>
<tr>
<td>4. 3-Chlorothiete-1,1-dioxide</td>
<td>-</td>
<td>-</td>
<td>3.38</td>
<td>17.2</td>
<td>21</td>
</tr>
<tr>
<td>5. Sulfolene</td>
<td></td>
<td>4.35</td>
<td>5.68</td>
<td>17.8</td>
<td>22</td>
</tr>
<tr>
<td>6. Sulfolane</td>
<td>285</td>
<td>4.35</td>
<td>5.68</td>
<td>19.6</td>
<td>24</td>
</tr>
</tbody>
</table>
Since the dipoles of many of the supercharging reagents were not available in the literature, each dipole was calculated in Gaussian03W from an optimized molecular structure using the B3LYP method with a 6-311+g(d,p) basis set. The B3LYP method was chosen for its comparatively good dipole calculations achieved at a relatively low computational cost. For compounds with known dipoles, a consistent overestimation averaging 16% was observed.

A clear trend of increasing supercharging ability with increasing dipole moment is shown in Figure 4.9. Both the average charge state (ACS) and the HOCS increase with increasing dipole moment of the supercharging reagent. Supercharged spectra of cyt c for these additional compounds are presented in Figure 4.10. With $R^2$ values of 0.74 for the HOCS and 0.76 for the ACS, the regression lines in Figure 4.9 demonstrate that even though reagent polarity is clearly correlated with supercharging, other factors likely also contribute to supercharging ability. The interaction of supercharging reagents with charged basic sites would depend on molecular structure, such as the type and location of...
Figure 4.10 Spectra of 10 μM cytochrome c in 50:50:0.1 MeOH:H₂O:FA supercharged with several reagents at 100 mM concentration.
functional groups. For example, the most effective supercharging reagents have one or more carbonyl, sulfinyl, sulfonyl, or nitro groups present in their structure which may be important for intermolecular interactions between the supercharging reagent and the charged basic site. Lomeli et al. demonstrated that for nitrobenzyl alcohol, supercharging increases in order of the para, meta, and ortho isomers. Interestingly, this is also the order of increasing dipole moment and decreasing distance between the hydroxyl and nitro groups for this series of compounds. In the present study, sulfolene was not as effective of a supercharging reagent as sulfolane, even though their calculated dipoles were nearly identical. The presence of the double bond on sulfolene may impact how this molecule interacts with charge sites as well as other sulfolene molecules. A plot of the Connolly solvent excluded volume was against the average charge state for cyt c upon the addition of each of the reagents had an R² value of 0.35 (n = 6, p = 0.216), suggesting that reagent size is not important. Intermolecular interactions might also explain the disparity of supercharging in the positive and negative modes. Functional groups capable of interacting favorably with positively charged sites may not interact favorably with negatively charged sites. All reagents were also analyzed at the same initial concentration but display a wide range of boiling points. Differences in evaporation rates may lead to different concentrations of the supercharging reagent at the time of ion production.

4.5 Conclusion

The addition of sulfolane to denatured solutions of cyt c can lead to a drastic increase of multiple charging. Upon comparing sulfolane’s effect on cyt c in both the positive and negative modes and confirming that supercharging only occurs in the positive mode, it seems that a more direct interaction between sulfolane and cyt c, other
than denaturing or increased charge availability, could be responsible for supercharging in this system. This claim is further supported by the strong relationship between the degree of reagent adduct formation and the enhancement of multiple charging. Since the presence of adducts is observed to increase concomitantly with increasing charge state, they appear to be important to the supercharging process. We propose that through a direct interaction of sulfolane with charged basic sites a means is provided for basic sites, proximal to existing charges, to overcome the electrostatic barrier and obtain or retain charges. The exact details of this possible mechanism are currently under investigation, but one possibility is that charge stabilization is provided through charge delocalizing through large scale solvent reordering away from the basic site due to the high dipole moment of sulfolane and other supercharging reagents. The clear correlation between the dipole moments of several supercharging reagents and the extent of supercharging observed for cyt c supports this claim. In the following chapter, a direct interaction model of supercharging is further investigated by comparing the gas-phase basicities of several reagents with their supercharging efficacies using three model proteins in both the positive and negative ion modes.

4.6 References


CHAPTER 5

EVIDENCE FOR A DIRECT INTERACTION MODEL OF SUPERCHARGING PROTEINS IN THE POSITIVE ION MODE

5.1 Introduction

The addition of certain reagents during the electrospray ionization mass spectrometry (ESI-MS) of proteins can shift the protein ion signal charge state distributions to higher average charge states, a phenomenon known as “supercharging”\(^1\) (Chapter 4). In this chapter, the role of reagent gas-phase basicity (GB) during this process in both the negative and positive ion modes is discussed. Reagents with known, or calculated, GBs were added individually in equimolar amounts to protein solutions which were subsequently electrosprayed for MS analysis. Shifts in the CSDs of the protein ion signals were monitored and related to the reagents’ GBs. Trends for this data were evaluated for possible insight into a supercharging mechanism involving the direct interaction between supercharging reagent and protein ion.

Reagent GB was confirmed to be directly related to the amount of supercharging observed in the negative ion mode as previously reported.\(^2\) Supercharging in the positive ion mode, on the other hand, showed a maximal trend. Interestingly, a loss of signal and supercharging efficacy was observed for reagents with GBs intermediate within the investigated range, between ~800 and ~840 kJ mol\(^{-1}\), at the 100 mM concentration used in the present study. The results indicate that supercharging might depend on the stability of a proton bridge formed between the reagent and a highly-charged protein ion.
5.2 Modification of Protein Multiple Charging and Reagent GB

There has been a strong, continuing interest in the mechanism of protein charging during ESI-MS since the initial report by Fenn, in particular the origin and extent of multiple charging. Protein ions with high charge states are desirable for electron transfer dissociation (ETD) or electron capture dissociation (ECD) fragmentation since this increases the information content of such analyses, and some mass analyzers perform better at lower m/z ratios, for example the orbitrap. Protein conformation plays an important role in protein multiple charging; the denatured form of a protein exhibits a higher extent of multiple charging than its native state. Reasons for this include the increased accessibility of charge sites to the solvent, the removal of internal charge neutralization interactions, and the minimization of the Coulombic repulsion resulting from many like charges supported on a single protein ion. For protein ions in the denatured state, multiple charging appears to be limited by differences in the gas-phase basicity (GB) of the protein ion and the solvent. In the positive ion mode, each successive protein charge state has a lower apparent GB (GB\text{app}) than the one preceding it due to the increasing Coulombic repulsion of the existing charges. The highest charge state with a GB\text{app} higher than that of the GB of the most basic solvent component will be the highest observed charge state (HOCS) in the protein charge state distribution (CSD).

The CSD observed for an electrosprayed protein solution is sensitive to factors other than protein tertiary structure and solvent GB, including changes to the instrumental settings and the addition of certain solvent additives. Some additives can shift the CSD to higher values in a process known as ‘supercharging’. Different
mechanisms for supercharging have been described, as discussed in Chapter 4. While there is evidence that supercharging reagents can increase protein multiple charging by acting as chemical denaturants for aqueous solutions \(^{25-30}\) or by increasing the surface tensions of denaturing solutions, \(^{31-32}\) it has been shown that supercharging can occur in the absence of these effects, \(^{24, 33}\) and so other explanations are required. For instance, it has previously been shown that a direct interaction between the supercharging reagents and the protein charge sites may also contribute to increased charging. \(^{24}\)

Few studies have investigated supercharging in the negative ion mode, and those have been limited to one or a few reagents typically used as additives for positive mode supercharging. \(^{24, 26, 32}\) Even from these limited studies, it is clear that supercharging in the positive and negative ion modes have different requirements, since reagents such as glycerol \(^{26, 32}\) or sulfolane \(^{24}\), which supercharge in the positive ion mode, either do not alter charge states in the negative ion mode or sometimes even reduce charging. Recently, Ganisl and coworkers demonstrated that protein charge state distributions can be manipulated in the negative ion mode with the addition of strong organic bases to a denaturing electrospray solvent. \(^2\) In general, the average charge state was shown to increase with increasing additive GB. It has been well documented that a protein’s solution phase charge state does not determine its gas phase charge state after electrospray ionization. \(^{34-37}\) Instead, protein charging during ESI-MS typically involves the loss or gain of protons based on differences between the solvent GB and protein GB\(_{\text{app}}\). \(^{15-16}\) For this reason, it was proposed that the increase in multiple charging was a result of the deprotonation of amino acids containing a carboxylic acid side chain by the
organic bases during the charge partitioning event during the final stages of gas phase ion formation.\textsuperscript{2}

If a similar correlation between additive GB and protein multiple-charging was to be observed in the positive ion mode in the absence of other charge modifying effects, it could be evidence for a direct interaction between the additive and protein basic amino acids. Although a supercharging mechanism involving protonation reactions between the additive and protein charge site has been challenged before,\textsuperscript{28} it has yet to be systematically studied, to our knowledge. In the work presented in the rest of this chapter, the GBs of several reagents were compared to their supercharging efficacies in the negative and positive ion modes to better understand the supercharging phenomenon and protein multiple charging during ESI-MS in general.

### 5.3 Experimental Details

All supercharging reagents, proteins, and ammonium hydroxide were purchased from Sigma-Aldrich (St. Louis, MO, USA) and used without further purification. The gas-phase basicities\textsuperscript{38-39} and boiling points, if available, of all reagents are summarized in Table 5.1 and all structures are presented in Figure 5.1. 18.2 M\(\Omega\) water was obtained in-house using a Purelab Ultra water purification system (Elga LLC, Woodridge, IL, USA). MeOH and formic acid were purchased from Fluka Analytical (St. Louis, MO, USA). Samples consisted of 10 \(\mu\)M protein in 80/19/1 MeOH/H\(_2\)O/NH\(_4\)OH in the negative ion mode or 80/20/0.1 MeOH/H\(_2\)O/formic acid in the positive ion mode. For samples containing a supercharging reagent, the reagent was present at 100 mM unless otherwise indicated and added as part of the MeOH fraction.
Table 5.1 Gas phase basicities and boiling points (if available) for all supercharging reagents investigated in this chapter. Values with a ‘*’ indicate that the value was calculated using the methods described in the experimental section and a '-' indicates that no data was found. For m-nitrobenzyl alcohol, the values obtained from ref were calculated using B3LYP and (full MP2) using the 6-311++G(2p,2d) basis set.

<table>
<thead>
<tr>
<th>Reagents</th>
<th>GB (kJ mol⁻¹)</th>
<th>BP (°C)¹⁰</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Negative Mode Reagents</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Pyrazole</td>
<td>861⁴¹</td>
<td>187</td>
</tr>
<tr>
<td>B. Imidazole</td>
<td>909⁴¹</td>
<td>257</td>
</tr>
<tr>
<td>C. Diethylamine</td>
<td>919⁴¹</td>
<td>56</td>
</tr>
<tr>
<td>D. Piperidine</td>
<td>921⁴¹</td>
<td>106</td>
</tr>
<tr>
<td>E. Triethylamine</td>
<td>951⁴¹</td>
<td>89</td>
</tr>
<tr>
<td>F. 1,5-Diazabicyclo[4.3.0]non-5-ene</td>
<td>1006⁴¹</td>
<td>-</td>
</tr>
<tr>
<td>G. 1,8-Diazabicyclo[5.4.0]undec-7-ene</td>
<td>1016⁴¹</td>
<td>-</td>
</tr>
<tr>
<td>H. 7-Methyl-1,5,7-triazabicyclo[4.4.0]dec-5-ene</td>
<td>1030⁴¹</td>
<td>-</td>
</tr>
<tr>
<td><strong>Positive Mode Reagents</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Ethylene sulfite</td>
<td>776³⁹</td>
<td>-</td>
</tr>
<tr>
<td>2. Dimethyl sulfone</td>
<td>780³⁹</td>
<td>-</td>
</tr>
<tr>
<td>3. Sulfolane</td>
<td>794³⁹</td>
<td>-</td>
</tr>
<tr>
<td>4. Dimethyl sulfoxide</td>
<td>849³⁹</td>
<td>-</td>
</tr>
<tr>
<td>5. Tetrahydrothiophene 1-oxide</td>
<td>864³⁹</td>
<td>-</td>
</tr>
<tr>
<td>6. Butadiene sulfone</td>
<td>782*</td>
<td>-</td>
</tr>
<tr>
<td>7. Diethyl sulfone</td>
<td>799*</td>
<td>-</td>
</tr>
<tr>
<td>8. m-nitrobenzyl alcohol</td>
<td>800 (779)²⁸</td>
<td>-</td>
</tr>
<tr>
<td>9. m-nitrobenzamide</td>
<td>823³⁸</td>
<td>-</td>
</tr>
<tr>
<td><strong>Solvent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>725⁴⁰</td>
<td>100</td>
</tr>
<tr>
<td>MeOH</td>
<td>660⁴⁰</td>
<td>65</td>
</tr>
</tbody>
</table>
A linear ion trap mass spectrometer (LTQ, Thermo Scientific, Waltham, MA, USA) equipped with an Ionmax ESI source was used for all MS analyses. Instrumental parameters included a spray voltage of +5 or -4 kV in the positive and negative ion modes, respectively, a capillary temperature of 250 °C, a capillary voltage of +/-20 V, and a tube lens voltage between +/-120 and +/-150 V. A 10 μM protein sample was continuously delivered to the mass spectrometer at a flow of 10 μL/min while samples containing both the protein and supercharging reagent were introduced into this flow as plugs using a 6 port Rheodyne injection valve equipped with a 10 μL sample loop. This

Figure 5.1 Chemical structures of the supercharging reagents investigated for both the positive (1-9) and negative (A-H) ion modes.
arrangement ensured approximately 1 minute of steady supercharged signal while maintaining a constant concentration of the protein.

The modeling software Gaussian 09W was used for the computational determination of gas phase basicities of butadiene sulfone and diethyl sulfone, compounds for which experimentally determined values could not be found. Energies of the optimized neutral and protonated structures were calculated at the MP2(full)/6-311++G(2d,2p) level with tight convergence criteria and the GBs were calculated using the formula

\[
\text{GB}_{\text{calc}} = \Delta G[M] + \Delta G[H^+] - \Delta G[MH^+] \\
\]
where ΔE is the calculated energy and ZPE is the zero point energy. The value for ΔG[H⁺] was obtained from the literature. Using the same method, the GBs of a set of 7 sulfuryl compounds for which experimentally determined values are known were also calculated and used to generate a curve of experimental vs. calculated values with an R² of 0.997 (Figure 5.2). The resulting curve was used to correct the calculated GBs of butadiene sulfone and diethyl sulfone to values directly comparable to the experimentally determined values of the other compounds (in kJ mol⁻¹):

\[
GB_{\text{calc}} = 1.2963 \times GB_{\text{exp}} - 241.69
\]  

(5.1)

5.4 Results and Discussion

5.4.1 Supercharging in the Negative Ion Mode

Ubiquitin and myoglobin were analyzed by ESI-MS in the negative ion mode with each of the negative ion mode reagents (Table 5.1 (A-H)) and the average charge states were determined from the CSDs as previously described. The reagents were selected based on the broad range of GB values they encompassed and have all been used in a similar study of protein charging in the negative ion mode. Plots of the average charge states (ACS) of the two proteins against the GBs of the additives are presented in Figure 5.3. For ubiquitin, points for the reagents pyrazole (861 kJ mol⁻¹) and imidazole (909 kJ mol⁻¹) were not included as their addition led to a reduction in multiple charging so that most of the charge states comprising the CSD were shifted outside the mass range (data not shown), and therefore ACS values could not be accurately calculated. Similar to the results previously reported, the ACS of both proteins increase linearly with increasing GB of the additives. Two of the reagents, diethylamine and triethylamine, have lower boiling points and higher vapor pressures than that of water.
concentration of the additive in the electrospray droplets increases as the more volatile solvent components preferentially evaporate.\textsuperscript{26, 28, 31} The markers for these volatile reagents are presented as unfilled diamonds to distinguish them from the other reagents, although they appear to follow the trend. The linear increase in protein multiple charging in the negative ion mode with increasing additive GB suggests that proton transfer occurs from the traditionally acidic residues glutamic acid and aspartic acid and the c-terminus to the neutral bases in the gas phase during charge partitioning at the final moments of ion formation.\textsuperscript{2} As the additive GB increases, so does both the probability of proton transfer to the base from the protein and the net negative charge of the protein ion population.

It should be noted that since the pK\textsubscript{b} values of these compounds roughly correlate with their GB values\textsuperscript{2}, a similar trend in the data would be observed for a plot of pK\textsubscript{b} against protein average charge state. It has been widely demonstrated that solution-phase

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.3.png}
\caption{Plots of the average charge state of (a) ubiquitin and (b) myoglobin supercharged in the negative ion mode with the reagents A-H in Table 1 as a function of the reagents’ GBs. Unfilled diamonds represent compounds with boiling points lower than that of water and the dashed lines represent the average charge states of the proteins without the addition of a supercharging reagent.}
\end{figure}
net charge does not reflect a protein’s charge state in the gas phase upon ionization by ESI,\textsuperscript{34-36} for reasons discussed in Chapter 2, which is why gas-phase basicities are considered in the present discussion.

5.4.2 Supercharging in the Positive Ion Mode

Ubiquitin, cytochrome \( c \), and myoglobin were analyzed by ESI-MS in the positive ion mode with the positive ion mode reagents 1-5 in Table 5.1. These reagents were selected because they all share a similar \( \text{S} = \text{O} \) or \( \text{S}(=\text{O})_2 \) moiety and their GBs have been experimentally determined.\textsuperscript{39} In Figure 5.4, panels (a) through (c) show the change in ACS for each protein as a function of the GBs of the reagents. While each additive resulted in some amount of increased multiple charging, the trend is not linear based on reagent GB, as it was for the negative ion mode. Instead, the trend appears to be maximal with no apparent dependence of the identity of the protein. For each protein, sulfolane resulted in the highest amount of supercharging.

Sulfuryl compounds with experimentally determined GBs in the range of 794 to 849 kJ mol\(^{-1}\) were not found in the literature, leading to a gap in the data around the apparent peak maximum. The GB of butadiene sulfone, which was used in a previous supercharging study,\textsuperscript{24} was calculated to be 782 kJ mol\(^{-1}\). To search for reagents with GBs in the desired range, the GBs of a large set of commercially available sulfuryl compounds were determined computationally using the methods described in the experimental section. Only one compound, diethyl sulfone, fell into this range with a GB of 799 kJ mol\(^{-1}\), which still left a gap in the data between 800 and 849 kJ mol\(^{-1}\). For this reason, other reagents without an \( \text{S} = \text{O} \) moiety were considered. m-Nitrobenzyl alcohol (m-NBA) is currently known to be one of the most effective positive ion mode
supercharging reagents, with a calculated GB around 800 kJ mol\(^{-1}\).\(^{28}\) A literature search for other compounds possessing a similar nitro moiety yielded m-nitrobenzamide, which has a GB of 823 kJ mol\(^{-1}\).\(^{38}\) Butadiene sulfone, diethyl sulfone, m-NBA, and m-nitrobenzamide were each tested, along with the other positive ion mode reagents again, for their supercharging efficacy with cytochrome c. These additional reagents have not
yet been tested for their supercharging efficacy with ubiquitin, myoglobin, or any other proteins.

Figure 5.4d shows the ACS of cytochrome c plotted against the GBs of the expanded set of positive ion mode reagents. It should be noted that while the calculated GBs of diethyl sulfone and butadiene sulfone were corrected to be comparable to the experimental values of the other sulfuryl compounds using the methods described in the experimental section, the nitro compounds were not since they are not structurally similar to the sulfuryl compounds. Nonetheless, the resulting data for these compounds are included in Figure 5.4d, but the data points are presented as unfilled diamonds instead of the filled ones used for the sulfuryl compounds. While the reagents 1-6 in Table 5.1 showed a similar trend as before, the new reagents 7-9, with GBs between 799 and 828 kJ mol\(^{-1}\), showed either little increase in the ACS (diethyl sulfone) or a reduction to the ACS (m-NBA and m-nitrobenzamide). This was likely due to the significant amount of adduct formation that was observed for these reagent solutions. It has been shown previously that significant adduct formation is observed for cyt c when m-NBA is added above a certain concentration.\(^1\) Interestingly, this limit was 0.7% v/v (~60 mM) in one study,\(^1\) while a later study reported effective supercharging at concentrations up to ~ 5% v/v (~400 mM) using identical solvent conditions.\(^31\) In the present report, m-NBA was present at 100 mM (1.2% v/v). In fact, effective supercharging was observed at the front and tail ends of the sample plugs for m-NBA and m-nitrobenzamide, where diffusion occurring during sample transfer from the sample delivery loop to the sprayer leads to a lower concentration of reagent (Figure 5.5). Thus, it appears that factors other than concentration alone, perhaps instrumental conditions, play a role in determining the
maximum amount of supercharging reagent that can be added before significant adduction occurs. This phenomenon requires further investigation.

5.4.3 The Mechanism of Supercharging Related to Reagent GB

In the negative ion mode, it is likely that supercharging results from a previously neutral acidic residue donating a proton to a neutral supercharging reagent, an organic base which is more basic than the standard solvent components of MeOH and H₂O. Due to the charge separation and the differences in the GB of the protein carboxylate (the intrinsic GBs for aspartate and glutamate have been estimated to be ~1429 kJ and ~1424 kJ mol⁻¹, respectively⁴⁴) relative to the base (the highest GB of the bases in the present study was 1030 kJ mol⁻¹, for MTBD), these reactions are not expected to be

Figure 5.5 (a) Combined extracted ion chromatogram (EIC) of the +19, +18, and +17 charge states of cyt c, which are typically enhanced upon supercharging by m-nitrobenzyl alcohol (m-NBA). (b) Averaged spectrum of cyt c with no m-NBA. (c) Averaged spectrum showing the supercharging observed at the tail of the reagent plug, where diffusion has reduced the concentration of reagent. An identical spectrum was obtained for the front. (d) Averaged spectrum showing the heavy adduction and loss of signal observed for the region of constant m-NBA concentration.
thermodynamically favorable at room temperature, indicating that the ions must be highly energized at the time of ionization.\textsuperscript{2} It is also possible that intramolecular charge solvation increases the stability of the negative charges supported by the protein.\textsuperscript{45}

The positive ion mode supercharging reagents investigated in the present study do not possess labile protons, so while in the negative ion mode a Brønsted base interacts with a Brønsted acid to increase charging, a similar interaction will not take place between the positive ion mode reagents and the basic amino acids. However, the trends in the data for the supercharging by the positive ion mode reagents (Figure 5.4) suggests that there is also direct interaction mechanism for supercharging in the positive ion mode which is dependent on the reagent GB, but different than what is observed in the negative ion mode. The difference in the correlation, a maximum versus linear trend, suggests that the interaction is more complex. As the GB of a positive ion mode reagent increases past a certain value, it is expected that multiple charging of the protein will be reduced because of the increased likelihood that the reagent will deprotonate charged residues, especially for higher charge states with lower GB\textsubscript{app} values. As seen in Figure 5.4a-c, supercharging was observed to decrease with increasing GB of the reagent after a maximum value of \(~800\) kJ mol\textsuperscript{-1}. On the other hand, with no labile protons, it is not immediately clear how the other compounds increase multiple charging, or why this effect increases with increasing GB up to \(~800\) kJ mol\textsuperscript{-1}.

Additional evidence for a direct interaction mechanism was previously reported.\textsuperscript{24} Under gentle instrumental conditions, sulfolane adducts are observed on charge states which emerge during the supercharging of cyt c, and the number of sulfolane adducts increases with charge state.\textsuperscript{24} Others have reported similar results.\textsuperscript{23, 46} It was recently
demonstrated by Cole and coworkers that the strength of the complex formed between a protonated protein residue and an ion is determined by how closely their GB match.\textsuperscript{47} Similarly, if a supercharging reagent has a GB similar to that of a charged residue it may form a strong proton-bound complex with it. The formation of stable proton bridges might explain why adducts are observed for the highest charge states of a supercharged protein,\textsuperscript{24} which have the lowest GB\textsubscript{app}, and why extensive adduction of some reagents occurs at concentrations as low as 60 mM.\textsuperscript{1} These proton-bound complexes may prevent the deprotonation of some basic residues which would otherwise lose their protons during charge partitioning during the final stages of ion formation.

In the acidic bulk solution that is typical of positive ion mode protein samples, the basic residues of the proteins are expected to be protonated. It is expected that neutral supercharging reagents, enriched in the droplet due to preferential evaporation of the other solvent components, would form stable complexes with charged basic residues at some point prior to final droplet evaporation. While some supercharging reagents may become protonated charge carriers during the electrospray process,\textsuperscript{24} the high Coulombic repulsion that would be experienced by a charged reagent molecule approaching a highly-charged protein ion would make the formation of a complex in the gas phase unlikely.

In fact, the positive ion mode reagents have GBs which span the range of GB\textsubscript{app} values calculated for the positive ion mode charge states of cyt c\textsuperscript{48-49} and other proteins.\textsuperscript{10} Schneir and coworkers used a bracketing method to experimentally determine the GB\textsubscript{app} values of denatured cyt c ions formed by electrospray ionization, up to the 15+ charge state.\textsuperscript{48} A linear decrease in GB\textsubscript{app} was observed with increasing charge state (CS). Assuming that the linearity continues with increased charging, linear regression of the
experimental data can be used to predict the GB$_{\text{app}}$ of charge states higher than 15+. The results (Figure 5.6) show that the range of GB values for the positive ion mode reagents used in the present study (776 – 864 kJ mol$^{-1}$) likely span the 14+ to 21+ charge states of cyt c. Interestingly, the average charge state of denatured cyt c is typically around 14+. Thus, the reagents have similar GB values as those predicted for the charge states that emerge or are enhanced when the reagents are added as supercharging reagents.

This data supports the theory that supercharging might involve the formation of stable proton-bound complexes between protein ions and supercharging reagents which protect these sites from deprotonation prior to final ion formation during the electrospray process. As the GB of the reagent approaches the maximum in Figure 5.4, increasing stability of the proton-bound complex promotes supercharging. However, these adducts are lost during transmission through the high vacuum region of the mass spectrometer prior to detection so that the proton preferentially stays with the protein ion. Around the
Table 5.2 Proton affinities for solvent species.

<table>
<thead>
<tr>
<th>Species</th>
<th>Wrobleski$^{52}$</th>
<th>Kawai$^{50}$</th>
<th>Knochenmus$^{51}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H$_2$O</td>
<td>660</td>
<td>695</td>
<td>699</td>
</tr>
<tr>
<td>(H$_2$O)$_2$</td>
<td>858</td>
<td>832</td>
<td>816</td>
</tr>
<tr>
<td>(H$_2$O)$_3$</td>
<td>920</td>
<td>888</td>
<td>862</td>
</tr>
<tr>
<td>(H$_2$O)$_4$</td>
<td>937</td>
<td>919</td>
<td>900</td>
</tr>
<tr>
<td>(H$_2$O)$_5$</td>
<td>950</td>
<td>926</td>
<td>904</td>
</tr>
<tr>
<td>(H$_2$O)$_6$</td>
<td>983</td>
<td>950</td>
<td>-</td>
</tr>
<tr>
<td>CH$_4$O</td>
<td>-</td>
<td>-</td>
<td>762</td>
</tr>
<tr>
<td>(CH$_4$O)$_2$</td>
<td>-</td>
<td>-</td>
<td>883</td>
</tr>
<tr>
<td>(CH$_4$O)$_3$</td>
<td>-</td>
<td>-</td>
<td>937</td>
</tr>
<tr>
<td>(CH$_4$O)$_4$</td>
<td>-</td>
<td>-</td>
<td>967</td>
</tr>
<tr>
<td>(CH$_4$O)$_5$</td>
<td>-</td>
<td>-</td>
<td>979</td>
</tr>
</tbody>
</table>

maximum in Figure 5.4, reagent GB becomes too well matched to that of the protein so that these adducts are stable enough to survive mass analysis and detection, leading to heavy and varied adduction and loss of signal. At reagent GB values above the maximum in Figure 5.4, the reagent begins to dissociate along with the proton, reducing any supercharging effect.

It has been argued$^{28}$ that supercharging cannot occur based on GB considerations in the positive ion mode as described above since each these supercharging compounds all have GB values higher than that of the typical solvent components, which should lead to additional deprotonation on that basis. Table 5.2 gives the calculated proton affinities (PA) of several solvent species, including multimers.$^{50-52}$ GB values for these species are expected to be around 30 kJ mol$^{-1}$ lower than their respective PA values since the entropy...
of a free proton at 1 bar (108.95 J mol$^{-1}$ K$^{-1}$) leads to a $T\Delta S$ of 32 kJ mol$^{-1}$ at 295 K.$^2$

Solvent species which are dimers or higher order multimers likely have GB values higher than that of the positive ion mode supercharging reagents. It seems reasonable that in the absence of supercharging reagent, a charge supported by the protein could be solvated by several solvent molecules and leave in the gas phase as part of a protonated cluster. Association of protein charges with supercharging reagents, promoted by a match in their GBs as described above, may in fact prevent deprotonation of the protein by these solvent clusters. Therefore, I do not believe that a mechanism for supercharging in the positive ion mode involving GBs can be immediately disqualified.

5.5 Conclusion

In the negative ion mode, supercharging is directly related to the GB of the reagent. As reagent GB increases, increased proton transfer from the carboxylic acid residues to the reagent molecules results in higher multiple charging of the protein ion. In the positive ion mode, a clear relationship between reagent GB and supercharging is also observed. However, it is not a direct relationship as in the negative ion mode. Instead, supercharging is maximized for reagents with GB values around 800 kJ mol$^{-1}$. The observation that supercharging is most effective for reagents with a particular GB, as well as the observation of supercharging reagent adducts for the higher charge states of supercharged protein ions, especially those with GB values in the most effective range, indicates that the formation of stable proton bridges between the supercharging reagents and basic protein residues is likely involved. How this proposed interaction increases multiple charging is still under investigation.
5.6 References


CHAPTER 6

AMBIENT IONIZATION: DESORPTION ELECTROSPRAY IONIZATION (DESI) AND RELATED METHODS

Reprinted (adapted) with permission from


Copyright © 2014 American Chemical Society

6.1 Introduction to Ambient Ionization

In the last decade, an increasingly large number of mass spectrometry sample introduction methods have emerged which attempt to reduce or remove sample preparation prior to ionization. These methods have become collectively known as the ambient ionization methods in reference to their ability to ionize samples from their immediate surroundings, bypassing the typical collection, extraction, and dilution procedures required for mass spectrometry analysis. It has commonly been stated that these methods require no sample preparation, but this is a misconception. The defining feature of the ambient ionization methods is that the sample preparation steps are coupled in time and proximity with the ionization steps. They are combinations of a much smaller set of well-defined ionization processes with traditional sample preparation methods conducted in a novel fashion.
Figure 6.1 shows some of the popular ambient ionization methods divided by their sampling and ionization processes. Liquid-liquid extraction, solid-liquid extraction, thermal desorption, and spallation are all common desorption processes which remove analyte material from its original location in preparation for further in-line sample processing and subsequent ionization. Ionization typically falls under two categories, processes resembling either electrospray ionization (ESI) or atmospheric pressure chemical ionization (APCI). Some methods are able to include additional processing steps in between the desorption and ionization events, including chromatographic separation, selective absorption/adsorption, and filtration. An obvious advantage of
ambient ionization is that both time and material resources can be saved by removing traditional sample collection and preparation steps.

Desorption electrospray ionization (DESI), introduced in 2004 by Graham Cooks and coworkers from Purdue University,\(^2\) was one of the original ambient ionization methods. It has since become one of the most popular ambient methods, in part because of the low cost of building a DESI source, its robustness, and its applicability to a wide range of analyte types.\(^9\) However, one analyte type which has been historically difficult to analyze using DESI is protein. A trend of decreasing instrumental response with increasing molecular mass has been observed for proteins,\(^10\) often attributed to the inability of protein molecules to effectively desorb from the sample surface.\(^10-12\) This mass-dependent instrumental response has limited protein analysis by DESI to proteins with molecular masses of around 25 kDa and below.

In the chapters following this one, work is presented which was undertaken in an effort to shed light on the problem of protein analysis by DESI. To accomplish this goal, the individual sampling and ionization processes of DESI were investigated independently using novel analytical techniques. Many of these processes are the same or similar to those used by the other spray desorption ambient methods. In order to separate and evaluate independently the sampling and ionization processes of DESI, it is useful to understand what is already known about these processes for DESI and related methods. In this chapter, an introduction to the mechanism of DESI and other spray desorption ambient methods is presented.
6.2 Mechanistic Aspects of Spray Desorption Ambient Mass Spectrometry

The ambient spray desorption methods are characterized by their use of a solvent spray as the sample processing agent for removing material from a surface for subsequent ionization and introduction into the mass spectrometer for analysis. The sample processing takes place in four steps: (1) Formation of a spray plume directed at the sample, (2) creation of a micro-localized liquid layer on the surface, (3) dissolution/extraction of the analyte into the liquid layer, (4) liberation of analyte containing droplets (secondary droplets) by momentum transfer between the pneumatically accelerated primary droplets and the liquid layer. In a final step, analyte ions are produced from the charged secondary droplets. Figure 6.2 shows a typical DESI source. The DESI experiment begins with a charged spray plume formed by pneumatically nebulizing the charged eluent of a central solvent capillary coaxial with a sheath gas capillary. These primary droplets initially wet the sample surface, forming a thin liquid layer into which the analyte is extracted. The secondary droplets which are

![Figure 6.2](image)

**Figure 6.2** A desorption electrospray ionization (DESI) source. A pneumatically-assisted electrospray is directed toward a sample present on a surface. Sample is extracted into the charged spray and carried to the mass spectrometer inlet in reflected droplets. Ionization proceeds via the typical electrospray mechanisms.
formed upon the subsequent impact of the primary droplets on this thin liquid layer carry analyte from the surface and analyte ions are then produced through traditional electrospray ionization mechanisms such as ion evaporation or charge residue. The other spray desorption methods, each similar to DESI, vary in the nature of the surface being sampled or the ionization agent used.

The most common implementation of the spray desorption methods is to analyze a continuous solid surface, for instance a deposited sample on glass or PTFE, a pharmaceutical tablet, tissue samples, plant material or a chromatography plate after separation. Samples can also be deposited on a mesh substrate and analyzed in transmission geometry (transmission mode DESI, TM-DESI), or liquid samples can be analyzed as they elute from a capillary (liquid-DESI) or are sprayed in a dual sprayer arrangement so that the charged DESI spray interacts with the sample spray (extractive electrospray ionization, EESI).

For traditional DESI and many of the other spray desorption methods, a high voltage is applied to the primary solvent spray, which is often an acidified methanol/water mixture, so that protonated solvent molecules or clusters serve as the main source of ionization for the resulting analyte ions. Molecular ions of the analyte can be formed if a charged solvent spray of toluene is used instead, where ionization occurs through charge exchange with solvent molecular ions (desorption ionization by charge exchange, DICE). Additionally, ionization by the spray methods can occur by means other than applying a high voltage to the primary spray. For example, a supersonic sheath gas flow 2-5 times higher than that used during traditional DESI can be used to form a primary spray where droplets are charged by a statistically unbalanced separation of
charge\textsuperscript{34–35} during formation (easy ambient sonic-spray ionization, EASI; formerly known as desorption sonic spray ionization, DeSSI).\textsuperscript{34} In another variation, gas-phase ions and neutrals formed during traditional DESI desorption are introduced to a metastable plume, where metastable-induced chemical ionization can occur for less polar molecules not amenable to electrospray ionization (desorption electrospray/metastable-induced ionization, DEMI).\textsuperscript{36}

Because of their close similarities, many mechanistic studies that were performed to investigate an individual technique are applicable to others as well. In addition, many parameters are interdependent and there is often a trade-off between the processes where the best settings for a particular process are not always optimal for the others.\textsuperscript{14} Further, different combinations of settings can produce the same results, complicating optimization.\textsuperscript{14} For example, Table 6.1 gives the averages and ranges for DESI parameters compiled from a considerable number of literature sources.\textsuperscript{14} The large ranges for each parameter demonstrate the wide parameter space available for successful analysis by DESI, and likely the other spray methods.

**Table 6.1** Descriptive statistics summarizing DESI parameters found in the literature.\textsuperscript{14}

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>n=</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voltage (kV)</td>
<td>1</td>
<td>5</td>
<td>3.9</td>
<td>26</td>
</tr>
<tr>
<td>Solvent flow rate (μL/min)</td>
<td>1.5</td>
<td>25</td>
<td>5.6</td>
<td>27</td>
</tr>
<tr>
<td>Incident angle (°)</td>
<td>35</td>
<td>90</td>
<td>56</td>
<td>24</td>
</tr>
<tr>
<td>Sprayer-to-surface distance (mm)</td>
<td>1</td>
<td>5</td>
<td>2.0</td>
<td>22</td>
</tr>
<tr>
<td>Gas pressure (psi)</td>
<td>100</td>
<td>261</td>
<td>136</td>
<td>18</td>
</tr>
</tbody>
</table>
6.2.1 The Primary Spray

Every spray desorption method begins with the generation of a stable solvent spray characterized by micrometer sized droplets driven by a nebulizing sheath gas. Droplet size and velocity are important parameters as they determine the interaction of the spray plume with the surface, affecting subsequent extraction and desorption. These sprays are comprised of droplets with sizes between 1 and 10 microns\textsuperscript{37-38} and velocities dependent on the sheath gas velocity.\textsuperscript{37} The homogeneity of the droplet size distribution depends on the source construction.\textsuperscript{37-38} An early study of pneumatically assisted electrosprays of methanol/water mixtures conducted using laser anemometry measurements showed a homogenous population of droplets\textsuperscript{38} while a later study showed up to three separate populations.\textsuperscript{37} The latter was attributed to the solvent capillary being off-axis within the sheath capillary,\textsuperscript{37} which is likely common for home-built sprayers comprised of flexible fused-silica tubing and Swagelok\textsuperscript{®} fittings.

The main factor affecting the initial droplet size (diameter) is the spray solvent composition according to the equation\textsuperscript{39}:

\[
d \propto (\gamma \varepsilon^2)^\frac{1}{7}
\]  

(6.1)

where $\gamma$ is the spray-solvent surface tension and $\varepsilon$ is its dielectric constant. A common DESI solvent system is acidified 50/50 MeOH/H\textsubscript{2}O. Initial droplet sizes decrease with increasing ionic strength,\textsuperscript{38} increasing organic fraction,\textsuperscript{39} and the addition of surfactants to the spray solution.\textsuperscript{40} Other factors which have a minimal effect on the initial droplet size are the sprayer design, spray voltage, and solvent flow rate.\textsuperscript{37} A smaller solvent capillary diameter, decreased protrusion of the solvent capillary from the sheath capillary, lower solvent flow rate, and increased voltage can lead to a decrease in the size of the
primary droplets.\textsuperscript{37} Droplet diameters do not appear to change significantly upon increasing gas pressures, except at low pressures where inefficient nebulization results in large droplets.\textsuperscript{37}

The main factor affecting initial droplet velocity is the sheath gas exit velocity.\textsuperscript{37} The surrounding sheath gas provides a shearing force to assist in nebulization of the solvent eluting from the coaxial solvent capillary.\textsuperscript{41} Solvent ionic strength can have a large effect on the initial droplet velocity.\textsuperscript{38} The addition of $5 \times 10^{-6}$ M KCl was found to significantly shift the velocity distribution of droplets from a 90\% MeOH solution to higher values, whereas the addition of $5 \times 10^{-3}$ KCl resulted in a distribution closer to the unmodified solution.\textsuperscript{38} Droplet velocity does not appear to change significantly with an increase in the spray solvent flow rate.\textsuperscript{37} Once the primary droplets are made, they begin to lose velocity due to collisions with the atmosphere, which also aids droplet evaporation. However, the distance traveled by droplets during DESI is usually on the order of a few millimeters, and the droplet diameters and velocities are similar to their original values when they impact the surface.\textsuperscript{37}

\textbf{6.2.2 Dissolution/Extraction}

In applications where the spray is used to sample a solid surface (e.g. DESI, EASI, DEMI, DICE), a thin liquid layer is formed when the surface is first bombarded by the solvent spray plume.\textsuperscript{37,42-43} Equilibrium exists between the deposition of solvent from the spray and the processes of solvent evaporation due to the strong sheath gas flow and secondary droplet emission.\textsuperscript{19} The formation of this liquid layer is important for stable and reproducible analyses. During EESI\textsuperscript{32} and liquid-DESI\textsuperscript{30} the sample is already in the
liquid state, excluding dissolution as part of the mechanism. Instead, mixing of the two liquid phases occurs.

Solubility plays an important role during ambient extraction. Generally, response monotonically increases with the solubility of analytes. However, samples with high solubility erode over time, which can lead to a loss of signal due to sample degradation. It has been shown that an analyte’s octanol-water partition coefficient can be used to choose an appropriate solvent system. In the liquid analysis technique EESI, relative solubility in the solvents of the primary and sample sprays is the predominant factor affecting extraction of analyte from sample droplets into the charged primary droplets. With DESI, an increase in the organic fraction of the spray solvent up to 90% leads to a linear improvement in the erosion diameter on surfaces primarily due to the formation of smaller droplets, which do not spread around the surface as much. Likewise, surface wettability effects the spreading of droplets, where increasing contact angle (decreasing wettability of the solvent) increases the erosion diameter. Surfaces for which the dry sample components have little affinity are not suitable, as the material can be blown or washed away before extraction and subsequent desorption. Microstructured surfaces can aid in the retention of the analyte in its original place and increase DESI efficiency.

6.2.3 Desorption

Desorption is the process whereby analyte is transferred from the surface, carried away in secondary droplets originating from the surface layer towards the mass spectrometer inlet. For spray desorption methods desorption proceeds by a droplet-pickup process. Numerical multiphase fluid-dynamics simulations demonstrated that
primary droplets impacting a thin liquid film result in the ejection of secondary droplets.\textsuperscript{42-43} Secondary droplets generated in this way are shown to be composed of liquid from both the primary spray and the liquid film, generated by momentum transfer from the impinging primary droplets.\textsuperscript{42} Secondary droplets may also arise from rivulets extending from the impact zone.\textsuperscript{46} The bulk of desorbed material remains near the surface layer\textsuperscript{47-48} and is limited to a narrow band extending toward the inlet.\textsuperscript{39, 47} During EESI, primary droplets collide with the liquid sample, resulting in turbulent mixing, and desorption of secondary droplets due to momentum transfer.\textsuperscript{32, 45} For TM-DESI, both the strand and mesh sizes were shown to be important characteristics for desorption, as large values of both decrease transmission and lead to solvent and analyte spreading, respectively.\textsuperscript{49} Studies of the spray desorption process used by DESI separated from the DESI ionization process, termed spray desorption collection (SDC),\textsuperscript{47} have shown that desorption is not dependent on the spray voltage. Geometric source parameters such as the distance of the spray from the surface and the sheath gas velocity can affect desorption,\textsuperscript{14} though the use of a sealed geometry-independent source design can reduce the influence of these factors.\textsuperscript{50}

The interaction of the spray with the surface forms an impact zone with three areas of interaction: a small elliptical area closest to the sprayer that is the point of highest effective desorption, a surrounding gradient of impact force extending to the solvent front, and an area of rivulets and slow moving droplets around the edges.\textsuperscript{28, 39, 46} The characteristics of the thin liquid layer formed within the impact zone depends on both 1) solvent and gas flows, 2) solvent physiochemical properties, and 3) the physiochemical properties of the surface.\textsuperscript{19} Optimum values for solvent and gas flows are
interdependent, since high gas flows enhance evaporation of the droplets, which can be somewhat compensated for by increasing the solvent flow rate. The optimum combination, or range of combinations, of solvent and flow rate produces a spray that effectively wets the surface and energetic enough to produce secondary droplets, yet dry enough and not so energetic as to cause sheeting of the solvent along the surface. 14

6.2.4 Ionization

The generation of charged secondary droplets during desorption ostensibly leads to an electrospray-like charging process, where small molecules become charged by ion evaporation 51 and larger molecules by charge residue. 52 Charged droplets undergoing evaporation, as well as ions formed in the atmospheric pressure region in front of the mass spectrometer inlet, are funneled into the MS assisted by the suction of the inlet. A lensing effect can be observed when a small negative potential is applied to the inlet capillary while analyzing positive ions. 19 The interaction of a charged spray with the surface can lead to a small loss of charge, 13, 44 as seen by the slight reduction in the multiply charged protein ions observed during DESI-MS as compared to ESI-MS. These charges can accumulate on the surface and impact the ion signal. 53 When the DESI spray initially interacts with the sample, analyte signal first appears with a rapid spike then decays a steady state. 53 The time constant for decay is dependent on the dielectric constant of the surface, where nonconductive surfaces like PTFE yield the most stable signal. 53

Similar to DESI, DICE begins with a charged primary spray. 33 Instead of protonated solvent molecules, the charge carriers are odd-electron molecular ions of toluene formed by electrochemical oxidation occurring at the spray capillary. 33 After
extraction and desorption, molecular ions of the analyte molecules are formed by a charge transfer process involving toluene molecular ions. At high spray potentials (>3.5 kV), spectra for cholesterol appeared similar to those produced by atmospheric pressure chemical ionization (APCI) mechanism.

During DEMI, analyte molecules that are first desorbed by a charged solvent spray, in a manner similar to DESI, are intercepted by a plume of metastable species en route to the mass spectrometer inlet. In the positive ionization mode, higher polarity molecules are ionized by sodiation and/or protonation following DESI mechanisms while neutrals are protonated by reactions with protonated water clusters formed from the interaction of the metastable plume with atmospheric gases. Because of this APCI component, DEMI can simultaneously ionize low polarity chemical species not typically amenable to DESI ionization, such as dibenzosuberone, dibromobenzosuberone, and cholesterol.

Not every spray method relies on the application of a high voltage to produce charged primary droplets in the desorbing spray for the generation of ions. During EASI, no voltage is applied to the primary spray and yet significant analyte ion signal can be produced. It has been proposed that charged droplets are produced by minuscule fluctuations in the ion concentrations of the bulk solution occurring immediately prior to nebulization and due to interactions with the material of the spray nozzle. Protonation in the positive-ionization mode is observed when acidified aqueous solutions are used. Interestingly, protonated ions are still observed when non-acidified dueterated solvents are used, leading to the speculation that atmospheric-pressure chemical ionization might also occur by interaction with charged intermediates formed during droplet collision with
Because EASI does not require a high applied potential to the primary spray, there is little chemical noise that is due to solvent clusters typically seen with DESI.  

### 6.3 Liquid Microjunction Methods

Liquid extraction of analyte material from solid samples can also be achieved with continuous liquid interfaces. In the ambient liquid microjunction methods, solvent is extruded from a capillary or pipet tip to form a solvent bridge, or microjunction, with the sample. Material is extracted as it distributes between the sample and the liquid phase, which is transferred to a position where the extracted sample can be ionized for analysis, often by electrospray. Three techniques have been reported that use a liquid microjunction for extraction followed by online ionization: Liquid microjunction surface sampling probe (LMJ-SSP), nano-DESI and liquid extraction surface analysis (LESA). These methods primarily differ in the way that the extraction solvent is delivered and removed from the surface.

Figure 6.3a shows a schematic of the LMJ-SSP source. The LMJ-SSP uses a pair of coaxial capillaries positioned 100 to 300 micrometers above the sample surface as illustrated in Figure 6.3a. Solvent is delivered from the outer capillary to form the liquid microjunction, which extracts material from the surface, and the inner capillary delivers this solution to an electrospray or APCI source. Aspiration rates, typically in the tens of μL/min, can be controlled by adjusting the inner and outer diameters of the source capillaries as well as the nebulizing gas flow. The syringe pump driving the solvent must be set to match the aspiration rate of the electrospray to maintain a constant liquid microjunction. If the inner capillary is retracted further, extracted
Material can get caught up in eddy-like flow patterns near the probe end which broaden and dilute extracts resulting in delayed detection.\textsuperscript{62} This phenomenon can be used advantageously, however, to mix multiple samples on a nanoliter scale during analysis.\textsuperscript{62}

The choice of solvent is an important consideration, not only for extraction of the analyte but also for controlling the wetting of the surface. Physiochemical properties of the solvent, such as surface tension and polarity, will determine the amount of spreading.
that occurs for a particular surface.\textsuperscript{61} Wettable or absorbant surfaces lead to a spreading of the analyte with the solvent, resulting in inefficient extraction\textsuperscript{61} and a loss of resolution. Wettable or absorbent surfaces can be treated by applying a coating to the surface to prevent solvent spreading\textsuperscript{63} or by transferring the analyte to a more suitable surface via blotting.\textsuperscript{64}

The LMJ-SSP can be operated in one of two modes, stepping or scanning.\textsuperscript{54} In the stepping mode, a single spot of the sample is analyzed at a time by moving the probe and creating a new microjunction at each site.\textsuperscript{54,59,65} This mode offers the benefits of discrete sampling and prolonged analyte signal.\textsuperscript{54} In the scanning mode, the liquid microjunction is dragged across the surface to form a 1-dimensional chemical image, or line scan, while 2-dimensional images can be formed by combining several line scans.\textsuperscript{27,65-66} During scanning, spatial resolution is affected by several factors including the scan speed and solvent flow rate. Increasing the scan rate results in a loss of spatial resolution at low flow rates\textsuperscript{61,67} but not at high flow rates.\textsuperscript{61} The loss of resolution at low flow rates is attributed to diffusion of the analyte within the dead volume of the inner transfer capillary.\textsuperscript{61,67} At high flow rates and scan speeds, however, a decrease in signal intensity is observed due to incomplete analyte extraction.\textsuperscript{61,65}

While in DESI spray desorption leads to liquid-solid extraction and momentum transfer to analyte removal from surfaces, in nano-DESI\textsuperscript{55} dissolution of surface compounds into a liquid phase occurs at a liquid microjunction similar to the LMJ-SSP method. Analyte extraction is slower and ionization is delayed during transport from the microjunction to the self-aspirating nanospray emitter. Because of its similarity to LMJ-
SSP, the two methods share many of the same mechanistic considerations, particularly solvent and surface effects.

In contrast to the nested capillary arrangement of the LMJ-SSP, nano-DESI uses an angled primary capillary to deliver solvent at flow rates of <1 μL/min to a nanospray capillary held vertically\textsuperscript{55} or angled\textsuperscript{68} above the surface, forming a liquid microjunction (Figure 6.3b). Flow through the junction is limited to the flow induced by the electrospray from the nanospray tip\textsuperscript{55} and assisted by suction from the vacuum system of the MS inlet.\textsuperscript{69} The size of the sampled spot is controlled by the size of the capillaries and the solvent flow rate.\textsuperscript{55} Spot sizes of less than 10 microns have been reported,\textsuperscript{69} giving nano-DESI the highest resolution among the liquid desorption methods.

Another technique in this family is liquid extraction surface analysis (LESA).\textsuperscript{56} Instead of forming a liquid microjunction with a continuous flow of solvent from one capillary to another, a small volume of solvent is extruded from a pipet tip held a small distance above the sample surface. After a designated amount of extraction time, the solvent is retracted and the pipet tip is repositioned behind an ESI-chip to initialize an electrospray. A new pipet tip is used for each analysis, removing the possibility of contamination and carryover. In addition, the whole process is fully automated removing any user error associated with solvent introduction/extraction and removal.

While upwards of 30 (or more) ambient ionization methods are currently being developed,\textsuperscript{70} this chapter has focused mostly on those whose primary desorption mechanism involves a solvent spray which is usually charged. DESI is particularly relevant for the remainder of this document, while the liquid microjunction techniques become relevant in Chapter 8.
6.4 References


11. Heaton, K.; Solazzo, C.; Collins, M. J.; Thomas-Oates, J.; Bergström, E. T., Towards the application of desorption electrospray ionisation mass spectrometry


7.1 Introduction

To study the challenge of protein analysis by desorption electrospray ionization mass spectrometry (DESI-MS), it is necessary to isolate the many parameters discussed in Chapter 6. In this chapter, the separation and independent investigation of the sub-processes which comprise DESI are presented. To accomplish this task, novel methods of analysis were used, including spray desorption collection (SDC) as a model for desorption process and reflective electrospray ionization (RESI) as a model for the ionization process in DESI. Both processes depend on several instrumental parameters, including the nebulizing gas flow rate, applied potential, and source geometries. Each of these parameters was optimized for desorption, as represented by the results obtained by SDC, and ionization, as represented by the results obtained by RESI. The optimized conditions were then compared to the optimization results for DESI to draw conclusions about the contribution of each parameter to desorption and ionization separately as well
as these processes combined within DESI. Findings confirm that optimal conditions for desorption and ionization are different and that in some cases the optimized DESI conditions are a compromise between both sets. Specifically, our results indicate that desorption efficiency is (1) independent of the applied potential and (2) the impact zone to inlet distance, and that (3) gas pressure settings and (4) sprayer to impact zone distances above optimal for DESI are detrimental to desorption but beneficial for ionization. In addition, possible interpretations for the observed trends are presented.

7.2 The Optimization of DESI

As presented in the previous chapter, the operational parameters of desorption electrospray ionization (DESI, Figure 7.1a) and other related methods using spray desorption as the primary sampling process have been previously studied.\(^1\)-\(^{13}\) It has been shown that DESI optimization depends on several variables including source geometry, source construction, interrogation surface material, instrumental settings, etc. Table 6.1 summarizes several of these operational parameters found in the literature compiled from over 25 references. As indicated by the large spread of the data, DESI is able to be performed using a wide range of operation values. Because the DESI parameter space is large and some parameters are correlated, optimization is not always straightforward. Many of the parameters are interrelated, so multiple combinations of settings are capable of providing stable and strong ion currents. In addition, it is well known that signal intensities are highly variable during DESI,\(^7\),\(^9\)-\(^{10}\),\(^{14}\)-\(^{17}\) although the use of an appropriate internal standard can result in inter-day variability of \(<10%\) RSD.\(^{16}\) These factors inevitably lead to the large variation in the optimization results presented in the literature.
Figure 7.1 Schematics of (a) desorption electrospray ionization (DESI), (b) reflective electrospray ionization (RESI), and (c) spray desorption collection (SDC).
Most molecules capable of analysis by DESI are desorbed by the droplet pickup mechanism. For droplet pickup to occur, analyte molecules are initially dissolved into a thin film of solvent formed on the interrogated surface. An initial 0.1 to 0.2 min is required to reach maximum signal intensity during DESI, which has been attributed to a delay in the formation of the liquid film. The addition of surfactant to the solvent spray reduces this initial delay, an effect attributed to the increased wettability and solubilizing power of the solvent. Once the film has formed, primary droplets colliding with the film generate two populations of secondary droplets that are ejected from the film. Larger, slower-moving droplets leave the surface close to their specular angle while smaller, faster-moving droplets follow the dynamics of the nebulizing gas and tend to skim the sample surface. After desorption, ionization is thought to occur by typical ESI mechanisms. The spectral features of DESI and ESI share many similarities, including multiply charged ions, the preservation of noncovalent complexes, and the observation of ions formed by metal adduction. In particular, protein spectra generated from both methods are remarkably similar, although charge state distributions are generally shifted to lower charge states in the case of DESI, suggesting that secondary droplets generated during DESI carry less charges than those generated during ESI. Ion mobility studies and the determination of internal energy distributions have also supported an ESI-like ionization mechanism for DESI. Analysis by DESI is comprised of two main processes, desorption of the analyte from the surface and subsequent ionization. It is likely that the optimum conditions found for DESI are a compromise between those conditions ideal for desorption and those ideal for ionization.
Despite the many studies which have investigated the DESI process, none have previously investigated the desorption and ionization processes of DESI completely separated from each other, since they are coupled during the DESI experiment. This challenge has resulted in ambiguity as to what factors are responsible for analyte effects observed during DESI analysis, such as the mass dependent response for proteins.\textsuperscript{28} By realizing that the ambient ionization methods, including DESI, are combinations of known sample processing and ionization methods,\textsuperscript{13} experiments were developed to model the individual processes. For this chapter, the relative contributions of desorption and ionization to the DESI process were separately investigated by decoupling desorption from ionization using novel methods of analysis. To study desorption independently, desorbed material that is typically analyzed immediately by mass spectrometry was instead collected using surface desorption collection (SDC, Figure 7.1c), a recently developed surface sampling technique\textsuperscript{29} that closely mimics the DESI desorption process. To study ionization separately, a system that closely mimics DESI but does not require desorption or analyte pickup was required. One consideration which had to be addressed was that the primary droplets produced by a DESI microsprayer and other typical ESI sources are substantially different from those produced after surface collision.\textsuperscript{5} Because ionization takes place from these droplets, a standard ESI source could not be used since it does not capture the interaction of the primary droplets with the sample surface. To better mimic the ionization conditions after droplet-pickup, the analyte was added to the electrospayed solvent directed at a clean surface and placed in front of the mass spectrometer, a process we call reflective electrospray ionization (RESI, Figure 7.1b). A similar method has been used previously for DESI (Figure 7.1a) optimization.\textsuperscript{28, 30} Using
these methods, as well as a standard DESI setup, several different DESI operational parameters were investigated. The results show how each parameter affects the desorption or ionization processes during DESI independently.

7.3 Experimental Details

Table 7.1 Standard operating conditions for SDC, RESI, and DESI used in the present chapter.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Base Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrospray voltage</td>
<td>5 kV</td>
</tr>
<tr>
<td>Solvent flow rate</td>
<td>3 μL/min</td>
</tr>
<tr>
<td>Nebulizing gas flow rate</td>
<td>~3 L/min</td>
</tr>
<tr>
<td>Incident angle</td>
<td>40°</td>
</tr>
<tr>
<td>Sprayer-to-surface distance</td>
<td>1.5 mm</td>
</tr>
<tr>
<td>Sprayer-to-inlet distance</td>
<td>5 mm</td>
</tr>
<tr>
<td>Inlet-to-sample distance</td>
<td>&lt;0.5 mm</td>
</tr>
</tbody>
</table>

Various parameters for DESI, SDC and RESI were optimized, including gas pressure, spray voltage, sprayer-to-inlet distance, and sprayer-to-surface distance. The standard operating conditions were inferred from the data in Table 6.1 considering our sprayer design, described below, and unless otherwise noted are listed in Table 7.1. A diagram of important experimental parameters, some of which were held constant in the present study, is presented in Figure 7.2.

7.3.1 Instrumental Details

A linear ion trap mass spectrometer (LTQ, Thermo Scientific, Waltham, MA, USA) was combined with a 3-dimensional translational stage (Purdue University, West Lafayette, IN, USA) for DESI and RESI analysis and SDC collection. For the generation of a pneumatically-assisted solvent spray, a typical spray ionization source was
A detailed description of DESI source design is presented in Appendix B. Briefly, the source is comprised of a 1/16” Swagelok T-piece, a fused silica solvent line (O.D. 190 μm, I.D. 75 μm), and a coaxial stainless steel nebulizing gas capillary (O.D. 1/16”, I.D. 250 μm, approximately 1.5 cm long). The spray potential was applied to the liquid junction on the stainless-steel syringe needle used to deliver the spray solvent. The tip of the extended mass spectrometer inlet capillary was bent at an angle of 10° to facilitate efficient droplet/ion transfer during DESI and RESI.\textsuperscript{10} Analysis of the samples collected by SDC was performed with a spectrofluorometer (FL920; Edinburgh Instruments, Reading, ENG) by measuring the fluorescence of rhodamine 6G at 547 nm at an excitation wavelength of 500 nm. Schematics of the various analytical techniques used in the current study are presented in Figure 7.2.
7.3.2 Surfaces and Samples

Two surfaces were required for this study, the sample surface used in all experiments and the collection surface used during SDC. Glass was chosen as the sample surface for DESI, SDC, and RESI experiments for reasons discussed below. For the collection surface used during SDC, silica gel TLC plates were found to give quantitative results after extraction of the analyte followed by fluorescence spectroscopy. Rhodamine 6G was used as the analyte for each study. It is a red colored dye that is readily found in some permanent markers and is easily visible to the naked eye. It has a high fluorescence quantum yield\(^{32}\) and has been used before in many DESI studies.\(^{6-7, 15, 22, 29, 33-34}\)

7.3.3 The Methods

7.3.3.1 Desorption Electrospray Ionization (DESI)

For each data point, five straight lines of rhodamine 6G dye were drawn on a clean glass surface using a red permanent marker. The sample was scanned through the charged solvent spray of pure MeOH at a speed of 150 μm/sec using the translational stage. The lines were drawn 1 cm apart to ensure that there was no carryover during the scans. For each line that was scanned, a peak was observed in the extracted ion chronogram for rhodamine 6G \((m/z 443.33)\). The area of each peak was calculated and used to generate the average response along with the standard deviation for the data point.

7.3.3.2 Spray Desorption Collection (SDC)

Similar to DESI, a red permanent marker was used to draw a single line of rhodamine 6G along the straight edge of a clean glass surface. A clean silica TLC plate was positioned immediately in front of and blocking the standard atmospheric pressure
inlet capillary. The sample edge was placed on the translational stage so that a 1 mm gap existed between the sample surface and the collection surface. This gap was left in place to allow excess nebulizing gas to escape in every direction away from the collection point. The sample was scanned horizontally at 150 µm/sec using the translational stage and a spray of pure methanol (MeOH). For each data point, material desorbed from the sample surface was collected in a single spot on the TLC plate for 133 seconds (2 cm of translation). Collection was typically observed to occur in a thin line just above the horizon of the sample surface. The visible lines were approximately 1 mm wide and 1 cm long. After changing the value of the parameter of interest the collection surface was moved so that sample collection occurred at a new spot approximately 5 mm from the previous spot. After sample collection was complete, 3 mm x 20 mm squares containing each collected spots were cut from the silica gel TLC plates and the silica was carefully scraped from the plastic substrate into labeled centrifuge tubes. 1 mL of MeOH was added to each tube and the tubes were agitated using a vortex mixer (Diagger Gebic 2™, Vernon Hills, Illinois) for 30 minutes to ensure complete extraction of rhodamine 6G. This step was followed by centrifugation of each tube to remove any suspended particles. The final solutions were analyzed by fluorescence spectroscopy. The average fluorescence measurements of each data set were then normalized for comparison between the data for DESI, SDC and RESI.

7.3.3.3 Reflective Electrospray Ionization (RESI)

A stock solution of rhodamine 6G in pure MeOH was serially diluted to a final concentration of 11.1 nM to be used as the working solution during RESI experiments. The solution was nebulized by the microspray source during the analysis and directed at a
clean glass sample surface. The surface was scanned horizontally at 150 µm/sec while the parameter of interest was changed at 30 second intervals. The average intensity of the rhodamine 6G peak was determined for each parameter value by averaging over the thirty seconds in the chronogram corresponding to that setting. This procedure was repeated three times for each parameter that was optimized. To allow comparison between trends observed for the mass spectral and fluorescence data, the average response values of the DESI, SDC and RESI data sets were normalized to their highest values.

7.4 Results and Discussion

7.4.1 Choosing the Sample Surface

Several authors have reported comparative studies between multiple DESI sample surfaces that have contributed to our current understanding of how surface properties affect the analytical performance of DESI. These surfaces include PMMA, PTFE, glass, paper, stainless steel, porous silicon, and TLC plates. “Rough” surfaces tend to show better stability and reproducibility than smooth ones, probably by preventing the washing and coffee-ring effects observed with the latter. The coffee-ring effect, sometimes referred to as the Marangoni effect, refers to the heterogeneous sample deposition often associated with samples prepared by allowing aliquots to air-dry on a surface. Capillary action within the drying droplets results in preferential analyte deposition along the periphery of the droplet. The DESI surface should also be nonconductive as charge losses can occur to the surface, which is exacerbated by solvent spreading. Thus, hydrophobic surfaces perform better with the polar solvents typically used during DESI. The ideal interrogation surface appears to rely heavily upon the nature of the analyte.

We investigated PTFE, glass, and plastic mesh for
instrumental response, signal stability and reproducibility for rhodamine 6G,\textsuperscript{15} an analyte typically used for DESI optimization. While PTFE showed the highest instrumental response, glass displayed the most stable signal for the analyte. Hence, glass was chosen as the sample surface for DESI, SDC and RESI experiments.

7.4.2 Optimization of the Applied Potential

![Graph showing the optimization of the applied potential for SDC (squares), RESI (triangles), and DESI (diamonds). Error bars represent one standard deviation of triplicate analyses. Lines connecting individual data points are presented only to guide the reader’s eye.]

Figure 7.3 Results for the optimization of the applied potential for SDC (squares), RESI (triangles), and DESI (diamonds). Error bars represent one standard deviation of triplicate analyses. Lines connecting individual data points are presented only to guide the reader’s eye.

Although a high potential is not required for obtaining a DESI signal,\textsuperscript{2, 11, 14, 33, 35} it has been demonstrated that instrumental response increases with increasing spray voltage up to 8 kV.\textsuperscript{2, 5, 7, 11} Increasing the voltage produces smaller droplets with larger velocities,\textsuperscript{5} leading to more efficient ionization. However, oxidative species in the electrospray also increase with increasing voltage, leading to the presence of oxidized products present in the spectra\textsuperscript{39-41} and increased spectral complexity. Figure 7.3 shows
optimization results for the applied potential between 0 to 8 kV, the instrumental limit, for DESI, SDC, and RESI.

The SDC curve in Figure 7.3 remains close to its optimum value across the range of voltages tested, indicating that desorption does not significantly depend on the high voltage applied during DESI. This result is consistent with the droplet-pickup model explained in Chapter 6 since desorption is thought to occur due to analyte solubility in the spray solvent followed by collisional ejection from the film, processes that aren’t expected to depend significantly on the applied potential. These results are also in agreement with previous studies where it was shown that an electrostatic contribution is not necessary to replicate the droplet impact characteristics observed during DESI\(^5\) using computational fluid dynamics\(^{24}\) and that a high voltage is not necessary for efficient desorption during desorption sonic spray ionization.\(^{14}\)

Unlike the SDC curve, the RESI and DESI curves both increase with increasing applied potential. This confirms that the applied potential dependence of DESI is based mostly on ionization or ion transfer processes and not desorption.

### 7.4.3 Optimization of the Nebulizing Gas Pressure

In DESI, instrumental response usually increases with increasing gas pressure until an optimum value is achieved, after which instrumental response plateaus or decreases.\(^2\), \(^5\), \(^7\), \(^9\) A similar result is presented in Figure 7.4 where gas pressure was optimized between 100 psi to 180 psi at the regulator, corresponding to nebulizing gas volumetric flow rates at the emitter of 0.96 L/min at 100 psi up to 1.72 L/min at 180 psi. The desorption and ionization contributions to the dependency of DESI on the nebulizing gas flow rate were studied by SDC and RESI respectively.
As indicated by the results shown for SDC in Figure 7.4, desorption efficiency initially improves up to 140 psi before decreasing again at higher pressures, following the same trend that is observed for DESI and optimizing at the same pressure. It was previously shown that with increasing gas pressure droplet velocities increase and a certain threshold velocity is required for the efficient production of secondary droplets after surface collision.\textsuperscript{5, 24-25} At gas flow rates lower and higher than the optimum value, sample re-deposition\textsuperscript{10} was also observed, which occurs when material desorbed from the spray deposits on the surface past the point of impact. This is likely caused by decreasing impacting droplet velocity for low gas flow rates and a larger force on the secondary droplets by the gas for high flow rates. It is further possible to explain the observed ion signal dependence on gas pressures by considering the effect of gas pressure on the size
of the impact zone or wetted spot. With higher gas pressures the solvent spray becomes more focused and evaporation increases, reducing the spot size somewhat. At very high gas pressures it is possible that the system becomes too dry, causing the wetted spot on the surface to disappear and a decrease in desorption efficiency. By probing the charge distribution on a nonconductive surface without the presence of an inlet providing suction, Cooks et al. demonstrated that a higher gas pressure focuses the charges arriving at the surface, a result that agrees with the smaller desorption spot sizes observed visually with higher gas pressures.

The dependence of ionization efficiency on nebulizing gas pressure settings was studied by RESI. At the lowest gas flow rate (≤ 100 psi), a high signal intensity is observed because there is not enough driving force exerted by the nebulizing gas to effectively direct droplets produced by the microsprayer toward the sample surface. These droplets and ions fly directly toward the atmospheric pressure inlet without impacting the surface, which leads to the initially high ion signal with RESI. On the other hand, in DESI, when no analyte ions are present in the spray solvent, no (or low) signal intensities are observed in this pressure regime. Ignoring this initial anomalously high response with RESI, under conditions where the primary droplets do collide with the surface, the signal continuously increases with increasing gas pressure. If the spray is sufficiently wet for the formation of the liquid film, secondary droplet production efficiency increases with increasing gas pressure. It has been previously shown that droplet diameters decrease with increasing gas pressure settings, confirming that, as is the case for ESI, smaller droplet sizes lead to more efficient ionization with RESI and thus with DESI.
While both desorption and ionization are affected by changes in gas flow rates, the close correlation between the SDC and DESI trend lines indicates that desorption effects dominate during optimization of the gas pressure. The mild increase in ionization efficiency observed at higher gas pressure settings does not compensate for the loss in desorption efficiency caused by sample re-deposition and the reduced size of the wetted spot at high gas pressure settings and the DESI signal intensity decreases, even as ionization efficiency improves.

7.4.4 Optimization of the Sprayer-to-Impact Zone Distance

![Figure 7.5](image.png)

Figure 7.5 Results for the optimization of the sprayer-to-impact zone distance for SDC (squares), RESI (triangles), and DESI (diamonds). Error bars represent one standard deviation of triplicate analyses. Lines connecting individual data points are presented only to guide the reader’s eye.

Increasing the distance between the sprayer and the impact zone was shown to decrease the velocity of droplets impacting the surface as well as the velocity of droplets leaving the surface. In the present study, the distance between the impact zone and mass
inlet or collection surface was kept constant while the distance between the sprayer and
the impact zone was varied from 1 mm to 8 mm by moving the sprayer diagonally away
from the surface. Figure 7.5 shows the results for the optimization of the sprayer-to-
impact zone distance for DESI, SDC, and RESI.

The dependence of the signal intensity observed for DESI on the distance between
the sprayer and the impact zone displays an interesting morphology. Instrumental
response initially increased with increasing distance up to 3 mm, followed by a ridge at 4
mm, after which signal intensity decreases. This shape of the DESI dependence on
sprayer-to-impact zone distance curve can be explained by the combined contributions of
desorption and ionization efficiencies.

Similar to the DESI curve, the SDC curve increases to a maximum at 3 mm
sprayer-to-impact zone distance, after which the curve continuously decreases. At short
distances, desorption efficiency may suffer as the nebulizing gas and droplet velocities
are high, which results in sheeting of the solvent along the surface instead of secondary
droplet formation. As the droplet velocities decrease with increasing distance, desorption
efficiency increases until an optimum value is achieved. After this value, desorption
efficiency decreases along with droplet size and velocity. At this point, droplets may
contain too little kinetic energy to effectively remove material from the surface and some
removed material may begin re-depositing past the impact zone instead of making it to
the inlet. As shown by these results, DESI response is initially low due to low desorption
efficiency. As desorption efficiency increases, so does DESI response, although DESI
response increases more steeply due to the concurrent increase in ionization efficiency as
indicated by the RESI results.

129
The RESI curve continuously increases with increasing sprayer-to-impact zone distance. Increasing the distance between the sprayer and the impact zone results in smaller droplets impacting the surface, which has been shown to increase the efficiency by which ions are transferred from the surface to the mass spectrometer inlet.\textsuperscript{42} This transfer efficiency appears to increase over the entire range of sprayer-to-impact zone distances that was investigated in the current study. As desorption efficiency decreases after 3 mm, DESI response continues to remain high due to the increasing ionization and ion transfer efficiencies. After 4 mm, however, the decreasing desorption efficiency overcomes the benefit of increasing ionization efficiency as there is little material desorbed and therefore nothing to ionize.

\textbf{7.4.5 Optimization of the Impact Zone-to-Inlet Distance}

For a fixed impact zone-to-inlet distance, progeny droplets initially remain close to the surface after formation at the impact zone. Their velocity decreases with increasing distance from the impact zone due to aerodynamic drag and eventually increase in height due to dispersion.\textsuperscript{5} We also observed collection to occur in a narrow line level with the horizon of the sample surface during SDC,\textsuperscript{29} and Kertesz et al. have shown that the ‘edge sampling’ geometry gives the best response for the same reason.\textsuperscript{34} Figure 7.6 shows the results for the optimization of the impact zone-to-inlet distance for DESI, SDC, and RESI.

While studying desorption efficiency using SDC, we could not optimize for distances shorter than 2 mm since it was necessary to leave a small space between the sample surface and the collector surface. The SDC response was not expected to change substantially with an increase in distance between impact zone and inlet since this region
reflects what happens to droplets and ions after desorption. From the SDC curve in Figure 7.6, it can be seen that desorption falls less than 20% from its optimal value at the largest distance. The observed decrease is likely due to dispersion of gas and desorbed material after leaving the sample surface, leading to a slight increase in the size of the collection zone. In DESI, this dispersion may also reduce the collection efficiency of the produced progeny droplets and ions by the atmospheric pressure interface.

The RESI curve was observed to increase with increasing distance until a maximum was achieved between 3 and 4 mm, after which the curve decreases. The smaller, faster moving progeny droplets formed during droplet collision with the surface\textsuperscript{24} tend to remain close to the surface.\textsuperscript{5} At very short distances, these droplets
might not achieve the height necessary to avoid hitting the bottom rim of the inlet capillary or re-depositing onto the sample surface. The dispersion of these small, fast-moving droplets increases with increasing distance from the impact zone,\(^5\) so as the impact zone-to-inlet distance increases, the sampling of progeny droplets and ions by the inlet also increases as they gain height from the surface. In addition, longer flight times will provide more time for evaporation and ion production from the progeny droplets. After the optimum distance, the decrease in the RESI curve is likely due to a decreasing sampling efficiency by the atmospheric pressure interface as the dispersion of the plume continues to increase, decreasing the fraction of the reflected plume that is sampled by the inlet.

The DESI curve follows the same trend as the RESI curve, although the DESI curve is lower initially. For both techniques, an increasing portion of the charges bypass the surface and head directly toward the inlet due to the increased electric field at very short distances.\(^7\) For DESI, this results in decreased desorption efficiency and low signal while for RESI the response is higher due to the presence of analyte ions already present in the spray solution. The DESI dependence on the impact zone-to-inlet distance appears to depend mostly on the ion transport and dispersion processes occurring post-desorption.

7.5 Conclusion

We have found that the optimum conditions for desorption do not always coincide with those for ionization during the concerted DESI process. Often, the optimum conditions for DESI are a compromise between those conditions that lead to efficient desorption and ionization respectively. In fact, the response trends for desorption and ionization were different within each parameter investigated, which demonstrates the
inherent complexity of the optimization process for DESI. In the present study, a single analyte and surface combination was used to optimize several parameters. Different analytes, such as proteins, may optimize differently based on their individual desorption and ionization requirements. However, the general optimization trends shown here for each parameter agree with published results.

7.6 References


CHAPTER 8

INVESTIGATING PROTEIN ANALYSIS BY DESORPTION ELECTROSPRAY IONIZATION MASS SPECTROMETRY

Reprinted (adapted) with permission from

K. A. Douglass and A. R. Venter.
Protein analysis by desorption electrospray ionization mass spectrometry and related methods.

Copyright © 2013 John Wiley & Sons, Ltd.

8.1 Introduction

Desorption electrospray ionization mass spectrometry (DESI-MS) requires little to no sample preparation and has been successfully applied to the study of biologically significant macromolecules such as proteins. However, DESI-MS, and likely other ambient methods that use spray desorption to process samples during ionization, appear limited to smaller proteins with molecular masses of 25 kDa or less\(^1\) and a decreasing instrumental response with increasing protein size has often been reported.\(^1,2\) It has been proposed that this limit results from the inability of some proteins to easily desorb from the surface during DESI sampling.\(^1,3,4\) In this chapter, the apparent mass dependency of the instrumental response observed during the DESI-MS analysis of proteins is investigated using spray desorption collection (SDC) and reflective electrospray ionization (RESI), methods which are described in Chapter 7. Proteins as large as 66kDa are shown to be quantitatively removed from surfaces using SDC. However, incomplete dissolution and the formation of protein-protein and protein-contaminant clusters appear to be responsible for the mass dependent loss in sensitivity for protein analysis.
Alternative ambient mass spectrometry approaches that address some of the problems encountered by spray desorption techniques for protein analysis are also discussed.

8.2 History of Protein Analysis by Desorption Electrospray Ionization

Desorption electrospray ionization mass spectrometry (DESI-MS) has become an attractive method for the study of biologically significant molecules such as lipids and proteins because of its softness and ambient nature. Extensive work by many research groups has shown the applicability of DESI-MS to the analysis and imaging of small molecules such as lipids, but protein imaging by DESI-MS, as well as the DESI-MS analysis of proteins in general, has seen very little development. Currently, MALDI-MS is the preferred method for the native analysis of proteins. MALDI-MS can be used to analyze proteins up to several hundred kilo Daltons and is capable of high spatial resolution during imaging experiments. However, MALDI requires the application of a matrix to the sample and the choice of matrix impacts the types of molecular ions that will be observed. In addition, the protein ions observed by MALDI are typically singly-charged, so analysis requires a mass spectrometer with a high mass range, although some reports have demonstrated extensive protein multiple charging. DESI has several advantages over MALDI, such as little to no sample preparation, unique chemistry that can occur due to the liquid sampling spray (e.g. reactive DESI), creation of multiply-charged ions, etc. For these reasons, ambient analysis of proteins remains of interest.

To perform a DESI experiment, a surface containing a sample of interest is placed in front of the mass spectrometer. Pneumatically accelerated solvent droplets produced by electrospray are directed at the surface. Under typical experimental settings, the primary DESI droplets generated at the sprayer tip have velocities around 100 to 120 m/s. After
the solvent spray collides with the surface, ions and droplets composed of the spray solvent and analyte materials from the surface are sampled into the atmospheric pressure interface of the mass spectrometer. Three possible mechanisms for the generation of analyte ions formed during DESI have been posited\textsuperscript{2, 11-13} and which mechanism occurs depends on the operating conditions and type of analyte being analyzed. The most common ionization method, and the one that is of interest for protein analysis, is known as the droplet pick-up mechanism.\textsuperscript{2, 10-11, 14-20} Droplets produced by the electrospray source collide with the sample and spread into a thin film covering its surface.\textsuperscript{19, 21} Analyte molecules are extracted or dissolved into this film and subsequent collisions by primary droplets result in the ejection of secondary droplets from the liquid surface layer containing the analyte. Two populations of these secondary droplets have been observed: Larger droplets which leave close to their specular angle and smaller droplets which get caught within the nebulizing gas flow close to the surface because of their smaller momentum.\textsuperscript{10, 15} Since the optimum collection angle for DESI has often been found to be $\leq 10^\circ$,\textsuperscript{2, 22} it is supposed that these smaller droplets are the primary source of detected ions. The narrow profile of collected material at the horizon of a sample surface, for a brightly colored rhodamine standard, also confirmed that the bulk of desorbed material is present in these low flying droplets.\textsuperscript{23} After desorption it is believed that typical electrospray ionization mechanisms lead to the generation of free gas phase ions. For proteins, DESI-MS spectra have characteristics, such as the observation of adducts, multimers and the conformational dependence on solvent composition, that suggest this ionization process is possibly similar to the charge residue model (CRM) proposed by Dole et al.\textsuperscript{2, 24} or possibly the other methods described in Chapters 1 and 2. The second DESI mechanism,
heterogeneous charge transfer\textsuperscript{2,12-13} occur for less polar small analyte molecules typically analyzed by atmospheric pressure chemical ionization (APCI) such as cholesterol or TNT while the final method, neutral extraction\textsuperscript{25}, occurs when neutral species in the headspace of the sample surface are captured into droplets and ionized by electrospray mechanisms.

In an early report by Takats et al.,\textsuperscript{2} several DESI-MS experimental parameters were optimized for the individual analysis of peptides and proteins varying in mass from 1 to 15 kDa, including the solvent system, source configuration, and sample surface. These parameters were then used to determine the detection limit for each peptide or protein. The samples consisted of 1 – 5 µL aliquots air-dried on various surfaces. Although there were slight variations in the optimal parameters for the individual molecules, the data showed that a smooth surface such as glass or PMMA, as well as a high angle of incidence, was suitable for DESI-MS analysis when MeOH/H\textsubscript{2}O was used as the solvent system. The sensitivity of DESI-MS towards protein analysis was observed to be slightly less than that of MALDI-MS, and the limits of detection for the proteins increased with increasing molecular mass.

In a later report by Shin et al., the DESI-MS detection limits were determined for several proteins spanning a mass range of 12 to 66 kDa.\textsuperscript{1} The protein solutions were spray-deposited onto PMMA substrates at various surface concentrations. The spectra were then deconvoluted and the detection limits were assigned to the smallest surface concentration for which a S/N > 5 could be obtained. Similar to the previous study by Takats et al., the limits of detection were observed to increase with increasing molecular mass, and a S/N > 5 was unable to be obtained from the largest proteins, ovalbumin and bovine serum albumin. For these two proteins, the authors instead reported the lowest
surface concentration from which ion peaks could be assigned to the respective proteins. A logarithmic plot of the detection limits (or lowest detectable surface concentrations) determined from both studies is presented in Figure 8.1. Although different methodologies and instruments were used in the two studies (for instance, the Thermo Scientific LTQ is about 10 times more sensitive than the Thermo LCQ, as seen by the lower limits of detection for identical proteins studied by both groups at around 12 kDa), there is a clear trend of increasing limit of detection with increasing molecular mass. This loss in sensitivity with protein size has often been attributed to the lack of efficient desorption of proteins from the sampled surface,\textsuperscript{1, 3−4} although this hypothesis has never been experimentally verified. In part, this may be due to the difficulty in separating the concerted desorption and ionization processes that are coupled during a DESI analysis.

In Chapter 7, the separate optimizations of desorption and ionization in DESI-MS were discussed.\textsuperscript{26} Desorption and ionization were found to optimize differently and the optimum parameter values for DESI were observed to generally be a compromise

\textbf{Figure 8.1} Logarithmic plot of protein DESI-MS detection limit vs. molecular mass collected on Thermo Fischer Scientific LTQ (♢)\textsuperscript{2} and LCQ (•)\textsuperscript{1}. 
between those optimum for desorption and those for ionization. To study desorption and ionization independently, it was necessary to develop novel methods of analysis which are used in the present study to specifically investigate protein desorption and ionization during the DESI analysis. Spray desorption collection (SDC) can be used to study desorption independently of ionization. In SDC, a typical DESI sprayer is used to sample a surface, but instead of real-time analysis by mass spectrometry the desorbed material is collected and analyzed at a later time. The collection efficiency of the desorbed material does not depend on whether it has been ionized or not. Subsequent post-collection analysis can be achieved by numerous methods, including ESI-MS, UV/VIS spectroscopy, and fluorescence spectroscopy. By choosing the appropriate post-collection analysis method, it is possible to characterize the desorbed material while removing possible biases introduced by the concerted DESI ionization process. Figure 8.2 illustrates the sequence of experiments used to investigate desorption separately from ionization: A mixture of proteins were characterized by direct analysis ESI. The same sample was also deposited onto a surface and collected by SDC. The collected sample was then analyzed under identical ESI conditions and the relative abundances of the proteins were compared between the original and desorbed samples.

We further investigated the ionization of proteins during DESI independently of desorption. One could attempt to apply the observations for standard ESI directly to the DESI ionization process by using the same sprayer and conditions as used for the generation of the primary solvent spray in a DESI experiment. However, the distributions of size and velocity, as well as the spatial distribution, of droplets from which ionization occurs during the DESI process are markedly different from those of droplets generated
During DESI, a droplet’s collision with the surface also reduces the amount of charge it carries due to neutralization on the surface.\textsuperscript{20} In an attempt to mimic these so-called secondary droplets that leave the surface after collision, without requiring desorption or incorporation of the analyte, the spray plume produced by electrospraying the analyte solution was bounced off of a clean surface under typical DESI spray...
conditions. Both ESI and RESI were used here to investigate the ionization and ion transport processes occurring during the DESI-MS analysis of proteins.

8.3 Experimental Details

8.3.1 Sample Preparation

Bovine cytochrome c (cyt c, 12.2 kDa), bovine hemoglobin (Hgb, 15.1 kDa subunit), and glacial acetic acid were purchased from Sigma-Aldrich (St. Louis, MO) and β-lactoglobulin (18.3 kDa) and α-chymotrypsinogen (25.7 kDa) were purchased from MP Biomedicals (Solon, OH). The reagents were used as received. A solution containing each protein at 20 μM was prepared in 50:47:3 MeOH:H₂O:acetic acid, from which a mixture containing 1 μM of each protein was prepared by dilution to 50:47:3 MeOH:H₂O:acetic acid. Similar solutions containing 1 μM and 20 μM of both cyt c and bovine serum albumin (BSA, 66.4 kDa, Sigma-Aldrich) were also prepared by dilution to 50:47:3 MeOH:H₂O:acetic acid.

8.3.2 Experimental Sequence

Each set of experiments was performed in the following order: DESI-MS, SDC collection, ESI-MS of the mix, and ESI-MS of the SDC-collected sample. RESI-MS was performed on a separate day. All spectra were acquired on a linear ion trap mass mass spectrometer (LTQ, Thermo Scientific, CA). Figure 8.2 describes the relationship between the different modes of analysis. The same spray emitter was used for each experiment and was constructed in-house. Additional details on DESI source construction are presented in Appendix B.
8.3.3 Conditions for DESI, RESI and SDC

For SDC collection and DESI-MS, 3 µL spots of the 20 µM mix were deposited in a straight line on the edge of a 1/16” porous PTFE sheet (SmallParts.com, Logansport, IN). Once the spots had dried, an automated 2-D sample positioning stage installed in front of the MS inlet was used to scan the sample through the solvent spray at a constant rate of 150 µm/s. For DESI-MS, 10 spots were scanned in a single acquisition and the spectrum was generated by averaging all scans. For SDC, a 1.5 mL Eppendorf tube was placed between the sample and the MS inlet to capture desorbed material for subsequent ESI analysis as shown in Figure 8.3a. The Eppendorf tube was modified with two off-axis holes in the rear (Figure 8.3b), which might allow excess nebulizing gas to escape, reducing the pressure barrier at the tube entrance. About 35 spots were scanned, followed by several rinses of the collection tube into a second tube using 200 µL of 50:47:3 MeOH:H₂O:acetic acid (Figure 8.3c).

Identical instrumental conditions and the same sprayer in an identical configuration as for DESI was used for the RESI experiments; only the solutions and surfaces were different. For DESI, the 1 µM protein mixture was deposited onto a porous Teflon surface, allowed to dry, and then scanned by the electrosprayed solvent, while for RESI the 1 µM protein mixture was sprayed while scanning across a clean porous PTFE surface.

8.3.4 Conditions for ESI

The same spray source used for DESI-MS and RESI-MS was used for the ESI-MS analysis of the 1 µM mix and SDC-collected samples. The experiment was run with front-end optics settings optimized consecutively for the efficient transfer of lower and
Figure 8.3 (a) An Eppendorf tube was used for collection following SDC. (b) The Eppendorf tube was modified by adding two small holes to the bottom for rinsing and squeezing it into an elongated shape to improve collection efficiency. (c) The collection tube was rinsed several times with the same aliquot of solvent to reconstitute the collected sample.
higher m/z values. These optimum settings were determined by tuning the system by automatic routine on either the +16 cyt c peak (m/z 765) or the +13 α-chymotrypsinogen peak (m/z 1974) of an infused sample of the protein mix. The system parameters for each experiment are presented in Table 8.1.

Table 8.1. Unless otherwise specified in the text, the experimental parameters used for DESI-MS, SDC-MS, RESI-MS, and ESI-MS.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrospray voltage</td>
<td>+5 kV</td>
</tr>
<tr>
<td>Tube lens voltage</td>
<td>+160 V or +250 V</td>
</tr>
<tr>
<td>Solvent flow rate</td>
<td>5 μL/min</td>
</tr>
<tr>
<td>Nebulizing gas pressure</td>
<td>120 psi</td>
</tr>
<tr>
<td>Sprayer angle of incidence</td>
<td>50°</td>
</tr>
<tr>
<td>Sprayer distance to inlet (DESI, RESI)</td>
<td>4 mm</td>
</tr>
<tr>
<td>Sprayer distance to surface (DESI, RESI, SDC)</td>
<td>1 mm</td>
</tr>
</tbody>
</table>

8.4 The Mass Dependent Instrumental Response and Desorption

We tested the apparent dependency of protein desorption on protein size observed during DESI-MS using five common proteins: cyt c (12.2 kDa), Hgb (15.1 kDa), β-lactoglobulin (18.3 kDa), α-chymotrypsinogen (25.7 kDa) and BSA (66.4 kDa). The charge state distributions for four of these proteins could be observed simultaneously using a linear ion trap mass spectrometer operating in high mass mode. In a separate experiment, BSA (66.4 kDa) was also analyzed under identical conditions as the other proteins to validate the observations for large proteins. It was not possible to add and analyze BSA into the mixture with the other proteins as the complexity of the extensive
multiple charged ion envelope of BSA obscured the charge states of the other proteins ions and complicated data analysis.

When a protein mixture containing equimolar concentrations of several proteins varying in molecular mass is analyzed by direct infusion ESI-MS, a mass dependent response bias is observed. There are a number of reasons leading to this overall trend, which includes differences in both ionization efficiencies\(^{30-32}\) and ion transfer during MS analysis. Such a result is shown in Figure 8.4 for proteins analyzed by a linear ion trap instrument for four proteins ranging in size from 12.2 kDa (cyt c) to 25.7 kDa (α-chymotrypsinogen). In ion trap instruments such as the Thermo Scientific LCQ and LTQ, ions enter the low-pressure region of the MS through a transfer capillary positioned off-axis from the center of a skimmer cone through which the ions must pass through to enter the vacuum region of the mass analyzer (Figure 4.6). The tube lens is a component of the ion optics that directs the entering ion beam towards the orifice on the skimmer cone by creating an electric field within the reduced pressure area. The optimum value of the tube lens voltage is proportional to the mass-to-charge ratio of the ion of interest. Due to the transient nature of the signal obtained from most samples prepared for ambient analyses
such as DESI-MS, the mass spectrometer ion optics are usually optimized while infusing a standard solution. For the results shown in Figure 8.4, the ion optics were optimized on the cyt c +16 peak (m/z 765) and this is a major reason behind the decreasing trend in relative intensities with increasing molecular mass observed for the highest intensity charge states of this set of proteins when analyzed directly by ESI-MS. The same mass dependency of ion sampling would also be in effect for DESI because the same ion optics are used to move ionized proteins through the atmospheric pressure interface (API) of the mass spectrometer.

In addition, with most previous DESI work, as is the case also for the aforementioned studies\textsuperscript{1,2}, the mass range was limited to a maximum of m/z 2000. However, the charge state distributions for β-lactoglobulin and α-chymotrypsinogen extend beyond this mass range in the data published by other researchers,\textsuperscript{1} which may have artificially decreased the sensitivity (increased the detection limits) for these higher molecular weight (MW) proteins. To circumvent this complication in the work presented here, the mass spectra were recorded between m/z 500 and 4000 using the high mass scan mode of the LTQ mass spectrometer.

These ion transfer and instrumental mass range contributions to the loss in sensitivity for larger proteins do not account for the more pronounced mass dependent loss in sensitivity observed during the DESI analysis of proteins. The biggest cause of the mass dependent loss in sensitivity has been ascribed to the increasing difficulty of desorbing larger proteins.\textsuperscript{1,3-4} However, this has never been experimentally verified.

We were able to test the hypothesis that during a DESI experiment smaller proteins are more effectively desorbed than larger ones by comparing the results for
DESI-MS with those for SDC-ESI-MS. As shown in Figure 8.4, when the samples collected during SDC were analyzed by ESI-MS it was found that the original equimolar ratio of proteins was conserved during the desorption and collection processes. The highest intensity charge states of each collected protein were observed at the same relative intensities as obtained by direct analysis of the original protein mixture by ESI-MS. This observation does not support the hypothesis that the efficiency of protein desorption is dependent on protein size for the range of proteins investigated in this study, and therefore there must be different reasons for the lack of sensitivity observed for larger proteins by DESI-MS.

8.5 Investigating Other Sources

To investigate other potential reasons for the loss of sensitivity or mass dependent bias in levels of detection observed for larger proteins, we analyzed the mixture of proteins by direct ESI-MS, DESI-MS, SDC-ESI-MS, and RESI-MS at two tube lens settings. As explained above, the key mass dependent tuning parameter for the mass spectrometer used most often for DESI analyses is the tube lens setting. All parameters of the MS system were first automatically tuned using an intermediate mass, the β-lactoglobulin +13 peak (m/z 1413) from a solution that was directly infused using the microspray ESI source described above. Thereafter, data was collected both at the optimum tube lens setting of +160 V for the +16 cyt c peak (m/z 765) and at the optimized tube lens setting of +250V for the +13 α-chymotrypsinogen peak (m/z 1974). These results are shown in Figure 8.5. Cyt c and BSA (66kDa) were also analyzed together in a similar fashion since BSA could not be included in the protein mixture due to its wide charge state distribution. These results are shown in Figure 8.6.
Figure 8.5 Direct ESI-MS (a,e), SDC-ESI-MS (b,f), DESI-MS (c,g), and RESI-MS (d,h) of a protein mixture at low (a-d) and high (e-h) tube lens settings. Triangles mark a peak attributed to one of the four proteins.
Figure 8.6 Cytochrome c (red triangles) and bovine serum albumin (unlabeled peaks under the brackets) analyzed simultaneously at the tube lens values of +160 V and +250 V for direct ESI-MS, SDC-ESI-MS, and DESI-MS.
For ESI-MS, the protein signal intensities decrease with increasing MW at the lower tube lens setting while intensities increase with increasing MW at the higher tube lens setting. This trend reversal upon adjustment of the tube lens value demonstrates that the tube lens setting plays a major role in the observed relative intensities of each protein, although other factors such as ionization efficiencies play a role as well.

The spectra obtained for SDC-collected protein subsequently analyzed by ESI-MS are nearly identical to the direct ESI-MS spectra, suggesting that desorption efficiencies of the proteins during DESI-MS are similar and therefore are not the major factor in the mass-dependent sensitivity seen during DESI-MS analysis. The RESI-MS spectra were also very similar to the direct ESI-MS spectra, and the slight shift in charge state distributions are consistent with a small amount of charge neutralization of the droplets following their collisions with the surface. This slight loss in charging observed by RESI is not enough to account for the decrease in detection limits for DESI by itself, though. Since the droplet interactions with the surface and the post-collisional spray distribution of DESI are mimicked by RESI, ionization and ion transmission events occurring post-desorption are also ruled out. However, the DESI-MS spectra are clearly different from the direct ESI-MS spectra. At the lower tube lens value of +160 V, β-lactoglobulin and α-chymotrypsinogen are not observed. Instead, several peaks corresponding to dimers of cytochrome c and hemoglobin are present. At the tube lens value of +250 V, α-chymotrypsinogen is observed, but at a lower intensity than cytochrome c and hemoglobin despite the high tube lens value. Therefore, there must be some component of DESI which fails to be captured in the SDC and RESI tests, i.e. a
process intrinsic to the overall process that is not solely due to protein removal, ionization, or ion transfer.

8.6 The Effect of Ineffective Protein Dissolution

During RESI, desorption is excluded by ensuring the analyte is already present in the spray solution prior to collision with the surface. Similarly, the ESI-MS analysis of samples after collection by SDC requires the analyte to be pre-dissolved in the spray solution. Neither of these methods captures the transition of the analyte from a dry solid present on the surface to an aqueous solution on the time-scale of DESI.

The important role of analyte dissolution into the microfilm of solvent, produced in situ during the DESI experiment, is illustrated by related ambient analysis methods that approach the dissolution step in a different fashion. For example, it has been shown that DESI-MS becomes amenable to the study of proteins up to 150 kDa if a liquid sample is analyzed directly without any preparative deposition and/or drying.\textsuperscript{33-34} Two other ambient surface-sampling techniques which can analyze dry solids and generate ions via electrospray, but differ from DESI in their desorption mechanism, are the liquid microjunction surface sampling probe (LMJ-SSP)\textsuperscript{35} and nano-DESI.\textsuperscript{36} Schematics of these techniques are presented in Figure 6.3. Instead of directing an electrospray at a surface as in DESI, the LMJ-SSP and nano-DESI methods use liquid microjunctions to desorb material. The LMJ-SSP uses two coaxial capillaries terminating a specific distance above the surface. The outer capillary delivers solvent to the surface, forming the liquid microjunction, while the inner capillary samples the liquid microjunction via aspiration generated by an electrospray formed at the other end. The configuration of nano-DESI is similar, with the main difference being that the solvent delivery capillary is
not coaxial with the sampling capillary. Both methods have demonstrated imaging capabilities\textsuperscript{37-39} as well as the ability to desorb proteins for successful analysis by mass spectrometry.\textsuperscript{35-36, 40}

During DESI-MS, the time between desorption and analysis is short and the sample is usually rapidly consumed\textsuperscript{41} due to the force of the pneumatically-assisted electrospray. This is not the case in the LMJ-SSP and nano-DESI techniques, where the sample is only removed from the surface via dissolution into the solvent and there is a considerable delay between extraction and ionization. When time-dependent response curves were obtained by nano-DESI for rhodamine 6G (443 g/mol), reserpine (608 g/mol), and cytochrome c (~12300 g/mol),\textsuperscript{36} the times required to reach maximum instrumental response were 7, 30, and 90 seconds, respectively. For comparison, the maximum instrumental response for rhodamine 6G from glass was observed immediately at the onset of DESI in a study by Green et al.\textsuperscript{41}

Recently, Trimpin et al. also demonstrated that highly multiply charged protein ions similar to those observed during ESI-MS (or by extension, DESI-MS) can be formed in the inlet of a mass spectrometer without the use of solvent or high voltage, provided that the sample was prepared in a suitable matrix.\textsuperscript{42-45} It has been proposed that highly charged droplets composed of the analyte and matrix compound are formed and transferred through the heated transfer capillary of a mass spectrometer and charging then proceeds similarly to ESI via desolvation of these droplets in the reduced pressure region of the ion source.\textsuperscript{42} It was found that the abundance of multiply-charged ions could be greatly increased with longer homogenization times of the analyte and solid matrix mixture before deposition.\textsuperscript{43} Similarly, the sensitivity of electrode-assisted desorption
electrospray ionization (EADASI-MS), a technique similar to DESI-MS, for larger proteins was improved upon with the addition of sinnapinic acid to the sample deposition.46

In both the liquid microjunction methods as well as inlet ionization, protein ions are formed from solid samples where the protein-protein interactions have been removed or minimized. The long time required for cytochrome c to achieve maximum signal intensity in the nano-DESI study was attributed to a slower dissolution rate for the protein into the solvent. This is the time that is required for the proteins to break the interactions between themselves as well as the surface and form solvated molecules. Similarly, the increased abundance of multiply charged protein ions with longer homogenization times observed during inlet ionization is likely due to the increased separation of protein molecules within the sample. During traditional DESI, proteins have considerably less time for this solvation process and the sample is rapidly removed from the surface. It is possible that during DESI the protein-protein and protein-contaminant interactions survive the DESI process after material is removed from the surface, causing a loss of signal. For example, dimers are observed in the DESI-MS spectra in Figure 8.5 at the lower tube lens value of +160 V for both cyt c and hemoglobin which were not observed in the corresponding RESI-MS, SDC-ESI-MS, or ESI-MS spectra. Dimers or multimers of larger proteins would likely fall outside of the mass window available to the instrument used in this study. The weak interactions holding the protein clusters together fall apart when the clusters experience higher-energy collisions with atmospheric gases in the intermediate region of the API at higher tube lens settings.
In addition to the protein-protein interactions, depositing and drying protein solutions onto a surface leads to a concentrating effect of salts and other impurities that are present in the protein samples and LC-MS grade solvents used to prepare these solutions. This effect can clearly be observed when comparing the ESI-MS and DESI-MS spectra obtained for β-lactoglobulin. Figure 8.7 shows the +12, +13, and +14 peaks for β-lactoglobulin obtained by ESI-MS (a) and DESI-MS (b). During DESI-MS, the degree of
adduct formation is much greater compared to ESI-MS, resulting in a large reduction in sensitivity since the signal is distributed across many more species. Also shown for comparison are the +13, +14, and +15 peaks for cyt c acquired by DESI-MS (c). The charge states of cyt c, are mostly comprised of a single peak even during DESI-MS, resulting in a strong signal. Some proteins such as cyt c appear to have a low propensity to form adduction with contaminant species during the DESI process for reasons possibly dependent on individual protein characteristics, although it appears from our limited study that larger proteins tend to suffer more from this effect.

8.7 The Addition of Solution Additives for Improved Protein DESI

To test whether incomplete dissolution is a factor during protein analysis by DESI, the protein mixture was deposited and analyzed as before, but also with the addition of either 200 mM ammonium acetate or 200 mM ammonium bicarbonate to the protein sample (Figure 8.8). The three samples were each deposited on the same substrate so that the blank (no additive), the sample with ammonium acetate added, and the sample with ammonium bicarbonate added were each analyzed in order during the same experiment. The individual spectra were then extracted from their respective times on the chromatogram. It was thought that as the sample dried, the protein in the additive samples would become encased in a matrix of the added component, aiding in dissolution. This idea was based on the improvement in instrumental response observed by Trimpin and coworkers with increasing homogenization times of the protein with the matrix during matrix-assisted inlet ionization. The compounds were chosen due to their known solubility in the DESI solvent system and their availability, while ammonium bicarbonate specifically was chosen based on work by Konermann et al. In addition to the solution
additives, the instrumental tube lens voltage was inadvertently set to +90 V instead of +160 V. Although this change was not intentional, a lower tube lens value reduces declustering and thus further exacerbates the DESI protein problem, which is not detrimental to the current study.

Figure 8.8a shows the DESI-MS analysis of the protein mixture without any additives. Although some peaks attributable to protein ions are observed, the spectrum is very noisy. This result was expected based on the low tube lens value and the typical difficulties associated with protein analysis by DESI-MS. With the addition of ammonium acetate (Figure 8.8b), the spectrum does not appear as noisy. However, the protein signal does not appear to be significantly improved. In fact, while protein signal above m/z 1500 appears stronger, the protein signal below m/z 1500 appears to have mostly disappeared. In stark contrast, with the addition of ammonium bicarbonate (Figure 8.8c) the spectrum appears very clean and the protein signal has been remarkably improved.

It is clear from Figure 8.8 that if easier dissolution upon the inclusion of a matrix compound is the source of improvement for protein analysis by DESI, not all matrix compounds are equally effective. It is interesting to note that the samples which included the matrix compound each dried more ‘flaky’ than the solution without. Upon analysis, much more of the sample was removed for these samples. The sample including ammonium acetate dried the flakiest, and during analysis flakes could be observed to blow away from the surface before analysis due to the high nebulizing gas flow rate. This may have contributed to the low signal observed during the analysis of the ammonium acetate sample. The sample including ammonium bicarbonate, although flakier than the
control, appeared to adhere to the surface better than the ammonium acetate sample. It is reasonable to think, then, that the ideal matrix compound would effectively adhere to the surface while still providing the benefit of intermolecular separation for the protein.

**Figure 8.8** (a) Protein mixture deposited in 50/47/3 MeOH/H$_2$O/AcOH (b) The same protein mixture with the addition of 200 mM ammonium acetate. (c) The same protein mixture with the addition of 200 mM ammonium bicarbonate.
molecules. The present study only investigated two compounds at a single concentration. Further studies will be required to determine the optimum concentrations to include with the protein sample, as well as which compounds work the best.

8.8 Signal Distribution and Sensitivity

Finally, some proteins are present in multiple variants, a factor which further distributes the signal over multiple peaks and artificially increases the limits of detection during analysis by mass spectrometry. For β-lactoglobulin, two forms of the protein, A and B, constitute the two major peaks observed for each charge state of the protein. Figure 8.9 shows representative DESI-MS spectra of cytochrome c and β-lactoglobulin when analyzed individually. Due to the heterogeneity of the β-lactoglobulin signal, the DESI-MS intensity for β-lactoglobulin is around 50x less than that of cyt c. Similarly, several peaks for each charge state are observed for α-chymotrypsinogen during ESI-MS and it is expected that extensive additional adduct formation occurred during DESI-MS as for β-lactoglobulin, greatly reducing its sensitivity in the same manner. No peaks corresponding to β-lactoglobulin or α-chymotrypsinogen were observed in either of the

![Figure 8.9 DESI-MS spectra for cytochrome c (a) and β-lactoglobulin (b) analyzed individually. The maximum signal intensity of the b-lactoglobulin is ~50x less than that of cytochrome c.](image)

162
DESI-MS spectra in Figure 8.5 because the sensitivities of the other proteins were greater to the extent that the signals for these two proteins were driven to the baseline of the combined spectra.

8.9 Conclusion

Proteins of all sizes are removed quantitatively from the sample surface during the DESI experiment as illustrated by analyzing solutions prepared after spray desorption collection. Some larger proteins have an increased propensity to form ions which are not only the protonated forms of the single protein molecule, leading to a loss of signal. These additional peaks may arise from solvent or alkali metal adduction, protein modification (i.e. oxidation), and dimers or multimers of the proteins. The origin of the increased adduction or modification of proteins during DESI stems from the manner in which the samples are deposited onto the sample surface. When spotted, the proteins accumulate in a high concentration as the spots evaporate, as do any impurities present in the deposited solution such as ubiquitous alkali metals. The rapid removal and immediate analysis of material from the surface during DESI ensures that these compounds are present at high concentrations in the secondary droplets from which ions are generated and without adequate time to disrupt protein-protein and protein-contaminant interactions. Other ambient analysis approaches, such as nano-DESI and LMJ-SSP, circumvent the slow dissolution problem by ionizing from re-dissolved solutions of the proteins where adequate time has elapsed between surface removal and ionization. Mass-dependent tuning parameters do not possess a single set of tuning parameters for which the analysis of a set of compounds with a large variation in molecular mass is optimal, further compounding the mass-dependent sensitivity observed during DESI-MS. This is
especially unfortunate considering that DESI typically consumes the sample during analysis, which precludes multiple analyses of the same sample or sample position. Developments in fast switching between tuning parameters may alleviate this problem, although this would impact the resolution during imaging applications. The use of mass spectrometers which minimize mass-dependent ion transmission is perhaps a more attractive solution.

8.10 References


40. Van, B. G. J.; Ford, M. J.; Doktycz, M. J.; Kennel, S. J., Evaluation of a surface-sampling probe electrospray mass spectrometry system for the analysis of surface-


In 1998,\textsuperscript{1} Fenn and coworkers predicted that electrospray ionization mass spectrometry (ESI-MS) would become a powerful tool for the analysis of biologically important macromolecules, namely proteins. Its ability to quickly and accurately provide protein identity along with sequence and structural information, while easily coupling with chromatographic separation prior to analysis, have indeed set it apart from any other analytical method. As the science of protein analysis progresses, it is doubtless to this researcher that ESI-MS will continue to play an increasingly prominent role. At the present time, however, there remains much that we don’t know about the electrospray phenomenon with regards to protein ionization. Continued research is required to increase our understanding of the ESI-MS of proteins so that improvements to current and future applications can materialize, especially as the new class of ambient ionization methods increase in popularity.

In the present work, I have described my efforts to improve protein analysis by ESI-MS and desorption electrospray ionization mass spectrometry (DESI-MS), two closely related techniques. For ESI-MS, the origin and modification of the multiple charging envelope was investigated. For DESI-MS, the challenge of protein analysis, where instrumental response decreases with increasing molecular mass of the protein, was investigated by separating the desorption and ionization events of DESI.\textsuperscript{2-5}
9.1 Protein Charge State Prediction and Modification

9.1.1 Charge State Prediction

In Chapter 3, a simple algorithm was presented which enables the prediction of the highest intensity charge state of denatured proteins by inspection of their amino acid sequences. It is assumed that in the positive ion mode the basic amino acids arginine, histidine, and lysine are potential charge carriers along with the N-terminus, while glutamic acid, aspartic acid, and the C-terminus are assumed to be potential charge carries in the negative ion mode. Starting from the N-terminus in the positive ion mode, or the C-terminus in the negative ion mode, charges are assigned to each potential charge carrier with the provision that at least three amino acids separate each consecutive charge. The number of assigned charges using this method was demonstrated to match well with the highest intensity charge state of many proteins, especially those with molecular masses of ~30 kDa or less.

In the present work, no distinction was made between the potential basic sites even though the individual amino acids have different intrinsic basicities as well as different structures which may impact the possible distances between charges. It would be interesting to see if the amino acid identity, along with its position along the protein chain, impacts the minimum separation of charges. Knowing this information could improve the accuracy of the predictions made by this method. One possible way to investigate this would be to have custom peptides synthesized with various numbers of non-chargeable residues separating chargeable amino acids of the same type. For example, polypeptides could be made where 1, 2, 3… etc. glycines separate 3 lysines. Similar polypeptides would be made for each different chargeable residue. Each
polypeptide would then be analyzed by ESI-MS and any differences in the average charge states of each spacing between the different polypeptides would indicate whether there are different requirements for the individual chargeable amino acids. It is expected that amino acids with more flexibility and/or higher intrinsic gas-phase basicities would require less separating residues between charges. These peptides can also be studied computationally using molecular dynamics calculations for proof-of-concept or corroboration with the experimental studies.

9.1.2 Charge State Modifications

In Chapters 4 and 5, the enhanced multiple charging of protein ions upon the addition of certain reagents, a phenomenon known as supercharging, was discussed. Using sulfolane, a common supercharging reagent, the average charge state of cytochrome c was significantly enhanced. It was shown that at low instrumental declustering settings, significant sulfolane adduction occurred preferentially for the higher charge states of the supercharged sample. This suggests that a direct interaction between the supercharging reagent and the protein charges sites occurs as part of the supercharging mechanism. Two possible interactions were proposed: (1) a diffusion of charge from the charge site through large scale dipole alignment of the reagent and (2) the formation of proton-bound complexes between the reagent and charge site. Both of these possibilities require further investigation.

For the first possibility, where supercharging is dependent on large scale dipole ordering diffusing charge from the charge site, further computational studies could be revealing. Similar to the approach pioneered by Konermann and coworkers, a charged amino acid could be surrounded by a mixture of supercharging reagent and solvent and
allowed to equilibrate. The alignment of the supercharging reagents relative to the charge could then be plotted as a function of the distance between them. For example, the alignment could be determined by the angle between the vector of the reagent dipole and the vector connecting the charge and the center of the reagent molecule. If there is ordering of reagent molecules surrounding the charge, the average alignment should increasingly deviate from being randomly distributed as the distance to the charge decreases. If not, then the alignment should be randomly distributed at all distances. Different reagents may be able to diffuse charge farther from the charge site, possibly dependent on factors such as the magnitude of the dipole and molecular size. In addition, there might be a minimum concentration of supercharging reagent necessary to result in supercharging by charge diffusion, a result which could be verified by experimentation.

For the second possibility, where proton-bound complex formation between reagents and charge sites results in supercharging, the detection of these adducts would be a good first step towards confirmation. So far, only sulfolane has been studied in this manner.\textsuperscript{4} One challenge is that these adducts, if formed, do not appear to readily survive until detection under normal operating conditions. It may be a good choice, then, to use a differential mobility analyzer (DMA) in-line with the mass spectrometer,\textsuperscript{12} which separates ions based on mobility through a drift gas similar to ion mobility spectrometry. However, the DMA detection would occur prior to transfer into the mass spectrometer, eliminating the possibility of adduct loss due to transmission through the high vacuum region of the mass spectrometer. It may also be possible to use the tube lens of certain mass spectrometers as described in Chapter 4. Another interesting possibility to investigate labile adducts would be to use super-atmospheric pressure electrospray
In 2011, Chen and coworkers introduced a novel ionization source to perform ESI at pressures up to 7 bar which was shown to be gentler than traditional ESI. These lower-energy ions may be less likely to lose adducts during transmission.

In my opinion, a satisfactory explanation of protein supercharging should predict the following observations:

1. In the positive ion mode, there is an increase in the highest observed charge state despite an increase in the gas-phase basicity of the solvent components.\textsuperscript{11}
2. A loss of signal and/or extensive adduction is sometimes observed at intermediate reagent gas-phase basicity values (Chapter 5).
3. Different reagents require different concentrations for supercharging.\textsuperscript{9,15-16}
4. Sometimes, non-Gaussian charge state distributions are observed during supercharging.
5. Supercharging reagent adducts can be observed at the highest charge states in
   the positive ion mode.\textsuperscript{4,9}
6. The dependence of supercharging on reagent gas-phase basicity is not linear in the positive ion mode.\textsuperscript{4}
7. Low gas-phase basicity reagents highly reduce the average charge state
   (Chapter 5).

At the current time, no one model accurately predicts each of these observations. In fact, it is certainly possible that supercharging can be achieved through more than one mechanism. For example, it is known that denatured proteins have higher charge states than those in the native state,\textsuperscript{17} yet supercharging can also occur in the absence of denaturing.\textsuperscript{12} The complexity of the supercharging phenomenon, and the importance of
protein charging in general, will ensure that these topics will remain popular areas of research for some time.

9.2 The Improved DESI-MS Analysis of Proteins

In Chapter 8, the challenge of protein analysis by DESI-MS was investigated using methods which separately studied the ionization and desorption events of DESI.\(^3\) The development of these methods is described in Chapter 7. It was discovered that protein molecules desorb equally from the sample surface regardless of protein size, excluding this factor from the loss in instrumental response observed with increasing protein size. Instead, it was shown that incomplete dissolution of the dried protein sample is the likely cause.\(^3\) This was confirmed using data from alternate methods of analysis, such as liquid DESI,\(^{18-19}\) nano-DESI,\(^{20}\) the liquid microjunction surface sampling probe (LMJ-SSP),\(^{21}\) and matrix assisted inlet ionization (MAII).\(^{22}\) In order to improve the analysis of proteins by DESI-MS, strategies must be developed to improve sample dissolution on the timescale of DESI, some of which were discussed in Chapter 8.

At the end of Chapter 8, it was shown that a remarkably improved DESI spectrum could be obtained by mixing the DESI sample with ammonium carbonate. However, the same improvement was not observed upon the addition of ammonium acetate. It appears that certain additives work better than others, and from the limited study that was conducted it is unclear what makes a good matrix compound. The optimum concentration of the matrix compound is also unknown since the presented work only investigated a single matrix concentration. It is possible that different matrix compounds are most effective at different concentrations. A study using several matrix compounds over a
range of concentrations should be conducted to determine which factors are important for the best matrix compounds.

Similarly, it is envisioned that supercharging reagents might find use as matrix compounds for protein analysis by DESI with two-fold benefit: cleaner spectra and increased multiple charging. Such a result could enable the combination of top-down proteomics with DESI-MS analysis. Unexpectedly, I was unable to find any studies during a literature search which have combined top-down proteomics with DESI. It is possible that the typical difficulties associated with protein DESI-MS analysis, such as low signal intensity and low applicability to proteins with high molecular masses, have limited the appeal of this approach. Unfortunately, many supercharging reagents reduce signal intensity at high concentrations\textsuperscript{4, 16} as well as increase adduction\textsuperscript{4, 9, 23} during supercharging by ESI-MS, factors which might exacerbate these problems. Furthermore, experiments where top-down proteomics would be most interesting, such as the analysis of complex biological samples like tissue sections, may not be amenable to the homogenous addition of matrix compounds, an apparent requirement for promoting cleaner spectra (Chapter 8).

9.3 References


A.1 Rationale

Mass spectra of proteins are complex due to the many peaks in their charge state distributions. Typically, the number of peaks present in a protein mass spectrum increase with increasing molecular mass. These individual charge states are usually represented by a base peak corresponding to the protonated ion. However, each charge state may be divided among several other peaks due to the attachment of one or more non-proton adducts. In the positive ion mode, these adducts are often solvent molecules (H₂O, MeOH, etc.), charge-neutralizing anionic species (HSO₄⁻, Cl⁻, etc.), metal cations (Na⁺, K⁺, etc.), or non-metal cations (i.e. NH₄⁺). Charge states can also be distributed across peaks corresponding to modified forms of the protein, which may have been present in the solution prior to ionization (glycosylation, methylation, etc.) or formed during ionization/ion transport (i.e. fragmentation). For example, Figure A1 shows the mass spectrum of cytochrome c from a supercharging solution containing sulfolane (left) as well as an exploded view of the +21 charge state (right). Besides the peak corresponding to [cyt c + 21H⁺]²⁺ (circles), peaks are also observed corresponding to [cyt c + 21H⁺ + x sulfolane]²⁺ (diamonds), where x = 1, 2, or 3, [cyt c + (21+y)H⁺ + yHSO₄⁻]²⁺ (squares), where y = 1, 2, or 3, and peaks with a mixture of both sulfolane and hydrogen sulfate (triangles). In the full mass spectrum, it can be seen that all charge states are comprised of many adduct peaks besides the base peaks.
179

When interpreting the mass spectra of proteins, it is often desirable to know the mass of a protein (especially if the identity of the protein is unknown), its total ion signal, and/or its average charge state. An estimate of a protein’s mass can easily be calculated by hand if the charge state of a protein peak is known, since the mass of an ion is simply its mass-to-charge ratio (m/z) multiplied by its charge, taking into account the mass of the extra proton(s) or other charged species if necessary. The charge state \( z_2 \) of a protein peak \((m/z)_2\) can be determined if the peak corresponding to one less charge state is known \((m/z)_1\):

\[
z_2 = \frac{(m/z)_2}{(m/z)_1 - (m/z)_2}
\]  

(A.1)

Some computer programs are capable of calculating the mass of a protein, or several proteins, present in a mass spectrum through a process called deconvolution. These programs are able to determine which peaks in a mass spectrum belong to the same m/z series, combining them and replotting the data in a mass-only dimension. For this purpose, I typically used a program called MagTran\(^1\), which was provided to me upon request by its author Zhongqi Zhang. Although deconvolution collapses the whole charge

Figure A1. A spectrum of cytochrome c supercharged with sulfolane (left) as well as an exploded view of the +21 charge state (right) showing several “adduct peaks”.
state distribution into a single mass peak it has been demonstrated that the intensities of these peaks can be inaccurate. To calculate by hand the total ion signal for a protein spanning many m/z values, the sum of all charge state intensities (I) must be calculated:

\[ \sum_{i=1}^{n} (I_i) \]  

(A.2)

For some information, such as the average charge state, there are no programs available to do the calculations, to my knowledge. To calculate the average charge state, the weighted contribution of each charge state must be determined by dividing the sum the products of each charge state (X) and their corresponding intensities (I) by the total intensity:

\[ \frac{\sum_{i=1}^{n} (I_i X_i)}{\sum_{i=1}^{n} (I_i)} \]  

(A.3)

While Equation A.1 is simple to solve quickly since only two m/z values are required, the evaluation of Equation A.2 or A.3 requires the knowledge of both the charge state and the intensity of each peak in the spectrum that belongs to the protein, a time-consuming and tedious task even for small proteins. This problem is even more daunting if large data sets need to be analyzed, such as when the average charge state is monitored for a change in instrumental parameters. To address this problem, I have taken advantage of Microsoft Excel and the ability to export mass spectrometry data as tables of x and y values to develop worksheets which greatly speed up the analysis of protein mass spectrometry data. Two worksheets were developed, one to calculate the average charge state of a protein and provide the individual intensity of each charge and one to map the intensities of each charge state across a third dimension, such as the tube lens value, which is an instrumental parameter.
A.2 Average Charge State and Intensities

Figure A2 shows the Microsoft Excel worksheet used to automatically calculate m/z values of protein charge states. Labels are described in the text.

Figure A2. The Microsoft Excel worksheet for automatically calculating m/z values of protein charge states. Labels are described in the text.

Figure A2 shows the Microsoft Excel worksheet used to automatically calculate the m/z value of each protein charge state with the addition of various adducts. (a) First, the raw data is pasted into columns A (mass) and B (intensity). In XCalibur, the software platform for Thermo Instruments, this data can be placed in the clipboard by right clicking a mass spectrum and choosing the exact mass export option. If the data is pasted into cell A1, it automatically arranges itself into the two columns. (b) Next, the mass of the protein is typed into cell D2. In order to calculate m/z values, the mass of the protein must be known beforehand. This mass can be determined by deconvolution (see above) or if the identity of the protein is known. (c) The mass of the desired adducts are then typed into the respective cells in column D. For instance, a value of 1 would be added to calculated the m/z values of the protonated peaks. (d) Finally, the mass values of the
protein and the first adduct are used to calculate the m/z value (column G) for each charge state (column F). Assuming that all charge comes from a single singly-charged source with a mass \( n \), the m/z value for the \( z^{th} \) charge state of a protein with mass \( m \) is:

\[
\frac{m}{z} = \frac{m + zn}{z} \tag{A.4}
\]

For example, cell G3 is calculated using the formula

\[=(D2+D5*F3)/F3\]

Using these m/z values, column H searches through column A for the values matching the m/z values in column G, then returns the corresponding intensities from column B, reporting the highest value within a range of m/z values matching the value typed into cell D23. For example, cell H3 is calculated using the formula

\[=IF(ISNUMBER(MAX(VLOOKUP((G3-(D2/2)),A:B,2,TRUE),VLOOKUP((G3-(D2/2)*4/5),A:B,2,TRUE),VLOOKUP((G3-(D2/2)*3/5),A:B,2,TRUE),VLOOKUP((G3-(D2/2)*2/5),A:B,2,TRUE),VLOOKUP((G3-(D2/2)*1/5),A:B,2,TRUE)),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*4/5),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*3/5),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*2/5),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*1/5),A:B,2,TRUE),VLOOKUP(G3,A:B,2,TRUE),VLOOKUP((G3+(D2/2)),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*4/5),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*3/5),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*2/5),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*1/5),A:B,2,TRUE),VLOOKUP(G3,A:B,2,TRUE),VLOOKUP((G3+(D2/2)),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*4/5),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*3/5),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*2/5),A:B,2,TRUE),VLOOKUP((G3+(D2/2)*1/5),A:B,2,TRUE)),0)\]

In this way, the intensities of all m/z values in the mass spectrum that could belong to the protein of interest are calculated. If the m/z values for species possessing more than a single adduct are desired, the additional adduct mass boxes in column D can be populated and the results are tabulated similarly to those in the example.

To check that the intensity values which were calculated are reasonable, the intensities of each charge state are plotted. An example showing data compiled from a
mass spectrum of cytochrome c (12 kDa) is shown in Figure A3. A mass of 12232 Da was used for the left panel (Figure A3a), which is the mass that MagTran reported after deconvolution. There is a clear peak centered on the +16 charge state and the charge states which are not typically observed for cytochrome c, those greater than +20, are at the baseline. At a glance, it can be confirmed that the calculated intensities are likely those belonging to the protein of interest. On the other hand, the panel on the right (Figure A3b) was generated using a mass of 12000 Da, which is 2% lower than the actual mass. There is no clear peak and the intensities appear random across the range of charge states. A result like this one would suggest that either the mass of the protein is incorrect (which is the case in this example) or that the protein is not present in the spectrum. In addition to enabling one to check the validity of the results of the automated peak detection, this page also reports the average charge state of the protein, including all peaks with intensities above the user-defined threshold, which can be decided based on the data in the plot.

**Figure A3.** The intensities of the charge states are plotted, providing a visual check of the validity of the results. (a) A strong peak is observed when the correct mass is used. (b) If the incorrect mass is used, or the protein is not present in the spectrum, the plot will appear noisy.
A3. Heat Maps

In order to monitor changes to a protein’s charge state distribution with changes in other experimental parameters, for instance instrumental source conditions or solution additive concentration, many protein spectra must be analyzed and the data presented in a useful way. One particularly useful way of presenting this kind of data is to create a heat map. A heat map is a three dimensional plot where two independent variables comprise the x and y axes while the dependent variable is presented as a color that is dependent on its position within the range of values found within the heat map. Figure A4 shows an example of the worksheet used to create heat maps from protein mass spectrometry data.

Figure A4. A heat map showing the change in the charge state distribution of cytochrome c as the instrument’s tube lens value is varied from 30 V to 250 V. On the right, the optimum tube lens value for each charge state is tabulated, while on the bottom the average charge states (ACS), highest intensity charge state (HICS), and sum of intensities are tabulated.

To generate the heat map, a mass spectrum must first be obtained for each value of the independent variable on the y axis (each of these mass spectra will automatically have the data for each charge state, the other independent variable, since these are all collected simultaneously). For each of these spectra, the intensities of each charge state are calculated using the charge states and intensities worksheet described above.
Intensities for each charge state are then pasted into their respective columns within the heat map. The cells are formatted so that the text is invisible and the color (fill) is determined by the magnitude of the value within the cell using the graded color scale found in the conditional formatting options.

From the rows in the heat map, the optimum value for each charge state can be tabulated (Figure A4, right). For example, the maximum tube lens value for the +7 charge state is calculated using the formula

\[
\text{INDEX($C6:Y6,1,MATCH(MAX(C7:Y7),C7:Y7,FALSE))}
\]

where the range C6:Y6 are the tube lens values and the range of C7:Y7 are the intensities of the +7 charge states at each tube lens value. In addition, the total intensity and average charge state can be calculated from the columns using Equations A.2 and A.3, respectively. For some columns, however, there are several charge states at near-baseline intensity which, although low intensity individually, may collectively skew the data. For this reason, a user-defined limit, representing the minimum intensity relative to the highest intensity in the heat map, determines which charge states are included in the calculations. All charge states with are included in calculations are formatted to display a black border. For example, the average charge state for the tube lens value of 150 V is calculated using the formula

\[
\frac{\text{IF}(C7>$AB$23*\text{MIN}(C7:C20),C7*$B$7,0)+\text{IF}(C8>$AB$23*\text{MIN}(C7:C20),C8*$B$8,0)+\text{IF}(C9>$AB$23*\text{MIN}(C7:C20),C9*$B$9,0)+\text{IF}(C10>$AB$23*\text{MIN}(C7:C20),C10*$B$10,0)+\text{IF}(C11>$AB$23*\text{MIN}(C7:C20),C11*$B$11,0)+\text{IF}(C12>$AB$23*\text{MIN}(C7:C20),C12*$B$12,0)+\text{IF}(C13>$AB$23*\text{MIN}(C7:C20),C13*$B$13,0)+\text{IF}(C14>$AB$23*\text{MIN}(C7:C20),C14*$B$14,0)+\text{IF}(C15>$AB$23*\text{MIN}(C7:C20),C15*$B$15,0)+\text{IF}(C16>$AB$23*\text{MIN}(C7:C20),C16*$B$16,0)+\text{IF}(C17>$AB$23*\text{MIN}(C7:C20),C17*$B$17,0)+\text{IF}(C18>$AB$23*\text{MIN}(C7:C20),C18*$B$18,0)+\text{IF}(C19>$AB$23*\text{MIN}(C7:C20),C19*$B$19,0)+\text{IF}(C20>$AB$23*\text{MIN}(C7:C20),C20*$B$20,0))/(\text{IF}(C7>$AB$23*\text{MIN}(C7:C20),C7,0)+\text{IF}(C8>$AB$23*\text{MIN}(C7:C20),C8,0)+\text{IF}(C9>$AB$23*\text{MIN}(C7:C20),C9,0)+\text{IF}(C10>$AB$23*\text{MIN}(C7:C20),C10,0)+\text{IF}(C11>$AB$23*\text{MIN}(C7:C20),C11,0)+\text{IF}(C12>
and the total intensity is calculated using the formula

$$=\text{IF}(C7>\text{ABS23*MIN(C7:C20),C7,0})+\text{IF}(C8>\text{ABS23*MIN(C7:C20),C8,0})+\text{IF}(C9>\text{ABS23*MIN(C7:C20),C9,0})+\text{IF}(C10>\text{ABS23*MIN(C7:C20),C10,0})+\text{IF}(C11>\text{ABS23*MIN(C7:C20),C11,0})+\text{IF}(C12>\text{ABS23*MIN(C7:C20),C12,0})+\text{IF}(C13>\text{ABS23*MIN(C7:C20),C13,0})+\text{IF}(C14>\text{ABS23*MIN(C7:C20),C14,0})+\text{IF}(C15>\text{ABS23*MIN(C7:C20),C15,0})+\text{IF}(C16>\text{ABS23*MIN(C7:C20),C16,0})+\text{IF}(C17>\text{ABS23*MIN(C7:C20),C17,0})+\text{IF}(C18>\text{ABS23*MIN(C7:C20),C18,0})+\text{IF}(C19>\text{ABS23*MIN(C7:C20),C19,0})+\text{IF}(C20>\text{ABS23*MIN(C7:C20),C20,0})$$

A4. References


APPENDIX B

CONSTRUCTION AND CALIBRATION OF DESI SPRAY SOURCES

B.1 Source Construction

The spray sources used for desorption electrospray ionization (DESI) are commonly constructed in-house from T unions and stainless steel or capillary tubing. Figure B1 shows a typical DESI source. Many different materials can be used, as well as materials with varying dimensions, leading to a variation in the performance characteristics between individual sprayers.

During my research, I commonly used a stainless steel (SS) 1/16” Swagelok T union as the base of the source. On one side of the T union (the right side in Figure B1), a sheath gas capillary was fixed in place. The sheath gas capillary was usually between 1 and 2 cm in length and constructed from either 1/16” SS tubing (I.D. 0.01 in, or 254 μm)
or fused silica capillary tubing (O.D. 430 μm, I.D. 320 μm). The sheath gas capillary length is determined by the constraints imposed by the geometry of the DESI stage, which includes an arm to hold the sprayer. The tip of the emitter usually needs to be a few mm from the sample surface and the arm can only float a few cm along each axis. If the sprayer needs additional reach, for instance if a small angle between the spray and the surface was being used, a longer sheath gas capillary needs to be used. To deliver the charged spray solvent, a solvent capillary was run through the length of the T union, exiting one side coaxial with the sheath gas capillary. The solvent capillary was made from fused silaca capillary tubing (O.D. 200 um, I.D. 75 um). The sheath gas (N₂) was fed from a cylinder with the gas flow controlled by a two-stage gas regulator. This was fed through the top of the T union using 1/16” flexible plastic tubing (I.D. 1/32”). SS and plastic tubing pieces were held in place using SS ferrules while graphite ferrules were used to hold fused silica pieces in place.

### B.2 Source Calibration

The flow rate of the sheath gas exiting a DESI sprayer depends on several characteristics, including the gas pressure at the regulator, the sheath gas capillary I.D., and the solvent capillary O.D. Small variations in any of these parameters can cause a change in the gas flow rate. For example, Table B1 shows the flow rate calibration data for two DESI sources constructed using identical materials with sheath gas capillary lengths of 1.5 cm. The data is plotted in Figure B2. Calibration was performed using a #11 Gilmont flowmeter. The DESI source was connected to the flowmeter using a section of flexible plastic tubing, where one end was pulled over the nut of the source holding the sheath gas capillary and the other end pulled over the bottom of the flow meter. A good
seal on both ends was confirmed using a detergent solution. The data confirms that the change in flow rate with change in pressure is linear for both sprayers with similar rates of 7.9 mL/min/psi. However, sprayer 2 has a flow rate around 80 mL/min higher than sprayer 1 at equivalent pressures, probably due a small difference in source dimensions.

**Table B1.** Calibration data for two DESI sources.

<table>
<thead>
<tr>
<th>Pressure (psi)</th>
<th>Flow Rate (mL/min) Source 1</th>
<th>Flow Rate (mL/min) Source 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>370</td>
<td>460</td>
</tr>
<tr>
<td>60</td>
<td>450</td>
<td>500</td>
</tr>
<tr>
<td>70</td>
<td>510</td>
<td>560</td>
</tr>
<tr>
<td>80</td>
<td>560</td>
<td>650</td>
</tr>
<tr>
<td>90</td>
<td>640</td>
<td>730</td>
</tr>
<tr>
<td>100</td>
<td>720</td>
<td>820</td>
</tr>
<tr>
<td>110</td>
<td>820</td>
<td>920</td>
</tr>
<tr>
<td>120</td>
<td>910</td>
<td>1010</td>
</tr>
<tr>
<td>130</td>
<td>1000</td>
<td>1090</td>
</tr>
<tr>
<td>140</td>
<td>1080</td>
<td>1170</td>
</tr>
<tr>
<td>150</td>
<td>1150</td>
<td>1250</td>
</tr>
<tr>
<td>160</td>
<td>1220</td>
<td>1310</td>
</tr>
<tr>
<td>170</td>
<td>1310</td>
<td>1390</td>
</tr>
<tr>
<td>180</td>
<td>1390</td>
<td>1460</td>
</tr>
<tr>
<td>190</td>
<td>1470</td>
<td>1540</td>
</tr>
<tr>
<td>200</td>
<td>1530</td>
<td>1610</td>
</tr>
<tr>
<td>210</td>
<td>1610</td>
<td>1700</td>
</tr>
<tr>
<td>220</td>
<td>1690</td>
<td>1770</td>
</tr>
<tr>
<td>230</td>
<td>1770</td>
<td>1860</td>
</tr>
<tr>
<td>240</td>
<td>1830</td>
<td>1930</td>
</tr>
</tbody>
</table>
Figure B2. Calibration data for two DESI sources constructed with similar materials.