

**Studying the Structure and Stability of Protein:Ligand Complexes  
using Ion Mobility-Mass Spectrometry**

**by**

**Jessica Rabuck**

**A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Chemistry)  
in the University of Michigan  
2016**

**Doctoral Committee:**

**Associate Professor Brandon T. Ruotolo, Chair  
Professor Robert T. Kennedy  
Professor Anna K. Mapp  
Assistant Professor Brent R. Martin  
Assistant Professor Matthew B. Soellner**

© Jessica Nicole Rabuck

All Rights Reserved  
2016

## **Dedication**

This thesis is dedicated to my husband Jon. Thank you for all of your love and support.

## Acknowledgements

First and foremost, I would like to thank my advisor, Dr. Brandon T. Ruotolo. You have allowed me to follow where my projects lead, even if the projects led me to areas outside of our laboratory. You are an excellent role model as a scientist and a mentor, and I strive to model my own mentoring and work ethic on yours.

I would also like to thank my committee members Dr. Robert Kennedy, Dr. Anna Mapp, Dr. Brent Martin, and Dr. Matthew Soellner. You have all been very supportive and insightful during meetings and discussion, and I appreciate how you have challenged me to increase my understanding of my research and how it plays a role in the larger communities.

During my dissertation research, I have had the privilege of working with a wonderful group of fellow scientists. Dr. Suk-Joon Hyung, thank you for having the patience to teach me about ion mobility and mass spectrometry, despite the fact that I came in only knowing how to draw a quadrupole. Dr. Billy Samulak, you were always available for helpful chats and advice—thank you for helping me walk through the postdoc process! Drs. Yueyang Zhong, Linjie Han, and Russel Bornschein, you have guided my knowledge of analytical chemistry and have helped me place my biochemistry knowledge in the appropriate context. Dr. Shui Niu, you were always available to talk about research and ligand binding. Dr. Molly Hopper, grad school would not have been the same without our networking opportunities and all of your advice and support through this entire process. Joseph Eschweiler, without you CIUSuite would not be the same. Thank you for all of your help with coding. My desk won't be the same without you sitting beside me! Daniel Polasky, this thesis also owes quite a bit to your coding expertise. Looking back at the data I analyzed at the beginning of my graduate work, I realize why the going was so slow. Yuwei Tian, I enjoyed our girl time together, and thank you for all of your helpful comments and insights. Sugyan Dixit, thank you for your enthusiasm and knowledge in the lab and for always having coffee prepared. Stephanie Godden, thank you for reading over my drafts and for lighting up the lab. James Keating,

thank you for being a wonderful mentee. Rachel Martini, you are an amazing young scientist, and I enjoyed rooming with you at ASMS. Dr. Richard Kerr, thank you for all of your helpful advice and insight about proteins and laboratory matters. Sarah Fantin and Chunyi Zhao, thank you for your enthusiasm and fresh eyes in the lab.

In addition to my coworkers, I was fortunate to work with many fantastic collaborators. Dr. Matthew Soellner, Dr. Meghan Breen, Dr. Kristin Ko, Michael Agius, and Christel Fox, thank you for all of your insights and the helpful discussions on the ins-and-outs of the ever-changing kinase project. My dissertation would not be what it is without any of you. Dr. Brian Coppola and the members of the CSIE|UM organizational committee, thank you for teaching me what it means to be a teacher. Daniel Polasky, Amy Bondy, and Nancy Wu, thank you for your support during my education project. Dr. Anna Mapp and Jean Lodge, you have been instrumental preparing KIX and peptide samples, and you have pushed me to expand native mass spectrometry into the context of traditional biophysical techniques. Additionally, thank you, Jean, for being an amazing friend outside of the lab. Our coffee dates were the highlight of my week. Ashley McQuarters and Alexander LaLonde, you two have made living in Ann Arbor so much fun! I could not have asked for better friends for experiencing the struggles and achievements of graduate school. Additionally, Abbigail Kuzmen, thank you for reading over my grant applications and deciding if you would fund me or not, even if you didn't always understand the science behind it, and for always being my cheerleader through my graduate school career.

I have an amazing family, and I'm grateful for both the family I was born into and the one I married into. Thank you all for your love and support and instilling in me the value of a great education, and for keeping me going when it seemed like graduate school would never end. Last, but definitely not least, I would like to thank my husband, Jon. Thank you for sharing this experience with me. Your love and support means more to me than you could imagine, and I'm fortunate to have you in my life.

## Table of Contents

Dedication .....	ii
Acknowledgements .....	iii
List of Figures .....	x
List of Tables .....	xiii
List of Appendices .....	xiv
Abstract .....	xv
Chapter 1. Introduction .....	1
1.1 The Current State of Drug Discovery .....	1
1.2 Targeting Kinases in Disease .....	2
1.2.1 BCR-Abl in Chronic Myeloid Leukemia .....	6
1.2.2 Normal Abl function .....	7
1.2.3 Abl in Brain Disorders .....	8
1.2.4 Src in Cancer .....	8
1.3 Current High Throughput Screening Methods for Kinase Inhibitor Discovery .....	9
1.4 Targeting Protein:Protein Interactions .....	15
1.5 Ion Mobility-Mass Spectrometry .....	17
1.5.1 Ion Generation .....	18
1.5.2 Maintaining and Detecting Native-Like Complexes in the Gas-Phase .....	20
1.5.3 Principles of Ion Mobility separation .....	22
1.6 Mass Spectrometry in Drug Discovery .....	23
1.6.1 Paradigms for Protein-Ligand Screening by IM-MS .....	23
1.6.2 Searching for Shifts in Protein-Ligand Stability .....	26
1.6.3 Combing for Conformational Transition .....	28

1.7 Summary .....	31
1.8 References .....	33
Chapter 2. CIUSuite: A Quantitative Analysis Package for Collision Induced	
Unfolding Measurements of Gas-Phase Protein Ions .....	48
2.1 Introduction to Collision Induced Unfolding.....	48
2.2 CIUSuite Overview.....	50
2.2.1 CIUSuite_plot .....	51
2.2.2 CIUSuite_stats .....	51
2.2.3 CIUSuite_compare.....	51
2.2.4 CIUSuite_detect.....	52
2.2.5 CIUSuite_analysis.....	52
2.2.5 CIUSuite_score .....	53
2.3 CIUSuite Applications .....	54
2.3.1 Rapid Analysis of Kinase Inhibitor Screens .....	54
2.4 Conclusions.....	57
2.5 References.....	57
Chapter 3. Activation State-Selective Kinase Inhibitor Assay Based on Ion	
Mobility-Mass Spectrometry .....	61
3.1 Introduction.....	61
3.2 Experimental Methods .....	64
3.2.1 General .....	64
3.2.2 Protein Purification, and Activity Assays .....	64
3.2.3 Ion Mobility-Mass Spectrometry .....	64
3.3 Results.....	66
3.3.1 Outlining the Challenges in Activation State-Selective Kinase Inhibitor Discovery .....	66
3.3.2 CIU Kinase Inhibitor Binding Mode Assay.....	66
3.3.3 Identifying the Expected CIU Response for Kinase Inhibitors in Different Binding Modes .....	68
3.3.4 Complete CIU Training Data Set from Known Kinase Inhibitors .....	69
3.3.5 Simple Scoring Approach Allows the Differentiation of Type I and	

Type II Inhibitors Based on CIU data.....	70
3.4 Discussion.....	73
3.5 Conclusions.....	75
3.6 References.....	76
Chapter 4. Collision Induced Unfolding and Dissociation Differentiates ATP-Competitive from Allosteric Protein Tyrosine Kinase Inhibitors.....	82
4.1 Introduction.....	82
4.2 Experimental Methods.....	84
4.2.1 General.....	84
4.2.2 Protein Expression and Purification.....	85
4.2.3 Ion Mobility-Mass Spectrometry.....	85
4.3 Comparison ATP-Competitive and Allosteric Abl Kinase Inhibitors.....	86
4.4 CIU vs. CID for Differentiating Kinase Inhibitors.....	87
4.5 Conclusions.....	90
4.6. References.....	92
Chapter 5. Ion Mobility-Mass Spectrometry-Based Assay Distinguishes between DFG-out and DFG-in Src Kinase Inhibitors.....	95
5.1 Introduction.....	95
5.2. Experimental.....	98
5.2.1 General.....	98
5.2.2 Protein Expression and Purification.....	99
5.2.3 Ion Mobility-Mass Spectrometry.....	99
5.3 Results and Discussion.....	99
5.3.1 IM-MS of Src Kinase Domain vs Three Domain Constructs.....	99
5.3.2 Identifying a CIU Fingerprint Region that Distinguishes Between Type I and Type II Src Inhibitors.....	101
5.3.3 CIU Accurately Classifies Weakly Bound and Unknown Inhibitors ...	102
5.4 Conclusions.....	105
5.5 References.....	106
Chapter 6. Collision Induced Unfolding Reveals Unique Fingerprints for Remote Protein Interaction Sites in the KIX Regulation Domain.....	113



6.1 Introduction.....	113
6.2 Materials and Methods.....	116
6.2.1 Protein Expression and Purification.....	116
6.2.2. Peptide Synthesis .....	117
6.2.3. Ion Mobility-Mass Spectrometry .....	117
6.3 Results.....	118
6.3.1 Each Peptide Binding Type has a Unique CIU fingerprints .....	118
6.3.2. Determining Differences in CIU Stability for Ligand Screening .....	121
6.3.3. Using CIU to Localize Ligand Binding .....	122
6.4 Conclusions.....	123
6.5 References.....	124
Chapter 7. Native Mass Spectrometry Exercises for the Undergraduate Laboratory .....	129
7.1 Introduction.....	129
7.1.1 Curriculum Overview .....	130
7.2 Mass Spectrometry of Multiprotein Complexes for the Undergraduate Laboratory.....	131
7.3 Protein-Ligand Complex Analysis Using Mass Spectrometry for the Undergraduate Laboratory .....	133
7.4. Potential Modifications.....	134
7.5 Discussion.....	134
7.6 References.....	135
Chapter 8. Conclusions and Future Directions .....	138
8.1 Conclusions.....	138
8.2 Immediate Future Directions .....	140
8.2.1 Screening type I, Type II, and Allosteric Inhibitors for the Abl Kinase .....	140
8.2.2 Exploring Src3D Kinase Inhibitors that Induce a Closed Conformation .....	140
8.2.3 Exploring KIX Allostery.....	141
8.2.4 Interactions between Abl and Dscam .....	141
8.4 Future Directions—Looking Ahead .....	141

8.3 References .....	142
Appendices .....	144

## List of Figures

Figure 1.1 Gene family distributions for current drugs per substance.....	1
Figure 1.2 Diagram of the kinase domain of the insulin receptor protein tyrosine kinase ....	2
Figure 1.3 Abl with type I, type II, and type IV inhibitors .....	3
Figure 1.4 Three domain structures of Src and Abl.....	4
Figure 1.5 Domain structure of Abl, BCR-Abl, and Src .....	6
Figure 1.6 Model for the signaling functions of Abl .....	7
Figure 1.7 Src localization to the membrane .....	8
Figure 1.8 The drug discovery process .....	10
Figure 1.9 Schematic of one type of PCR-based output.....	11
Figure 1.10 Kinase assay using a Dap-pyrene substrate.....	13
Figure 1.11 Activation-state dependent assay using acrylodan .....	13
Figure 1.12 Schematic of ITC, DSC, and DSF.....	14
Figure 1.13 Representation of KIX.....	16
Figure 1.14 Schematic of the Waters Synapt G2 HDMS .....	17
Figure 1.15 Diagram of ESI.....	20
Figure 1.16 Basic screening strategies for intact protein:ligand complexes by IM-MS.....	24
Figure 1.17 IM-MS stability measurements differentiate ligand binding in a tetrameric complex .....	26
Figure 1.18 CIU fingerprinting differentiates Abl kinase inhibitor binding modes .....	29
Figure 2.1 Schematic representation of CIUSuite modules.....	51
Figure 2.2 Activation state-selective kinase inhibitors are classified by CIU fingerprints and CIUSuite .....	55
Figure 3.1 A kinase activity assay illustrates the challenges associated with differentiating between type I and type II inhibitors.....	67
Figure 3.2 Schematic diagram describing the basic steps involved in the CIU assay .....	69

Figure 3.3 CIU fingerprints for indicated 11+ Abl:inhibitor complex ions.....	70
Figure 4.1 Experimental workflow for differentiating between Abl allosteric and ATP-competitive kinase inhibitors .....	83
Figure 4.2 Comparison of the average ATP-competitive and allosteric CIU fingerprints ....	86
Figure 4.3 Scaled deviation score analysis of ATP-competitive and allosteric inhibitors ....	87
Figure 4.4 Comparison of all inhibitors using CIUSuite .....	88
Figure 4.5 Simulated results from a high throughput screen using the averages and standard deviations for ATP-competitive and allosteric inhibitors from Table 4.1 .....	91
Figure 4.6 Simulated results from a high throughput screen using the averages and standard deviations for ATP-competitive and allosteric inhibitors from Table 4.2 .....	91
Figure 5.1 Results of CIU fingerprints for nilotinib-bound and dasatinib-bound Abl and Src .....	96
Figure 5.2 Results from CIUSuite_analysis.....	101
Figure 5.3 An illustration of a theoretical HTS using the scores from Table 1.....	104
Figure 6.1 Structure of KIX with peptides and corresponding mass spectra.....	114
Figure 6.2 CIU fingerprints for KIX:peptide complexes.....	119
Figure 6.3 RMSD values computed from CIUSuite.....	120
Figure 6.4 Histograms of the RMSD values and Stabilities of the CIU features observed...	121
Figure 6.5 SDS vs collision voltage for inhibitors.....	122
Figure 6.6 Z-scores for each KIX:peptide complex.....	123
Figure 7.1 MS analysis of ConA.....	132
Figure 7.2 Overview of the lysozyme experiment.....	133
Figure I.1 Individual mass-resolved components of the CIU fingerprints for Abl:inhibitor complex .....	144
Figure I.2 X-ray structures for Abl bound dasatinib and imatinib with the corresponding CIU fingerprints and drift time plots .....	145
Figure I.3 Averaged IM Spectra generated from all type I and type II datasets for selected regions of the CIU fingerprint .....	146
Figure I.4 Ion mobility spectra for apo Abl, Abl:datasatinib, and Abl:imatinib .....	149
Figure I.5 An illustration using the scores from Table 1 .....	150
Figure I.6 ATP Km Curve of Non-phosphorylated Abl .....	151

Figure I.7 APT Km Curve of phosphorylated Abl .....	151
Figure I.8 Non-phosphorylated Abl data .....	152
Figure I.9 Phosphorylated Abl data .....	153
Figure I.10 Mass spectra of Abl and P-Abl .....	154

## List of Tables

Table 3.1 CIU-based similarity scores for the current training data set of type I and II kinase inhibitors .....	72
Table 4.1 CIU-based similarity scores for the current training data set of ATP-competitive and allosteric inhibitors .....	89
Table 4.1 CID-based similarity scores for the current training data set of ATP-competitive and allosteric inhibitors .....	90
Table 5.1 Scores resulting from the CIUSuite analysis from Figure 5.2 .....	103
Table I.1 Table of CIU scores against type I and type II kinase inhibitors at three different collision voltages .....	147
Table I.2 Table of the ratios of type I and type II kinase inhibitor scoring at three different collision voltages .....	148

## **List of Appendices**

Appendix I. Chapter 3 Supporting Information.....	144
Appendix II. Chapter 7 Supporting Information.....	156

## Abstract

The majority of drug discovery efforts have primarily been focused on the discovery of potent small molecule inhibitors using screening methodologies are based on enzymatic turnover rates to determine compound potency. However, not all targets align with the classical enzyme/inhibitor paradigm. Many of these difficult targets have unique conformational changes that can be targeted for therapeutic design, but such information about the conformation consequences of ligand binding is not always provided by classical screening technologies. Described here is the application of ion mobility-mass spectrometry combined with collision induced unfolding (CIU) to develop screening protocols for peptides and small molecules that induce a conformation change when bound to a target protein.

First, a series of Python modules were designed to maximize the data information content extracted from CIU data sets. This analysis was used to develop a screen for type I and type II kinase Abl kinase inhibitors, which stabilize the active and inactive kinase conformations, respectively. Analysis suggests that the information content of a CIU screen can be maximized by finding a region in the CIU datasets where the differences between type I and type II inhibitors are maximized. Following these studies, this workflow was applied to differentiate between ATP-competitive and allosteric kinase inhibitors. It was found that charge stripping, where the ligand dissociates from the protein with a 1+ charge, plays a central role in distinguishing between these two inhibitor types. CIU was also used to explore type I and type II inhibitors for the Src protein tyrosine kinase. While differences between the two types of inhibitors were discovered using CIU, data suggest that the larger open/closed kinase conformation also plays a role in the CIU data. Additionally, the workflow to study kinase inhibitors was applied to KIX:peptide complexes. For the first time, unique CIU features were identified for peptides that bind to remote binding sites on a protein target. All of this work falls under the umbrella of native mass spectrometry, which was introduced to a biomedical analysis undergraduate laboratory using two well-studied protein systems.



# Chapter 1

## Introduction

### 1.1 The Current State of Drug Discovery

Out of 10,000 protein folds<sup>1</sup> and 16,000 protein families<sup>2</sup>, only 120 families have been successfully targeted by drug discovery efforts. Of these families, the top ten compose 69% of

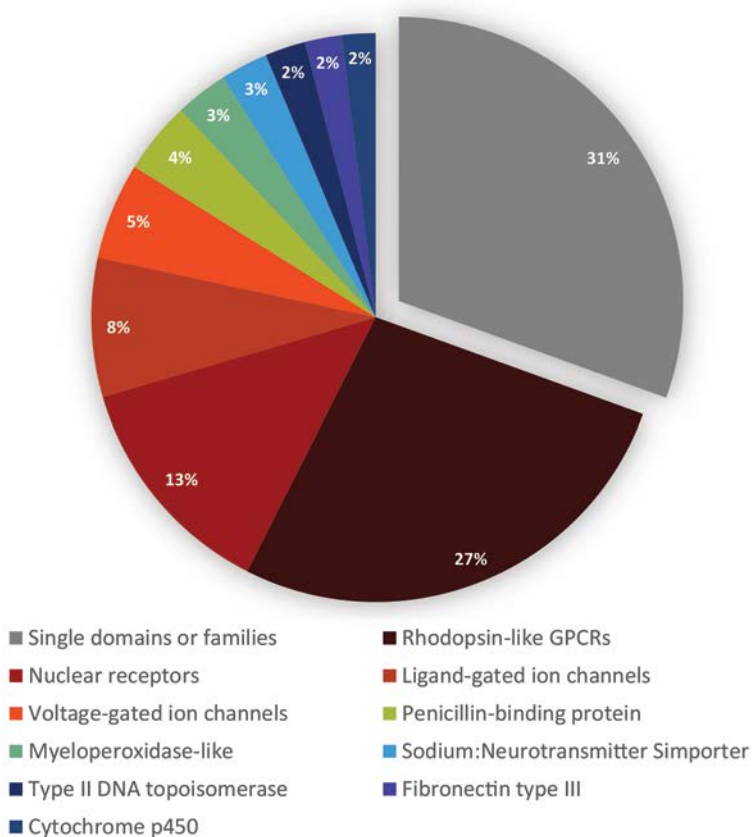


Figure 1. 1.1 Gene family distribution for current drugs per drug substance. The gray area represents 120 domain families or single domains for which only a few drugs have been successfully launched as of 2005. Included in the gray area are kinases, while the colored regions are a family share of a percentage of all FDA-approved drugs for the top ten families<sup>1</sup>.

the percentage of all FDA-approved drugs in 2005 (Figure 1.1)<sup>1</sup>. The narrow focus of current target selection stems in part from the challenges associated with designing potentially efficacious small molecules, and rapidly assessing their potency, for targets that do not align with the classical enzyme/inhibitor paradigm<sup>3</sup>. Many of these targets were previously thought to be ‘undruggable’, as most high throughput screening (HTS) technologies are based on enzymatic turnover rates for determining compound potency. A second class of previously ‘undruggable’ targets contains members of highly similar enzymes (such as kinases), which

have critical amino acids that are highly conserved, making them difficult to selectively target. Clearly, there is a need for new biophysical screening approaches capable of evaluating new targets that do not follow the traditional enzyme/inhibitor standard.

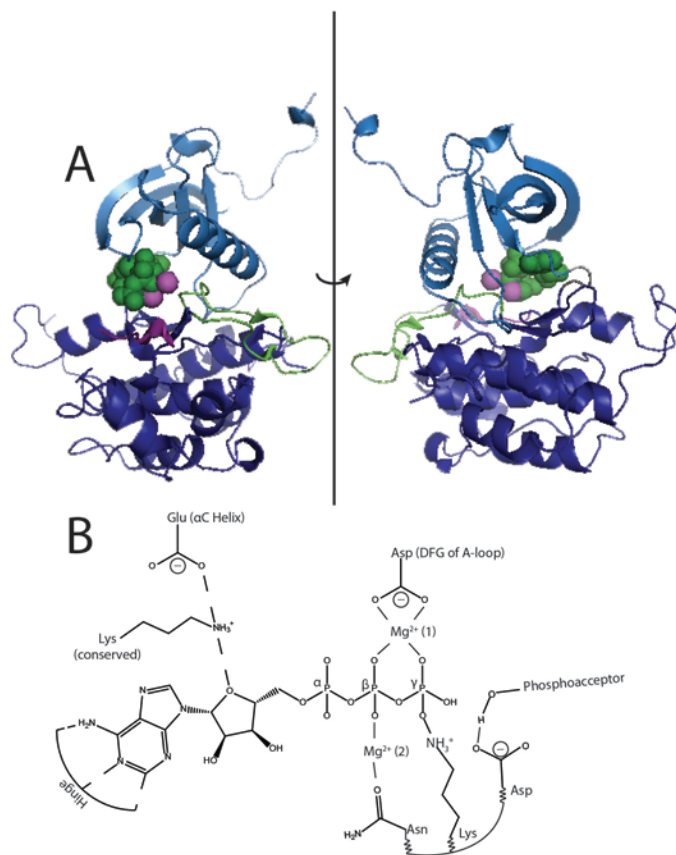


Figure 1.2 (A) The kinase domain of the insulin receptor PTK (PDB ID 3IRK). The right-side is rotated 180° relative to the left. The N-lobe is colored in blue, the C-lobe is in purple, and the activation loop is in the green ribbon representation. AMP-PMP is represented by the green spheres, and the two coordinating  $Mg^{2+}$  are represented by the pink balls. (B) Representative diagram of ATP with the amino acids that interact to complete phosphate transfer.

An example of this paradigm is the chimeric oncoprotein BCR-Abl, which causes chronic myeloid leukemia (CML). BCR-Abl is generated by a chromosomal translocation that fuses the breakpoint cluster region protein (BCR) with the Abelson protein tyrosine kinase (Abl). For many years, kinases were thought to be ‘undruggable’ targets due to the structural homology observed among members of the kinase family, their high affinity for ATP, and their broad roles in cell signaling<sup>4</sup>. Discovery of the kinase inhibitor imatinib<sup>5</sup> demonstrated that a conformationally selective inhibitor, preferentially targeting the unique inactive conformations of the Abl kinase<sup>6</sup>, could be a successful strategy for the design of cancer therapeutics. While many kinase inhibitors based on imatinib have been discovered (type II inhibitors), the applicability of the approaches used for inhibitor discovery are limited to scaffolds

that resemble the original small molecule, and the robustness of such strategies for the other members of the kinase family are largely unknown.

## 1.2 Targeting Kinases in Disease

Cancer is a disease resulting from the loss of normal cellular controls. Many types of cancer are associated with mutations or overexpression of kinases<sup>7</sup>, which function in many important

cellular processes such as T-cell and B-cell activation, responses to stimuli, mitogenesis, differentiation and development, angiogenesis, platelet activation, cell shape and attachment, neurotransmitter signaling, cell cycle control, growth control, oncogenesis, cell survival and apoptosis, transcriptional regulation, and glucose uptake<sup>8</sup>. The basic function of a kinase is to transfer the  $\gamma$ -phosphate of adenosine triphosphate (ATP) to the protein alcohol group of serine or threonine and/or the protein phenolic groups on tyrosines. While 96.8% of all phosphorylated residues are Ser or Thr, only 1.8% of phosphorylated residues are Tyr<sup>9</sup>.

The kinase domain (KD) is composed of two subdomains: an N-lobe that is mainly comprised of beta-sheets with one alpha helix (typically called the  $\alpha$ C-helix, Figure 1. 2A, blue), and the C-lobe, which is primarily alpha-helical (Figure 1. 2A, purple). ATP binds to a cleft formed between the two subdomains (Figure 1.2A, green and 2B). The phosphates of the ATP are aligned by interactions

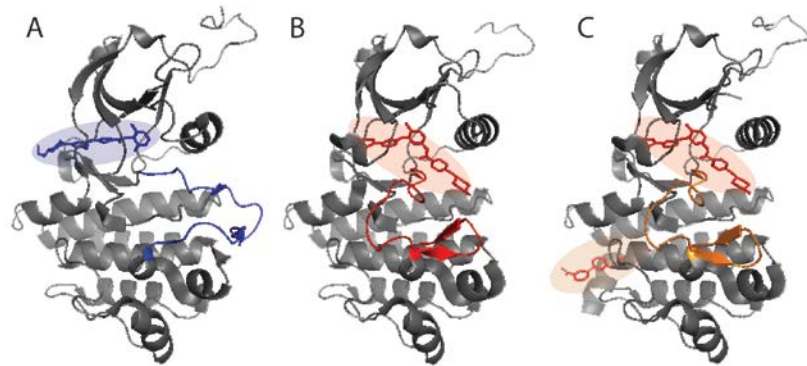


Figure 1.3. Abl with type I, type II, and type IV inhibitors. (A) Abl in complex with dasatinib (blue sticks, in blue circle). This causes the activation loop to extend outward (blue cartoon) (PDB ID 2GQG). (B) Abl in complex with imatinib (red sticks, red circle). Type II inhibitors like imatinib cause the activation loop (red cartoon) to tuck into the kinase domain, preventing substrate binding and partially occluding the ATP binding site (PDBID 1IEP). (C) Abl in complex with imatinib (red sticks, red circle) and GNF-2 (orange sticks, orange circle). Like the imatinib complex, the activation loop (orange) is in the inactive position (PDB ID 3K5V).

with the glycine-rich loop and with  $Mg^{2+}$  ions (Figure 1.2A, pink, Figure 1. 2B). Within the ATP-binding pocket, there is a hydrophobic pocket that is controlled by a “gatekeeper” residue which can either block or allow access to the back pocket of the binding site for the proper positioning of ATP. A flexible activation loop contains a conserved DFG sequence which is critical for catalysis<sup>10</sup>. The DFG-motif is flipped depending on the activation state of the kinase. In the DFG-in conformation, the protein is in an active form, serving as a dock for potential substrates (Figure 1.2A, green cartoon, Figure 1.2B, Figure 1.3A), whereas in the DFG-out conformation, the protein is in an inactive form (Figure 1.3B,C). It has been thought that the active PTKs are all very structurally similar; however, the inactive structural forms of the PTKs are conformational diverse<sup>4</sup>.

Kinase activation and inactivation rely on a series of phosphorylation and dephosphorylation events that act as molecular switches to turn target proteins “on” or “off.” As kinases often act in cellular signaling cascades, these molecular switches are a part of kinase regulation, as well. For example, the Rous sarcoma virus protein tyrosine kinase (Src) requires a dephosphorylation event on Tyr527 on the C-terminal tail of the kinase domain as well as a phosphorylation event on Tyr416 of the activation loop in order to activate the kinases<sup>11</sup>. Another form of internal regulation relies upon the Src-homology 2 (SH2) and Src-homology 3 (SH3) domains docking to

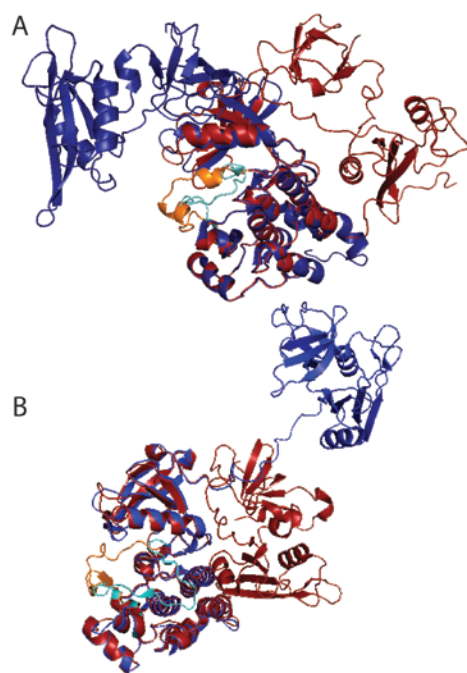


Figure 1.4. Three domain structure of Src and Abl. (A) Open form of Src (blue) vs closed form of Src (red). The activation loops are colored in cyan (active) and orange (inactive) from PDB IDs 1y57 and 2SRC, respectively. (B) Open form of Abl (blue) adapted from PDBID 1OPK. The activation loop is in cyan. In red is the closed form of the kinase from 2FO0 with the activation loop from 1IEP in orange.

the KD, which inactivates the kinase (Figure 1.4A,B red structures)<sup>12, 13</sup>. The structure of the docking interactions varies between PTK families.

The first PTK inhibitors were developed in the early 1980s<sup>14</sup>. However, the landmark FDA approval of Gleevec (imatinib) in May 2001 inspired the search for structurally-specific protein tyrosine kinases<sup>15</sup>. The approval of Gleevec ushered in a new era for cancer treatment. Previously, chemotherapy had consisted of interferons, which activate immune cells<sup>16</sup>. However, interferon treatment came with a myriad of flu-like symptoms, and often these side effects were severe. Since imatinib is selective for Abl, the number of side effects is greatly decreased comparatively to interferon treatment<sup>17</sup>.

Gleevec was found to bind in the ATP-binding site of the catalytic kinase domain. Further structural probes into kinase drug discovery revealed three types of ATP-competitive inhibitors: type II, type I, and  $\alpha$ C helix-out.

Type II inhibitors, like Gleevec, stabilize the inactive DFG-out conformation, in which the conserved Asp-Phe-Gly (DFG) at the start of the activation loop is flipped outward, displacing the catalytic aspartic acid residue and phenylalanine side chain<sup>18</sup> (Figure 1.3B). These movements partially occlude the ATP-binding site. Type I inhibitors contrastingly stabilize the active DFG-in kinase conformation in which the DFG triad

flips in toward the center of kinase (Figure 1.3A). These changes in the kinase domain have consequences for the global kinase structure. Notably, in recent studies using a three domain construct of Src (SH3-SH2-kinase domain, Src3D, Figure 1.1I), type II inhibitors were found to stabilize an elongated, string-like Src3D structure<sup>19, 20</sup> (Figure 1.4A, blue structure). Type I inhibitors were found to stabilize structures that ranged from fully-extended, string like conformations to a more neutral position between the open and closed 3D structure<sup>19</sup>. Conversely,  $\alpha$ C helix-out inhibitors have been observed to stabilize a conformation similar to the autoinhibited<sup>20</sup>, closed Src3D structure, where the SH3 and SH2 domains pack against the side of the kinase domain<sup>19, 21</sup>. This conformation is stabilized by the movement of the  $\alpha$ C helix in the N-lobe of the KD, which swings out to disrupt a salt bridge between a catalytic lysine (Lys 295 in Src) and a conserved glutamic acid (Glu310) on the  $\alpha$ C helix. Similarly, I and type II inhibitors open the three domain form of Abl (Abl3D) to the same extent<sup>22</sup>. It is hypothesized that binding of type II inhibitors to Abl3D can capture the kinase in a transient, pre-activation form, where the regulatory domains are dis-engaged, but the kinase domain remains in the inactive conformation. While type II inhibitors were thought to target a unique inactive conformation, and therefore had a greater inherent specificity than type I inhibitors, the discovery of promiscuous type II inhibitors and specific type I inhibitors has proved this thought to be false<sup>23</sup>.

Further frustrating drug discovery efforts, mutations in the kinase domain often abrogate drug binding either by introducing bulky, hydrophobic amino acids into the drug binding pocket or by causing a conformational change in the kinase that is unfavorable for inhibitor binding. Additionally, most kinase inhibitors are somewhat promiscuous, even if they are developed to target a unique inactive kinase conformation. Recent kinase drug discovery efforts have therefore started to focus on small molecule inhibitors that target unique allosteric sites remote from the ATP binding site. However, the discovery of most allosteric inhibitors, also known as type IV inhibitors, have occurred serendipitously<sup>23</sup>. For example, two allosteric inhibitors for Abl were discovered after showing activity in an *in vivo* screen against BCR-Abl but failed in a follow up *in vitro* assay against the kinase domain of Abl. Crystal structures revealed that these two inhibitors, GNF-2 and GNF-5, bind to the allosteric myristoyl binding site of the Abl kinase domain (Figure 1.3C)<sup>24, 25</sup>. Further studies indicated that allosteric inhibitors can rescue the inactive conformation of the kinase<sup>25, 26</sup>, and when used in combination with type II inhibitors,

have been shown to suppress resistance of mutations *in vitro*<sup>23, 27</sup>. NMR studies have revealed that GNF-2 and GNF-5 are  $\alpha$ C helix inhibitors, which also have the ability to rescue the closed kinase formation in Abl. As these inhibitors target nonconserved binding sites in the kinase, they are highly selective and remain unaffected by mutations local to the primary ATP-binding site.

### 1.2.1 BCR-Abl in Chronic Myeloid Leukemia

Abl was the first PTK to specifically be targeted by a small molecule inhibitor. In chronic myeloid leukemia (CML), a cancer of the bone marrow that affects 15-20% of the western population<sup>28</sup>, a chromosomal translocation t(9;22)(q34;q11) fuses the C-terminal region of the breakpoint cluster region (BCR) to the N-terminus of Abl (Figure 1.5A), causing a constitutively active kinase<sup>28</sup> (Figure 1.5B). This fusion leads to breakdowns in the PTK regulation of Abl. It has been hypothesized that since BCR contains a tetramerization domain on its N-terminus<sup>29</sup>, BCR-Abl oligomerizes into a tetramer *in vivo*. It is thought that the close proximity of the kinase domains in the tetramer auto-phosphorylate each other, despite the lack of auto phosphorylation in *in vitro* Abl kinase assays. Furthermore, the N-terminus of Abl is truncated during the translocation, which causes the loss of the myristoylation site. Normally, the myristate moiety binds to a pocket in the C-lobe of the kinase domain, which stabilizes the protein in an inactive conformation (Figure 1.3C). On a cellular level, BCR-Abl lacks the ability to bypass the nuclear pore and therefore cannot pass into the cell nucleus. As BCR-Abl is trapped in the cytoplasm<sup>30</sup>, it binds to and phosphorylates several proteins in the Ras signaling pathway, including the GTPase-activating protein, Rac, and the growth receptor-bound protein 2.

There are currently five FDA-approved kinase inhibitors target the Abl kinase domain in BCR-Abl: imatinib (Gleevec, Novartis), dasatinib (Sprycel, Bristol-Myers Squibb), nilotinib

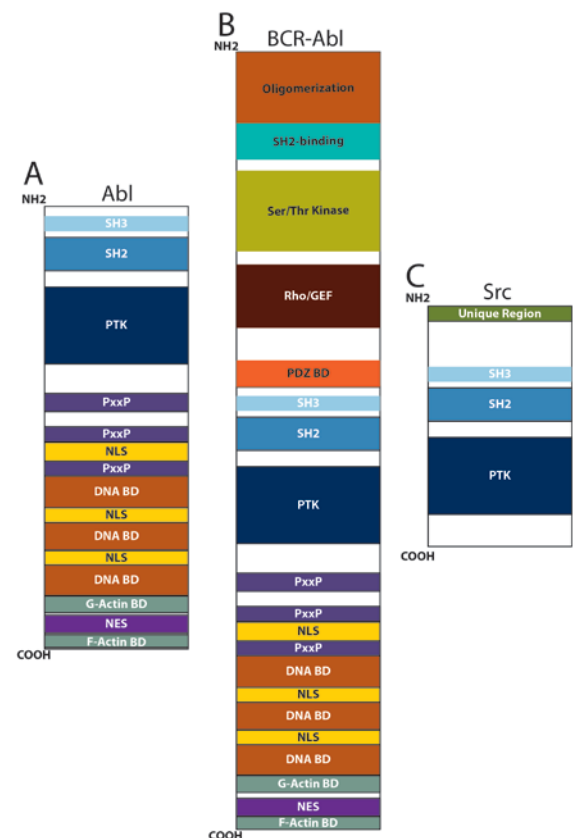


Figure 1.5. Domain structures. (A) Domain structure of Abl. (B) Domain structure of BCR-Abl. (C) Domain structure of Src.



(Tasigna, Novartis), ponatinib (Iclusig, Ariad), and bosutinib (Bosulif, Pfizer). All of these inhibitors are ATP-competitive, and fall either into the type I or type II binding categories. So far, no allosteric inhibitors have been approved to treat CML, although *in vivo* studies have shown cooperative effects between GNF-2 and several kinase inhibitors<sup>31</sup>.

### 1.2.2 Normal Abl Function

Abl serves as an important hub in cell regulation (Figure 1.6) by regulating growth factors<sup>32, 33</sup>, chemokines<sup>34, 35</sup>, DNA damage<sup>36</sup>, oxidative stress<sup>37-39</sup>, adhesion receptors<sup>40, 41</sup>, and microbial pathogens<sup>42-44</sup> in order to modulate downstream functions such as DNA damage-induced apoptosis<sup>45-47</sup> (Figure 1.6, green proteins) and DNA repair<sup>48, 49</sup> (Figure 1.6, blue proteins), actin polymerization and processing<sup>50-52</sup>, and normal cell proliferation<sup>53</sup>. Abl also contains a DNA-binding domain, several nuclear localization signals (NLS) and nuclear export signals (NES) (Figure 1.5A)<sup>54</sup>. While the functions of nuclear Abl have been explored<sup>55, 56</sup>, the role of cytoplasmic Abl has yet to be completely defined<sup>57</sup>. Further confusing c-Abl function is the N-terminal variants, termed 1a and 1b. 1b is myristoylated<sup>58</sup>, where 1a is not; however, Abl substrates are not specific to one isoform or the other.

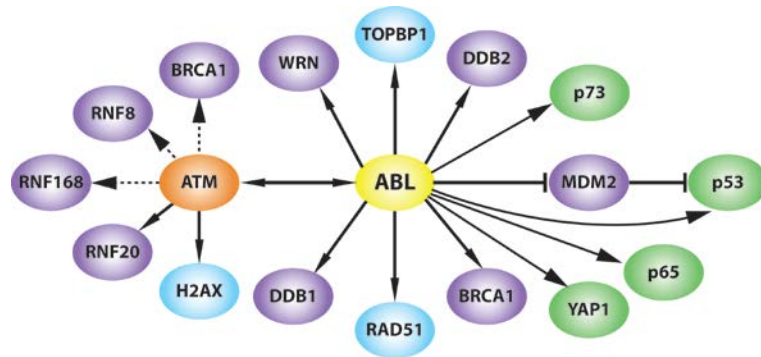


Figure 1.6. Model for the signaling functions of c-Abl. Abl responds to DNA damage and may act in a cascade with enzymes that affect whether the cell lives or dies. Solid line: direct interaction. Dashed line: indirect interaction. Ubiquitin related proteins are in violet, modulators of DNA-damage-induced apoptosis are in green, DNA repair and DNA signaling proteins are in blue.

Abl activity is positively or negatively regulated by a number of protein:protein interactions<sup>36, 53, 56</sup>. Due to the number of interaction partners, it is suggested that Abl is exhibits restricted activation, meaning that each signal activates a fraction of cellular Abl proteins to selectively phosphorylate location-dependent substrates<sup>53</sup>. As an example, BCR-Abl is localized in

the cytoplasm, where the phosphorylation events lead to cancer, whereas nuclear entrapment of BCR-Abl causes cell death<sup>59, 60</sup>. Therefore, Abl may be pre-partitioned into multiple distinct protein complexes in one cell, and in each complex, Abl is inactive by binding to one of its *trans* inhibitors that determines the signal function of that Abl molecule. This argument is strengthened

by the fact that Abl does not recognize specific DNA sequences, but it does have specific interactions with transcriptional regulators such as p53<sup>61, 62</sup>.

### 1.2.3 Abl in Brain Disorders

Recently, Abl has been implicated in brain disorders such as Down Syndrome (DS), Fragile X Syndrome (FXS), Alzheimer's Disease, and Parkinson's Disease<sup>56, 63, 64</sup>, as Abl plays multiple roles in neuronal development. Mutations in Abl lead to complications in neurulation, dendrogenesis, and axonal guidance, while Abl hyper-activity can cause other damaging neurological phenotypes<sup>63</sup>. In Alzheimer's Disease, phosphorylated Abl has been found to co-localize with amyloid plaques, neurofibrillary tangles, and granulovacuolar degeneration<sup>65, 66</sup>. Abl activation has also been linked to an increase in the amount of tyrosine phosphorylation of the parkin protein in patients with Parkinson's disease<sup>67, 68</sup>. It has been hypothesized that Abl is activated by neuroinflammation, either by cleavage of Abl at the C terminus by caspases<sup>69</sup> or by the overexpression of Abl<sup>70</sup>, which can further contribute to neuronal damage. Inhibition of Abl by small molecule kinase inhibitors has been shown to rescue neuroprotective function in Parkinson's disease<sup>71</sup>, and treatment of Alzheimer's disease mouse models with kinase inhibitors has been shown to clear beta-amyloid plaques<sup>72</sup> and reduce astrocyte and dendritic cell numbers, improving cognitive performance<sup>73</sup>. Similarly, treatment of Abl with small molecule inhibitors has been shown to rescue uncontrolled growth of presynaptic terminals in *Drosophila* larvae models of DS and FXS<sup>64</sup>.

### 1.2.4 Src in Cancer

While Abl is known as the first kinase specifically targeted by small molecule inhibitors, the Src protein tyrosine kinase was the first protein to be identified as a "proto-oncogene" in normal mammalian cells<sup>10</sup>. Src regulates apoptosis<sup>74</sup>, proliferation<sup>75, 76</sup>, cell adhesion<sup>77, 78</sup>, cell migration, invasion<sup>79, 80</sup>, angiogenesis, and metastasis<sup>10</sup>. High levels of Src are found in a number of tumors, including breast, colon, pancreatic, lung, and prostate cancers<sup>81</sup>, as well as in

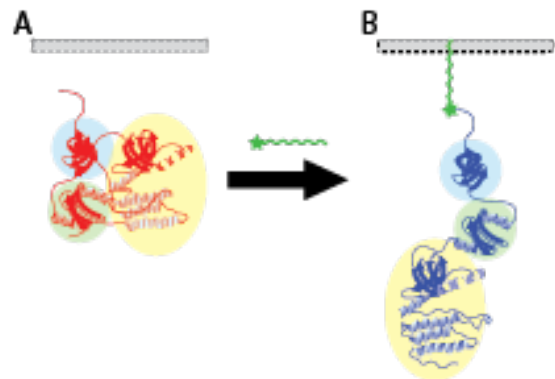


Figure 1.7. Src localization to the membrane promotes activation. The SH3 domain is represented highlighted in blue, the SH2 domain in green, and the KD in yellow. (A) In the inactivated form of Src, where the SH2 domain docks on the C-terminal tail of the KD, Src membrane localization is not promoted. (B) Src localization to the membrane by the myristoylated N-terminus (green star) promotes activation.



neurodegenerative diseases<sup>82</sup>, epilepsy<sup>83</sup>, and HIV/AIDS<sup>84</sup>. Src has also been shown to be the major resistance factor to Herceptin, the first line therapy for Her2+ breast cancer. It has been proposed that Src activity is increased by direct or indirect interactions with receptor tyrosine kinases, reductions in the levels of negative regulatory enzymes, or enhanced activity of phosphatases that dephosphorylate the regulatory Tyr on the Src C-terminal tail<sup>11-13</sup>.

Full-length Src is a less massive kinase than full-length Abl, weighing in at ~60kDa, and is composed of a unique region at the myristoylated N-terminus, followed by the regulatory SH3 and SH2 domains and the KD<sup>12</sup> (Figure 1.5C). Unlike Abl, where the myristoylated N-terminus binds in the C-lobe of the kinase domain as a form of negative regulation, Src myristoylation is used to promote membrane localization and activation<sup>85, 86</sup> (Figure 1.7). In the closed, inactive Src3D conformation, the SH2 domain recognizes and binds to the phosphorylated Y530 residue of the KD C-terminus<sup>20</sup> (Figure 1.4A, red). Binding of SH2 and SH3 ligands can affect the activity of Src through alternating conformational change, either permitting or access to the catalytic KD. Furthermore, displacement of the SH3 and SH2 domains are hypothesized to support Src substrate selection<sup>11</sup>. Therefore, modulating the conformation of Src could be a fruitful pathway for influencing downstream signaling effects.

### **1.3 Current High Throughput Screening Methods for Kinase Inhibitor Discovery**

Drug discovery efforts can be categorized into two extremes: HTS, where speed is prioritized over information content, and high content screening (HCS), where information is prioritized over speed. However, there are many technologies that fall between these two extremes. Due to space constraints, this section is not meant to be comprehensive of all HTS and HCS technologies. Instead, this section will focus on technologies that are specific to kinase inhibitor discovery, and readers are directed to several reviews of current, general HTS<sup>87, 88</sup> and HCS<sup>89, 90</sup> approaches.

As the number of potential therapeutic targets increase, there have been a number of scientific advancements in HTS technologies that can rapidly identify hits in a large compound library. Many of the assays rely on scintillation proximity assays (SPA) or fluorescence detection techniques. In SPAs, radiolabeled ligands are bound to an immobilized protein target before a

washing step to remove any unbound ligand. The light from the bound ligand is used to identify a binding constant<sup>91</sup>. Fluorescence assays are also popular in HTS, as most fluorophores have a short duty cycle and have high emitted photon fluxes that require modest excitation light sources. Furthermore, there are a variety of fluorescence techniques at the disposal of HTS applications, such as fluorescence resonance energy transfer (FRET), fluorescence polarization (FP) assays, homogeneous time resolved fluorescence (HTRF) and fluorescence correlation spectroscopy (FCS). FRET relies on the non-radiative transfer of energy between a donor and acceptor fluorophore. This has been used extensively in enzymatic assays where a short peptide is labeled at both ends with a donor and acceptor fluorophore. The resulting cleavage from the enzyme causes a decrease in fluorescence due to the decrease in proximity of the fluorophores<sup>92</sup>. In an FP assay<sup>93</sup>, the rotational diffusion coefficient of a small, labeled probe changes when it is bound to a larger molecule. HTRF takes advantage of the long fluorescence lifetimes of europium chelates to increase the sensitivity of fluorescence-based assays<sup>94</sup>. Another popular screening technology is the AlphaScreen. In this immunoassay, ligands and proteins are attached to donor or receptor beads. When the beads are brought into close proximity by the interaction between a ligand and a protein, causing a fluorescent signal<sup>95</sup>. Additionally, NMR<sup>89</sup> and X-ray crystallography<sup>90, 96</sup> fragment-based screens have been used to directly identify a potential inhibitor for a protein target and characterize its mode of binding in a HCS-type setting.

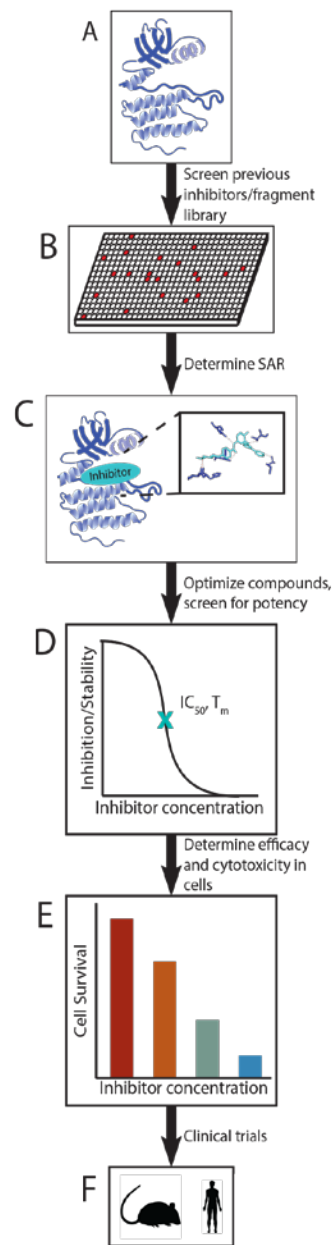


Figure 1.8. The drug discovery process. (A) A kinase is selected for screening. (B) The target is screened against a panel of known kinase inhibitors or against a fragment library. (C) NMR, X-ray crystallography, or computational efforts are used to determine the SAR of the inhibitor:target interaction. (D) After optimization, inhibitors are screened for potency, typically using some sort of  $IC_{50}$  or  $T_m$  as a measure of inhibitor strength. (E) The inhibitor is tested in cancer cell lines to determine efficacy and cytotoxicity. (F) Inhibitors are moved onto animal studies and clinical trials.

However, this dissertation will focus on tools that are specific to the kinase community. There are a myriad of screening technologies that are used to discover new kinase inhibitors *in vivo* and *in vitro*, but the steps that need to be taken from the discovery of a new kinase target to a clinical inhibitor follow the same basic steps<sup>97</sup>. For a new kinase target (Figure 1.8A), the first step of a screen often involves the re-screening of inhibitors or fragment libraries that were developed in previous kinase assays (Figure 1.8B). The scaffolds of any identified lead compounds are typically diversified to search for selective, highly-potent inhibitor compounds. Once the initially broad range of inhibitors is narrowed down, structure-informed design is often used to determine key interactions between the inhibitor and the active site of the kinase (Figure 1.8C). This information is used to modulate potency or selectivity. Crystal structures and homology models are often used for virtual ligand screening approaches. The resulting lead compounds are then optimized, and their potency is tested, usually in terms of an  $IC_{50}$  or  $T_m$  measurement for the protein:inhibitor complex (Figure 1.8D). From there, the best compounds are screened in cells to determine efficacy and cytotoxicity (Figure 1.8D). Finally, the surviving inhibitor(s) are sent to animal studies and eventually clinical trials in human subjects (Figure 1.8E).

Despite the relative simplicity of the workflow above, the goal of achieving a highly-selective kinase inhibitor is not a trivial task, as there are over 2,000 nucleotide-dependent enzymes in addition to the 518 kinases encoded in the human genome<sup>98</sup>. Ideally, the selectivity of an

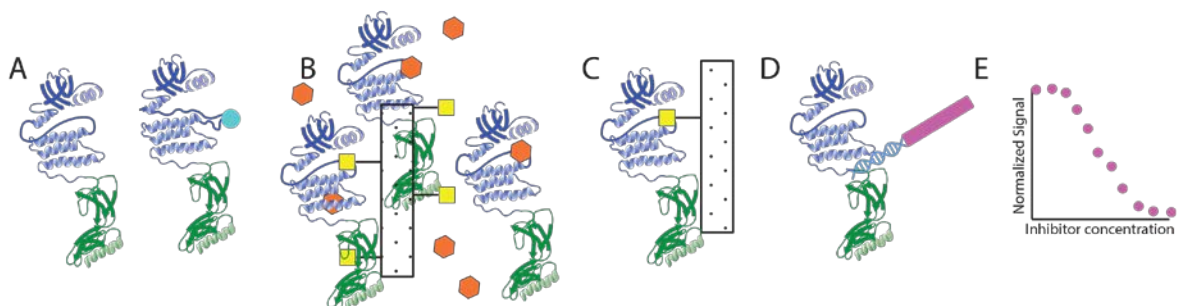


Figure 1.9. One type of PCR-based output. (A) Kinases (purple) are fused to a DNA binding domain of a protein such as NFκB (green) and either phosphorylated (represented by the blue ball) or dephosphorylated. (B) Kinase fusions are incubated with a kinase inhibitor probe (yellow) loaded onto affinity beads as well as test inhibitors (orange). (C) Kinases that are not bound to the probe are washed off. (D) Any remaining kinase fusions that are bound to the affinity beads through an interaction with an inhibitor probe are eluted off of the column and incubated with a chimeric double-stranded DNA tag fused to a qPCR amplicon (pink). (E) qPCR readout is then used to create a binding curve to determine a  $K_d$  value for the test inhibitor binding to the phosphorylated and non-phosphorylated kinase fusions.

inhibitor is first screened at the protein level, before moving to the cell and whole organism levels<sup>97</sup>. There are over 400 diverse enzymatic or binding assays that are available commercially for kinases<sup>97</sup>. As these commercial services are often high cost, typically compounds are screened at either 1 or 10  $\mu$ M concentration before being evaluated in dose-response curves against the kinase of interest. This means that despite the many sources of ligand binding data available to researchers in principle, only a very limited set of selectivity data are available for the most common kinase inhibitors in practice. To this end, there are several approaches that can be used to generate such kinase:inhibitor selectivity data.

In one output, kinase constructs are fused to the DNA binding domain of Nf $\kappa$ B in infected HEK293 cells. The kinases are then either endogenously phosphorylated in the presence of phosphatase inhibitors, or de-phosphorylated by incubating cell extracts at 45 min at 30°C to promote endogenous phosphatase activity. Kinases are then labeled with a chimeric double-stranded DNA tag construct containing the Nf $\kappa$ B binding site fused to an amplicon for qPCR readout (Figure 1.9A). Competition assays are then performed by loading affinity beads with a kinase inhibitor probe molecules and a test compound (Figure 1. 9B) before removing any kinases not bound to the affinity beads with wash steps and eluting the bound kinase from the affinity column.  $K_d$  values can be measured from the qPCR readouts<sup>99, 100</sup> (Figure 1.9D, E).

Alternatively, fluorescence assays can be designed that are sensitive to the activation state of the kinase. In one type of assay, an environmentally sensitive fluorophore like acrylodan can be installed within the activation loop region of the KD through a cysteine mutation<sup>101</sup> (Figure 1.10A). When the activation loop is in the inactive conformation, there is no change in the emission maxima at 475 and 505 nm. When the activation loop is in the active conformation, an increase in the maximum emission at 505 nm compared to 475 nm is observed. The fluorescence output can be used to determine an  $IC_{50}$  value (Figure 1.10B), as well as determining if an inhibitor binds in a type I or type II manner. In another fluorescence assay, a tailored fluorescent constituent like 2,3-diaminopropionic acid (Dap)-pyrene can be incorporated into a substrate kinase peptide (Figure 1.11A). When the Tyr of the activation loop is phosphorylated, the pi-pi interactions of the pyrene and phenolic ring of the Tyr are interrupted, thus shifting the overall fluorescent signature of the peptide and increasing the emission following excitation (Figure 1.11B)<sup>102</sup>. However, upon inhibition, phosphorylation of the substrate Tyr is decreased.  $IC_{50}$

values can be extracted from such data by measuring the fluorescence output as a function of inhibitor concentration, and then compared across all inhibitors (Figure 1.11C) among kinases that have been activated by phosphorylation or those that remain unactivated.

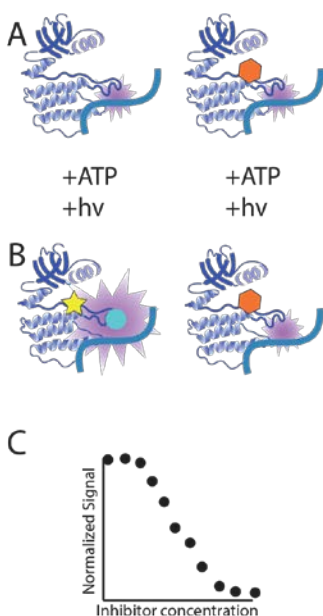


Figure 1.10. Kinase assay using a Dap-pyrene substrate. (A) Kinase can be incubated with or without an inhibitor. The Dap-pyrene forms a pi-pi interaction with the tyrosine phenol, quenching Dap=pyrene fluorescence (B) ATP and an excitation wavelength can be added. The inhibitor prevents ATP from binding, therefore preventing kinase autophosphorylation. (C) The fluorescence signal against inhibitor concentration is used to create  $IC_{50}$  curves for the inhibitor of interest.

Phosphorylation of the tyrosine breaks the pi-pi interactions, increasing fluorescence. (C) The fluorescence signal against inhibitor concentration is used to create  $IC_{50}$  curves for the inhibitor of interest.

Another approach to determining compound potency is measuring the melting temperature of the kinase:ligand complex. For these experiments, technology such as isothermal calorimetry (ITC)<sup>105</sup>,<sup>106</sup>, differential scanning fluorimetry (DSF)<sup>107</sup>, and differential scanning calorimetry (DSC)<sup>108</sup>,<sup>109</sup> have been successfully utilized. In an ITC experiment, sensitive thermopile or thermocouple circuits are used to detect small temperature differences between a reference and a sample cell. The amount of power that is needed to maintain the sample and reference cell at the

Additionally, affinity selection-mass spectrometry (AS-MS) techniques<sup>103</sup>, which can identify protein-bound components from complex mixtures, have also been described. In an AS-MS experiment, the protein is first incubated with ligands of interest to promote complex formation before separation of the protein:ligand complexes from non-binding components. The ligands can then be identified by measuring the molecular weight or through unique collision induced fragmentation patterns. This technique has been used has been used to rank the binding affinity of multiple ligands for a receptor protein while also determining if inhibitors bound in a competitive or allosteric binding site<sup>104</sup>.

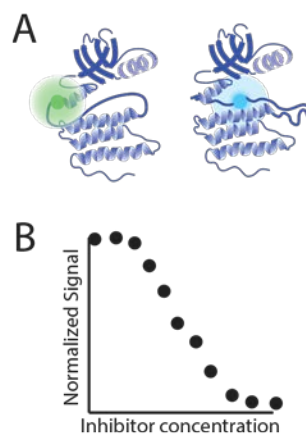


Figure 1.11. Activation-state dependent inhibitor assay. (A) Acrylodan, an environmentally-sensitive fluorophore, can be used to label a cysteine mutant on the kinases's activation loop. (B) The normalized fluorescence readout can be used to determine  $IC_{50}$  for the kinase and inhibitor. A decrease in the normalized fluorescence can also be used to determine if the inhibitor is a type I or a type II.

same temperature is measured as a function of inhibitor binding. ITC can measure binding affinity and the thermodynamics of binding. DSC works in a manner similar to ITC, but the output is the amount of heat that is needed to increase the temperature of both a sample and a reference cell as a function of temperature. In a typical DSF experiment, fluorescence from Trp or Tyr residues in the protein are measured to look for a diagnostic shift to indicate protein unfolding. Typically in kinase assays, SYPRO Orange, a thermofluor that binds nonspecifically to hydrophobic surfaces, is added to the protein or protein:ligand complex to more easily measure the unfolding of the complex. As the protein unfolds, the dye thermofluor binds to the exposed hydrophobic surfaces, thus increasing the fluorescence of SYPRO Orange<sup>110</sup> by decreasing its interaction with water. Thus, the unfolding curves can be used to calculate the melting temperature of protein:ligand complexes.

After the initial validation in an *in vitro* screen, the cellular selectivity of the potential inhibitor must be assessed. For this purpose, many cell lines have been engineered to determine the inhibition of a kinase<sup>97</sup>. For example, in the pro-B-cell murine line, Ba/F3 cell line, transformation by a kinase leads to interleukin-3-independent proliferation. Non-specific cytotoxicity can be “rescued” through the introduction of interleukin-3. This approach can be combined with traditional affinity chromatography or competitive displacement assays in an unbiased proteomics-based approach. Non-specific interactions can also be determined at this level by using capture compound mass spectrometry (CCMS)<sup>111</sup>. In this approach, a photo-

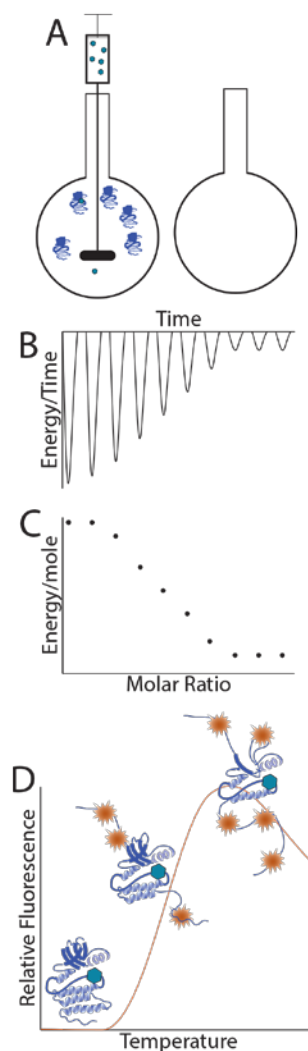


Figure 1.12. Schematic of ITC, DSC, and DSF. (A) In ITC and DSC, a reference cell of buffer is maintained while ligand is titrated into a constantly-stirring sample chamber. (B) The amount of energy it takes to maintain the same temperature as ligand is added. (C) The final output is a binding curve. (D) In DSF, the sample is heated with SYPRO orange, a fluorophore that binds nonspecifically to hydrophobic surfaces. As the protein is heated, it unfolds, and SYPRO orange binds to the now-exposed hydrophobic surfaces. The result is a melting curve.

activatable reactive moiety and a sorting moiety are added to a compound of interest (termed capture compound). Capture compounds are then incubated with cell lysates, and photolysis leads to the generation of a carbene or nitrene, which forms a covalent crosslink between the capture compound and the target proteins. The sorting mechanism for these compounds relies upon interactions with biotin, which allows the implementation of streptavidin-coated magnetic beads. The captured proteins are then enzymatically digested and analyzed using LC-MS.

In the final steps of inhibitor screening, kinase inhibitor selectivity must be evaluated at the organism level. However, this is often difficult, as different cell lines can have dramatically different responses to a given inhibitor. Furthermore, evaluating inhibitors in non-human models can have drastically different effects, and the on- and off- target effects must be evaluated by globally monitoring the changes in phosphorylation. While CC-MS can also be applied to this problem, it remains a novel technique that has not fully been validated in the context of kinase inhibitor screening.

While the approaches described above are individually important, successful inhibitor discovery workflows typically favor increased speed and chemical space coverage over the ability to differentiate the binding modes accessed by the screened small molecules. The final verification of inhibitor binding mode is most often determined through the high-resolution X-ray structure of the protein:ligand complex, but typically only for high-value lead compounds. Such structural information can be extremely important in evaluating the molecular details of the protein:inhibitor interaction formed, the details of any conformational changes caused by ligand binding, and predicting downstream signaling effects<sup>21, 112-117</sup>. Thus, in the continuum of drug discovery screening that ranges from HTS to high content screening, there is a need to operate beyond a pure drug discovery mode, where structural characteristics of the protein:inhibitor interaction can be probed. This new methods of drug discovery that access the full range of both structure and binding information content described above would be highly valuable in optimizing the efficiency of current inhibitor discovery paradigms. Here, ion mobility-mass spectrometry (IM-MS) can operate in a mode that provides high information content in a medium-throughput screening space.

## **1.4 Targeting Protein:Protein Interactions**

Another class of “undruggable” targets are modulators of transcription. Transcriptional activators play a key role in human disease, as they bind to DNA and stimulate gene transcription. Misregulation of transcription has been implicated in most diseases including many cancers and inflammatory diseases<sup>118-120</sup>. However, targeting transcriptional activators are difficult due to the

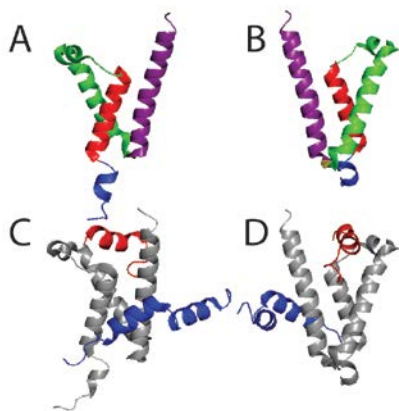


Figure 1.13. Representation of KIX. (A) KIX colored by helix:  $\alpha$ -helix 1 in red, 3-10 helix 1 in orange,  $\alpha$ -helix 2 in green,  $\alpha$ -helix three in purple, and 3-10 2 in blue. (B) 180° view from (A). (C) KIX is colored in gray. MLL is colored in red, and pKID is colored in blue. (D) 180° view of (C).

transient nature of the protein:protein interactions involved in the transcription process and the lack of structure in many of the key protein targets. One of the domains that has been the focus of recent drug discovery efforts is the kinase-inducible domain (KIX) of the CREB (cAMP response element binding protein)-binding protein (CBP). While CBP is a large multi-domain protein that coordinates several transcription factors, its KIX domain serves as a hub that can bind over 12 different coactivators through two binding sites: the pKID binding site and the MLL binding site<sup>119, 120</sup>. Although KIX is a conformationally dynamic protein, the NMR

structures of the binary and ternary KIX complexes reveal that KIX is composed of three alpha-helices and two 3-10 helices<sup>121</sup> (Figure 1.13A,B). The MLL binding site is formed between the first and third alpha helix (Figure 1.13C,D, red), while the pKID binding site is formed in a groove between the first and third alpha helices, with pKID wrapping around the third alpha helix (Figure 1.13 C,D, blue). pKID is phosphorylated at Ser133, which directly participates in binding to KIX. pKID is an example of an inducible activator; upon phosphorylation, pKID interacts more strongly with KIX than its counterpart, c-Myb<sup>122, 123</sup>. C-Myb and pKID share an overlapping binding site, and pKID can out-compete c-Myb, leading to changes in gene instruction. NMR studies have revealed that the pKID and MLL peptides are unstructured, and upon binding to KIX, undergo significant conformational changes into alpha helices<sup>124</sup>.

Peptides and peptidomimetics have been found that can inhibit these protein:protein interactions that function in cells, and these inhibitors have been extremely useful in defining the



characteristics of these activators<sup>125, 126</sup>. However, it can be difficult to localize the effect of such inhibitors, especially as the pKID and MLL binding sites are connected by an allosteric network in KIX—in that whenever MLL is bound, pKID binds with a two-fold higher affinity, and vice versa<sup>124, 127</sup>. Two natural products have also been discovered that act as orthosteric inhibitors of the KIX:MLL complex: sekikaic and lobaric acid. These inhibitors also have been observed to exert an allosteric effect on the pKID binding site of KIX<sup>128</sup>.

## 1.5 Ion Mobility-Mass Spectrometry

The advantages of using techniques such as 2D nuclear magnetic resonance (NMR) in drug discovery include direct information about the binding site, and dual experiments with orthosteric

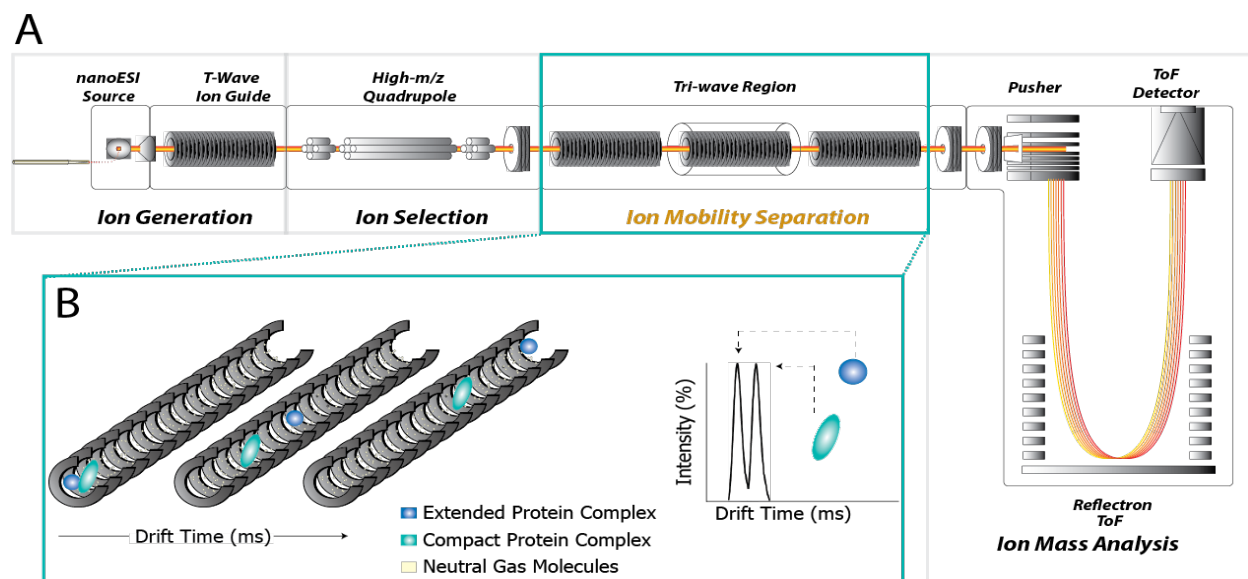


Figure 1.14 Schematic of the Waters Synapt G2 HDMS. A) Instrument diagram of the main regions of the instrument: ion generation by nESI, ion selection up to 32,000  $m/z$  in the quadrupole, the Tri-wave region, where IM separation occurs, and the TOF mass analyzer. (B) Details of IM separation. Ions are introduced into the drift tube filled with inert, neutral gas molecules under the influence of a weak electric field. Small, more compact ions (blue) will experience fewer collisions than larger, more extended ions (teal); thus, the ions will be separated by size.

inhibitors can identify second-site binders. However, the requirements of NMR, e.g. highly soluble, stable, and homogeneous solutions at high concentrations, isotopically enriched media, and relatively small proteins (usually less than 40kDa, depending on the experiment) can frustrate the drug discovery and structural determination process. The knowledge gap generated

by such challenges can only be surmounted through development of new HTS technologies, capable of assessing the structure of dynamic, heterogeneous protein systems at low levels. Ion mobility-mass spectrometry (IM-MS) is a technology that promises to alleviate some of these problems, as this technique can simultaneously measures the size and molecular weight of intact protein complexes from small amounts of complex mixtures. For the IM-MS experiments described here, the commercially-available Waters Synapt G2 HDMS was used. (Figure 1.14)<sup>129, 130</sup>.

### 1.5.1 Ion Generation

The experiments described herein fall under the umbrella of native MS, where protein and protein complexes are introduced into the gas-phase under conditions that maintain the native-like structure and stoichiometry *in vacuo*<sup>131</sup>. These types of experiments are possible largely due to the advent of electrospray ionization (ESI). The scale of ESI has been reduced to lower flow rates and smaller capillary diameters. This form of ESI, referred to as nano-ESI (nESI) benefits from increased sensitivity and lower sample consumption as compared to traditional ESI. Furthermore, nESI is more tolerant to salt and other impurities, which is advantageous for native MS experiments. ESI and nESI are achieved by applying a potential difference between the inlet of the mass spectrometer and a conductive capillary containing the analyte solution. Once an electric field between the capillary and the counter electrode of the mass spectrometry reaches a high enough field strength (Figure 1.15A), charge will accumulate at the liquid surface at the end of the capillary, a Taylor cone is formed at the tip of the capillary, and highly-charged droplets will be generated (Figure 1.15B)<sup>132</sup>. These droplets will undergo a series of evaporation and fission events until the final stage of ionization.

There are many models that describe the ion formation mechanism in ESI and nESI. One of the most prevalent, called the charged residue mechanism (CRM)<sup>133, 134</sup> proposes that ions are created following complete solvent evaporation, which deposits charge on the analyte in a manner correlated with its solvent accessible surface area (Figure 1.15C). In another important ion formation mechanism, termed the ion evaporation model (IEM)<sup>135, 136</sup>, droplet evaporation occurs until the field strength of the droplet's surface is large enough for solvated, surface active ions to be ejected into the gas-phase (Figure 1.15D). Generally, it has been proposed that smaller analytes follow the ion evaporation model<sup>136, 137</sup>, and larger analyte follow the charged

residue model<sup>133</sup>. However, neither model fully explains ion formation. To this end, in 2009, Hogan *et al.* proposed a model of ESI that combined the CRM and IEM into a mechanism called charged residue-field emission model. In this mechanism, proteins are proposed to be charged residues, but the final charge state on the proteins is dependent on emission of small buffer ions from ESI droplets before complete solvent evaporation. Each ion has an  $E^*$ , or the electric field at which an ion will be emitted from a droplet, and charged species with the lowest  $E^*$  will emit from the droplets first. After this process occurs, the species with the next lowest  $E^*$  will emit from the droplet until the protein is finally emitted from the droplet. Thus, the final charge on a droplet is determined by the protein surface area and the amount of charge carried away by other, smaller ions<sup>138, 139</sup>.

An alternative to these models, especially for unfolded proteins, proteins that are partially hydrophobic, or proteins that are capable of binding excess charge carriers is the chain ejection model (CEM), developed by Konermann *et al.* in 2012. The CEM proposes that when unfolded proteins, or proteins with largely exposed hydrophobic regions are inside a charged droplet, unfolded chains will move to the surface of the droplet, followed by the expulsion of the unfolded chain. This leads to the ejection of rest of the protein and separation from the droplet<sup>140</sup>. Another mechanism, proposed by the Loo lab, suggests that ions evaporate from protrusions of decomposing droplets, where the charge of the protein is defined by the partitioning of charge between the analyte and the solvent, based on the titrateable groups on large bio-ion surfaces<sup>141</sup>. However, it is unlikely that any of these processes occur independently or exclusively; rather, a combination of the IEM and CRM with other modes of ion formation is likely at play during most nESI experiments.

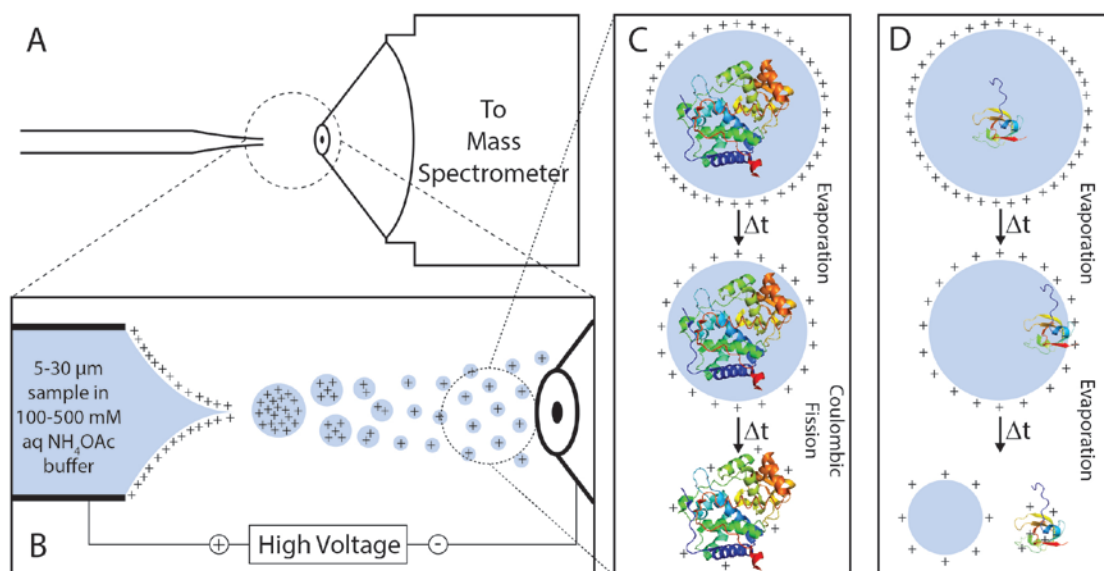


Figure 1.15. Diagram of ESI. (A) Position of the capillary relative to the inlet of the mass spectrometer. (B) Formation of the Taylor cone, which generates highly charged droplets. Typically, protein molecules follow the (C) charge residue model, where droplets evaporate until their diameter has reached the point at which the analyte is thoroughly evaporated. (E) Smaller proteins are predicted to follow the ion evaporation mechanism (IEM). In this model, as the droplet evaporated, the analyte will travel to the edge of the droplet. At the point at which the electric field on the Rayleigh-charged droplet is sufficiently high, the protein ions will be ejected from the droplet surface.

### 1.5.2 Maintaining and Detecting Native-Like Complexes in the Gas-Phase

The first reports of using MS to study protein complexes in their native state were met with skepticism, as it was thought that the hydrophobic, apolar vacuum of the mass spectrometer may turn proteins inside-out. Katt and Chait<sup>142</sup>, and Ganem, Li and Henion<sup>143</sup> were the first to demonstrate that specific noncovalent complexes could be detected by ESI-MS. Since then, many observations of specific protein complexes, identical to the selectivity and stoichiometry of those known in solution, have been observed in the gas-phase. Furthermore, the structures of proteins in the gas-phase have been studied extensively, revealing surprising similarities to their solution-state analogues<sup>144</sup>.

During these native MS experiments, proteins are typically first buffer exchanged into a native-like (pH 7), MS-compatible buffer such as 100-500mM ammonium acetate. The ionization conditions of the mass spectrometer are then optimized so that residual water and buffer molecules will be stripped away by ion-neutral collisions within the beamline of the instrument without causing protein complex unfolding or dissociation. This is done by increasing the

pressure of neutrals within the initial ion guide regions of the instrument, which acts to focus larger ions by reducing their radial velocities<sup>145</sup>.

The transmission of large biomolecular complexes to the IMS portion of the instrument is made possible by a high  $m/z$  quadrupole to allow for tandem MS experiments. A quadrupole is made up of two perpendicular rod pairs, where the individual rods within each pair are parallel to each other. Each pair of rods have an applied radio frequency (RF,  $V\cos(\omega t)$ ) with a DC offset ( $U$ ), with one pair having a positive applied potential defined as  $U+V\cos(\omega t)$  and the other pair having a negative applied potential defined as  $U-V\cos(\omega t)$ . Only ions of certain  $m/z$  values can be transmitted for given DC and RF voltages; therefore, ions are transmitted through the quadrupole based on their stabilities in the oscillating electric fields that are applied to the rods. Typically, quadrupole mass analyzers operate below 4000  $m/z$ <sup>146</sup>, and whenever only an RF field is applied, can act as a broad-band mass filter. However, by reducing the frequency of the RF voltage applied to the rod set, it has been possible to mass analyze and transmit ions up to 32,000  $m/z$  and beyond<sup>147</sup>.

Unlike a quadrupole, a time-of-flight (ToF) mass analyzer has no theoretical upper  $m/z$  limitation, making it an ideal choice for native MS experiments. In a ToF mass analyzer, ions are given a fixed kinetic energy and allowed to traverse a field-free vacuum. The transit time of the ions through this field free region is related to the  $m/z$  of the ion through the classical description of kinetic energy. The principle behind ToF relies upon producing a packet of ions moving in the same direction and having constant kinetic energy in order to achieve high mass resolving power. The mass resolving power of a ToF is further limited by differences in the initial energy, positions, and collisions the ions undergo during ToF separation. Some of these limitations can be overcome using a reflectron ToF<sup>148</sup>, where the ions are typically time-domain focused by an electrostatic ion mirror. The reflectron is composed of a series of equally-spaced rings or grid electrodes to redirect the ions toward the detector or another reflectron. Ions with greater kinetic energy will penetrate the electric field more deeply, therefore taking longer to reach the detector. Therefore, two ions with the same  $m/z$  but with different kinetic energies will reach the detector at the same time. Faster ion detectors are necessary for ToF-MS, and relatively recent advancements in this area have sought to increase the sensitivity and dynamic range of the ToF measurement<sup>149</sup>.

### 1.5.3 Principles of Ion Mobility Separations

Langevin first described measuring the mobility of an ion in the gas-phase in 1903<sup>150</sup>, with Bradbury Later applying Langevin's theory to measure the mobility of ions in the gas-phase<sup>151</sup>. In the 1950s and 1960s, Mason and McDaniel laid much of the theoretical and experimental framework used by modern IM instrumentation to study gas-phase ion molecule reactions<sup>152</sup>. Although IM is not a new analytical technique, it has only been used in the analysis of bio-ions since Bowers *et al.* showed its utility in separating protein conformers in 1995<sup>153</sup>. Since then, several forms of IM have been used in the analysis of bio-ions: drift tube IM, high-field asymmetric waveform IM (FAIMS), and traveling-wave (TWIMS).

In the most simple form of IM, a packet of ions is pulsed into a drift tube where they are separated by mobility through a pressurized drift cell in the presence of a weak axial electric field<sup>154</sup>. The ion drift times depend on both size and charge. These drift times can easily be converted into collision cross section (CCS), which is an orientationally averaged ion-neutral size parameter that can be used to compare the global sizes of IM ions to sizes derived from other biophysical techniques such as x-ray crystallography, NMR, and cryo-electron microscopy.

The IM-MS platform used to conduct all the experiments described in this thesis utilizes TWIMS to separate ions<sup>129, 155, 156</sup>. In TWIMS, alternating phases of RF voltage are applied to a stacked ring ion guide. Ions are carried along the crests of the potential waves produced, while a static background of neutral gas resists their transit to the MS stage of the instrument. Larger ions will experience more collisions with the gas molecules, and as a result, will roll over the crests of the waves more frequently and will thus be the last to exit the TWIMS separator. As ion movement in a TWIMS instrument is currently not completely understood at this time, CCS is typically determined by = calibrating the drift time data for ions of interest against the drift times of ions that have known CCS values, separated under identical TWIMS conditions<sup>157-161</sup>. For protein CCS determination by TWIMS, it is important that the ions that are used for calibration bracket the expected CCSs and mobility values of an unknown ion. For proteins and protein complexes, databases of standard native proteins and protein complexes have been curated, and are commonly used to generate CCS values for proteins with average precisions of 3%.<sup>156, 158, 160</sup>.

## 1.6 Mass Spectrometry in Drug Discovery

MS of large biomolecules is a versatile tool used within both drug discovery and development<sup>162</sup>. Technologies such as electrospray ionization (ESI)<sup>163</sup> enable MS to contribute to proteomic analyses<sup>164</sup>, purity assessments of isolated targets, structural determinations of protein–ligand complexes<sup>165, 166</sup>, and biomarker discoveries<sup>167, 168</sup> that link candidate molecules to critical metabolic processes in pre-clinical evaluations. The ability of MS to gain a large amount of both qualitative and quantitative information from complex, dynamic biological mixtures is its chief advantage over other analytical tools. When labeling chemistries<sup>169, 170</sup>, chemical cross-linking<sup>171</sup>, hydrogen-deuterium exchange<sup>172-174</sup>, and other technologies are combined with MS, it becomes clear that the range of this information content can include significant structural data on the protein–ligand interaction that can rapidly inform the discovery of lead compounds.

### 1.6.1 Paradigms for Protein-Ligand Screening by IM-MS

Computational methods are typically used in conjunction with IM data to generate atomic models of peptide and protein structure<sup>175</sup>, and have advanced significantly over the past few years in their ability to generate such models for larger systems<sup>176</sup>. Smaller protein–ligand systems can be analyzed by IMMS deduce the binding locations for small molecules within protein targets<sup>177-179</sup> and, in some cases, produce atomic models of protein–ligand complexes<sup>180, 181</sup>. Larger protein–ligand complexes are currently beyond the scope of such detailed computational methods, and instead often involves the observation of a key protein conformation shift as a function of a known binding event that can be linked directly to compound efficacy. Subsequent experiments can then be constructed to search a broader library of compounds for similar conformation shifts upon binding the same target (Figure 1.16). This general mode of operation is currently the most-commonly employed approach for IM-MS in the context of protein–ligand analysis and screening.

In addition to the above-noted charge state dependence for gas-phase protein CCS, early studies noted other critical variables that affect the gas-phase structure of desolvated protein ions<sup>182, 183</sup>. Among these, altering the internal temperature of the ions produced had a dramatic influence on the size of the protein ion recorded by IM, primarily leading to a positive correlation between protein ion CCS and their internal temperature, with protein ions of high internal temperatures adopting large, string-like conformational states<sup>184</sup>. Subsequent data have extended these observations to include protein–protein<sup>185–188</sup> and protein–ligand<sup>189, 190</sup> complexes, each of which display similar yet distinct unfolding properties upon gas-phase activation. Most contemporary

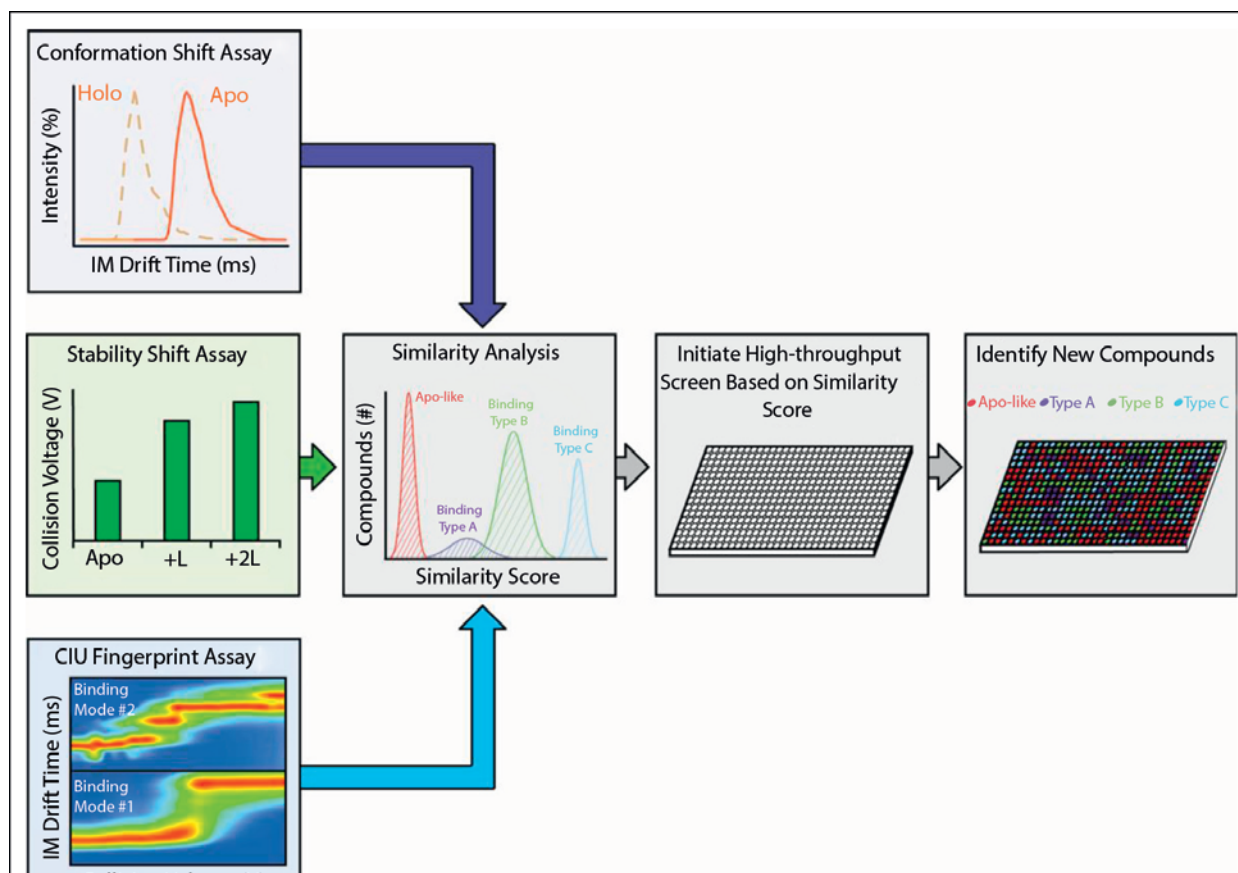


Figure 1.166. Basic screening strategies for intact protein–ligand complexes by IM-MS. Several different modes are available to assess the consequences of small molecule binding within intact proteins using IM-MS screens. Binding may result in a clear conformation shift by IM, and can thus be used as the basis for a conformational shift assay. Ligand binding may also alter the stability of the protein ion when compared to control data, enabling a stability shift assay. Finally, detailed protein unfolding data can be recorded by tracking the sizes occupied by protein–ligand complexes upon activation, and the differences observed between protein–ligand complexes of known binding modes or conformations can be used to construct CIU Fingerprint-based assays. Once known binders are analyzed and metrics assembled that allow for sufficiently accurate scoring of known data, resulting in clustered responses that differentiate a desired class of binder from other potential ligands, a library can be assembled from previously untested molecular scaffolds and used to search for new compounds that replicate the stability or conformation shifts observed in efficacious molecules.



experiments utilize collisional activation to initiate unfolding<sup>185, 189-194</sup>, however other activation methodologies have been shown to elicit conformational change<sup>175, 182, 195-197</sup>, although to a lesser extent. Collision induced unfolding (CIU) can be used in two basic modes in the context of protein–ligand screening experiments (Figure 1.16). Firstly, the surviving population of the most-compact form of the protein, typically that which is most-highly correlated to its solution structure, is tracked as a function of the voltage used to accelerate ions and initiate unfolding. Differences recorded in protein–ligand complex stabilities primarily relate to the stability of the gas-phase complex, and can be compared to both solution measurements and apoprotein CIU data to provide a workable screening methodology<sup>131, 188, 190-192, 194</sup>. In addition to measuring the survival of a single conformational form of the protein–ligand complex upon activation, the unfolding pathway of the protein can be followed in detail to generate additional points of comparison between either apo-states or alternate conformational families of the protein. Since many possible tertiary structures project identical ion CCS values, the detection of subtle conformational shifts in protein–ligand complexes is often challenging for IM-MS methods. CIU fingerprints can be a useful tool in circumventing such limitations, as the unfolding intermediates accessed by proteins during CIU can be uniquely related to specific protein–ligand binding modes<sup>188, 190-192, 194</sup>.

All of the above modes of operation can be combined into metrics that define the structural stability and conformation changes that occur upon binding an efficacious molecule to a protein target, the properties for which are sought to be replicated in new molecular scaffolds (Figure 1.16). Alternatively, IM-MS results can be interpreted using other data, including NMR and X-ray structure information, or computational models in an attempt to link specific conformational shifts observed to desired ligand binding modes<sup>154, 177, 178, 186, 198-201</sup>. Once sufficiently descriptive scoring algorithms are established, the developed screen can be applied to larger libraries to search for molecules that bear similar effects on target protein conformation and stability.

### **1.6.2 Searching for Shifts in Protein-Ligand Stability**

Gas-phase protein–ligand stability measurements by ESI-MS have a long history<sup>131, 144, 154</sup>. While relative ion intensities can be used to generate binding constant information<sup>175, 202</sup>, CID<sup>203, 204</sup> and other tandem MS technologies have been used for many years to study the stability and dissociation of protein–protein and protein–ligand complexes<sup>205, 206</sup>. More recently, ESI-CID-MS has been applied to protein–ligand complex systems of potential pharmaceutical interest. For example, a recent study investigated the dimeric monocyte chemoattractant protein-1 (MCP-1), and found CID thresholds for the complex to be relatively low in the absence of Arixtra, a glycosaminoglycan analog binder. The results, which included IM-MS, indicated that the dimeric MCP-1 is significantly stabilized upon Arixtra binding, and that Arixtra interacts with both of the subunits within the MCP-1 complex<sup>207</sup>.

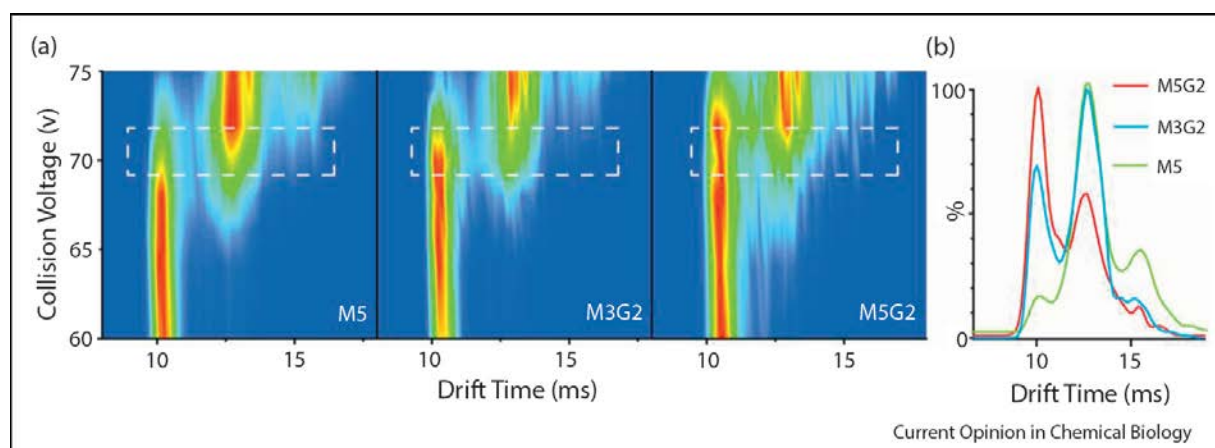


Figure 1.17. IM-MS stability measurements differentiate ligand binding in a tetrameric protein complex. (a) CIU fingerprints of the Concanavalin A (ConA) tetramer acquired on a Synapt G2 quadrupole-IM-MS instrument (Waters, Milford MA), bound in a 1:4 protein tetramer:ligand stoichiometry for three different mannosyl sugars: 3a,6a-mannopentaose (M5, MW = 828.74), 3a,6a-mannotriose-di-(N-acetyl-D-glucosamine) (M3G2, MW = 910.84) and 3a,6a-mannopentaose-di-(N-acetyl-D-glucosamine) (M5G2, MW = 1235.10). Regions of the CIU fingerprints marked with a dashed box shown for each complex are selected for detailed analysis over a range of voltages. (b) Drift time data for the selected regions from (a) show differential stabilities for the ligand complexes that relate primarily to the strength of the protein–ligand complex isolated in the gas phase.

Such IM measurements have appeared with increasing frequency in conjunction with MS-based stability measurements of protein–ligand complexes. For instance, IM-MS was used to study the stability of ubiquitin-cis-[Pd(en)(H<sub>2</sub>O)<sub>2</sub>]<sup>2+</sup> complexes, and indicated that Pd-bound ubiquitin exhibits diminished gas-phase unfolding when compared to the apo protein. Furthermore, it was found that Pd<sup>2+</sup> binding aided conformational stability to a greater extent than Pd(en)<sub>2</sub><sup>2+</sup><sup>208</sup>. IM measurements of protein ions bound to extensive anion and cation populations have been used to

deduce a Hofmeister series analog for gas-phase protein structure<sup>191, 194</sup>. These studies identified several gas-phase specific mechanisms by which proteins can gain differential stability from bound ligands in the absence of solvent. Bound anions, for example, tend toward evaporation upon collisional activation of the complex, thus allowing the protein to dissipate excess internal energy and retain its shape over a broader array of IM-MS conditions. In contrast, cation adducts tend to stabilize complexes by remaining bound to the protein, serving to tether regions of the biomolecule through multi-dentate interactions<sup>192</sup>. More recently, IM-MS data for crown ether (CE) – protein complexes have suggested new modes of stabilizing protein structures in the gas-phase upon ligand binding<sup>209</sup>. The CE compounds studied noncovalently bind preferentially to primary amines, e.g. lysine side chains, and serve to solvate the ionic charge present. The IM-MS data collected showed that CE binding can compensate for rearrangements local to the charge site in a manner potentially similar to solvent molecules in the condensed phase, and thus suggests future routes for tuning and manipulating protein structures in the gas phase through ligand attachment.

Detailed CIU datasets have also been used to study protein–ligand stabilities, and serve as powerful tools to investigate the consequences of small molecule attachment in larger protein systems. In a recent study, ESI-IMMS was used to evaluate the structural stability of natively compact protein ions (FK-binding protein, hen egg-white lysozyme, and horse heart myoglobin) as a function of small molecule binding<sup>189</sup>. The results show clear shifts in the CIU stabilities of ligand bound complexes relative to apo protein, shifting the onset of CIU by up to 21 eV. CIU datasets were also used to assess the stability changes produced in a familial amyloid polyneuropathy (FAP)-associated variant form of the tetrameric protein transthyretin (TTR) upon binding its natural ligand, thyroxine<sup>190</sup>. By combining CID and CIU datasets, it was found that thyroxine binding stabilizes the L55P disease-associated form of TTR to a greater extent than the wild type protein. Furthermore, CIU fingerprints were shown not to depend on the L55P point mutation, and that ligand binding primarily influenced the stability of the most compact tetramer conformations, rather than significantly unfolded forms of the complex. An example of this type of analysis is shown in Figure 1. 19, a dataset selected from ongoing IM-MS work in our laboratory. Concanavalin A (Con A) is a 103 kDa lectin tetramer with well-known structure and sugar binding properties<sup>208</sup>. Our data shows CIU datasets recorded for three ConA-manosyl sugar complexes, and indicates strong shifts in CIU stability for different ligand bound

populations (Figure 1.17). Importantly, if selected areas of the CIU plot are interrogated as shown, the relative stabilities recorded for the most compact form of the protein track precisely with the relative binding strengths of the manosyl sugars used in our experiments. While more data will be required to validate this result, it also illustrates the potential utility of CIU based stability measurements for protein-ligand complexes.

### 1.6.3 Combing for Conformational Transitions

Protein structure and function are intimately related, and conformational changes often herald alterations in the activity of a protein. IM-MS is being used increasingly to supplement NMR and crystallography in order to provide a dynamic view of the structural changes caused by protein-ligand binding. In one study, a new model for the activation of guanosine-monophosphate dependent protein kinases (PKGs) by guanylate cyclases was developed based on both IM-MS and previously available X-ray data<sup>198</sup>. Following on from this work, the calcium binding protein calmodulin (CaM), which mediates the Ca<sup>2+</sup>/CaM kinase II (CaMKII), was studied extensively using IM-MS<sup>210</sup>. A comparison of high-resolution structures acquired for CaM while bound to Ca<sup>2+</sup> and in its apo-form have previously revealed the presence of an extended conformation for the former and a more compact structure for the latter<sup>211</sup>. IM-MS studies confirmed these findings, and also concluded that while shorter peptide ligands are able to bind to both structures, longer sequences bind the globular CaM structure preferentially. This binding trend was not observed in preceding X-ray or NMR datasets. IM-MS has also been used to probe the structure of protein:DNA complexes. Recent work has used IM-MS to investigate the interactions between the central glycolytic gene repressor (CggR), fructose-1,6-biphosphate (FBP), and the DNA target binding sequence of CggR. Previous reports had presented evidence supporting a mechanism involving the cooperative binding of CggR dimers to DNA<sup>212</sup>. IM-MS data confirmed this mechanism, and also observed that FBP addition abolished CggR-DNA binding.

Protein targets that remain relatively challenging for traditional biochemical methods have become a fruitful area for IM-MS experimentation. Ligand binding targets such as intact viral capsids<sup>199, 201</sup>, chaperonin complexes<sup>199</sup>, intrinsically disordered proteins<sup>213</sup> (IDPs), and amyloidogenic proteins (e.g. amyloid  $\beta$  (A $\beta$ ) and  $\beta$ 2-microglobulin ( $\beta$ 2m))<sup>200, 214</sup> have all been

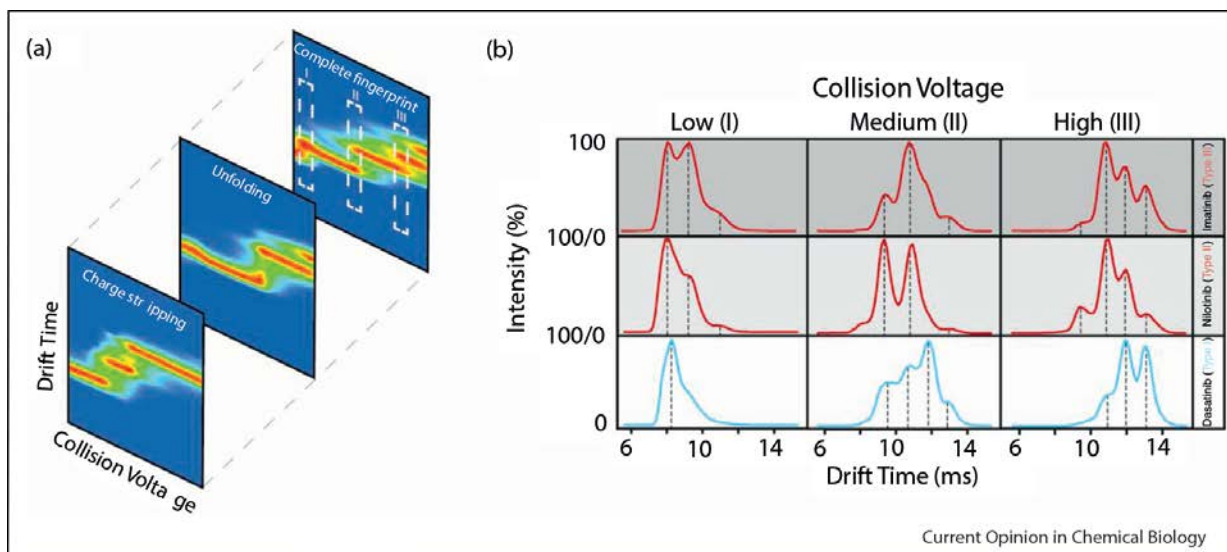


Figure 1.18. CIU fingerprinting differentiates Abelson (Abl) kinase inhibitor binding modes. (a) Components of the CIU fingerprints acquired on a Synapt G2 quadrupole-IM-MS instrument for the tyrosine kinase inhibitor nilotinib, a type II binder (binds only the inactive form of the enzyme). Dashed boxes shown on the complete fingerprint represent three collision voltage regions selected for detailed analysis (I, II, and III). (b) IM drift time spectra at low (I), medium (II), and high (III) collision voltages for three kinase inhibitors bound to the Abl kinase: imatinib (dark gray, top), nilotinib (light gray, middle), and dasatinib (white, bottom), with the first two inhibitors are known to bind in a type II mode (red data) and the latter inhibitor known to bind in a type I mode (blue data). Significant differences are observed between the CIU data derived from Abl-inhibitor complexes of each type, and dashed lines mark all of the unique features with relative intensities and centroid drift times that allow for the similarity analysis and classification of each dataset.

studied in the context of ligand binding by IM-MS. Little is known about the pathways and regulation of viral capsid assembly because intermediates are of low abundance, and therefore elucidation of their exact structure and composition is difficult. IM-MS was used to reveal that viral capsids are dynamic entities that can freely exchange with building blocks in solution<sup>201</sup>. In another study, folding of the bacteriophage T4 capsid protein gp23 by the GroEL chaperonin complex revealed that even in the gas phase, the target protein is enclosed within the GroEL cavity. This study showcases the ability of IMMS to determine the sizes and compositions of unknown chaperonin complexes bound to client proteins<sup>199</sup>. Despite being smaller, typically, than the targets discussed above, IDPs present structural populations of unprecedented flexibility making them a tremendously challenging class of targets for traditional structural biology tools. Recent IM-MS experiments have been applied to the cyclic kinase inhibitor Sic1, an IDP of

central importance in cell cycle control and implicated in the etiology of many cancers<sup>213</sup>. IM-MS data showed that Sic1 exists in three distinct monomeric conformations, as well as two dimeric conformations, with most of the population existing in a highly collapsed monomeric state that is distinct from what is typically expected for a highly unfolded protein. In related work, IM-MS studies have shown that ligand binding can prevent the dimerization of c-Myc and c-MAX, two IDP oncoproteins, via the formation of a leucine zipper-type interaction critical for DNA binding<sup>215</sup>. IM-MS measurements have also been used to study the conformational distribution of the IDP IB5, a protein that is responsible for taste astringency, while bound to multiple tannin ligands<sup>216</sup>. This study presents a convincing mechanism for dynamic IDP function, in which several conformations of the protein serve in a functional role.

Amyloid forming proteins, and their interactions with potential therapeutics, have been extensively studied by IM-MS. A generally accepted route for therapeutic discovery in diseases such as Alzheimer's involves small molecule inhibitors for the early stages of A $\beta$  aggregation. Detailed IM-MS experiments on short peptide segments<sup>217</sup> and full-length A $\beta$ <sup>178, 214</sup> have been used to elucidate the general features of peptide aggregation. More recently, IM-MS was used to study the impact of short peptide fragment ligands derived the hydrophobic c-terminus of A $\beta$  (c-terminal fragments, CTFs) on the early aggregation pathways of the full-length protein<sup>218, 219</sup>. In addition to deducing the structures for CTFs in isolation, the data collected revealed that CTF binding modulates extensively the oligomeric states of the A $\beta$  aggregates detected, eliminating entirely the populations of larger aggregate species found in control samples without changing the morphologies of the insoluble amyloid fibrils. In a similarly themed study, a panel of small molecule inhibitors was screened to find those that interrupt  $\beta$ 2m aggregation and fibril formation<sup>200</sup>. Out of the 44 ligands studied by IM-MS, rifamycin, a well-known antibiotic, was identified to be the most potent protein aggregation inhibitor. Furthermore, IM-MS data clearly linked the potency of rifamycin to its direct interactions with  $\beta$ 2m monomers. Another recent example involves the polyphenol (—)-epigallocatechin-3-gallate (EGCG), a small molecule with known anti-amyloidogenic properties<sup>220</sup>. IM-MS data, integrated with NMR and other biophysical information, revealed that EGCG preferentially binds to compact forms of A $\beta$  monomers and dimers<sup>177</sup>. Furthermore, increased anti-amyloidogenic activity for EGCG along with similar conformational changes were observed upon Cu(II) and Zn(II) binding to A $\beta$ , and the conformational distributions recorded by IM-MS for ternary metal-small molecule-peptide

complexes suggest that metal binding may pre-organize A $\beta$  for EGCG interactions and thus facilitate binding. An excellent recent review summarizes IM-MS experiments in the context of amyloidogenic proteins, covering many cases where the technology has been used to study and evaluate small molecule interactions in such systems<sup>179</sup>.

## 1.7 Summary

While measurements of protein:inhibitor binding strengths are extremely important to the drug discovery pipeline, the effect of the inhibitor on protein structure is becoming increasingly vital in the first steps of drug discovery. The discovery of Gleevec, which binds specifically to the inactive conformation of Abl, launched a search for small molecule kinase inhibitors that can stabilize the inactive form of the kinase. Later, inhibitors were found that bound to an allosteric site remote from the ATP-binding site of Abl. These allosteric inhibitors were discovered serendipitously after discrepancies were found between *in vivo* and *in vitro* assays. Following a similar trend, Src

IM-MS promises to ameliorate the challenges in the analysis of difficult protein targets faced by traditional biophysical characteristic techniques, such as rapidly determining small conformational changes in a flexible protein system. While IM-MS is limited by the sensitivity of the CCS measurement, a combination of CIU and CID and elucidate small structural details to create a high-throughput screening paradigm that can rapidly differentiate between modes of inhibitor binding, as well as distinguishing between allosteric and competitive inhibitors.

Currently, the IM-MS protein–ligand screening methods discussed here are primarily limited by: the detection limits associated with ESI-MS and the software solutions currently available to rapidly extract scored responses and computational models of protein structure from IM-MS data. Significant development challenges also surround the throughput of IM-MS screening technology, which is currently limited to hundreds of samples-per-day, primarily through deficiencies in rapid sample introduction methods and post-analysis software tools. Despite these challenges, the advantages of IM-MS based screens, which are capable of detecting minor conformational changes in protein targets within mixtures at relatively low concentrations without the need for chemical labeling, are enabling a growing number of studies involving protein–ligand complexes of pharmaceutical interest. This trend is likely to continue in the

future, and lead to both the support and the acceleration of continued efforts in the pharmaceutical sciences.

In chapter 2, a suite of Python modules, called CIUSuite, is developed in the analysis of CIU data. As CIU datasets are dense, the adaptation of CIU as a HTS has been limited by the lack of software that can rapidly parse the data and provide parameters for a screen. **(Published, Jessica N. Rabuck-Gibbons, Joseph D. Eschweiler, Yuwei Tian, and Brandon T. Ruotolo (2015) CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions, *Analytical Chemistry* 87(22), 11516-11522.)**

In chapter 3, we determine a set of CIU parameters that distinguish between type I and type II kinase inhibitors using the Abl kinase domain as a model system. CIU and CID data are used to create a unique set of CIU fingerprints, where the CIU of Abl:inhibitor complexes is unable to differentiate between the inhibitor types. **(Published, Jessica N. Rabuck, Suk-Joon Hyung, Kristin S. Ko, Christel C. Fox, Matthew B. Soellner, and Brandon T. Ruotolo (2013) Activation State-Selective Kinase Inhibitor Assay Based on Ion Mobility-Mass Spectrometry, *Analytical Chemistry*, 85(15), 6997-7002.)**

In chapter 4, we expand our data to include GNF-2, GNF-5, and myristic acid. We find that CIU and CID are capable of distinguishing between ATP-competitive and allosteric kinase inhibitors using Abl as a model system. **(Manuscript in Progress.)**

In chapter 5, this work is extended to a new target: the Src protein tyrosine kinase. This screen uses the three domain construct of Src to distinguish between type I and type II kinase inhibitors, and the parameters from this screen are used to identify weak type II inhibitors and three previously unknown kinase inhibitors developed in the Soellner lab. **(Manuscript in Progress.)**

In chapter 6, we apply our CIU technology to the KIX transcriptional activator system. We determine a unique CIU fingerprint for MLL and pKID-like binders. We also use mass spectrometry data to probe the allosteric interactions between the two binding pockets.

In chapter 7, native mass spectrometry techniques are introduced to undergraduates in a Biochemical Analysis laboratory setting. Following a review of the current native mass spectrometry literature, our laboratory protocols for determining the pH dependence of the



concanavalin A tetramer and the binding affinities of a ligand to hen egg white lysozyme are described. **(Manuscript in Progress.)**

Part of the content in this introduction was published as a review. **(Jessica N. Rabuck, Shuai Niu, and Brandon T. Ruotolo (2013) Ion Mobility-Mass Spectrometry of Intact Protein-Ligand Complexes for Pharmaceutical Drug Discovery and Development, *Current Opinion in Chemical Biology*, 17(5), 809-817.)**

## 1.8 References

- [1] Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) How many drug targets are there?, *Nature reviews Drug discovery* 5, 993-996.
- [2] Vitkup, D., Melamud, E., Moulton, J., and Sander, C. (2001) Completeness in structural genomics, *Nature Structural & Molecular Biology* 8, 559-566.
- [3] Makley, L. N., and Gestwicki, J. E. (2013) Expanding the Number of 'Druggable' Targets: Non-Enzymes and Protein-Protein Interactions, *Chemical biology & drug design* 81, 22-32.
- [4] Cohen, P. (2002) Protein kinases—the major drug targets of the twenty-first century?, *Nature reviews Drug discovery* 1, 309-315.
- [5] Druker, B. J., Talpaz, M., Resta, D. J., Peng, B., Buchdunger, E., Ford, J. M., Lydon, N. B., Kantarjian, H., Capdeville, R., and Ohno-Jones, S. (2001) Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia, *N Engl J Med* 2001, 1031-1037.
- [6] Nagar, B., Bornmann, W. G., Pellicena, P., Schindler, T., Veach, D. R., Miller, W. T., Clarkson, B., and Kuriyan, J. (2002) Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571), *Cancer research* 62, 4236-4243.
- [7] Baselga, J. (2006) Targeting tyrosine kinases in cancer: the second wave, *Science* 312, 1175-1178.
- [8] Al-Obeidi, F. A., Wu, J. J., and Lam, K. S. (1998) Protein tyrosine kinases: structure, substrate specificity, and drug discovery, *Peptide Science* 47, 197-223.
- [9] Olsen, J. V., Blagoev, B., Gnädig, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks, *Cell* 127, 635-648.
- [10] Brown, M. T., and Cooper, J. A. (1996) Regulation, substrates and functions of src, *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* 1287, 121-149.
- [11] Roskoski, R. (2005) Src kinase regulation by phosphorylation and dephosphorylation, *Biochemical and biophysical research communications* 331, 1-14.
- [12] Roskoski, R. (2004) Src protein-tyrosine kinase structure and regulation, *Biochemical and biophysical research communications* 324, 1155-1164.
- [13] Xu, W., Doshi, A., Lei, M., Eck, M. J., and Harrison, S. C. (1999) Crystal structures of c-Src reveal features of its autoinhibitory mechanism, *Molecular cell* 3, 629-638.
- [14] Tamaoki, T., Nomoto, H., Takahashi, I., Kato, Y., Morimoto, M., and Tomita, F. (1986) Staurosporine, a potent inhibitor of phospholipid  $Ca^{++}$  dependent protein kinase, *Biochemical and biophysical research communications* 135, 397-402.

- [15] Druker, B. J. (2002) STI571 (Gleevec™) as a paradigm for cancer therapy, *Trends in molecular medicine* 8, S14-S18.
- [16] Ferrantini, M., and Belardelli, F. (2000) Gene therapy of cancer with interferon: lessons from tumor models and perspectives for clinical applications, In *Seminars in cancer biology*, pp 145-157, Elsevier.
- [17] Shawver, L. K., Slamon, D., and Ullrich, A. (2002) Smart drugs: tyrosine kinase inhibitors in cancer therapy, *Cancer cell* 1, 117-123.
- [18] Mol, C., Fabbro, D., and Hosfield, D. (2004) Structural insights into the conformational selectivity of STI-571 and related kinase inhibitors, *Current opinion in drug discovery & development* 7, 639-648.
- [19] Krishnamurty, R., Brigham, J. L., Leonard, S. E., Ranjitkar, P., Larson, E. T., Dale, E. J., Merritt, E. A., and Maly, D. J. (2013) Active site profiling reveals coupling between domains in SRC-family kinases, *Nature chemical biology* 9, 43-50.
- [20] Cowan-Jacob, S. W., Fendrich, G., Manley, P. W., Jahnke, W., Fabbro, D., Liebetanz, J., and Meyer, T. (2005) The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation, *Structure* 13, 861-871.
- [21] Leonard, S. E., Register, A., Krishnamurty, R., Brighty, G. J., and Maly, D. J. (2014) Divergent modulation of Src-family kinase regulatory interactions with ATP-competitive inhibitors, *ACS chemical biology* 9, 1894-1905.
- [22] Skora, L., Mestan, J., Fabbro, D., Jahnke, W., and Grzesiek, S. (2013) NMR reveals the allosteric opening and closing of Abelson tyrosine kinase by ATP-site and myristoyl pocket inhibitors, *Proceedings of the National Academy of Sciences* 110, E4437-E4445.
- [23] Müller, S., Chaikuad, A., Gray, N. S., and Knapp, S. (2015) The ins and outs of selective kinase inhibitor development, *Nature chemical biology* 11, 818-821.
- [24] Fabbro, D., Manley, P. W., Jahnke, W., Liebetanz, J., Szyttenholm, A., Fendrich, G., Strauss, A., Zhang, J., Gray, N. S., and Adrian, F. (2010) Inhibitors of the Abl kinase directed at either the ATP-or myristate-binding site, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1804, 454-462.
- [25] Liu, Y., and Gray, N. S. (2006) Rational design of inhibitors that bind to inactive kinase conformations, *Nature chemical biology* 2, 358-364.
- [26] Iacob, R. E., Zhang, J., Gray, N. S., and Engen, J. R. (2011) Allosteric interactions between the myristate-and ATP-site of the Abl kinase, *PLoS One* 6, e15929.
- [27] Zhang, J., Adrián, F. J., Jahnke, W., Cowan-Jacob, S. W., Li, A. G., Iacob, R. E., Sim, T., Powers, J., Dierks, C., and Sun, F. (2010) Targeting Bcr–Abl by combining allosteric with ATP-binding-site inhibitors, *Nature* 463, 501-506.
- [28] Weisberg, E., Manley, P., Mestan, J., Cowan-Jacob, S., Ray, A., and Griffin, J. (2006) AMN107 (nilotinib): a novel and selective inhibitor of BCR-ABL, *British Journal of Cancer* 94, 1765-1769.
- [29] Zhao, X., Ghaffari, S., Lodish, H., Malashkevich, V. N., and Kim, P. S. (2002) Structure of the Bcr-Abl oncoprotein oligomerization domain, *Nature Structural & Molecular Biology* 9, 117-120.
- [30] Lawrence, D. S., and Niu, J. (1998) Protein kinase inhibitors: the tyrosine-specific protein kinases, *Pharmacology & therapeutics* 77, 81-114.
- [31] Khateb, M., Ruimi, N., Khamisie, H., Najajreh, Y., Mian, A., Metodiev, A., Ruthardt, M., and Mahajna, J. (2012) Overcoming Bcr-Abl T315I mutation by combination of GNF-2 and ATP competitors in an Abl-independent mechanism, *BMC cancer* 12, 563.

- [32] Gu, J. J., Ryu, J. R., and Pendergast, A. M. (2009) Abl tyrosine kinases in T-cell signaling, *Immunological reviews* 228, 170-183.
- [33] Plattner, R., Kadlec, L., DeMali, K. A., Kazlauskas, A., and Pendergast, A. M. (1999) c-Abl is activated by growth factors and Src family kinases and has a role in the cellular response to PDGF, *Genes & development* 13, 2400-2411.
- [34] Smith-Pearson, P. S., Greuber, E. K., Yogalingam, G., and Pendergast, A. M. (2010) Abl kinases are required for invadopodia formation and chemokine-induced invasion, *Journal of Biological Chemistry* 285, 40201-40211.
- [35] Gu, J. J., Lavau, C. P., Pugacheva, E., Soderblom, E. J., Moseley, M. A., and Pendergast, A. M. (2012) Abl Family Kinases Modulate T Cell-Mediated Inflammation and Chemokine-Induced Migration Through a HEF1-Rap1 Signaling Module, *Science signaling* 5, ra51.
- [36] Gonfloni, S. (2010) DNA damage stress response in germ cells: role of c-Abl and clinical implications, *Oncogene* 29, 6193-6202.
- [37] Cao, C., Leng, Y., and Kufe, D. (2003) Catalase activity is regulated by c-Abl and Arg in the oxidative stress response, *Journal of Biological Chemistry* 278, 29667-29675.
- [38] Sun, X., Wu, F., Datta, R., Kharbanda, S., and Kufe, D. (2000) Interaction between protein kinase C  $\delta$  and the c-Abl tyrosine kinase in the cellular response to oxidative stress, *Journal of Biological Chemistry* 275, 7470-7473.
- [39] Sanguinetti, A. R., and Mastick, C. C. (2003) c-Abl is required for oxidative stress-induced phosphorylation of caveolin-1 on tyrosine 14, *Cellular signalling* 15, 289-298.
- [40] Aplin, A. E., and Juliano, R. L. (2001) Regulation of nucleocytoplasmic trafficking by cell adhesion receptors and the cytoskeleton, *The Journal of cell biology* 155, 187-192.
- [41] Lanier, L. M., and Gertler, F. B. (2000) From Abl to actin: Abl tyrosine kinase and associated proteins in growth cone motility, *Current opinion in neurobiology* 10, 80-87.
- [42] Burton, E. A., Plattner, R., and Pendergast, A. M. (2003) Abl tyrosine kinases are required for infection by *Shigella flexneri*, *The EMBO Journal* 22, 5471-5479.
- [43] Pielage, J. F., Powell, K. R., Kalman, D., and Engel, J. N. (2008) RNAi screen reveals an Abl kinase-dependent host cell pathway involved in *Pseudomonas aeruginosa* internalization, *PLoS Pathog* 4, e1000031.
- [44] Backert, S., Feller, S. M., and Wessler, S. (2008) Emerging roles of Abl family tyrosine kinases in microbial pathogenesis, *Trends in biochemical sciences* 33, 80-90.
- [45] Yuan, Z.-M., Huang, Y., Ishiko, T., Kharbanda, S., Weichselbaum, R., and Kufe, D. (1997) Regulation of DNA damage-induced apoptosis by the c-Abl tyrosine kinase, *Proceedings of the National Academy of Sciences* 94, 1437-1440.
- [46] Cong, F., and Goff, S. P. (1999) c-Abl-induced apoptosis, but not cell cycle arrest, requires mitogen-activated protein kinase kinase 6 activation, *Proceedings of the National Academy of Sciences* 96, 13819-13824.
- [47] Yoshida, K., and Miki, Y. (2005) Enabling death by the Abl tyrosine kinase: mechanisms for nuclear shuttling of c-Abl in response to DNA damage, *Cell Cycle* 4, 777-779.
- [48] Madhani, H. D., Bohr, V. A., and Hanawalt, P. C. (1986) Differential DNA repair in transcriptionally active and inactive proto-oncogenes: c-abl and c-mos, *Cell* 45, 417-423.
- [49] Bernstein, C., Bernstein, H., Payne, C. M., and Garewal, H. (2002) DNA repair/pro-apoptotic dual-role proteins in five major DNA repair pathways: fail-safe protection against carcinogenesis, *Mutation Research/Reviews in Mutation Research* 511, 145-178.

- [50] Huang, Y., Comiskey, E. O., Dupree, R. S., Li, S., Koleske, A. J., and Burkhardt, J. K. (2008) The c-Abl tyrosine kinase regulates actin remodeling at the immune synapse, *Blood* 112, 111-119.
- [51] Van Etten, R. A., Jackson, P. K., Baltimore, D., Sanders, M. C., Matsudaira, P. T., and Janmey, P. A. (1994) The COOH terminus of the c-Abl tyrosine kinase contains distinct F- and G-actin binding domains with bundling activity, *The Journal of cell biology* 124, 325-340.
- [52] Woodring, P. J., Hunter, T., and Wang, J. Y. (2003) Regulation of F-actin-dependent processes by the Abl family of tyrosine kinases, *Journal of cell science* 116, 2613-2626.
- [53] Maiani, E., Diederich, M., and Gonfloni, S. (2011) DNA damage response: the emerging role of c-Abl as a regulatory switch?, *Biochemical pharmacology* 82, 1269-1276.
- [54] Wang, J. Y. (2000) Regulation of cell death by the Abl tyrosine kinase, *Oncogene* 19.
- [55] Cottom, J., Hofmann, G., Siegfried, B., Yang, J., Zhang, H., Yi, T., Ho, T. F., Quinn, C., Wang, D.-Y., and Johanson, K. (2011) Assay development and high-throughput screening of small molecular c-Abl kinase activators, *Journal of biomolecular screening* 16, 53-64.
- [56] Gonfloni, S., Maiani, E., Di Bartolomeo, C., Diederich, M., and Cesareni, G. (2012) Oxidative stress, DNA damage, and c-Abl signaling: at the crossroad in neurodegenerative diseases?, *International journal of cell biology* 2012.
- [57] Sirvent, A., Benistant, C., and Roche, S. (2008) Cytoplasmic signalling by the c-Abl tyrosine kinase in normal and cancer cells, *Biology of the Cell* 100, 617-631.
- [58] Hantschel, O., Nagar, B., Guettler, S., Kretzschmar, J., Dorey, K., Kuriyan, J., and Superti-Furga, G. (2003) A myristoyl/phosphotyrosine switch regulates c-Abl, *Cell* 112, 845-857.
- [59] Wetzler, M., Talpaz, M., Van Etten, R., Hirsh-Ginsberg, C., Beran, M., and Kurzrock, R. (1993) Subcellular localization of Bcr, Abl, and Bcr-Abl proteins in normal and leukemic cells and correlation of expression with myeloid differentiation, *Journal of Clinical Investigation* 92, 1925.
- [60] Vigneri, P., and Wang, J. Y. (2001) Induction of apoptosis in chronic myelogenous leukemia cells through nuclear entrapment of BCR-ABL tyrosine kinase, *Nature medicine* 7, 228-234.
- [61] Levine, A. J. (1997) p53, the cellular gatekeeper for growth and division, *cell* 88, 323-331.
- [62] Goga, A., Liu, X., Hambuch, T. M., Senechal, K., Major, E., Berk, A. J., Witte, O. N., and Sawyers, C. L. (1995) p53 dependent growth suppression by the c-Abl nuclear tyrosine kinase, *Oncogene* 11, 791-799.
- [63] Schlatterer, S. D., Acker, C. M., and Davies, P. (2011) c-Abl in neurodegenerative disease, *Journal of Molecular Neuroscience* 45, 445-452.
- [64] Sterne, G. R., Kim, J. H., and Ye, B. (2015) Dysregulated Dscam levels act through Abelson tyrosine kinase to enlarge presynaptic arbors, *Elife* 4, e05196.
- [65] Jing, Z., Caltagarone, J., and Bowser, R. (2009) Altered subcellular distribution of c-Abl in Alzheimer's disease, *Journal of Alzheimer's Disease* 17, 409-422.
- [66] Derkinderen, P., Scales, T. M., Hanger, D. P., Leung, K.-Y., Byers, H. L., Ward, M. A., Lenz, C., Price, C., Bird, I. N., and Perera, T. (2005) Tyrosine 394 is phosphorylated in Alzheimer's paired helical filament tau and in fetal tau with c-Abl as the candidate tyrosine kinase, *The Journal of neuroscience* 25, 6584-6593.
- [67] Imam, S. Z., Zhou, Q., Yamamoto, A., Valente, A. J., Ali, S. F., Bains, M., Roberts, J. L., Kahle, P. J., Clark, R. A., and Li, S. (2011) Novel regulation of parkin function through

- c-Abl-mediated tyrosine phosphorylation: implications for Parkinson's disease, *The Journal of Neuroscience* 31, 157-163.
- [68] Mahul-Mellier, A.-L., Fauvet, B., Gysbers, A., Dikiy, I., Oueslati, A., Georgeon, S., Lamontanara, A. J., Bisquertt, A., Eliezer, D., and Masliah, E. (2014) c-Abl phosphorylates  $\alpha$ -synuclein and regulates its degradation: implication for  $\alpha$ -synuclein clearance and contribution to the pathogenesis of Parkinson's disease, *Human molecular genetics* 23, 2858-2879.
  - [69] Barilà, D., Rufini, A., Condò, I., Ventura, N., Dorey, K., Superti-Furga, G., and Testi, R. (2003) Caspase-dependent cleavage of c-Abl contributes to apoptosis, *Molecular and cellular biology* 23, 2790-2799.
  - [70] Schlatterer, S. D., Tremblay, M. A., Acker, C. M., and Davies, P. (2011) Neuronal c-Abl overexpression leads to neuronal loss and neuroinflammation in the mouse forebrain, *Journal of Alzheimer's Disease* 25, 119-133.
  - [71] Gaki, G. S., and Papavassiliou, A. G. (2014) Oxidative stress-induced signaling pathways implicated in the pathogenesis of Parkinson's disease, *Neuromolecular medicine* 16, 217-230.
  - [72] Cancino, G. I., Toledo, E. M., Leal, N. R., Hernandez, D. E., Yévenes, L. F., Inestrosa, N. C., and Alvarez, A. R. (2008) STI571 prevents apoptosis, tau phosphorylation and behavioural impairments induced by Alzheimer's  $\beta$ -amyloid deposits, *Brain* 131, 2425-2442.
  - [73] Lonskaya, I., Hebron, M. L., Desforges, N. M., Franjie, A., and Moussa, C. E. H. (2013) Tyrosine kinase inhibition increases functional parkin-Bec1 interaction and enhances amyloid clearance and cognitive performance, *EMBO molecular medicine* 5, 1247-1262.
  - [74] Cursi, S., Rufini, A., Stagni, V., Condò, I., Matafora, V., Bachi, A., Bonifazi, A. P., Coppola, L., Superti-Furga, G., and Testi, R. (2006) Src kinase phosphorylates Caspase-8 on Tyr380: a novel mechanism of apoptosis suppression, *The EMBO journal* 25, 1895-1905.
  - [75] Cooper, J. A., Gould, K. L., Cartwright, C. A., and Hunter, T. (1986) Tyr527 is phosphorylated in pp60c-src: implications for regulation, *Science* 231, 1431-1434.
  - [76] Read, R. D., Bach, E. A., and Cagan, R. L. (2004) Drosophila C-terminal Src kinase negatively regulates organ growth and cell proliferation through inhibition of the Src, Jun N-terminal kinase, and STAT pathways, *Molecular and Cellular Biology* 24, 6676-6689.
  - [77] Vuori, K., Hirai, H., Aizawa, S., and Ruoslahti, E. (1996) Introduction of p130cas signaling complex formation upon integrin-mediated cell adhesion: a role for Src family kinases, *Molecular and Cellular Biology* 16, 2606-2613.
  - [78] Sanjay, A., Houghton, A., Neff, L., DiDomenico, E., Bardelay, C., Antoine, E., Levy, J., Gailit, J., Bowtell, D., and Horne, W. C. (2001) Cbl associates with Pyk2 and Src to regulate Src kinase activity,  $\alpha\beta3$  integrin-mediated signaling, cell adhesion, and osteoclast motility, *The Journal of cell biology* 152, 181-196.
  - [79] Bourguignon, L. Y., Zhu, H., Shao, L., and Chen, Y.-W. (2001) CD44 interaction with c-Src kinase promotes cortactin-mediated cytoskeleton function and hyaluronic acid-dependent ovarian tumor cell migration, *Journal of Biological Chemistry* 276, 7327-7336.
  - [80] Carragher, N., Walker, S., Carragher, L. S., Harris, F., Sawyer, T., Brunton, V., Ozanne, B., and Frame, M. (2006) Calpain 2 and Src dependence distinguishes mesenchymal and amoeboid modes of tumour cell invasion: a link to integrin function, *Oncogene* 25, 5726-5740.

- [81] Irby, R. B., and Yeatman, T. J. (2000) Role of Src expression and activation in human cancer, *Oncogene* 19.
- [82] Lee, G., Newman, S. T., Gard, D. L., Band, H., and Panchamoorthy, G. (1998) Tau interacts with src-family non-receptor tyrosine kinases, *Journal of cell science* 111, 3167-3177.
- [83] Sanna, P. P., Berton, F., Cammalleri, M., Tallent, M. K., Siggins, G. R., Bloom, F. E., and Francesconi, W. (2000) A role for Src kinase in spontaneous epileptiform activity in the CA3 region of the hippocampus, *Proceedings of the National Academy of Sciences* 97, 8653-8657.
- [84] Tribble, R. P., Emert-Sedlak, L., and Smithgall, T. E. (2006) HIV-1 Nef selectively activates Src family kinases Hck, Lyn, and c-Src through direct SH3 domain interaction, *Journal of Biological Chemistry* 281, 27029-27038.
- [85] Song, K., Sargiacomo, M., Galbiati, F., Parenti, M., and Lisanti, M. (1997) Targeting of a G alpha subunit (Gi1 alpha) and c-Src tyrosine kinase to caveolae membranes: clarifying the role of N-myristoylation, *Cellular and molecular biology (Noisy-le-Grand, France)* 43, 293-303.
- [86] Patwardhan, P., and Resh, M. D. (2010) Myristoylation and membrane binding regulate c-Src stability and kinase activity, *Molecular and cellular biology* 30, 4094-4107.
- [87] Liu, B., Li, S., and Hu, J. (2004) Technological advances in high-throughput screening, *American Journal of Pharmacogenomics* 4, 263-276.
- [88] Sundberg, S. A. (2000) High-throughput and ultra-high-throughput screening: solution- and cell-based approaches, *Current opinion in biotechnology* 11, 47-53.
- [89] Wu, B., Barile, E., K De, S., Wei, J., Purves, A., and Pellicchia, M. (2015) High-throughput screening by Nuclear Magnetic Resonance (HTS by NMR) for the identification of PPIs antagonists, *Current topics in medicinal chemistry* 15, 2032-2042.
- [90] Blundell, T. L., and Patel, S. (2004) High-throughput X-ray crystallography for drug discovery, *Current opinion in pharmacology* 4, 490-496.
- [91] Lundstrom, K. (2016) New winds in GPCR-based drug discovery, *Future medicinal chemistry* 8, 605-608.
- [92] Boute, N., Jockers, R., and Issad, T. (2002) The use of resonance energy transfer in high-throughput screening: BRET versus FRET, *Trends in pharmacological sciences* 23, 351-354.
- [93] Hall, M. D., Yasgar, A., Peryea, T., Braisted, J. C., Jadhav, A., Simeonov, A., and Coussens, N. P. (2016) Fluorescence polarization assays in high-throughput screening and drug discovery: a review, *Methods and Applications in Fluorescence* 4, 022001.
- [94] Degorce, F., Card, A., Soh, S., Trinquet, E., Knapik, G. P., and Xie, B. (2009) HTRF: a technology tailored for drug discovery—a review of theoretical aspects and recent applications, *Current Chemical Genomics and Translational Medicine* 3.
- [95] Eglen, R. M., Reisine, T., Roby, P., Rouleau, N., Illy, C., Bossé, R., and Bielefeld, M. (2008) The use of AlphaScreen technology in HTS: current status, *Current chemical genomics* 1, 2.
- [96] Nienaber, V. L., Richardson, P. L., Klighofer, V., Bouska, J. J., Giranda, V. L., and Greer, J. (2000) Discovering novel ligands for macromolecules using X-ray crystallographic screening, *Nature biotechnology* 18, 1105-1108.
- [97] Zhang, J., Yang, P. L., and Gray, N. S. (2009) Targeting cancer with small molecule kinase inhibitors, *Nature Reviews Cancer* 9, 28-39.

- [98] Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., and Holt, R. A. (2001) The sequence of the human genome, *science* 291, 1304-1351.
- [99] Karaman, M. W., Herrgard, S., Treiber, D. K., Gallant, P., Atteridge, C. E., Campbell, B. T., Chan, K. W., Ciceri, P., Davis, M. I., and Edeen, P. T. (2008) A quantitative analysis of kinase inhibitor selectivity, *Nature biotechnology* 26, 127-132.
- [100] Davis, M. I., Hunt, J. P., Herrgard, S., Ciceri, P., Wodicka, L. M., Pallares, G., Hocker, M., Treiber, D. K., and Zarrinkar, P. P. (2011) Comprehensive analysis of kinase inhibitor selectivity, *Nature biotechnology* 29, 1046-1051.
- [101] Simard, J. R., Getlik, M. u., Grütter, C., Pawar, V., Wulfert, S., Rabiller, M., and Rauh, D. (2009) Development of a fluorescent-tagged kinase assay system for the detection and characterization of allosteric kinase inhibitors, *Journal of the American Chemical Society* 131, 13286-13296.
- [102] Wang, Q., Cahill, S. M., Blumenstein, M., and Lawrence, D. S. (2006) Self-reporting fluorescent substrates of protein tyrosine kinases, *Journal of the American Chemical Society* 128, 1808-1809.
- [103] Annis, D. A., Nickbarg, E., Yang, X., Ziebell, M. R., and Whitehurst, C. E. (2007) Affinity selection-mass spectrometry screening techniques for small molecule drug discovery, *Current opinion in chemical biology* 11, 518-526.
- [104] Comess, K. M., Trumbull, J. D., Park, C., Chen, Z., Judge, R. A., Voorbach, M. J., Coen, M., Gao, L., Tang, H., and Kovar, P. (2006) Kinase drug discovery by affinity selection/mass spectrometry (ASMS): application to DNA damage checkpoint kinase Chk1, *Journal of biomolecular screening* 11, 755-764.
- [105] Leavitt, S., and Freire, E. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry, *Current opinion in structural biology* 11, 560-566.
- [106] Young, P. R., McLaughlin, M. M., Kumar, S., Kassis, S., Doyle, M. L., McNulty, D., Gallagher, T. F., Fisher, S., McDonnell, P. C., and Carr, S. A. (1997) Pyridinyl imidazole inhibitors of p38 mitogen-activated protein kinase bind in the ATP site, *Journal of Biological Chemistry* 272, 12116-12121.
- [107] Fedorov, O., Niesen, F. H., and Knapp, S. (2012) Kinase inhibitor selectivity profiling using differential scanning fluorimetry, *Kinase Inhibitors: Methods and Protocols*, 109-118.
- [108] Kroe, R. R., Regan, J., Proto, A., Peet, G. W., Roy, T., Landro, L. D., Fuschetto, N. G., Pargellis, C. A., and Ingraham, R. H. (2003) Thermal denaturation: A method to rank slow binding, high-affinity P38 $\alpha$  MAP kinase inhibitors, *Journal of medicinal chemistry* 46, 4669-4675.
- [109] Szucova, L., Trávníček, Z., Zatloukal, M., and Popa, I. (2006) Novel platinum (II) and palladium (II) complexes with cyclin-dependent kinase inhibitors: synthesis, characterization and antitumour activity, *Bioorganic & medicinal chemistry* 14, 479-491.
- [110] Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, *Nature protocols* 2, 2212-2221.
- [111] Fischer, J. J., Dalhoff, C., Schrey, A. K., neé Baessler, O. Y. G., Michaelis, S., Andrich, K., Glinski, M., Kroll, F., Sefkow, M., and Dreger, M. (2011) Dasatinib, imatinib and staurosporine capture compounds—Complementary tools for the profiling of kinases by Capture Compound Mass Spectrometry (CCMS), *Journal of proteomics* 75, 160-168.

- [112] Hatzivassiliou, G., Song, K., Yen, I., Brandhuber, B. J., Anderson, D. J., Alvarado, R., Ludlam, M. J., Stokoe, D., Gloor, S. L., and Vigers, G. (2010) RAF inhibitors prime wild-type RAF to activate the MAPK pathway and enhance growth, *Nature* 464, 431-435.
- [113] Poulikakos, P. I., Zhang, C., Bollag, G., Shokat, K. M., and Rosen, N. (2010) RAF inhibitors transactivate RAF dimers and ERK signalling in cells with wild-type BRAF, *Nature* 464, 427-430.
- [114] Chan, T. O., Zhang, J., Rodeck, U., Pascal, J. M., Armen, R. S., Spring, M., Dumitru, C. D., Myers, V., Li, X., and Cheung, J. Y. (2011) Resistance of Akt kinases to dephosphorylation through ATP-dependent conformational plasticity, *Proceedings of the National Academy of Sciences* 108, E1120-E1127.
- [115] Andraos, R., Qian, Z., Bonenfant, D., Rubert, J., Vangrevelinghe, E., Scheufler, C., Marque, F., Régnier, C. H., De Pover, A., and Ryckelynck, H. (2012) Modulation of activation-loop phosphorylation by JAK inhibitors is binding mode dependent, *Cancer discovery* 2, 512-523.
- [116] Wang, L., Perera, B. G. K., Hari, S. B., Bhattacharai, B., Backes, B. J., Seeliger, M. A., Schürer, S. C., Oakes, S. A., Papa, F. R., and Maly, D. J. (2012) Divergent allosteric control of the IRE1 $\alpha$  endoribonuclease using kinase inhibitors, *Nature chemical biology* 8, 982-989.
- [117] Koppikar, P., Bhagwat, N., Kilpivaara, O., Manshouri, T., Adli, M., Hricik, T., Liu, F., Saunders, L. M., Mullally, A., and Abdel-Wahab, O. (2012) Heterodimeric JAK-STAT activation as a mechanism of persistence to JAK2 inhibitor therapy, *Nature* 489, 155-159.
- [118] Haura, E. B., Turkson, J., and Jove, R. (2005) Mechanisms of disease: Insights into the emerging role of signal transducers and activators of transcription in cancer, *Nature clinical practice Oncology* 2, 315-324.
- [119] Lee, L. W., and Mapp, A. K. (2010) Transcriptional switches: chemical approaches to gene regulation, *Journal of Biological Chemistry* 285, 11033-11038.
- [120] Mapp, A. K. (2003) Regulating transcription: a chemical perspective, *Organic & biomolecular chemistry* 1, 2217-2220.
- [121] Radhakrishnan, I., Pérez-Alvarado, G. C., Parker, D., Dyson, H. J., Montminy, M. R., and Wright, P. E. (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator: coactivator interactions, *Cell* 91, 741-752.
- [122] Mayr, B., and Montminy, M. (2001) Transcriptional regulation by the phosphorylation-dependent factor CREB, *Nature reviews Molecular cell biology* 2, 599-609.
- [123] Radhakrishnan, I., Pérez-Alvarado, G. C., Dyson, H. J., and Wright, P. E. (1998) Conformational preferences in the Ser 133-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB, *FEBS letters* 430, 317-322.
- [124] Brüscheiler, S., Konrat, R., and Tollinger, M. (2013) Allosteric communication in the KIX domain proceeds through dynamic repacking of the hydrophobic core, *ACS chemical biology* 8, 1600-1610.
- [125] Rowe, S. P., and Mapp, A. K. (2008) Assessing the permissiveness of transcriptional activator binding sites, *Biopolymers* 89, 578-581.



- [126] Gee, C. T., Koleski, E. J., and Pomerantz, W. C. (2015) Fragment Screening and Druggability Assessment for the CBP/p300 KIX Domain through Protein-Observed 19F NMR Spectroscopy, *Angewandte Chemie International Edition* 54, 3735-3739.
- [127] Palazzesi, F., Barducci, A., Tollinger, M., and Parrinello, M. (2013) The allosteric communication pathways in KIX domain of CBP, *Proceedings of the National Academy of Sciences* 110, 14237-14242.
- [128] Majmudar, C. Y., Højfeldt, J. W., Arevang, C. J., Pomerantz, W. C., Gagnon, J. K., Schultz, P. J., Cesa, L. C., Doss, C. H., Rowe, S. P., and Vásquez, V. (2012) Sekikaic acid and lobaric acid target a dynamic interface of the coactivator CBP/p300, *Angewandte Chemie International Edition* 51, 11258-11262.
- [129] Pringle, S. D., Giles, K., Wildgoose, J. L., Williams, J. P., Slade, S. E., Thalassinou, K., Bateman, R. H., Bowers, M. T., and Scrivens, J. H. (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument, *International Journal of Mass Spectrometry* 261, 1-12.
- [130] Giles, K., Williams, J. P., and Campuzano, I. (2011) Enhancements in travelling wave ion mobility resolution, *Rapid Communications in Mass Spectrometry* 25, 1559-1566.
- [131] Ruotolo, B. T., and Robinson, C. V. (2006) Aspects of native proteins are retained in vacuum, *Current opinion in chemical biology* 10, 402-408.
- [132] Whitehouse, C. M., Dreyer, R., Yamashita, M., and Fenn, J. (1989) Electrospray ionization for mass-spectrometry of large biomolecules, *Science* 246, 64-71.
- [133] De La Mora, J. F. (2000) Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism, *Analytica chimica acta* 406, 93-104.
- [134] Dole, M., Mack, L., Hines, R., Mobley, R., Ferguson, L., and Alice, M. d. (1968) Molecular beams of macroions, *The Journal of Chemical Physics* 49, 2240-2249.
- [135] Thomson, B., and Iribarne, J. (1979) Field induced ion evaporation from liquid surfaces at atmospheric pressure, *The Journal of Chemical Physics* 71, 4451-4463.
- [136] Iribarne, J., and Thomson, B. (1976) On the evaporation of small ions from charged droplets, *The Journal of Chemical Physics* 64, 2287-2294.
- [137] Labowsky, M., Fenn, J., and de la Mora, J. F. (2000) A continuum model for ion evaporation from a drop: effect of curvature and charge on ion solvation energy, *Analytica Chimica Acta* 406, 105-118.
- [138] Hogan Jr, C. J., Carroll, J. A., Rohrs, H. W., Biswas, P., and Gross, M. L. (2008) Charge carrier field emission determines the number of charges on native state proteins in electrospray ionization, *Journal of the American Chemical Society* 130, 6926-6927.
- [139] Hogan Jr, C. J., Carroll, J. A., Rohrs, H. W., Biswas, P., and Gross, M. L. (2008) Combined charged residue-field emission model of macromolecular electrospray ionization, *Analytical chemistry* 81, 369-377.
- [140] Konermann, L., Ahadi, E., Rodriguez, A. D., and Vahidi, S. (2012) Unraveling the mechanism of electrospray ionization, *Analytical chemistry* 85, 2-9.
- [141] Loo, R. R. O., Lakshmanan, R., and Loo, J. A. (2014) What protein charging (and supercharging) reveal about the mechanism of electrospray ionization, *Journal of The American Society for Mass Spectrometry* 25, 1675-1693.
- [142] Katta, V., and Chait, B. T. (1991) Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry, *Journal of the American Chemical Society* 113, 8534-8535.

- [143] Ganem, B., Li, Y. T., and Henion, J. D. (1991) Detection of noncovalent receptor-ligand complexes by mass spectrometry, *Journal of the American Chemical Society* 113, 6294-6296.
- [144] Loo, J. A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry, *Mass Spectrometry Reviews* 16, 1-23.
- [145] Hernández, H., and Robinson, C. V. (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry, *Nature protocols* 2, 715-726.
- [146] Hiroki, S., Abe, T., and Murakami, Y. (1991) Development of a quadrupole mass spectrometer using the second stable zone in Mathieu's stability diagram, *Review of scientific instruments* 62, 2121-2124.
- [147] Sobott, F., Hernández, H., McCammon, M. G., Tito, M. A., and Robinson, C. V. (2002) A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies, *Analytical chemistry* 74, 1402-1407.
- [148] Cornish, T. J., and Cotter, R. J. (1993) A curved-field reflectron for improved energy focusing of product ions in time-of-flight mass spectrometry, *Rapid communications in mass spectrometry* 7, 1037-1040.
- [149] Wollnik, H., Grüner, U., and Li, G. (1992) Time-of-flight mass spectrometers, In *Mass Spectrometry in the Biological Sciences: A Tutorial*, pp 117-131, Springer.
- [150] Langevin, P. (1903) Ionization of gases, *Ann. Chim. Phys* 28, 433.
- [151] Bradbury, N. E. (1931) The Mobility of Aged Ions in Air in Relation to the Nature of Gaseous Ions, *Physical Review* 37, 1311.
- [152] McDaniel, E., Martin, D., and Barnes, W. (1962) Drift tube-mass spectrometer for studies of low-energy ion-molecule reactions, *Review of Scientific Instruments* 33, 2-7.
- [153] von Helden, G., Wyttenbach, T., and Bowers, M. T. (1995) Conformation of macromolecules in the gas phase: use of matrix-assisted laser desorption methods in ion chromatography, *Science* 267, 1483.
- [154] Uetrecht, C., Rose, R. J., van Duijn, E., Lorenzen, K., and Heck, A. J. (2010) Ion mobility mass spectrometry of proteins and protein assemblies, *Chemical Society Reviews* 39, 1633-1655.
- [155] Giles, K., Pringle, S. D., Worthington, K. R., Little, D., Wildgoose, J. L., and Bateman, R. H. (2004) Applications of a travelling wave-based radio-frequency-only stacked ring ion guide, *Rapid Communications in Mass Spectrometry* 18, 2401-2414.
- [156] Bush, M. F., Campuzano, I. D., and Robinson, C. V. (2012) Ion mobility mass spectrometry of peptide ions: effects of drift gas and calibration strategies, *Analytical chemistry* 84, 7124-7130.
- [157] Fenn, L. S., Kliman, M., Mahsut, A., Zhao, S. R., and McLean, J. A. (2009) Characterizing ion mobility-mass spectrometry conformation space for the analysis of complex biological samples, *Analytical and bioanalytical chemistry* 394, 235-244.
- [158] Bush, M. F., Hall, Z., Giles, K., Hoyes, J., Robinson, C. V., and Ruotolo, B. T. (2010) Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology, *Analytical Chemistry* 82, 9557-9565.
- [159] Campuzano, I., Bush, M. F., Robinson, C. V., Beaumont, C., Richardson, K., Kim, H., and Kim, H. I. (2011) Structural characterization of drug-like compounds by ion mobility mass spectrometry: comparison of theoretical and experimentally derived nitrogen collision cross sections, *Analytical chemistry* 84, 1026-1033.

- [160] Salbo, R., Bush, M. F., Naver, H., Campuzano, I., Robinson, C. V., Pettersson, I., Jørgensen, T. J., and Haselmann, K. F. (2012) Traveling-wave ion mobility mass spectrometry of protein complexes: accurate calibrated collision cross-sections of human insulin oligomers, *Rapid Communications in Mass Spectrometry* 26, 1181-1193.
- [161] Henderson, S. C., Li, J., Counterman, A. E., and Clemmer, D. E. (1999) Intrinsic size parameters for Val, Ile, Leu, Gln, Thr, Phe, and Trp residues from ion mobility measurements of polyamino acid ions, *The Journal of Physical Chemistry B* 103, 8780-8785.
- [162] Geoghegan, K. F., and Kelly, M. A. (2005) Biochemical applications of mass spectrometry in pharmaceutical drug discovery, *Mass spectrometry reviews* 24, 347-366.
- [163] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1990) Electrospray ionization—principles and practice, *Mass Spectrometry Reviews* 9, 37-70.
- [164] Scholten, A., Aye, T. T., and Heck, A. J. (2008) A multi-angular mass spectrometric view at cyclic nucleotide dependent protein kinases: In vivo characterization and structure/function relationships, *Mass spectrometry reviews* 27, 331-353.
- [165] Congreve, M., Murray, C. W., and Blundell, T. L. (2005) Keynote review: Structural biology and drug discovery, *Drug discovery today* 10, 895-907.
- [166] Wang, R., Fang, X., Lu, Y., and Wang, S. (2004) The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures, *Journal of medicinal chemistry* 47, 2977-2980.
- [167] Benesch, J. L., Ruotolo, B. T., Simmons, D. A., and Robinson, C. V. (2007) Protein complexes in the gas phase: technology for structural genomics and proteomics, *Chemical reviews* 107, 3544-3567.
- [168] Rifai, N., Gillette, M. A., and Carr, S. A. (2006) Protein biomarker discovery and validation: the long and uncertain path to clinical utility, *Nature biotechnology* 24, 971-983.
- [169] Leite, J. F., and Cascio, M. (2002) Probing the topology of the glycine receptor by chemical modification coupled to mass spectrometry, *Biochemistry* 41, 6140-6148.
- [170] Liu, F., and Goshe, M. B. (2010) Combinatorial electrostatic collision-induced dissociative chemical cross-linking reagents for probing protein surface topology, *Analytical chemistry* 82, 6215-6223.
- [171] Sinz, A. (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions, *Mass spectrometry reviews* 25, 663-682.
- [172] Green, M. K., and Lebrilla, C. B. (1997) Ion-molecule reactions as probes of gas-phase structures of peptides and proteins, *Mass spectrometry reviews* 16, 53-71.
- [173] Zhu, M. M., Chitta, R., and Gross, M. L. (2005) PLIMSTEX: a novel mass spectrometric method for the quantification of protein–ligand interactions in solution, *International Journal of Mass Spectrometry* 240, 213-220.
- [174] Zhu, M. M., Rempel, D. L., and Gross, M. L. (2004) Modeling data from titration, amide H/D exchange, and mass spectrometry to obtain protein-ligand binding constants, *Journal of the American Society for Mass Spectrometry* 15, 388-397.
- [175] Mao, Y., Woenckhaus, J., Kolafa, J., Ratner, M. A., and Jarrold, M. F. (1999) Thermal unfolding of unsolvated cytochrome c: experiment and molecular dynamics simulations, *Journal of the American Chemical Society* 121, 2712-2721.

- [176] Alber, F., Dokudovskaya, S., Veenhoff, L. M., Zhang, W., Kipper, J., Devos, D., Suprpto, A., Karni-Schmidt, O., Williams, R., and Chait, B. T. (2007) Determining the architectures of macromolecular assemblies, *Nature* 450, 683-694.
- [177] Hyung, S.-J., DeToma, A. S., Brender, J. R., Lee, S., Vivekanandan, S., Kochi, A., Choi, J.-S., Ramamoorthy, A., Ruotolo, B. T., and Lim, M. H. (2013) Insights into anti-amyloidogenic properties of the green tea extract (–)-epigallocatechin-3-gallate toward metal-associated amyloid- $\beta$  species, *Proceedings of the National Academy of Sciences* 110, 3743-3748.
- [178] Teplow, D. B., Lazo, N. D., Bitan, G., Bernstein, S., Wytttenbach, T., Bowers, M. T., Baumketner, A., Shea, J.-E., Urbanc, B., and Cruz, L. (2006) Elucidating amyloid  $\beta$ -protein folding and assembly: a multidisciplinary approach, *Accounts of chemical research* 39, 635-645.
- [179] Woods, L., Radford, S., and Ashcroft, A. (2013) Advances in ion mobility spectrometry–mass spectrometry reveal key insights into amyloid assembly, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1834, 1257-1268.
- [180] Politis, A., Park, A. Y., Hyung, S.-J., Barsky, D., Ruotolo, B. T., and Robinson, C. V. (2010) Integrating ion mobility mass spectrometry with molecular modelling to determine the architecture of multiprotein complexes, *PloS one* 5, e12080.
- [181] Taverner, T., Hernández, H., Sharon, M., Ruotolo, B. T., Matak-Vinković, D., Devos, D., Russell, R. B., and Robinson, C. V. (2008) Subunit Architecture of Intact Protein Complexes from Mass Spectrometry and Homology Modeling†, *Accounts of chemical research* 41, 617-627.
- [182] Shelimov, K. B., and Jarrold, M. F. (1997) Conformations, unfolding, and refolding of apomyoglobin in vacuum: An activation barrier for gas-phase protein folding, *Journal of the American Chemical Society* 119, 2987-2994.
- [183] Suckau, D., Shi, Y., Beu, S. C., Senko, M. W., Quinn, J. P., Wampler, F., and McLafferty, F. W. (1993) Coexisting stable conformations of gaseous protein ions, *Proceedings of the National Academy of Sciences* 90, 790-793.
- [184] Shelimov, K. B., Clemmer, D. E., Hudgins, R. R., and Jarrold, M. F. (1997) Protein structure in vacuo: Gas-phase conformations of BPTI and cytochrome c, *Journal of the American Chemical Society* 119, 2240-2248.
- [185] Benesch, J. L. (2009) Collisional activation of protein complexes: picking up the pieces, *Journal of the American Society for Mass Spectrometry* 20, 341-348.
- [186] Heck, A. J. (2008) Native mass spectrometry: a bridge between interactomics and structural biology, *Nature methods* 5, 927-933.
- [187] Ruotolo, B. T., Giles, K., Campuzano, I., Sandercock, A. M., Bateman, R. H., and Robinson, C. V. (2005) Evidence for macromolecular protein rings in the absence of bulk water, *Science* 310, 1658-1661.
- [188] Ruotolo, B. T., Hyung, S. J., Robinson, P. M., Giles, K., Bateman, R. H., and Robinson, C. V. (2007) Ion Mobility–Mass Spectrometry Reveals Long-Lived, Unfolded Intermediates in the Dissociation of Protein Complexes, *Angewandte Chemie International Edition* 46, 8001-8004.
- [189] Hopper, J. T., and Oldham, N. J. (2009) Collision induced unfolding of protein ions in the gas phase studied by ion mobility-mass spectrometry: the effect of ligand binding on conformational stability, *Journal of the American Society for Mass Spectrometry* 20, 1851-1858.

- [190] Hyung, S.-J., Robinson, C. V., and Ruotolo, B. T. (2009) Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes, *Chemistry & biology* 16, 382-390.
- [191] Han, L., Hyung, S. J., and Ruotolo, B. T. (2012) Bound cations significantly stabilize the structure of multiprotein complexes in the gas phase, *Angewandte Chemie International Edition* 51, 5692-5695.
- [192] Han, L., and Ruotolo, B. T. (2013) Traveling-wave ion mobility-mass spectrometry reveals additional mechanistic details in the stabilization of protein complex ions through tuned salt additives, *International Journal for Ion Mobility Spectrometry* 16, 41-50.
- [193] Klassen, J. S., and Kebarle, P. (1997) Collision-induced dissociation threshold energies of protonated glycine, glycineamide, and some related small peptides and peptide amino amides, *Journal of the American Chemical Society* 119, 6552-6563.
- [194] Han, L., Hyung, S.-J., Mayers, J. J., and Ruotolo, B. T. (2011) Bound anions differentially stabilize multiprotein complexes in the absence of bulk solvent, *Journal of the American Chemical Society* 133, 11358-11367.
- [195] Badman, E. R., Hoaglund-Hyzer, C. S., and Clemmer, D. E. (2001) Monitoring Structural Changes of Proteins in an Ion Trap over~ 10-200 ms: Unfolding Transitions in Cytochrome c Ions, *Analytical chemistry* 73, 6000-6007.
- [196] Zhou, M., Dagan, S., and Wysocki, V. H. (2012) Protein subunits released by surface collisions of noncovalent complexes: natively compact structures revealed by ion mobility mass spectrometry, *Angewandte Chemie International Edition* 51, 4336-4339.
- [197] Zhou, M., Dagan, S., and Wysocki, V. H. (2013) Impact of charge state on gas-phase behaviors of noncovalent protein complexes in collision induced dissociation and surface induced dissociation, *Analyst* 138, 1353-1362.
- [198] Alverdi, V., Mazon, H., Versluis, C., Hemrika, W., Esposito, G., van den Heuvel, R., Scholten, A., and Heck, A. J. (2008) cGMP-binding prepares PKG for substrate binding by disclosing the C-terminal domain, *Journal of molecular biology* 375, 1380-1393.
- [199] Duijn, E. v., Barendregt, A., Synowsky, S., Versluis, C., and Heck, A. J. (2009) Chaperonin complexes monitored by ion mobility mass spectrometry, *Journal of the American Chemical Society* 131, 1452-1459.
- [200] Woods, L. A., Platt, G. W., Hellewell, A. L., Hewitt, E. W., Homans, S. W., Ashcroft, A. E., and Radford, S. E. (2011) Ligand binding to distinct states diverts aggregation of an amyloid-forming protein, *Nature chemical biology* 7, 730-739.
- [201] Uetrecht, C., Barbu, I. M., Shoemaker, G. K., van Duijn, E., and Heck, A. J. (2011) Interrogating viral capsid assembly with ion mobility-mass spectrometry, *Nature chemistry* 3, 126-132.
- [202] Benesch, J. L., and Ruotolo, B. T. (2011) Mass spectrometry: come of age for structural and dynamical biology, *Current opinion in structural biology* 21, 641-649.
- [203] Afonso, C., Hathout, Y., and Fenselau, C. (2002) Qualitative characterization of biomolecular zinc complexes by collisionally induced dissociation, *Journal of mass spectrometry* 37, 755-759.
- [204] Tabb, D. L., Huang, Y., Wysocki, V. H., and Yates, J. R. (2004) Influence of basic residue content on fragment ion peak intensities in low-energy collision-induced dissociation spectra of peptides, *Analytical chemistry* 76, 1243-1248.

- [205] Jurchen, J. C., Garcia, D. E., and Williams, E. R. (2004) Further studies on the origins of asymmetric charge partitioning in protein homodimers, *Journal of the American Society for Mass Spectrometry* 15, 1408-1415.
- [206] Jurchen, J. C., and Williams, E. R. (2003) Origin of asymmetric charge partitioning in the dissociation of gas-phase protein homodimers, *Journal of the American Chemical Society* 125, 2817-2826.
- [207] Schenauer, M. R., and Leary, J. A. (2009) An ion mobility–mass spectrometry investigation of monocyte chemoattractant protein-1, *International journal of mass spectrometry* 287, 70-76.
- [208] Goldstein, I., Hollerman, C., and Smith, E. (1965) Protein-carbohydrate interaction. II. Inhibition studies on the interaction of concanavalin a with polysaccharides\*, *Biochemistry* 4, 876-883.
- [209] Warnke, S., von Helden, G., and Pagel, K. (2013) Protein structure in the gas phase: the influence of side-chain microsolvation, *Journal of the American Chemical Society* 135, 1177-1180.
- [210] Wytenbach, T., Grabenauer, M., Thalassinou, K., Scrivens, J. H., and Bowers, M. T. (2009) The effect of calcium ions and peptide ligands on the relative stabilities of the calmodulin dumbbell and compact structures, *The Journal of Physical Chemistry B* 114, 437-447.
- [211] Zhang, M., Tanaka, T., and Ikura, M. (1995) Calcium-induced conformational transition revealed by the solution structure of apo calmodulin, *Nature Structural & Molecular Biology* 2, 758-767.
- [212] Atmanene, C. d., Chaix, D., Bessin, Y., Declerck, N., Van Dorsselaer, A., and Sanglier-Cianferani, S. (2010) Combination of noncovalent mass spectrometry and traveling wave ion mobility spectrometry reveals sugar-induced conformational changes of central glycolytic genes repressor/DNA complex, *Analytical chemistry* 82, 3597-3605.
- [213] Brocca, S., Testa, L., Sobott, F., Šamalikova, M., Natalello, A., Papaleo, E., Lotti, M., De Gioia, L., Doglia, S. M., and Alberghina, L. (2011) Compaction properties of an intrinsically disordered protein: Sic1 and its kinase-inhibitor domain, *Biophysical journal* 100, 2243-2252.
- [214] Bernstein, S. L., Dupuis, N. F., Lazo, N. D., Wytenbach, T., Condron, M. M., Bitan, G., Teplow, D. B., Shea, J.-E., Ruotolo, B. T., and Robinson, C. V. (2009) Amyloid- $\beta$  protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease, *Nature Chemistry* 1, 326-331.
- [215] Harvey, S. R., Porrini, M., Stachl, C., MacMillan, D., Zinzalla, G., and Barran, P. E. (2012) Small-molecule inhibition of c-MYC: MAX leucine zipper formation is revealed by ion mobility mass spectrometry, *Journal of the American Chemical Society* 134, 19384-19392.
- [216] Canon, F., Ballivian, R., Chiot, F., Antoine, R., Sarni-Manchado, P., Lemoine, J., and Dugourd, P. (2011) Folding of a salivary intrinsically disordered protein upon binding to tannins, *Journal of the American Chemical Society* 133, 7847-7852.
- [217] Bleiholder, C., Dupuis, N. F., Wytenbach, T., and Bowers, M. T. (2011) Ion mobility–mass spectrometry reveals a conformational conversion from random assembly to  $\beta$ -sheet in amyloid fibril formation, *Nature chemistry* 3, 172-177.
- [218] Gessel, M. M., Wu, C., Li, H., Bitan, G., Shea, J.-E., and Bowers, M. T. (2011) A $\beta$  (39–42) modulates A $\beta$  oligomerization but not fibril formation, *Biochemistry* 51, 108-117.

- [219] Wu, C., Murray, M. M., Bernstein, S. L., Condron, M. M., Bitan, G., Shea, J.-E., and Bowers, M. T. (2009) The structure of A $\beta$ 42 C-terminal fragments probed by a combined experimental and theoretical study, *Journal of molecular biology* 387, 492-501.
- [220] Ehrnhoefer, D. E., Bieschke, J., Boeddrich, A., Herbst, M., Masino, L., Lurz, R., Engemann, S., Pastore, A., and Wanker, E. E. (2008) EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers, *Nature structural & molecular biology* 15, 558-566.

## Chapter 2.

# CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions

Jessica N. Rabuck-Gibbons, Joseph D. Eschweiler, Yuwei Tian, and Brandon T. Ruotolo

(2015) CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions, *Analytical Chemistry* 87(22), 11516-1522

Ion mobility-mass spectrometry (IM-MS) is a technology of growing importance for structural biology, providing complementary 3D structure information for biomolecules within samples that are difficult to analyze using conventional analytical tools through the near-simultaneous acquisition of ion collision cross sections (CCSs) and masses. Despite recent advances in IM-MS instrumentation, the resolution of closely related protein conformations remains challenging. Collision induced unfolding (CIU) has been demonstrated as a useful tool for resolving isocrosssectional protein ions, as they often follow distinct unfolding pathways when subjected to collisional heating in the gas phase. CIU has been used for a variety of applications, from differentiating binding modes of activation state-selective kinase inhibitors to characterizing the domain structure of multidomain proteins. With the growing utilization of CIU as a tool for structural biology, significant challenges have emerged in data analysis and interpretation, specifically the normalization and comparison of CIU data sets. Here, we present CIUSuite, a suite of software modules designed for the rapid processing, analysis, comparison, and classification of CIU data. We demonstrate these tools as part of a series of workflows for applications in comparative structural biology, biotherapeutic analysis, and high throughput screening of kinase inhibitors. These examples illustrate both the potential for CIU in general protein analysis as well as a demonstration of best practices in the interpretation of CIU data.

## 2.1 Introduction to Collision Induced Unfolding



Native mass spectrometry (MS) is now a widespread technique in the structural biology community due to its ability to study the stoichiometry and connectivity of heterogeneous biomolecules while having lesser requirements on concentration and purity for such samples than other techniques<sup>1, 2</sup>. Coupling native MS with ion mobility spectrometry (IM-MS) allows for simultaneous interrogation of the mass, charge, and size of biological macromolecules, which has proven invaluable in the structural analysis of complex biological systems<sup>3</sup>. Recently, IM-MS has been successfully utilized to solve the structures of important macromolecular complexes<sup>4-7</sup>, probe structural changes upon ligand binding<sup>8, 9</sup>, examine the polydispersity of protein complexes<sup>10, 11</sup>, and study the effects of small molecules on amyloid formation in disease models<sup>12-14</sup>.

A key feature of IM-MS for structural and pharmaceutical applications is the ability to measure the orientationally averaged collision cross section (CCS) of an ion in addition to its mass and charge. The CCS is a coarse-grained size parameter that is limited in information content when viewed alone but can become information rich when measured as a function of stoichiometry<sup>15</sup>, ligand binding,<sup>16</sup> or ion activation<sup>2, 17</sup>. Additionally, the experimental CCS is an extremely important scoring metric for modeling complex systems, as it can be compared to CCSs calculated from other known or inferred structures<sup>18</sup>. A long-term challenge for IM-MS has been the resolution of closely related protein conformations, commonly observed by X-ray and NMR analyses. Despite recent enhancements to IM resolving power, IM still faces significant challenges when separating protein conformations that differ by less than 2% in CCS. The information content of an IM-MS experiment can be greatly enriched by the addition of gas-phase ion activation, as some differences in protein structure are too subtle to be separated by ground-state CCS alone. Early experiments that utilized gas-phase protein unfolding to both study and differentiate protein structures focused on small, single domain proteins and detected stability difference for proteins as a function of charge state, and for those with intact disulfide bonds<sup>19</sup>. Subsequent experiments extended these observations to the ligand-bound forms of wild-type (WT) and disease associated variants of tetrameric transthyretin (TTR)<sup>2</sup>. In this study, a 3D contour plot of ion intensity as a function of activation voltage and drift time, termed a collision induced unfolding (CIU) fingerprint, was used to perform an in-depth analysis of subtle differences in the unfolding and dissociation pathways of TTR variants, identifying additional ligand-based protein stabilization in mutant TTR forms not detectable by IM-MS alone.

Since these earlier experiments, CIU fingerprints have been used in the context of various applications. These efforts include studying the influence of bound anions and cations on gas-phase protein stability<sup>20, 21</sup>, distinguishing between inhibitors that stabilize either the active or the inactive form of the Abelson protein tyrosine kinase (Abl)<sup>22</sup>, measuring stability enhancements and cooperativity effects in proteins upon ligand-binding<sup>17, 23</sup>, probing the selectivity of lipid binding in membrane proteins<sup>24</sup>, determining the domain structures for a broad range proteins with varying molecular weights and domain structures<sup>25</sup>, and differentiating between disulfide binding isoforms in antibodies<sup>26</sup>.

Despite these varied and potentially impactful applications, CIU has not reached its full potential as a tool for structural biology and drug discovery. Key challenges for the technique include the general underutilization of structural information content of CIU data as well as a lack of high-throughput experimental frameworks and data analysis tools. Recent advances have been made in analysis and interpretation of other IM-MS data types, including deconvolution algorithms<sup>27, 28</sup> and an array of methods for prediction of CCSs from experimental or model structures<sup>29-31</sup>, but thus far CIU data has not been the focus of any such data analysis packages.

In order to move forward in the use of CIU as a tool for general structural biology, as well as for high-throughput pharmaceutical applications specifically, data analysis tools and strategies for handling the large amount of data that is produced by CIU fingerprints must be developed and implemented. In this report, we describe such software tools, collectively named CIUSuite, designed to ameliorate many of the challenges described above. Additionally, through detailed discussions of three diverse applications of the CIU technique, we illustrate workflows and best practices for extracting maximal information content from CIU data.

## 2.2 CIUSuite Overview

To facilitate interpretation of CIU data for a variety of applications, we designed CIUSuite, a series of Python<sup>32</sup> modules for the generation and manipulation of CIU fingerprints. CIUSuite consists of six modules that allow the user to readily access statistical and structural information from CIU experiments by designing user-defined CIUSuite workflows (Figure 2.1). The main data structure in CIUSuite is the three-dimensional size, activation energy, and intensity matrix that forms the CIU fingerprint. The size axis is IM-MS drift time (ms) by default; however, the

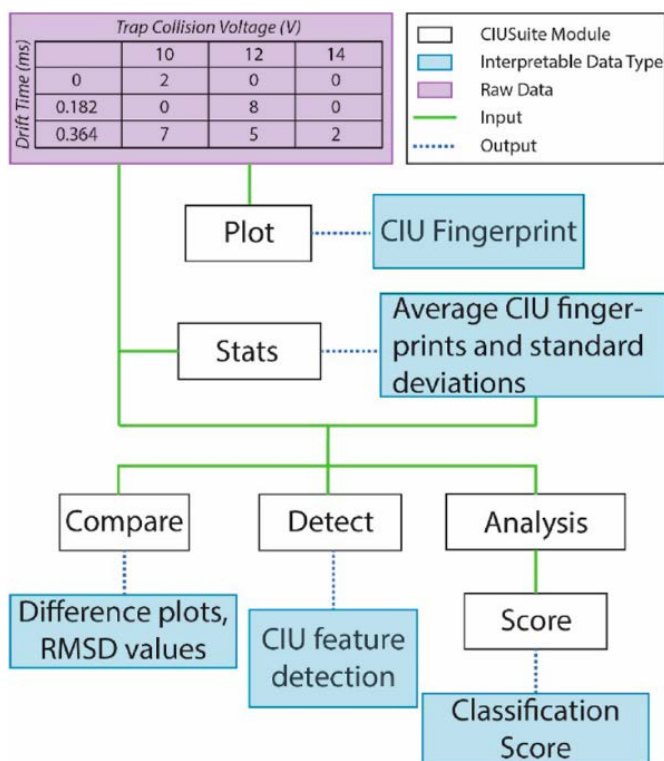


Figure 2.1 Schematic representation of CIUSuite modules. All modules (shown as white boxes) take as input raw data in the form of a 2D matrix (purple) formatted such that ion intensity is collected a function of drift time and trap collision voltage. Additionally, modules can accept outputs from CIU\_stats for groupwise comparisons using average and standard deviation measurements. Example outputs from each of these modules are also shown (blue). Both the modules and their outputs are discussed in detail in the text.

writes the corresponding contour plot to a .png file.

## 2.2.2 CIUSuite\_stats

CIUSuite\_stats outputs both visual information that can be interpreted by the user and numerical matrices that can be used for downstream analysis. This calculation is performed for every data point in the CIU fingerprint, and both the average and standard deviation matrices are output as a .csv matrix as well as a .png figure in the same fashion as CIU\_plot. This module requires at least 3 data sets to calculate the standard deviation fingerprint.

## 2.2.3 CIUSuite\_compare

CIUSuite\_compare allows for the facile comparison of CIU fingerprints by matrix subtraction and visualization of the difference matrix. Inputs for CIUSuite\_compare can be raw data

user is able to substitute CCS values when available. The activation energy axis can be expressed in volts, laboratory-frame energy, or center-of-mass frame energy. In an attempt to obviate potential problems arising from signal intensity variation between samples, the ion intensities observed at each activation energy typically are normalized to a maximum value of 1 and smoothed using a Savitsky–Golay filter with a window length of three and polynomial order of two; however, these parameters can also be easily adjusted by the user.

### 2.2.1 CIUSuite\_plot

CIUSuite\_plot forms the basis of the CIUSuite. CIUSuite\_plot batch processes any CIU data in its working directory tagged with the suffix “\_raw.csv” and

matrices or the average matrices output from CIUSuite\_stats. CIUSuite\_compare also utilizes the root-mean-square deviation (RMSD) parameter to report the absolute difference between two matrices and prints the RMSD on the difference plot. Here, RMSD is defined in eq. 1 as

$$RMSD = \sqrt{\frac{\sum(A-B)^2}{m \times n}} \times 100\% \quad (1)$$

where A and B are both  $m \times n$  CIU matrices. Additionally, the pairwise RMSD matrix (distance matrix) is output as a .csv if the user desires to utilize other clustering algorithms.

## 2.2.4 CIUSuite\_detect

CIUSuite\_detect is a simple feature detection algorithm that allows for quantitative analysis of CIU data. The algorithm utilizes the first derivative test to identify local maxima in the data, before refining the shape of the feature using user-defined data scaling and intensity thresholds. After features are identified and refined, their stabilities in collision voltage space as well as their centroid drift times (or CCSs) are output to a file summarizing the data set.

## 2.2.5 CIUSuite\_analysis

CIUSuite\_analysis was developed as a tool for adaptation of CIU fingerprinting for high-throughput ligand screening and structural biology. CIU\_analysis allows the user to identify areas within the CIU fingerprint that are useful for categorizing data sets into groups, such as type I or type II kinase inhibitors (vide infra). The current implementation of CIU\_analysis takes as input a training data set, where each file is annotated as either a type I (\_typeI\_raw.csv) or type II (\_typeII\_raw.csv) fingerprint. CIU\_analysis utilizes a scaled deviation score where each fingerprint in the data set is compared to the average fingerprints for both the type I and type II groups. Here, the type I scaled deviation score (SDS) is defined in eq 2:

$$SDS_1 = \sum_{j=0}^{j=m} \frac{(X_{ij} - A_{ij}^1) \times A_{ij}^1}{S_{ij}^1} \quad (2)$$

where X is a CIU matrix,  $A^1$  is the average type I matrix,  $S^1$  is the type I standard deviation matrix, i is a given collision energy, j a given drift time, and m is the total number of drift time bins. The primary outputs are two plots of SDS vs Collision Voltage, one corresponding to the type I average SDS value, the other corresponding to the type II average SDS value. These plots display the average SDS value (with 2 standard deviations as the error bars) for both type I and

type II fingerprints compared to the corresponding average value. The information contained in these two plots is extremely valuable for the accurate classification of unknown fingerprints as well as targeting CIU workflows toward optimal regions of dissimilarity between the two data classes, increasing the throughput of the experiment. CIUSuite\_analysis also outputs a plot of SDS vs Collision Voltage for each component of the training data set, allowing the user to identify outliers or other anomalies that may bias the analysis.

### 2.2.6 CIUSuite\_score

CIUSuite\_score is predicated on data from CIUSuite\_analysis. CIUSuite\_score accepts as input a training data set and “unknown” data that is tagged with “\_uk\_raw.csv.” Previous analyses of CIU fingerprint data<sup>22</sup> have shown that focusing on specific collision voltages, rather than using the entire CIU fingerprint, can increase the throughput and robustness of the resulting screen. After identifying the voltage ranges or drift times in the CIU fingerprints that yield significant group-wise deviation values using CIUSuite\_analysis, the user can enter these values into the scoring module to calculate classification scores based only on these regions. Each training data set is grouped according to a user-defined tag, and corresponding SDS values are calculated. SDS values are then summed over all of the collision energies to be scored, assigning a single scaled deviation value for each fingerprint. For example, type I fingerprints should have low overall deviation relative to the type I average, whereas they should have higher deviation scores relative to the type II average. The type I z-score with respect to the type I training data for an unknown fingerprint is simply the z-score of its SDS compared to the average SDS for a type I compared to the type I average, as described in eq 3:

$$z - score^1 = \frac{SDS^1 - \bar{x}^1}{s^1} \quad (3)$$

where  $SDS^1$  is the SDS of the unknown compared to the type I average,  $\bar{x}^1$  is the average SDS of a type I compared to the type I average, and  $s^1$  is the standard deviation of SDS values around  $\bar{x}^1$ . The output for CIUSuite\_score comprises of a .csv file that contains the type I and type II z-scores for each data set and a graph showing the type II classification-score vs type I classification-score. The resulting plot displays type I training data in blue, type II training data in red, and unknown scores in cyan. A blue and red box around the data sets indicates two standard deviations from the type I and type II training data, respectively.

## 2.3 CIUSuite Applications

CIUSuite has successfully been used to analyze the similarities of the monoclonal antibody (mAB) IgG4 for four samples from four different suppliers. Standard deviations equal to or smaller than the baseline values were observed, further illustrating the excellent reproducibility of the CIU method. Taken together, the data shown in Figure 2.2 illustrate the capabilities of CIUSuite to evaluate CIU data for potential applications in biopharmaceutical characterization and quality control. CIUSuite was also used to quantify differences in the unfolding pathways of homologous albumins. The differences in the stability of CIU features between homologous proteins may indicate potential for domain or interface-specific stability measurements to be used in biopharmaceutical or protein engineering applications.

### 2.3.1 Rapid Analysis of Kinase Inhibitor Screens

CIU fingerprint assays can be used to differentiate between ligands that cause conformational changes in proteins that are too subtle to be differentiated by IM-MS alone. Previously, we published a CIU fingerprint assay capable of differentiating between two different classes of ATP-competitive kinase inhibitors<sup>22</sup>. In this previous study, we validated our assay using ions comprised of the Abelson protein tyrosine kinase (Abl) domain bound individually to six type I inhibitors and five type II inhibitors. Type I inhibitors are ATP competitive ligands that stabilize the active conformation of the kinase, in which the activation loop is in an extended conformation. Type II ligands are also ATP-competitive but stabilize the inactive conformation of the kinase, in which the activation loop covers the substrate binding site.

Previously, we used a manual approach to identify those regions within the CIU data for type I

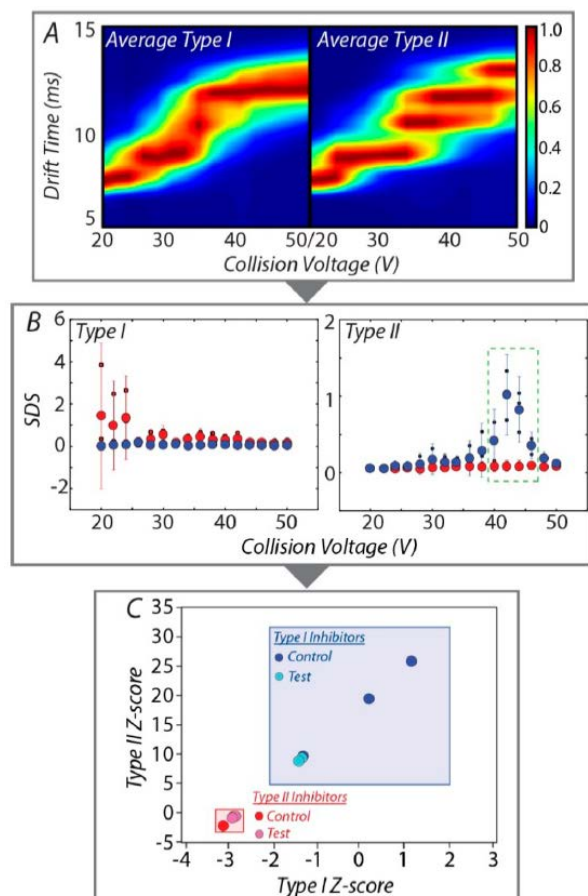


Figure 2.2 Activation state-selective kinase inhibitors are classified by CIU fingerprints and CIUSuite analysis. (A) CIUSuite\_stats is used to create average type I and type II CIU fingerprints from a training set of known type I and type II inhibitors. (B) CIUSuite\_analysis is then employed to analyze the SDS to determine the best collision voltage for inhibitor analysis. Here, type I and type II inhibitors are represented by blue and red dots, respectively. The dashed green box indicates the area of greatest difference that is used to subsequently score the inhibitors. (C) CIUSuite\_score is implemented to rank the inhibitors based on the type II characteristic of each inhibitor. Control points (dark blue and dark red) are used to train the scoring algorithm, whereas test points (light cyan and light pink) are analyzed to validate the procedure.

and II bound Abl ions that were most differentiating. This workflow led the development of a  $\chi^2$ -based score, which allowed us to calculate the agreement of individual data with both the average type I and type II CIU fingerprint. When compared, the region between 40 and 44 V was found to be the optimal region within the CIU data to find those data that had the strongest agreement with type II fingerprints. CIUSuite has improved dramatically our ability to rapidly assess the regions of dissimilarity within CIU fingerprints and construct validated workflows for drug screening assays.

The CIUSuite protocol for screening for conformationally selective ligands utilizes the CIUSuite\_stats, CIUSuite\_analysis, and CIUSuite\_score modules. First, we use CIUSuite\_stats to determine the average type I and average type II CIU fingerprints from known type I and type II inhibitors bound to protein ions (Figure 2.2A). This training data set is then given as input to CIUSuite\_analysis, which computes SDS values and plots these against the calculated type I and type II averages as a function of collision voltage (Figure 2.2B). Once constructed, the SDS plots allow the user to readily identify those regions of the composite

CIU data that provides the greatest ability to distinguish between the protein-inhibitor complexes sampled. In the case of Abl kinase, the CIUSuite\_analysis results shown in Figure 2.2B agrees well with our slower, manual approach applied previously, both identifying the

voltage range of 40–44 V as most differentiating for I and type II kinase inhibitors (green dashed box). Using CIUSuite\_score, we used the CIU data extracted from this range of voltages to compute z-scores using eq 3. A plot detailing the results of these CIUSuite calculations is shown in Figure 2.2C, which effectively clusters type I and II CIU signals, along with colored boxes which indicate the standard deviations of these clusters. Importantly, this plot includes points that both result from CIU data that was used to train our z-score analysis to detect type I and II responses as well as points computed from test CIU data that possess both a known kinase binding mode and was not used to train the screening protocol shown. For both type I and II test points, we achieve excellent clustering, with strong statistically significant differentiation on the type II z-score axis, further validating our CIU kinase inhibitor assay.

As indicated above, CIUSuite dramatically improves the speed of CIU data analysis for potential high-throughput applications in drug discovery and development through the accurate identification of differentiating CIU fingerprint regions. For example, typical collection times for complete kinase-inhibitor complex CIU fingerprints range as high as 50 min per inhibitor (3 min per voltage step, 16 total steps).

In contrast, CIU data ranging from 40 and 44 V can be collected in less than 3 minutes/sample, assuming sufficient signal intensity is available for the complex. This predicts a throughput of ~480 protein–inhibitor complexes for a 24 h screen. CIUSuite also improves the speed at which these differentiating regions can be identified. Our previous manual workflow required several weeks to identify the 40–44 V region discussed above, whereas the same region was identified by CIUSuite in seconds. However, a time of 1 second per inhibitor can be achieved from the above data.. The limit of quantification for a CIU experiment is limited by the sample; for example, in the experiments described here, in a 1 second collection window, a S/N of 22 could be achieved. The S/N was calculated by selecting the signal for the protein in both mass and drift time and comparing it to the remaining signal. If the improved time scales of both data analysis and collection are summed, we estimate that a screen involving 86,400 inhibitors, including assay development, could conceivably be performed in 2 days using CIUSuite, whereas manual methods would have required over 1 month of analysis and acquisition time for a highly trained IM-MS operator. As brighter ion sources and faster sample manipulation techniques become



available, CIUSuite will likely enable the development and implementation of such screens on the time scale of a few hours.

## 2.4 Conclusions

Analysis of CIU data is an emerging challenge for those seeking to expand the information content of typical IM-MS experiments. Rapid and robust procedures for CIU fingerprint analysis are necessary for the continued development and application of such gas-phase unfolding experiments. Emerging applications, such as protein engineering and high-throughput screening, which involve the rapid analysis of large numbers of samples, require streamlined quantitative analysis, provided by CIUSuite, in order to achieve realistic analysis capacities. Forthcoming challenges in this field include the integration of CIU fingerprint data into databases, allowing for analysis of variability across instruments, as well as the comparison of CIU data for quality control applications. Although the CIU analysis workflows contained herein overcome major bottle necks in experiment and analysis time, hurdles still exist in enabling CIU for high-throughput applications. One exciting area will surely involve the integration of adaptive, data-dependent algorithms within CIU analysis workflows. These algorithms will likely include the ability to dynamically focus on fingerprint regions that are the most information-rich and to rapidly tune instrumental conditions to ensure that adequate signal-to-noise is achieved for each measurement.

This software and the mathematical procedures contained within CIUSuite represent a framework for the continued study of gas-phase protein unfolding, and we anticipate that its application will lead to further discoveries regarding the basic biophysics of proteins in the absence of bulk solvent. In addition, as the study of protein unfolding analysis in the gas-phase is a relatively new area, the authors encourage modification and expansion of CIUSuite capabilities, so that the base approaches described here can be applied to data structures not yet conceived.

## 2.5 References

- [1] Benesch, J. L. P., and Ruotolo, B. T. (2011) Mass Spectrometry: an Approach Come-of-Age for Structural and Dynamical Biology, *Current opinion in structural biology* 21, 641-649.

- [2] Hyung, S.-J., Robinson, C. V., and Ruotolo, B. T. (2009) Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes, *Chemistry & biology* 16, 382-390.
- [3] Uetrecht, C., Rose, R. J., van Duijn, E., Lorenzen, K., and Heck, A. J. R. (2010) Ion mobility mass spectrometry of proteins and protein assemblies, *Chemical Society Reviews* 39, 1633-1655.
- [4] Politis, A., Park, A. Y., Hall, Z., Ruotolo, B. T., and Robinson, C. V. (2013) Integrative Modelling Coupled with Ion Mobility Mass Spectrometry Reveals Structural Features of the Clamp Loader in Complex with Single-Stranded DNA Binding Protein, *Journal of Molecular Biology* 425, 4790-4801.
- [5] Ruotolo, B. T., Benesch, J. L., Sandercock, A. M., Hyung, S.-J., and Robinson, C. V. (2008) Ion mobility-mass spectrometry analysis of large protein complexes, *Nature Protocols* 3, 1139-1152.
- [6] Sharon, M., Mao, H., Boeri Erba, E., Stephens, E., Zheng, N., and Robinson, C. V. (2009) Symmetrical Modularity of the COP9 Signalosome Complex Suggests its Multifunctionality, *Structure* 17, 31-40.
- [7] Uetrecht, C., Barbu, I. M., Shoemaker, G. K., van Duijn, E., and Heck, A. J. R. (2011) Interrogating viral capsid assembly with ion mobility-mass spectrometry, *Nat Chem* 3, 126-132.
- [8] Niu, S., Rabuck, J. N., and Ruotolo, B. T. (2013) Ion mobility-mass spectrometry of intact protein-ligand complexes for pharmaceutical drug discovery and development, *Current Opinion in Chemical Biology* 17, 809-817.
- [9] Xing, W., Busino, L., Hinds, T. R., Marionni, S. T., Saifee, N. H., Bush, M. F., Pagano, M., and Zheng, N. (2013) SCFFBXL3 ubiquitin ligase targets cryptochromes at their cofactor pocket, *Nature* 496, 64-68.
- [10] Pagel, K., Natan, E., Hall, Z., Fersht, A. R., and Robinson, C. V. (2013) Intrinsically Disordered p53 and Its Complexes Populate Compact Conformations in the Gas Phase, *Angewandte Chemie International Edition* 52, 361-365.
- [11] Shepherd, D. A., Marty, M. T., Giles, K., Baldwin, A. J., and Benesch, J. L. P. (2015) Combining tandem mass spectrometry with ion mobility separation to determine the architecture of polydisperse proteins, *International journal of mass spectrometry* 377, 663-671.
- [12] Do, T. D., Economou, N. J., Chamas, A., Buratto, S. K., Shea, J.-E., and Bowers, M. T. (2014) Interactions between Amyloid- $\beta$  and Tau Fragments Promote Aberrant Aggregates: Implications for Amyloid Toxicity, *The Journal of Physical Chemistry B* 118, 11220-11230.
- [13] Soper, M. T., DeToma, A. S., Hyung, S.-J., Lim, M. H., and Ruotolo, B. T. (2013) Amyloid- $\beta$ -neuropeptide interactions assessed by ion mobility-mass spectrometry, *Physical Chemistry Chemical Physics* 15, 8952-8961.
- [14] Susa, A. C., Wu, C., Bernstein, S. L., Dupuis, N. F., Wang, H., Raleigh, D. P., Shea, J.-E., and Bowers, M. T. (2014) Defining the Molecular Basis of Amyloid Inhibitors: Human Islet Amyloid Polypeptide-Insulin Interactions, *Journal of the American Chemical Society* 136, 12912-12919.
- [15] Pukala, T. L., Ruotolo, B. T., Zhou, M., Politis, A., Stefanescu, R., Leary, J. A., and Robinson, C. V. (2009) Subunit architecture of multiprotein assemblies determined using restraints from gas-phase measurements, *Structure* 17, 1235-1243.

- [16] Brocca, S., Testa, L., Sobott, F., Šamalikova, M., Natalello, A., Papaleo, E., Lotti, M., De Gioia, L., Doglia, Silvia M., Alberghina, L., and Grandori, R. (2011) Compaction Properties of an Intrinsically Disordered Protein: Sic1 and Its Kinase-Inhibitor Domain, *Biophysical Journal* 100, 2243-2252.
- [17] Niu, S., and Ruotolo, B. T. (2015) Collisional unfolding of multiprotein complexes reveals cooperative stabilization upon ligand binding, *Protein Science*.
- [18] Politis, A., Park, A. Y., Hyung, S.-J., Barsky, D., Ruotolo, B. T., and Robinson, C. V. (2010) Integrating ion mobility mass spectrometry with molecular modelling to determine the architecture of multiprotein complexes.
- [19] Shelimov, K. B., Clemmer, D. E., Hudgins, R. R., and Jarrold, M. F. (1997) Protein structure in vacuo: Gas-phase conformations of BPTI and cytochrome c, *Journal of the American Chemical Society* 119, 2240-2248.
- [20] Han, L., Hyung, S.-J., Mayers, J. J., and Ruotolo, B. T. (2011) Bound anions differentially stabilize multiprotein complexes in the absence of bulk solvent, *Journal of the American Chemical Society* 133, 11358-11367.
- [21] Han, L., Hyung, S. J., and Ruotolo, B. T. (2012) Bound cations significantly stabilize the structure of multiprotein complexes in the gas phase, *Angewandte Chemie International Edition* 51, 5692-5695.
- [22] Rabuck, J. N., Hyung, S.-J., Ko, K. S., Fox, C. C., Soellner, M. B., and Ruotolo, B. T. (2013) Activation state-selective kinase inhibitor assay based on ion mobility-mass spectrometry, *Analytical chemistry* 85, 6995-7002.
- [23] Hopper, J. T., and Oldham, N. J. (2009) Collision induced unfolding of protein ions in the gas phase studied by ion mobility-mass spectrometry: the effect of ligand binding on conformational stability, *Journal of the American Society for Mass Spectrometry* 20, 1851-1858.
- [24] Laganowsky, A., Reading, E., Allison, T. M., Ulmschneider, M. B., Degiacomi, M. T., Baldwin, A. J., and Robinson, C. V. (2014) Membrane proteins bind lipids selectively to modulate their structure and function, *Nature* 510, 172-175.
- [25] Zhong, Y., Han, L., and Ruotolo, B. T. (2014) Collisional and Coulombic Unfolding of Gas-Phase Proteins: High Correlation to Their Domain Structures in Solution, *Angewandte Chemie* 126, 9363-9366.
- [26] Tian, Y., Han, L., Buckner, A. C., and Ruotolo, B. T. (2015) Collision induced unfolding of intact antibodies: rapid characterization of disulfide bonding patterns, glycosylation, and structures, *Analytical chemistry* 87, 11509-11515.
- [27] Marty, M. T., Baldwin, A. J., Marklund, E. G., Hochberg, G. K., Benesch, J. L., and Robinson, C. V. (2015) Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles, *Analytical chemistry* 87, 4370-4376.
- [28] Sivalingam, G. N., Yan, J., Sahota, H., and Thalassinou, K. (2013) Amphitrite: A program for processing travelling wave ion mobility mass spectrometry data, *International journal of mass spectrometry* 345, 54-62.
- [29] Bleiholder, C., Wyttenbach, T., and Bowers, M. T. (2011) A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (I). Method, *International journal of mass spectrometry* 308, 1-10.
- [30] Larriba, C., and Hogan, C. J. (2013) Free molecular collision cross section calculation methods for nanoparticles and complex ions with energy accommodation, *Journal of Computational Physics* 251, 344-363.

- [31] Marklund, E. G., Degiacomi, M. T., Robinson, C. V., Baldwin, A. J., and Benesch, J. L. (2015) Collision Cross Sections for Structural Proteomics, *Structure* 23, 791-799.
- [32] Van Rossum, G., and Drake Jr, F. L. (1995) *Python tutorial*, Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands.

## Chapter 3.

# Activation State-Selective Kinase Inhibitor Assay Based on Ion Mobility-Mass Spectrometry

**Jessica N. Rabuck**, Suk-Joon Hyung, Kristin S. Ko, Christel C. Fox, Matthew B. Soelner, and Brandon T. Ruotolo (2013) Activation State-Selective Inhibitor Assay Based on Ion Mobility-Mass Spectrometry, *Analytical Chemistry* 85(15), 6995-7002.

The discovery of activation state dependent kinase inhibitors, which bind specifically to the inactive conformation of the protein, is considered to be a promising pathway to improved cancer treatments. Identifying such inhibitors is challenging, however, because they can have  $K_d$  values similar to molecules known to inhibit kinase function by interacting with the active form. Furthermore, while inhibitor induced changes within the kinase tertiary structure are significant, few technologies are able to correctly assign inhibitor binding modes in a high-throughput fashion based exclusively on protein-inhibitor complex formation and changes in local protein structure. We have developed a new assay, using ion mobility-mass spectrometry, capable of both rapidly detecting inhibitor binding and classifying the resultant kinase binding modes. Here, we demonstrate the ability of our approach to classify a broad set of kinase inhibitors, using micrograms of protein, without the need for protein modification or tagging.

### 3.1 Introduction

Kinase regulation plays a central role in multiple biochemical pathways and several disease states, most notably, cancer<sup>2</sup>. For example, tyrosine kinase inhibitors are a prominent treatment approach for chronic myelogenous leukemia (CML), where fusion between the Abelson (Abl) kinase gene and the break point cluster (BCR) at chromosome 22 results in a chimeric BCR-Abl tyrosine kinase implicated in the disease<sup>4</sup>. As such, there are many ongoing efforts aimed at designing small molecules capable of influencing the function of this broad class of proteins. Small molecule kinase inhibitors can fall into at least four general categories, with two of the most important being

those that bind to the adenosine triphosphate (ATP) binding site within the kinase domain (type I) and those that extend into a nearby “allosteric” site outside the ATP binding pocket (type II)<sup>6</sup>. The tertiary structure of kinases bound to these two small molecule classes are known to undergo small yet significant changes<sup>7</sup>. Type I molecules are conformationally nonspecific and thus will bind all states of the kinase including the open or active conformation. In contrast, type II binders interact preferentially with an inactive or closed conformation, where the flexible “activation loop” region of the protein refolds to cover the substrate binding site. While the active kinase form is broadly conserved, inactive forms can vary considerably between kinases. Thus, while many small molecule drugs are available for type I binding, these inhibitors typically lead to less-selective control over kinase function<sup>6</sup>. Type II inhibitors are, therefore, generally preferred for therapeutic purposes, as they provide higher degrees of kinase selectivity. However, the widespread use of type II kinase inhibitors as cancer therapies has, in some cases, led to drug resistance in many cell lines and CML patients<sup>8</sup>; thus, new type II inhibitors are needed to counteract such effects.

The main technology underpinning our approach to this problem is ion mobility-mass spectrometry (IM-MS), where ions produced by nano-electrospray ionization (nESI) can be filtered first by a quadrupole according to their  $m/z$ , separated according to their orientationally averaged size (collision cross section, CCS) on the millisecond time scale, and can then be analyzed by time-of-flight mass spectrometry<sup>9, 10</sup>. IM-MS has been used extensively to characterize the structures of small biomolecules in the gas phase<sup>11-13</sup> and has begun to be used broadly to analyze the structure of larger proteins and protein complexes<sup>14, 15</sup>, in many cases revealing high degrees of correlation between solvated and solvent-free data sets<sup>16</sup>. Many past IM-MS experiments have focused on protein and peptide systems where alterations in IM data could be related to significant structural changes in the gas-phase biomolecules of interest<sup>17</sup>. For example, IM-MS experiments are capable of discerning helical and globular peptide conformations<sup>18</sup>, as well as the calcium dependent conformational shifts of calmodulin<sup>19, 20</sup>, at modest IM resolution values. To assess finer protein tertiary structure details, IM-MS data sets must be combined with sophisticated MD simulations<sup>21, 22</sup>. Since many protein folds project identical CCS values, the information content carried by the IM-MS experiment necessarily decreases as the size of the protein increases, and the structural filtering requirements of the MD simulations utilized are greatly enhanced. Despite this inherent limitation, the structures of many small proteins have been determined in this fashion, including the desolvated structures for

ubiquitin<sup>23</sup> and A $\beta$ 1–42<sup>24</sup>. However, it is also clear from these previous reports that the inherent limitations of CCS as a lone constraint in structure determinations are a key challenge for the application of IM-MS in structural biology.

In addition to simple CCS measurements, IM-MS is also capable of recording protein CCS as a function of ion internal energy, thus enabling the technology to record protein unfolding as well as static protein structure. The first observations of protein ion unfolding predate the application of IM-MS to gas-phase biomolecules<sup>25</sup> and are related to the influences of Coulombic forces on gas-phase protein structure. Following these observations, IM<sup>26</sup> and IM-MS<sup>27</sup> were coupled with ESI, enabling the observation of protein unfolding both as a function of ion charge and internal temperature. Though these observations appeared throughout the early IM-MS literature, they were rarely interpreted relative to solution-phase protein structures. Recently, we, and others, introduced collision induced unfolding (CIU) as a means of distinguishing between the subtle differences in protein tertiary structure that result as a function of small molecule binding<sup>28, 29</sup>. For example, CIU data has been used to assess the different binding modes that exist between tetrameric transthyretin (TTR) and its natural ligand, thyroxine<sup>29</sup>. Significant differences in CIU response were detected in wild type TTR and an amyloidogenic mutant form of the protein, indicating that different thyroxine binding modes are operative in the two protein forms. While these data allowed for a critical demonstration of the CIU method as a means of detecting local stability shifts within proteins upon ligand binding, such data had not yet been correlated with the more subtle changes that occur within protein tertiary structure upon small molecule incorporation.

In this report, we develop a new method using the basic architecture of CIU, aimed at protein kinases. We have begun by differentiating type I and II inhibitors using the protein kinase domain of Abl as a model system. Key innovations in this new IM-MS and CIU-based method include using both collision induced dissociation (CID) and CIU data as a means of creating a more discriminating “fingerprint” for kinase– inhibitor complexes, as well as employing CIU “fingerprints” to assign regions of interest capable of streamlining the CIU methodology. Most critically, we find that, despite the relatively small structural changes that exist between active and inactive kinases, CIU is an excellent method for differentiating type I and II inhibitors, requiring relatively small amounts of unmodified protein. In addition, since the approach is built using an

MS-based technology, it functions in a manner that can be extended, in principle, to any kinase/inhibitor system in a relatively high-throughput mode (up to 100s of samples per day).

## **3.2 Experimental Methods**

### **3.2.1 General**

Imatinib, dasatinib, ponatinib, nilotinib, tozasertib, staurosporine, saracatinib, and sorafenib were purchased from LC Laboratories (Woburn, MA); PP2 was purchased from Sigma (St. Louis, MO). DCC-2036 was purchased from SelleckChem (Houston, TX), and bosutinib was purchased from Tocris Bioscience (Bristol, United Kingdom). Protein samples were buffer exchanged into 100 mM ammonium acetate at pH 7 using Micro Bio-Spin 6 columns (Bio-Rad, Hercules, CA) and prepared to a final concentration of 8.8  $\mu$ M. The inhibitors were added in a 1:1 mol ratio of Abl/inhibitor and incubated at room temperature for 1 h, after which all samples were moved to ice until analysis

### **3.2.2 Protein Purification and Activity Assays**

c-Abl kinase domain was synthesized by GeneArt (Life Technologies, Grand Island, NY) using E. coli modified codons. The kinase domain was subcloned into pET28a, modified with a TEV protease cleavable N-terminal 6x-His tag. The plasmid was transformed by electroporation into BL21DE3 electrocompetent cells containing YopH in pCDFDuet-1. Cell growth, expression, and protein purification were performed using modified literature protocols previously reported for expression of the wild-type c-Src kinase domain<sup>30</sup>. For the experimental details of our kinase activity assays, see Appendix I.

### **3.2.3 Ion Mobility-Mass Spectrometry**

Sample aliquots (~5  $\mu$ L) were analyzed by ion mobility-mass spectrometry on a quadrupole-ion mobility-time-of-flight mass spectrometer (Q-IM-TOF MS) instrument (Synapt G2 HDMS, Waters, Milford, Ma) and ionized using a nESI source, as described previously<sup>10, 31</sup>. The capillary voltages ranged from 1.2 to 2.0 kV, with the source operating in positive mode and the sample cone operating at 50 V. The trap traveling-wave ion guide was pressurized to  $2.2 \times 10^{-2}$  mbar of argon gas. The TOF-MS was operated over the m/z range of 1000–8000 and at a pressure of  $1.6 \times 10^{-6}$ , and the quad profile was set to dwell on 3000 m/z. The wave height was set to 40 V, and



the wave velocity was set to 900 m/s. The concentration of both inhibitor and protein are kept sufficiently low so as to avoid the formation of artifact complexes, and all data collected conforms to the expected inhibitor binding stoichiometry<sup>32</sup>.

Prior to the ion mobility separator, ions were activated by collisions in the ion trap traveling-wave ion guide in order to perform CIU of protein complexes and investigate the differences in the unfolding pathways of unphosphorylated Abl stabilized by type I and type II inhibitors<sup>29</sup>. Charge states were chosen on the basis of their relative intensity as a function of ligand binding and according to the number of intermediate conformations that could be observed during CIU experiments. Each mass-selected ion was activated by increasing the trap collision voltage (TCE), from 20 to 50 V in 2 V increments

All mass spectra were calibrated externally using a solution of cesium iodide (100 mg mL<sup>-1</sup>) and were processed using Masslynx 4.1 software (Waters, UK)<sup>33</sup>. The data for previously identified type I and type II drift time spectra were averaged to produce their corresponding average reference for high-throughput analysis. Similarity scores were found by first determining the  $\chi^2$  correlation between the averaged data and individual data sets.

$$\xi^2 = \xi_{i=1}^r \xi_{i=1}^c \frac{(A_{ij} - E_{IJ})^2}{E_{IJ}} \quad (1)$$

In eq 1 above,  $A_{ij}$  and  $E_{ij}$  are the actual and expected frequency of the  $i^{\text{th}}$  row,  $j^{\text{th}}$  column, respectively,  $r$  is the number of data rows, and  $c$  is the number of data columns. The  $\chi^2$  value is used here as a measure of the goodness of fit between IM data acquired from a given inhibitor and either type I or II averaged data and could be further used in conjunction with knowledge of the critical values from the  $\chi^2$  probability distribution and the degrees of freedom for a data set to determine the confidence interval of such assignments (as in the Pearson test)<sup>34</sup>. Determining  $\chi^2$  is a common approach used to analyze MS and separations data sets for similarity<sup>35-37</sup>. The similarity score discussed at length here was derived by normalizing the  $\chi^2$  value obtained for an imatinib/average type II comparison to a value of 100. The same normalization factor was then applied to all other  $\chi^2$  values, thus creating a scale for ranking the type II character of a given Abl-inhibitor CIU response. Smaller values indicate a poorer fit to the averaged type II data (the inverse of the original  $\chi^2$  metric). The errors reported here result from standard deviations of three replicate

experiments and are 1.9% on average. Complete  $\chi^2$  and similarity scores for three different CIU fingerprint regions are shown below in Appendix I. White colored values indicate those that would not cluster correctly during an analysis based only on the similarity score shown.

### **3.3 Results**

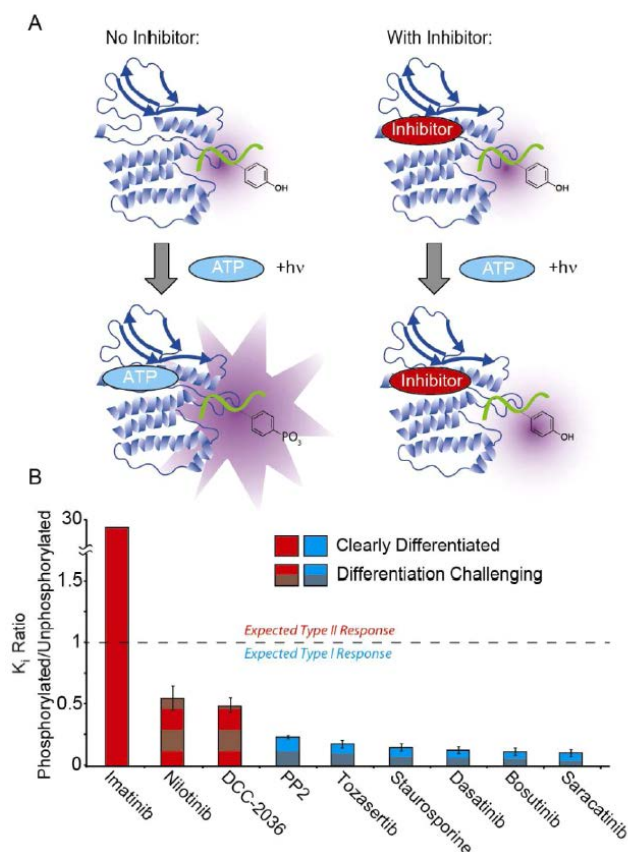
#### **3.3.1 Outlining the Challenges in Activation State-Selective Kinase Inhibitor Discovery**

Screening for type II inhibitors with current measurement technologies is a major challenge, as the structural changes that occur within the kinase domain upon binding are small and mainly limited to the activation loop within the protein fold. Enzyme activity assays, while effective in limited cases, are exceptionally difficult to utilize within the framework of a broad inhibitor screen. For example, we began our experiments by employing a fluorescence assay, where the activity of the Abl kinase domain (33.2 kDa) is measured in the presence of an inhibitor. Measurements are then made for Abl where the activation loop has been phosphorylated, causing the enzyme to favor its active conformation (Figure 3.1). Since the primary mode for differentiating kinase inhibitor binding types involves similar activity assays<sup>38-40</sup>, the results shown in Figure 3.1, likely represent a good benchmark for state-of-the-art high-throughput methods in kinase inhibitor binding mode identification.

While our results indicate the expected disparity in enzymatic activities for the known type II inhibitors imatinib (Gleevec), and dasatinib (Sprycel), all other inhibitors tested had a statistical preference for the active form of Abl, despite representing different known Abl binding modes (Figure 3.1B). Type I and II inhibitors appear to bind with similar strengths to the kinase if the activation state of the protein is not considered, making them difficult to distinguish using classical molecular biology tools. Screening methods that compare compound binding strengths to the different phosphorylation-controlled forms of Abl are available<sup>41</sup>, as are those that use fluorescence resonance energy transfer (FRET) to track the conformational form of the protein once bound to unknown inhibitors<sup>42</sup>, but they typically involve protein modifications, covalent tagging, and limited dynamic range, making them non-optimal for high-throughput screening efforts.

#### **3.3.2 CIU Kinase Inhibitor Binding Mode Assay**

Many results have demonstrated that, when collisionally activated in the gas phase, protein ions of



**Figure 3.1** A kinase activity assay illustrates the challenges associated with differentiating between type I and II inhibitors. (A) Abl kinase in the absence of inhibitor is exposed to the peptide substrate AEXIYAAPF-OH (green), where X is 2,3-diaminopropionic acid (Dap)-pyrene, a tailored fluorescent constituent. When ATP is added, the tyrosine at position 4 is phosphorylated and shifts the overall fluorescent signature of the peptide, so that emission at 405 nm is enhanced, which can be measured following excitation at 350 nm (indicated as  $h\nu$  in the illustration). In the presence of an inhibitor, phosphorylation of the peptide is reduced and recorded as attenuated fluorescence emission. (B) Ratios of the inhibitor binding constants ( $K_i$ ) for the phosphorylated (active) and nonphosphorylated (inactive) forms of the Abl kinase, derived using the assay shown in part A. Positive ratios in excess of 1 indicate type II character. While imatinib produces the expected type II response (28.99,  $2\sigma = 6.28$ ), all other known type II inhibitors produce responses that cluster more closely with type I binders (all error bars shown represent  $2\sigma$ ). A color code is used to indicate the known binding mode for each inhibitor: red for type II and blue for type I. Those responses shown in either red/gray or blue/gray stripes are difficult to differentiate based on kinase activity measurements alone, despite known differences in the binding mode (same basic color code as above).

sufficiently low charge state can unfold<sup>14, 43, 44</sup>. Gas-phase protein unfolding is distinct from protein denaturation in solution and is primarily characterized by the adoption of multiple long-lived (>100 ms) intermediately unfolded species that are likely unique to the solvent free environment. Unfolding processes have been studied in both monomeric<sup>27, 28, 45</sup> and multimeric<sup>14, 44</sup> protein model systems as a means of providing information on local structure and stability changes that occur upon ligand binding<sup>29</sup>. The CIU strategy employed here is significantly modified from these previous reports, as we have optimized our ability to distinguish between different kinase inhibitors by including both unfolding and dissociative transitions in our CIU fingerprints (Figure 3.2). Ions are first generated by nESI in a range of charge states, which are then filtered with a quadrupole mass analyzer. For much of our final data set, we have chosen to focus on 11+ Abl:inhibitor complex ions because they typically provide a larger number of unfolding transitions due to the increased Coulombic strain on the gas-phase protein<sup>27</sup>. Following  $m/z$  selection, activation is achieved by accelerating ions into an ion trap pressurized with argon. Energetic collisions increase the internal temperature of the protein–ligand

complex and illicit unfolding transitions. At coincident energies, the protein also undergoes charge-stripping (through the dissociation of small, loosely bound, positively charged counterions) and ligand dissociation events that we also collect into our CIU fingerprint. Since the unfolding transitions for these systems still dominate the fingerprints recorded, we term these data “CIU Fingerprints”. We have extensively tracked the signals within these data sets to assign each transition observed (Appendix I), and including all three types of transitions shown in Figure 3.2 is critically important for differentiating kinase inhibitors using our approach. The total three-dimensional data set is then compiled by plotting the IM drift time features observed against the acceleration voltage used to generate them. We use a contour plot representation, as it allows us to track the relative intensities of features in our CIU fingerprints across the entire data set for a given protein–inhibitor complex, and thus allows us to easily focus on areas within a given data set that provide a maximized ability to distinguish between inhibitors of interest.

### **3.3.3 Identifying the Expected CIU Response for Kinase Inhibitors in Different Binding Modes**

In order to build, test, and evaluate our CIU method in the first instance, we selected two kinase inhibitors, one from each of the two types we intended to differentiate using our approach. For these initial experiments, we chose dasatinib and imatinib as our archetypal type I and II kinase inhibitors, respectively. Both molecules are approved for leukemia treatment, and previous structural biology and screening data has shown that they are among the clearest examples of their respective binding modes (Appendix I)<sup>46</sup>. CIU fingerprints for 10+ complex ions for both inhibitors in the presence of apo-Abl reveal striking differences (Appendix I). Three or four main CIU features are observed over the acceleration voltages shown, ranging from 9 to 13 ms, with imatinib-bound complexes displaying patterns where the feature at ~12 ms is entirely absent. Strikingly, the ligand-bound proteins are both destabilized relative to the apo form in the gas phase, which is unique among similar gas-phase protein–ligand stability measurements reported in the literature<sup>28, 29, 47, 48</sup>. Most importantly, we observe clear differences in the energetics of CIU, such that both ligand bound states can be differentiated from one another and the apo form simply by measuring the intensity of the compact state observed over a narrow window of collision voltages (dashed box regions, Appendix I). The signal intensity difference observed between imatinib and dasatinib-bound complexes for the most-compact conformer is approximately 2.5-fold (Appendix

I) and thus provides us with a promising level of dynamic range and differentiating power relative to the two inhibitor–kinase binding modes targeted in these experiments

### 3.3.4 Complete CIU Training Data Set from Known Kinase Inhibitors

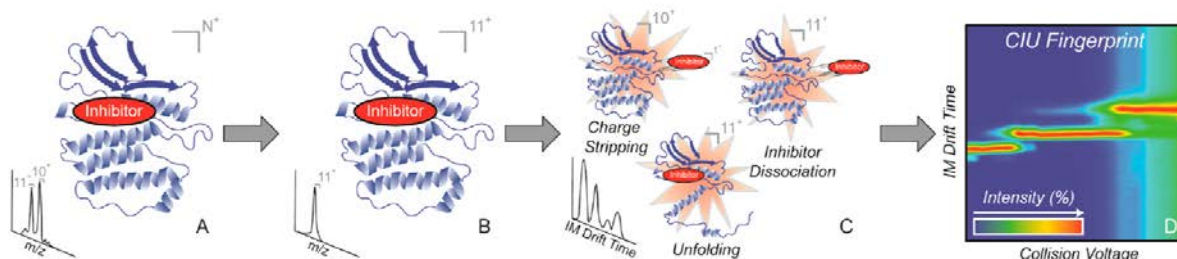


Figure 3.2 Schematic describing the basic steps involved in the CIU assay to determine kinase inhibitor binding modes. Protein ligand complexes are ionized by nESI in a range of charge states (A); a single charge state is selected for activation in a quadrupole mass filter (B). Following selection, collisional activation with argon (varying the amount of accelerating voltage) is used to initiate three different processes: charge stripping, inhibitor dissociation, and gas-phase protein unfolding (C). IM drift time is monitored and recorded over a broad range of collision voltages to create a complete CIU fingerprint, which creates a contour plot of the intensity of ion populations as a function of these two parameters (D).

Building on the initial data set shown in Appendix I, we have compiled a more complete CIU data set for  $11^+$  Abl-inhibitor complexes that includes 7 inhibitors previously classified by X-ray crystallography in complex with the Abl kinase (Figure 3.3). Within this data set, we note that there are a number of CIU regions that may be optimal to differentiate type I and II inhibitors. For example, the region between 22 and 26 V provides an IM spectrum analogous to the region between 32 and 42 V shown in Appendix I, where the low-energy transitions of compact kinase–inhibitor complexes are probed. We also note that the region between 30 and 36 V provides at least 3 drift time features, appearing between 10 and 14 ms that may be utilized for effective differentiation. We have extensively analyzed these two areas, along with the region between 40 and 44 V, highlighted in Figure 3.3A. This final region provides  $\sim 8$  drift time features between 8 and 14 ms for type II-bound Abl ions, while exhibiting 2 main drift time features for type I bound proteins. We have computed average drift time spectra for this region of the CIU fingerprints (Figure 3.3B) and, in addition to the dramatic differences in the number of peaks observed, all of the centroid values for the drift time peaks recorded within this range are exclusive to either type I or type II data sets. Thus, while not all IM drift time peaks observed in our data set are resolved, the dramatic differences in the number and centroid values for the features recorded between the

two binding types allow us to use this region alone to differentiate type I and II kinase inhibitors. This result is critically important for using CIU in a high-throughput screening mode, as while complete CIU fingerprint data collected over all the acceleration voltage ranges shown in Figures 3.3 and Appendix I, can take many minutes to acquire, drift time data over a narrow acceleration voltage range can be acquired in seconds.

### 3.3.5 Simple Scoring Approach Allows the Differentiation of Type I and Type II Inhibitors based on CIU Data

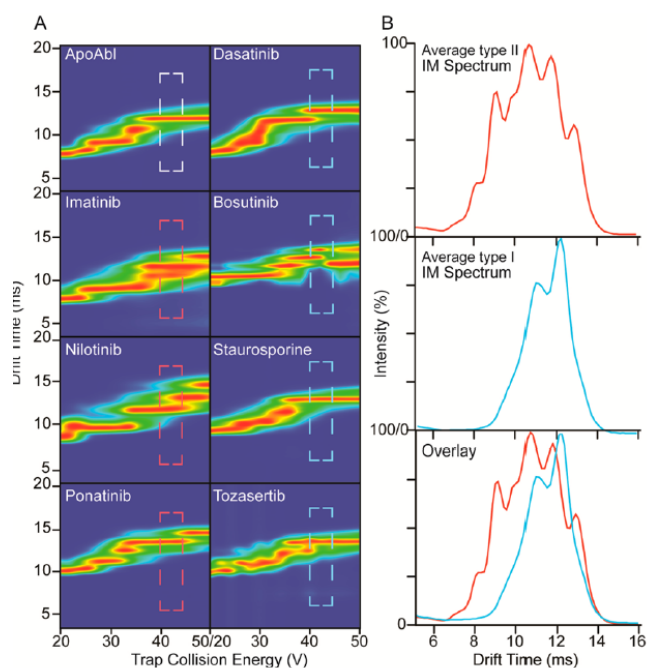


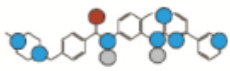
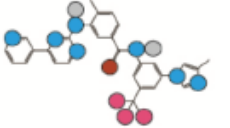
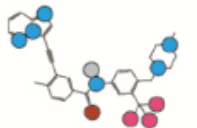

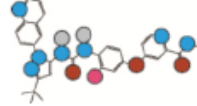
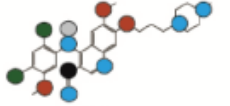


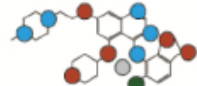
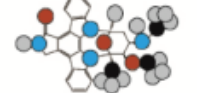

Figure 3.3 CIU fingerprints for indicated 11+ Abl-inhibitor complex ions (A). Dashed areas are between 40 and 44 V of acceleration voltage and are color-coded to correspond with the known binding mode of the inhibitor (as previously). Average IM drift time data derived from integrating the dashed-box regions from A (B). Data from selected types (I and II as indicated) are averaged, displayed, and overlaid below. Average IM drift time data in this region provide the greatest dissimilarity between inhibitor binding types, revealing sufficient detail to classify all of the inhibitors studied in this report.

While the data above points toward the ability of a CIU-based IM-MS approach to engage in a high-throughput screen containing hundreds or thousands of potential inhibitors, simplified metrics that measure the similarity or difference of an experimental data set when compared to reference data for either type I or type II inhibitors is necessary to enable such applications. We have evaluated the CIU data collected for 11 inhibitor-kinase complexes, comparing multiple CIU regions and applying a wide range of approaches to evaluate their similarity (Appendix I)<sup>1, 3, 5, 38</sup>. The optimized approach that we have developed uses the selected region shown in Figure 3.3 and a simple  $\chi^2$ -based score. In this way, we can evaluate the similarity of the IM drift time data recorded between 40 and

44 V for an individual inhibitor complex and the averaged IM drift time data for type II inhibitors shown in Figure 3.3B. Despite the broad chemical space that we have probed using our CIU approach, the simple  $\chi^2$ -based similarity score shown effectively clusters all type II and type I

binders based on their similarity to the average type II CIU response. The similarity scores shown in Table 3.1 range from 100 (a normalized score) for imatinib, which has the greatest similarity to the average type II CIU response within the selected acceleration voltage range, to 21.8 for dasatinib, which exhibits the least similarity to the type II response and, therefore, the most type I character in our CIU data set. The breadth of different scores, dispersed relatively evenly in the range defined by the two extreme responses recorded for imatinib and dasatinib-bound complexes, allows us to project that CIU data would allow for type I and II differentiation in an even broader chemical screen than shown here. Replicate experiments were performed and analyzed to provide standard deviations for the similarity scores shown, and these were found to represent 1.9% of the values on average.

Table 3.1 CIU-Based Similarity Scores for the Current Training Data Set of Type I and II Kinase Inhibitors

Drugs	Mass (Da)	Structure <sup>1</sup>	K <sub>d</sub> (nM) <sup>2</sup>	PDB ID	Type <sup>3</sup>	Type II Similarity Score
Imatinib	493.6		1.1	1IEP	II	100 +/-1.2
Nilotinib	529.2		10	3CS9	II	73.9 +/-0.5
Ponatinib	532.4		--	3OXZ	II	61.9 +/-0.9
Sorafenib	637.0		130	N/A	II	48.6 +/-1.0
DCC-2036	553.6		--	3QRI	II	43.4 +/-0.2
Bosutinib	529.2		0.12	3UE4	I	42.5 +/-1.2
Tozasertib	464.6		13	2F4J	I	42.2 +/-0.7
PP2	301.1		--	N/A	I	38.9 +/-0.9
Saracatinib	541.4		--	N/A	I	35.2 +/-0.7
Staurosporine	466.5		37	2HZ4	I	24.5 +/-1.5
Dasatinib	487.2		0.029	2GQG	I	21.8 +/-0.1

1 Atom types shown as colored circles: blue = nitrogen, gray = hydrogen, red = oxygen, pink = fluorine, green = chlorine, black = carbon, and yellow = sulfur. 2 Ligand dissociation constants can be found in ref 26. 3 Binding type derived from PDB entries showing the indicated small molecule bound to the Abl kinase. In cases where no Abl structure was available, data for similar kinases were used to confirm the binding type shown. References: PP2<sup>4</sup>; Saracatinib<sup>3</sup>; Sorafenib<sup>5</sup>.



### 3.4 Discussion

Most previous methodologies aimed at the rapid discovery of novel kinase inhibitors have not focused on differentiating the binding modes accessed by the small molecules screened and have favored instead increased speed and chemical space over additional information content on the bound species created<sup>49</sup>. More recently, viral-fusion  $K_d$  analysis<sup>41</sup>, fluorescence<sup>42</sup>, and affinity selection MS (AS-MS) approaches<sup>50</sup> have emerged with the ability to differentiate the binding modes of potential kinase inhibitors, with final verification of the inhibitor binding mode determined ultimately by solving the high-resolution structure of the complex<sup>7</sup>. For example, on the basis of the concentration-dependent binding curves observed for dissociated ligands captured following multi-stage chromatography, AS-MS can differentiate between inhibitors that bind in either a competitive, independent, or cooperative manner to the kinase. Generating such data, however, is time-consuming due to the chromatographic separation typically employed and the multiple concentration dependent runs that must be generated, making such detailed AS-MS data difficult to acquire in a screening mode<sup>50</sup>. While fluorescence-based approaches are capable of generating binding mode information rapidly during screening, they require covalent modification of the kinase and typically have a low (10%) dynamic range with respect to the type I and II kinase binding modes discussed here<sup>42</sup>. Similarly, viral fusion-based approaches<sup>39</sup> while enabling vast data sets of kinase–inhibitor interaction profiling and  $K_d$  measurements<sup>40</sup> require genetic fusion of the kinase to bacteriophage (or, in principle, other signal-amplifying tags) and display limited discriminating power for many type II inhibitors<sup>41</sup>. The IM-MS and CIU methodologies disclosed in this report mitigate many of the limitations described above. CIU provides information that is clearly correlated to the kinase–inhibitor binding mode, can operate from mixtures or in competitive assays, requires little starting material, the data can be accumulated rapidly if regions of interest are identified (as above) that adequately differentiate between the structural states of interest, and does not require covalent modification of the protein substrate prior to screening. Despite the small difference in CCS between the active and inactive forms of the kinase predicted computationally ( $\sim 1\%$ )<sup>51</sup>, CIU data allows for their clear differentiation and thus circumvents many of the traditional limitations of IM-MS where many protein tertiary structures overlap in

drift time and are difficult to resolve by their CCS alone (Appendix I; IM resolution of ~100 is required).

A comparison of the results shown in Tables 3.1 and Appendix I, reveals a clear advantage for performing CIU experiments using more highly charged ions. This observation is not surprising, as previous protein complex CID and CIU data sets have suggested similar trends. For example, charge amplification protocols applied to the 396 kDa 24-mer of HSP16.5 from *Methanococcus jannaschii*<sup>43</sup> and the boiling stable protein 12-mer from *Populus tremula*<sup>52</sup> have revealed substantially enhanced unfolding and dissociation. Detailed computational and IM studies have linked the unfolding transitions observed in CIU to charge migration, which in turn relies upon the total amount of charge on the protein surface. As such, charge amplification agents (i.e., sulfolane) may be useful in enhancing the CIU information content presented in this report<sup>53</sup> while acknowledging the potential for such additives to alter the ability of protein ions to access compact starting structures<sup>54</sup>. Critically, the 11+ charge state of Abl:inhibitor complexes is sufficient to effectively differentiate between the activation states of the protein complexes probed here, as well as demarcate a strong correlation between inhibitor binding mode and CIU response

The training data set used to validate the CIU method presented in this report contains 11 total inhibitors (5 type II, and 6 type I) selected in order to represent a broad range of structures, binding constants, and molecular masses. Inhibitors were also selected for this initial set based on the availability of high-resolution structural data confirming the binding mode accessed with either Abl or other kinase domains that possess a high degree of sequence homology (Table 1). The 11 compounds included here represent all the commercially available compounds that fit the above criteria. It is worth noting that the total number of compounds we have used in proof-of-principle experiments in this report is significantly greater than that used in previous technology development efforts. We have undertaken these extra steps primarily due to the novel nature of the CIU technology employed in our assay. A detailed analysis of the data presented in Table 1 reveals a number of inhibitors clustered around a similarity score of 43, and this value would be a likely cutoff point for future screening applications that utilize the CIU protocol described here. It is worth noting that, while using this similarity score as a cutoff may engender a significant false positive rate, on the basis of the fact that DCC-2036, bosutinib, and tozasertib all possess scores within error of this value, it is also clear that this rate can be adjusted to easily accommodate a

more-stringent screen without a significant loss of dynamic range. The average type II similarity score of 65.6 is 1.4 standard deviations away from the average type I value of 34.2. We find that type I values are more-tightly clustered in our data set, having a standard deviation of 9.0, which places the mean type I response 3.5 standard deviations from that of the average type II. We have used our current data set to extrapolate the potential results in a larger library screen (Appendix I) as well as investigate alternate scoring functions for our data set (Appendix I). In all cases, we find strong evidence of correlation between the CIU response described here and the known kinase–inhibitor binding modes for the complexes within our data set.

Limitations of the CIU approach currently revolve around the potential universality of the technology. In order to measure a CIU response, kinase–inhibitor complexes must first be generated by nESI, which is a property that may not be constant across all kinases and inhibitors. We note, however, that Abl and other related kinases have, so far, provided excellent signal intensities and dynamic range for our experiments. In addition, while the correlations reported here remain robust, the mechanism of CIU depends, in a relatively unknown way, on the structure of the intermediates generated during gas-phase activation. As such, example inhibitors must be identified in order to train the CIU methodology prior to screening a new protein target, and complete CIU fingerprints of the desired binding targets must be obtained initially so that regions of discriminatory power can be identified.

### **3.5 Conclusions**

In summary, we present strong evidence indicating a marked correlation between the gas-phase unfolding of kinase–inhibitor complexes and their known activation states when bound to the same inhibitors in solution. Our assay integrates elements of CID and CIU for the first time to maximize the discriminatory power of the IM-MS data observed. We utilized CIU fingerprint analysis to identify regions of maximum difference between Abl kinase ions bound to type I and II inhibitors, and developed a simple scoring metric that clusters these data in a manner precisely correlated with their known binding modes. Overall, our data suggests that the CIU approach presented will likely be a highly effective screening tool that obviates many of the limitations of current technologies. Future CIU development work will seek to use larger screening data sets to accurately define the discrimination power and confidence intervals associated with type I and II

kinase binders, investigate the ability of CIU to screen for allosteric (type IV) kinase inhibitors (e.g., GNF-2)<sup>49</sup>, and develop screens for other kinases linked to cancer (e.g., Src)<sup>30</sup>.

### 3.6 References

- [1] Muratore, K. E., Seeliger, M. A., Wang, Z., Fomina, D., Neiswinger, J., Havranek, J. J., Baker, D., Kuriyan, J., and Cole, P. A. (2009) Comparative Analysis of Mutant Tyrosine Kinase Chemical Rescue<sup>†‡</sup>, *Biochemistry* 48, 3378-3386.
- [2] Krause, D. S., and Van Etten, R. A. (2005) Tyrosine kinases as targets for cancer therapy, *New England Journal of Medicine* 353, 172-187.
- [3] Hennequin, L. F., Allen, J., Breed, J., Curwen, J., Fennell, M., Green, T. P., Lambert-van der Brempt, C., Morgentin, R., Norman, R. A., and Olivier, A. (2006) N-(5-chloro-1, 3-benzodioxol-4-yl)-7-[2-(4-methylpiperazin-1-yl) ethoxy]-5-(tetrahydro-2 h-pyran-4-yloxy) quinazolin-4-amine, a novel, highly selective, orally available, dual-specific c-Src/Abl kinase inhibitor, *Journal of medicinal chemistry* 49, 6465-6488.
- [4] Deininger, M. W., Goldman, J. M., and Melo, J. V. (2000) The molecular biology of chronic myeloid leukemia, *Blood* 96, 3343-3356.
- [5] Namboodiri, H. V., Bukhtiyarova, M., Ramcharan, J., Karpusas, M., Lee, Y., and Springman, E. B. (2010) Analysis of imatinib and sorafenib binding to p38 $\alpha$  compared with c-Abl and b-Raf provides structural insights for understanding the selectivity of inhibitors targeting the DFG-out form of protein kinases, *Biochemistry* 49, 3611-3618.
- [6] Liu, Y., and Gray, N. S. (2006) Rational design of inhibitors that bind to inactive kinase conformations, *Nature chemical biology* 2, 358-364.
- [7] Thaimattam, R., Banerjee, R., Miglani, R., and Iqbal, J. (2007) Protein kinase inhibitors: structural insights into selectivity, *Current pharmaceutical design* 13, 2751-2765.
- [8] Bixby, D., and Talpaz, M. (2009) Mechanisms of resistance to tyrosine kinase inhibitors in chronic myeloid leukemia and recent therapeutic strategies to overcome resistance, *ASH Education Program Book 2009*, 461-476.
- [9] Giles, K., Williams, J. P., and Campuzano, I. (2011) Enhancements in travelling wave ion mobility resolution, *Rapid Communications in Mass Spectrometry* 25, 1559-1566.

- [10] Zhong, Y., Hyung, S.-J., and Ruotolo, B. T. (2011) Characterizing the resolution and accuracy of a second-generation traveling-wave ion mobility separator for biomolecular ions, *Analyst* 136, 3534-3541.
- [11] Clemmer, D. E., and Jarrold, M. F. (1997) Ion mobility measurements and their applications to clusters and biomolecules, *Journal of Mass Spectrometry* 32, 577-592.
- [12] McLean, J. A., Ruotolo, B. T., Gillig, K. J., and Russell, D. H. (2005) Ion mobility–mass spectrometry: a new paradigm for proteomics, *International Journal of Mass Spectrometry* 240, 301-315.
- [13] von Helden, G., Wyttenbach, T., and Bowers, M. T. (1995) Conformation of macromolecules in the gas phase: use of matrix-assisted laser desorption methods in ion chromatography, *Science* 267, 1483.
- [14] Ruotolo, B. T., Giles, K., Campuzano, I., Sandercock, A. M., Bateman, R. H., and Robinson, C. V. (2005) Evidence for macromolecular protein rings in the absence of bulk water, *Science* 310, 1658-1661.
- [15] Zhong, Y., Hyung, S.-J., and Ruotolo, B. T. (2012) Ion mobility–mass spectrometry for structural proteomics, *Expert review of proteomics* 9, 47-58.
- [16] Benesch, J. L., and Ruotolo, B. T. (2011) Mass spectrometry: come of age for structural and dynamical biology, *Current opinion in structural biology* 21, 641-649.
- [17] Ruotolo, B. T., and Robinson, C. V. (2006) Aspects of native proteins are retained in vacuum, *Current opinion in chemical biology* 10, 402-408.
- [18] Jarrold, M. F. (2007) Helices and Sheets in vacuoThe HTML version of this article has been enhanced with colour images, *Physical Chemistry Chemical Physics* 9, 1659-1671.
- [19] Faull, P. A., Korkeila, K. E., Kalapothakis, J. M., Gray, A., McCullough, B. J., and Barran, P. E. (2009) Gas-phase metalloprotein complexes interrogated by ion mobility-mass spectrometry, *International Journal of Mass Spectrometry* 283, 140-148.
- [20] Wyttenbach, T., Grabenauer, M., Thalassinou, K., Scrivens, J. H., and Bowers, M. T. (2009) The effect of calcium ions and peptide ligands on the relative stabilities of the calmodulin dumbbell and compact structures, *The Journal of Physical Chemistry B* 114, 437-447.
- [21] Mao, Y., Ratner, M. A., and Jarrold, M. F. (1999) Molecular dynamics simulations of the charge-induced unfolding and refolding of unsolvated cytochrome c, *The Journal of Physical Chemistry B* 103, 10017-10021.

- [22] Tao, L., Dahl, D. B., Pérez, L. M., and Russell, D. H. (2009) The contributions of molecular framework to IMS collision cross-sections of gas-phase peptide ions, *Journal of the American Society for Mass Spectrometry* 20, 1593-1602.
- [23] Segev, E., Wytttenbach, T., Bowers, M. T., and Gerber, R. B. (2008) Conformational evolution of ubiquitin ions in electrospray mass spectrometry: molecular dynamics simulations at gradually increasing temperatures, *Physical Chemistry Chemical Physics* 10, 3077-3082.
- [24] Baumketner, A., Bernstein, S. L., Wytttenbach, T., Bitan, G., Teplow, D. B., Bowers, M. T., and Shea, J. E. (2006) Amyloid  $\beta$ -protein monomer structure: A computational and experimental study, *Protein Science* 15, 420-428.
- [25] Suckau, D., Shi, Y., Beu, S. C., Senko, M. W., Quinn, J. P., Wampler, F., and McLafferty, F. W. (1993) Coexisting stable conformations of gaseous protein ions, *Proceedings of the National Academy of Sciences* 90, 790-793.
- [26] Wittmer, D., Chen, Y. H., Luckenbill, B. K., and Hill Jr, H. H. (1994) Electrospray ionization ion mobility spectrometry, *Analytical Chemistry* 66, 2348-2355.
- [27] Shelimov, K. B., Clemmer, D. E., Hudgins, R. R., and Jarrold, M. F. (1997) Protein structure in vacuo: Gas-phase conformations of BPTI and cytochrome c, *Journal of the American Chemical Society* 119, 2240-2248.
- [28] Hopper, J. T., and Oldham, N. J. (2009) Collision induced unfolding of protein ions in the gas phase studied by ion mobility-mass spectrometry: the effect of ligand binding on conformational stability, *Journal of the American Society for Mass Spectrometry* 20, 1851-1858.
- [29] Hyung, S.-J., Robinson, C. V., and Ruotolo, B. T. (2009) Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes, *Chemistry & biology* 16, 382-390.
- [30] Seeliger, M. A., Nagar, B., Frank, F., Cao, X., Henderson, M. N., and Kuriyan, J. (2007) c-Src binds to the cancer drug imatinib with an inactive Abl/c-Kit conformation and a distributed thermodynamic penalty, *Structure* 15, 299-311.
- [31] Ruotolo, B. T., Benesch, J. L., Sandercock, A. M., Hyung, S.-J., and Robinson, C. V. (2008) Ion mobility-mass spectrometry analysis of large protein complexes, *Nature Protocols* 3, 1139-1152.

- [32] Benesch, J. L., Ruotolo, B. T., Simmons, D. A., and Robinson, C. V. (2007) Protein complexes in the gas phase: technology for structural genomics and proteomics, *Chemical reviews* 107, 3544-3567.
- [33] Hernández, H., and Robinson, C. V. (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry, *Nature protocols* 2, 715-726.
- [34] Pearson, K. (1900) X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 50, 157-175.
- [35] Park, J. W., Song, J. Y., Lee, S. G., Jun, J. S., Park, J. U., Chung, M. J., Ju, J. S., Nizamutdinov, D., Chang, M. W., and Youn, H. S. (2006) Quantitative analysis of representative proteome components and clustering of *Helicobacter pylori* clinical strains, *Helicobacter* 11, 533-543.
- [36] Roth, O., Spreux-Varoquaux, O., Bouchet, S., Rousselot, P., Castaigne, S., Rigauddau, S., Ragueneau, V., Therond, P., Devillier, P., and Molimard, M. (2010) Imatinib assay by HPLC with photodiode-array UV detection in plasma from patients with chronic myeloid leukemia: Comparison with LC-MS/MS, *Clinica Chimica Acta* 411, 140-146.
- [37] Xue, R., Dong, L., Zhang, S., Deng, C., Liu, T., Wang, J., and Shen, X. (2008) Investigation of volatile biomarkers in liver cancer blood using solid-phase microextraction and gas chromatography/mass spectrometry, *Rapid communications in mass spectrometry* 22, 1181-1186.
- [38] Davis, M. I., Hunt, J. P., Herrgard, S., Ciceri, P., Wodicka, L. M., Pallares, G., Hocker, M., Treiber, D. K., and Zarrinkar, P. P. (2011) Comprehensive analysis of kinase inhibitor selectivity, *Nature biotechnology* 29, 1046-1051.
- [39] Fabian, M. A., Biggs, W. H., Treiber, D. K., Atteridge, C. E., Azimioara, M. D., Benedetti, M. G., Carter, T. A., Ciceri, P., Edeen, P. T., and Floyd, M. (2005) A small molecule–kinase interaction map for clinical kinase inhibitors, *Nature biotechnology* 23, 329-336.
- [40] Karaman, M. W., Herrgard, S., Treiber, D. K., Gallant, P., Atteridge, C. E., Campbell, B. T., Chan, K. W., Ciceri, P., Davis, M. I., and Edeen, P. T. (2008) A quantitative analysis of kinase inhibitor selectivity, *Nature biotechnology* 26, 127-132.

- [41] Wodicka, L. M., Ciceri, P., Davis, M. I., Hunt, J. P., Floyd, M., Salerno, S., Hua, X. H., Ford, J. M., Armstrong, R. C., and Zarrinkar, P. P. (2010) Activation state-dependent binding of small molecule kinase inhibitors: structural insights from biochemistry, *Chemistry & biology* 17, 1241-1249.
- [42] Simard, J. R., Klüter, S., Grütter, C., Getlik, M., Rabiller, M., Rode, H. B., and Rauh, D. (2009) A new screening assay for allosteric inhibitors of cSrc, *Nature chemical biology* 5, 394-396.
- [43] Benesch, J. L. (2009) Collisional activation of protein complexes: picking up the pieces, *Journal of the American Society for Mass Spectrometry* 20, 341-348.
- [44] Ruotolo, B. T., Hyung, S. J., Robinson, P. M., Giles, K., Bateman, R. H., and Robinson, C. V. (2007) Ion Mobility–Mass Spectrometry Reveals Long-Lived, Unfolded Intermediates in the Dissociation of Protein Complexes, *Angewandte Chemie International Edition* 46, 8001-8004.
- [45] Badman, E. R., Hoaglund-Hyzer, C. S., and Clemmer, D. E. (2001) Monitoring Structural Changes of Proteins in an Ion Trap over~ 10-200 ms: Unfolding Transitions in Cytochrome c Ions, *Analytical chemistry* 73, 6000-6007.
- [46] Vajpai, N., Strauss, A., Fendrich, G., Cowan-Jacob, S. W., Manley, P. W., Grzesiek, S., and Jahnke, W. (2008) Solution conformations and dynamics of ABL kinase-inhibitor complexes determined by NMR substantiate the different binding modes of imatinib/nilotinib and dasatinib, *Journal of biological chemistry* 283, 18292-18302.
- [47] Han, L., Hyung, S. J., and Ruotolo, B. T. (2012) Bound cations significantly stabilize the structure of multiprotein complexes in the gas phase, *Angewandte Chemie International Edition* 51, 5692-5695.
- [48] Han, L., Hyung, S.-J., Mayers, J. J., and Ruotolo, B. T. (2011) Bound anions differentially stabilize multiprotein complexes in the absence of bulk solvent, *Journal of the American Chemical Society* 133, 11358-11367.
- [49] Adrián, F. J., Ding, Q., Sim, T., Velentza, A., Sloan, C., Liu, Y., Zhang, G., Hur, W., Ding, S., and Manley, P. (2006) Allosteric inhibitors of Bcr-abl–dependent cell proliferation, *Nature chemical biology* 2, 95-102.
- [50] Annis, D. A., Nazef, N., Chuang, C.-C., Scott, M. P., and Nash, H. M. (2004) A general technique to rank protein-ligand binding affinities and determine allosteric versus direct



- binding site competition in compound mixtures, *Journal of the American Chemical Society* **126**, 15495-15503.
- [51] Mesleh, M., Hunter, J., Shvartsburg, A., Schatz, G. C., and Jarrold, M. (1996) Structural information from ion mobility measurements: effects of the long-range potential, *The Journal of Physical Chemistry* **100**, 16082-16086.
- [52] Erba, E. B., Ruotolo, B. T., Barsky, D., and Robinson, C. V. (2010) Ion mobility-mass spectrometry reveals the influence of subunit packing and charge on the dissociation of multiprotein complexes, *Analytical chemistry* **82**, 9702-9710.
- [53] Lomeli, S. H., Peng, I. X., Yin, S., Loo, R. R. O., and Loo, J. A. (2010) New reagents for increasing ESI multiple charging of proteins and protein complexes, *Journal of the American Society for Mass Spectrometry* **21**, 127-131.
- [54] Hogan Jr, C. J., Loo, R. R. O., Loo, J. A., and de la Mora, J. F. (2010) Ion mobility-mass spectrometry of phosphorylase B ions generated with supercharging reagents but in charge-reducing buffer, *Physical Chemistry Chemical Physics* **12**, 13476-13483.

## **Chapter 4.**

### **Collision Induced Unfolding and Dissociation Differentiates ATP-Competitive from Allosteric Protein Tyrosine Kinase Inhibitors**

Chronic myeloid leukemia is caused by a chimeric oncoprotein comprised of the breakpoint cluster region protein and the Abelson protein tyrosine kinase, and is often treated using protein tyrosine kinase inhibitors that target the ATP-binding domain of the kinase. A more recently-discovered class of inhibitors target the myristoylation site in the C-terminal region of the kinase domain, and offer allosteric control over kinase activity. The discovery process surrounding such inhibitors, however, is exceptionally challenging as allosteric sites are frequently not apparent with X-ray structures and, if they are known, their role in kinase function is not always obvious. As such, there is currently a general lack of assays capable of identifying allosteric kinase inhibitors. In this report, we describe an assay for detecting allosteric ligand binding in the Abelson protein kinase domain using gas-phase protein unfolding and dissociation. Our data strongly differentiates between ATP-competitive and myristate pocket-binding ligands through analyses of both gas-phase unfolding patterns and dissociative charge stripping observed in kinase-inhibitor complex ions produced by electrospray ionization. Furthermore, we discuss and quantify these results, as well as project the utility of this technique for future efforts in kinase inhibitor discovery.

#### **4.1 Introduction**

The Abelson protein tyrosine kinase (Abl) is a prime pharmacological target due to its propensity to form an oncogenic chimeric protein with the breakpoint cluster region protein (BCR-Abl). BCR-Abl causes chronic myeloid leukemia (CML), which accounts for ~15% of all leukemias worldwide<sup>1</sup>. There are currently five kinase inhibitors approved by the FDA to treat CML. These inhibitors, while all ATP-competitive, fall into two classes: type I and type II, which respectively

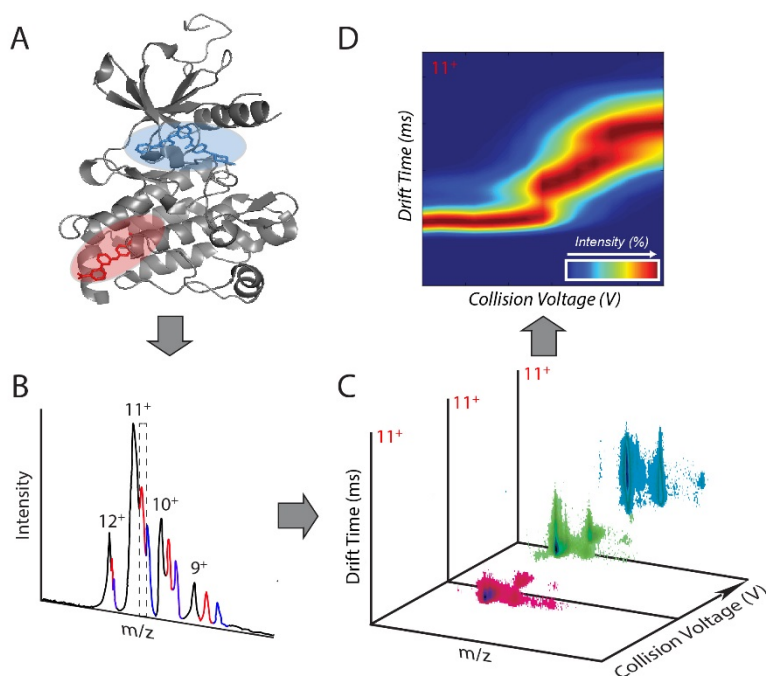


Figure 4.1. Figure 4.1 illustrates the experimental workflow for these experiments. We selected a panel of type I and type II inhibitors from our previous screen<sup>[1]</sup>. We then selected GNF-2, GNF-5, and myristic acid for our model allosteric inhibitors. While the availability of allosteric inhibitors is limited, this inhibitor set exhibits a range of molecular weights and structures. Ligand binding was confirmed by MS analysis (Figure 4.1B), after which the 11+ charge state was selected for further analysis. The 11+ charge state was selected based on the number of conformational transitions observed in our previous kinase dataset, in which data extracted from kinase-inhibitor complexes of this charge state were determined to be optimal for the differentiation of ATP-competitive inhibitors<sup>[6]</sup>. The selected ions were then activated by sequentially increasing the acceleration voltage used to propel kinase-inhibitor complex ions into the trap region that precedes the IM separator (Figure 4.1C). Through increasing this collision voltage, we were able to raise the internal energy of the ions and observe their unfolding. The drift time distributions of activated ions were then separated by IM, informing on the number of and relative populations of unfolded states created at each activation level. At lower collision energies, the drift time distribution appeared narrow, and at higher collision energies, the drift time distribution took on broadened and multi-modal qualities, indicating gas-phase unfolding. We created CIU fingerprints by plotting IM drift times against collision voltages (Figure 4.1D)<sup>[3]</sup>. The CIU fingerprints reported here are composed of 3 species that are generated from the 11+ ions originally isolated: the 11+ ligand:protein complex, the 11+ apo protein caused by neutral ligand loss, and the 10+ apo protein that results from charge stripping as the inhibitor dissociates from the protein as a singly charged ion

stabilize an active or an inactive kinase conformation (Figure 4.1A, blue). Mutations in the kinase domain of Abl often abrogate drug binding by introducing bulky, hydrophobic amino acids into the binding site of the drug or cause a conformational change in the kinase that is unfavorable for inhibitor binding. Understandably, drug discovery efforts have therefore started to focus on small molecule inhibitors that target unique allosteric sites within specific kinases in order to drive selectivity and potency.

The discovery of most allosteric inhibitors have occurred serendipitously<sup>2</sup>. For example, two allosteric inhibitors for Abl were discovered after exhibiting activity in an *in vivo* screen against BCR-Abl but failing in a follow up *in vitro* assay against the kinase domain of Abl. Crystal structures revealed

that these two inhibitors, GNF-2 and GNF-5, bind to the allosteric myristoyl binding site of the Abl kinase domain (Figure 4.1A, red)<sup>3, 4</sup>. Further studies indicated that these allosteric inhibitors

stabilize the  $\alpha$ C-helix out conformation, in which a salt bridge is formed between the catalytic lysine and a conserved glutamic acid on the  $\alpha$ C helix<sup>4, 5</sup>. This movement has been shown to rescue the closed conformation in the three-domain structure of Abl, which is composed of the regulatory SH3 and SH2 domains in addition to the kinase domain. When used in combination with type II inhibitors, have been shown to suppress resistance of mutations *in vitro*<sup>2, 6</sup>. Furthermore, as these inhibitors target non-conserved binding sites unique to the Abl kinase, they are highly selective and remain unaffected by mutations local to the primary ATP-binding site. Therefore, recent efforts in kinase drug discovery have focused on finding additional allosteric inhibitors<sup>2</sup>.

Previously, we developed a conformationally-sensitive ion mobility-mass spectrometry (IM-MS) assay that differentiates type I and type II ATP-competitive kinase inhibitors<sup>7</sup>. IM-MS measures the mass and orientationally-averaged size (collision cross section, CCS) of analyte ions as generated by nano-electrospray ionization (nESI) from native-like, aqueous buffers. This technique has the benefits of low limits of detection for protein complexes, minimal requirements for sample purity, and no requirements for protein labeling or modification. Protein conformations that differ by as little as 2-3% in terms of their CCS can be resolved by IM-MS<sup>8</sup>, but by introducing gas-phase activation in the form of collision induced unfolding (CIU, analogous to a gas-phase calorimetry experiment), subtle differences among protein:ligand complexes undetectable by CCS measurement alone become apparent<sup>7, 9, 10</sup>. IM-MS and CIU have been extensively used to distinguish binding modes of ligands to proteins<sup>10-12</sup>, probe the polydispersity of intrinsically disordered proteins<sup>11, 12</sup>, in drug discovery efforts in amyloid formation<sup>13-16</sup>, and in biosimilar analysis<sup>17-19</sup>. Here, we report the first use of a CIU-based screening method to discriminate between ATP-competitive and allosteric inhibitors using the Abl kinase domain as a model system.

## 4.2 Experimental Methods

### 4.2.1 General

Imatinib, dasatinib, and tozasertib were purchased from LC laboratories (Woburn, MA). GNF-5, GNF-2, and DCC-2036 were purchased from SelleckChem (Houston, TX), and myristic acid was purchased from Sigma (St. Louis, MO). Protein samples were buffer exchanged into 200

mM ammonium acetate at pH 7 (Sigma, St. Louis, MO) using Micro Bio-Spin 6 columns (Bio-Rad, Hercules, CA) for a final concentration of 6  $\mu$ M. Inhibitors were added at a 1:1 molar ratio of Abl:inhibitor, and the samples were incubated at room temperature for 1 hr. After incubation, the samples were moved to ice until analysis.

#### **4.2.2 Protein Expression and Purification**

Abl DNA was synthesized by GeneArt (Life Technologies, Grand Island, NY) using *E. coli* modified codons and subcloned into pET28a with a modified TEV-protease cleavable N-terminal 6x-His tag. The plasmid was transformed by electroporation into BL21 DE3 electrochemically competent cells with a YopH in pCDFDuet-1. Cell growth, protein expression, and purification were performed using modified literature protocols previously developed for the c-Src kinase domain<sup>20</sup> without cleavage of the His-tag.

#### **4.2.3 Ion Mobility-Mass Spectrometry**

Sample aliquots (5-7  $\mu$ l) were analyzed by IM-MS on a nano-electrospray ionization-quadrupole-ion mobility-time-of-flight mass spectrometer (nESI-Q-IM-TOF MS) instrument (Synapt G2 HDMS, Waters, Milford, Ma) as described previously. The capillary voltage ranged from 1.5-1.7 kV. The source was operated in positive ion mode with a sample cone operating at 50V and the quad profile was set to dwell on 3400. The traveling wave ion guide was pressurized at  $2.88 \times 10^{-2}$  mbar of argon gas, and the wave height and wave velocity were set to 40V and 900 m/s respectively. The TOF-MS was operated at a pressure of at  $1.48 \times 10^{-6}$  mbar over an  $m/z$  range of 1000-8000. The protein and inhibitor concentrations were sufficiently low so as to not induce artifact complexes, and all data collected conformed to the expected 1:1 inhibitor binding stoichiometry.

Prior to ion mobility separation, the ions were activated in the trap traveling-wave ion guide to perform CIU of protein complexes as previously described. Each mass selected ion was activated by increasing the trap collision voltage in steps of 2V from 20V to 50V. All mass spectra were calibrated externally using a solution of cesium iodide (20 mg/ml) and were processed using the Python-based CIUSuite software package<sup>21</sup>, which employs a z-score-based scoring metric to classify inhibitors according to user-defined groups.

### 4.3 Comparison of the CIU Response of ATP-Competitive and Allosteric Abl Kinase Inhibitors

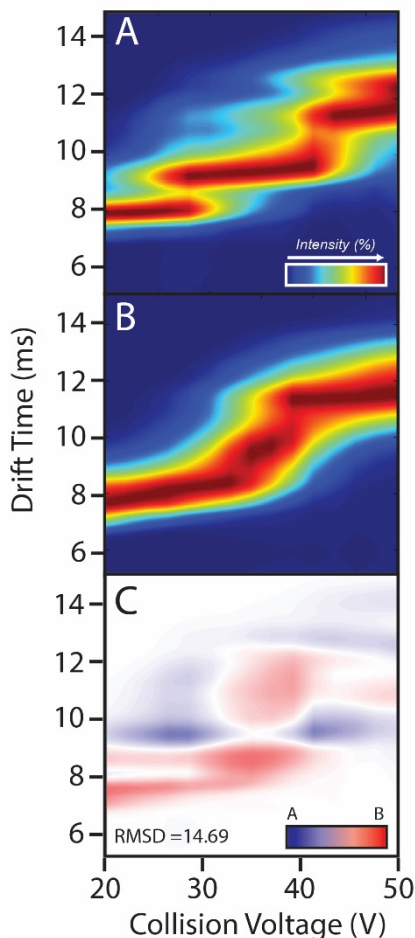


Figure 4.2. (A) Average ATP-competitive and (B) average allosteric CIU fingerprints. (C) RMSD analysis of the average fingerprints in (A). Red indicates areas where allosteric inhibitors have data density, and blue indicates areas where ATP-competitive inhibitors.

A comparison of the average CIU fingerprints for ATP-competitive and allosteric inhibitor datasets yield striking differences. The ATP-competitive inhibitors broadly exhibit four CIU features, centered at 8.09, 9.36, 11.55, and 12.67 ms, whereas allosteric inhibitors exhibit three structural families during CIU, centered on 8.73, 9.29, 11.48 ms (Figure 4.2B). A difference plot reveals many areas of difference between the two average CIU fingerprints, with a RMSD of 14.69%. This is almost three times the 5% RMSD obtained for individual sample replicates ( $n=3$ ) (Figure 4.2B). These data indicate that the average CIU fingerprints are different.

In order to demonstrate the potential of CIU data in downstream screens for allosteric inhibitors, we created a scaled deviation score (SDS) for all the inhibitors

used in this study normalized against the allosteric inhibitor CIU average plotted against the collision voltage used to achieve the indicated level of CIU (Figure 4.3C). Each point within the figure represents an average value for all inhibitors measured in the indicated class, and the error bars shown represent the standard deviation of SDS scores represented in each class. Our data indicate that the two sets of inhibitors are maximally differentiated at 38-40V (Figure 4.3A green dashed box). In addition to SDS, we computed z-scores for the dataset shown in Figure 4.3A,

finding that allosteric inhibitors produce a z-score of greater than 3, while data for all ATP-competitive inhibitors results in a z-score of near zero (Figure 4.4).

#### 4.4 CIU versus CID for Differentiating Kinase Inhibitors

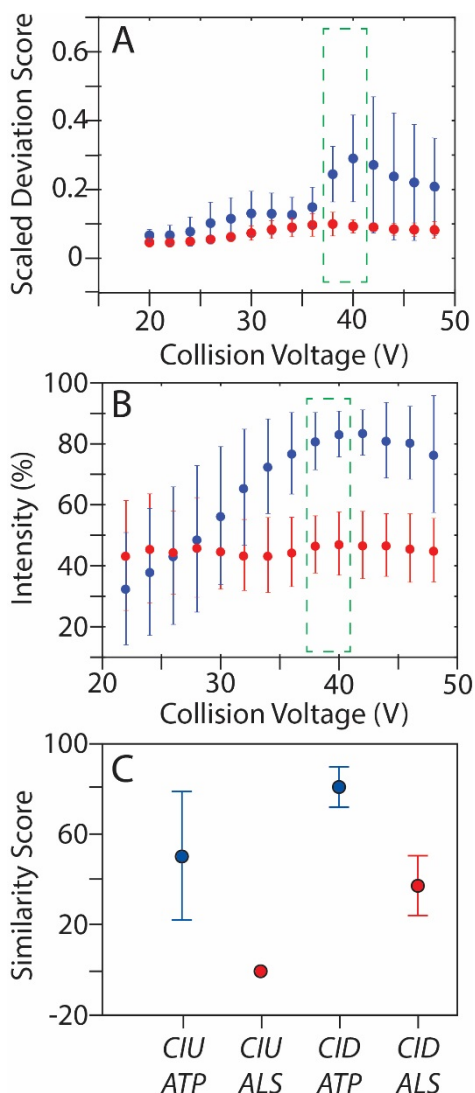


Figure 4.3. (A) Scaled deviation score (SDS) analysis of ATP-competitive and allosteric inhibitors against an allosteric average score. By plotting the SDS against TCV, areas where the two inhibitor types are easily differentiated (green box). (D) Comparison of the z-scores compared to an allosteric z-score and ATP-competitive z-score further shows differentiation and good grouping between the two inhibitor types. (E) Analysis of the amount of 10+ charge stripped product reveals reproducible differences between the two types of inhibitors and is partially responsible for the large differences between the two types of inhibitors.

The large differences between the two classes of CIU fingerprints studied here is surprising, given that the allosteric inhibitors and type II inhibitors cause a similar, but not identical, conformational change in the Abl kinase domain<sup>3, 5</sup>. These differences can be attributed, in part, to the different propensities that the two classes of inhibitors promote charge stripping due to collision induced ejection of the charged ligand upon activation. Following the selection of 11+ complex ions, this process results in lesser amounts of the 10+ apo species in mass spectra for allosteric inhibitors than for Abl bound to ATP-competitive inhibitors (Figure 4.3). For example, at a collision voltage value of 34V, the charge stripped species associated with ATP-competitive inhibitor-complexes occupies, on average, >60% of the total ion intensity. Conversely, the charged stripped species of the allosteric inhibitors only occupies ~40% of the total ion intensity at that same voltage (Figure 4.3B). These results likely reflect the generally greater hydrophobicity of the allosteric inhibitors

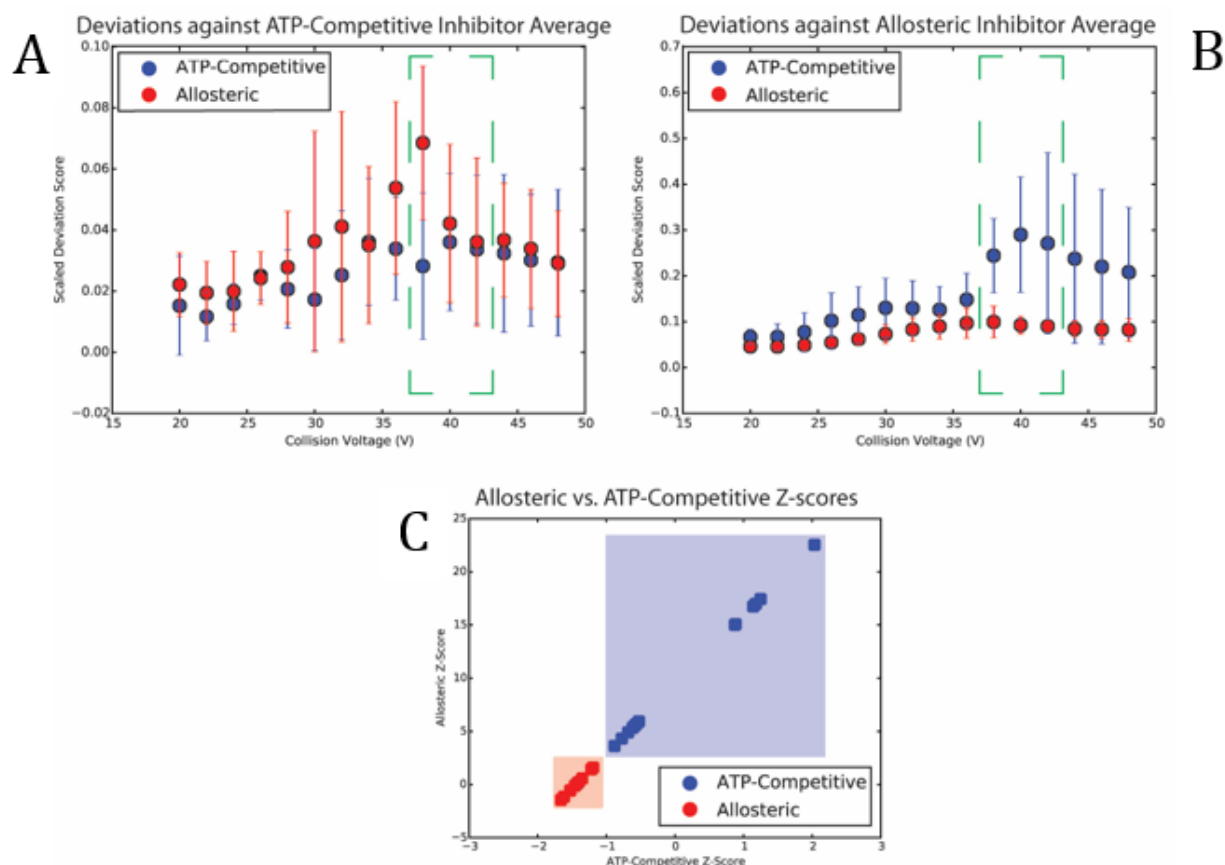


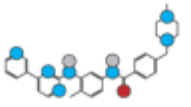
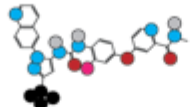



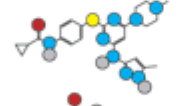

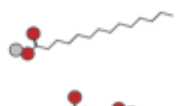

Figure 4.4. Comparison of all inhibitors using CIUSuite. (A) SDS against the average ATP-competitive inhibitor average. (B) SDS against an allosteric inhibitor average. (C) Z-scores against an ATP-competitive average (x-axis) and an allosteric z-score (y-axis). The allosteric and ATP-competitive inhibitors cluster well together, although the two classes of inhibitors are the best separated along the allosteric z-score axis. The boxes represent the standard deviations for each class of inhibitor along each axis. All ATP-competitive inhibitors are in blue, and the allosteric inhibitors are in red.

studied here when compared to those that are ATP-competitive, an extension of the hydrophobic myristate site itself, rendering the latter class of inhibitors more able to accept charge for ejection via CID.

As the two classes of inhibitors were clearly differentiated using both CIU and CID for charge stripped complexes, we developed final scores for inhibitors according to both a normalized CIU z-score corresponding to data acquired at collision voltages between 38 and 40V and a normalized score based on the percentage of the total MS signal intensity carried by 10+ apo ions at a collision voltages between 38 and 40V (CID score). The average CIU score for ATP-competitive inhibitors is  $50.01 \pm 28.02$  and  $-0.34 \pm 1.89$  for allosteric (Table 4.1).



Table 4.1 CIU-based similarity scores for current training dataset of ATP-Competitive (type ATP) and allosteric (type IV) inhibitors.

Drugs	Mass (Da)	Structure <sup>1</sup>	K <sub>d</sub> (nM) <sup>2</sup>	PDB ID	type	Normalized ATP similarity score
Imatinib	493.60		1.1	1IEP	ATP	98.69±1.28
DCC-2036	553.59		--	3QRI	ATP	89.53±5.64
Dasatinib	487.16		0.029	2GQG	ATP	33.14±1.08
Ponatinib	532.56		--	3OXZ	ATP	31.07±4.17
Bosutinib	529.16		0.12	3UE4	ATP	31.27±0.84
Tozasertib	464.59		13	2F4J	ATP	23.46±2.29
GNF-2	374.32		--	3K5V	IV	8.04±1.10
Myristic Acid	228.37		--	2OPJ	IV	1.98±1.84
GNF-5	418.37		--	--	IV	2.58±1.11


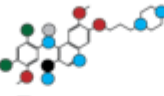


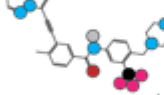


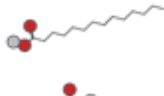

1. Hetero-atoms and some functional groups shown as colored circles: blue = nitrogen, grey = hydrogen, red= oxygen, pink = fluorine, green = chlorine, black = carbon, yellow = sulfur.

Equally, ATP-competitive inhibitor complexes give rise to a CID score of  $80.32 \pm 8.83$ , whereas complexes comprised of allosteric inhibitors generate a CID score of  $37.02 \pm 13.01$  (Table 4.2). Based on this analysis, as above, both CID and CIU scores generate significantly

different clusters of data correlated with the binding site of the inhibitors within the Abl kinase ions probed. Interestingly, the CIU results produce a very narrow distribution of scores for allosteric inhibitor complexes, while revealing a broader distribution for ATP-competitive inhibitor complexes. This result seems to mirror the expected distributions of these two general classes of protein ligand complexes, as ATP-competitive inhibitors are known to occupy at least two significantly different binding modes. As such, CIU scores seem to better reflect the subtle conformational dynamics that underscore the kinase-ligand binding process (Table 4.1). Such differences appear to be absent in our CID scores, likely reflecting a reliance on ligand hydrophobicity, which happens to be correlated with allosteric binding in Abl (Table 4.2).

## 4.5 Conclusions

Table 4. 2. CID-based similarity scores for current training dataset of ATP-competitive (type ATP) and Allosteric (type IV) inhibitors.

Drugs	Mass (Da)	Structure <sup>1</sup>	Kd (nM) <sup>2</sup>	PDB ID	type	Percent 10+ at 42V
Tozasertib	464.59		13	2F4J	ATP	90.78±0.11
Bosutinib	529.16		0.12	3UE4	ATP	88.95±0.47
DCC-2036	553.59		--	3QRI	ATP	84.50±3.04
Dasatinib	487.16		0.029	2GQG	ATP	74.63±0.48
Ponatinib	532.56		--	3OXZ	ATP	72.26±9.05
Imatinib	493.60		1.1	1IEP	ATP	70.79±5.19
GNF-5	418.37		--	--	IV	51.57±12.16
Myristic Acid	228.37		--	2OPJ	IV	33.24±15.36
GNF-2	374.32		--	3K5V	IV	26.24±27.18

1. Hetero-atoms and some functional groups shown as colored circles: blue = nitrogen, grey = hydrogen, red= oxygen, pink = fluorine, green = chlorine, black = carbon, yellow = sulfur.

In this report, we have shown that both CIU and CID distinguish between ATP-competitive and allosteric kinase inhibitors within the Abl kinase domain. Our dataset for this study includes three type I inhibitors, three type II inhibitors, and three type IV inhibitors. By identifying a highly differentiating area in both CIU and CID space of 38-40V, it is likely that future high-throughput screening efforts can be greatly abbreviated

to a rate of 1-2 minutes per sample. Based on the observed separation between the normalized CIU and CID scores, a high-throughput screen for allosteric inhibitors would have a false positive rate of 1% (Figure 4.5, 4.6) with a score cut-off of one standard deviation. A cut-off score of 3 standard deviations would increase the false positive rate to 16%. We anticipate that as brighter ion sources and better sample introduction technologies are created, screening rates can be increased to seconds per sample. More broadly, the CIU/CID method described in this report can likely be extended to other kinases where allosteric sites are more poorly defined than the myristate pocket within Abl. While in principle both charge stripping based CID analysis and

CIU can readily be deployed in the context of any kinase-based screen for improved confidence in identifying myristate pocket binders in Abl, CIU is likely the more general of the two technologies, as it directly reflects the conformational preferences of the protein-ligand target ions. This conformational information is reported both the in CIU data itself, as well as in the range of CIU-based scores associated with both ATP-competitive and allosteric binders in this study. Our future work will pursue both efforts in Abl allosteric inhibitor discovery, as well as broadening our methods to include other kinase targets<sup>22, 23</sup>.

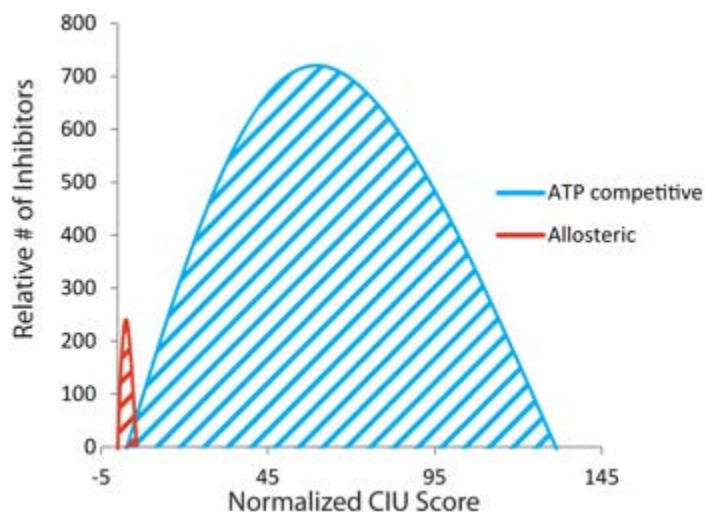


Figure 4.5 Simulated results from a high throughput screen using the averages and standard deviations for ATP-competitive and allosteric inhibitors from Table 4.1. A very small amount of overlap is predicted between the two classes of inhibitors, but the normalized CIU scores can be used to create a more or less stringent cutoff.

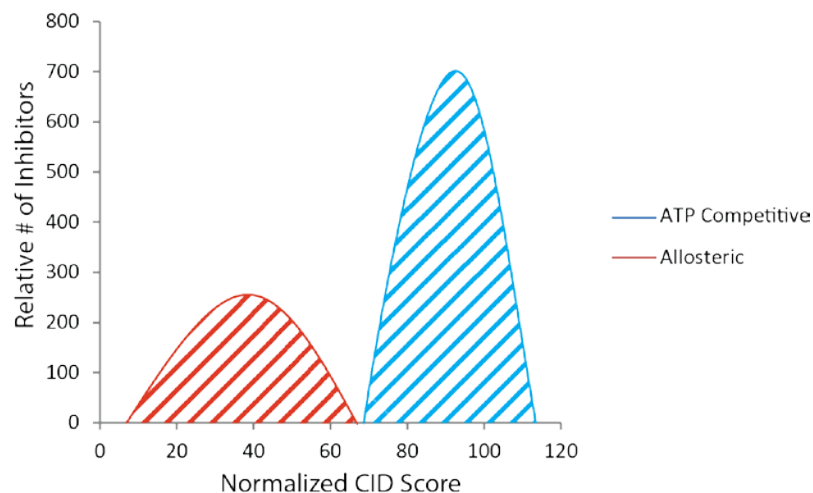


Figure 4.6. Simulated results from a high throughput screen using the averages and standard deviations for ATP-competitive and allosteric inhibitors from Table 4.2

## 4.6 References

- [1] Deininger, M. W., Goldman, J. M., and Melo, J. V. (2000) The molecular biology of chronic myeloid leukemia, *Blood* 96, 3343-3356.
- [2] Müller, S., Chaikuad, A., Gray, N. S., and Knapp, S. (2015) The ins and outs of selective kinase inhibitor development, *Nature chemical biology* 11, 818-821.
- [3] Fabbro, D., Manley, P. W., Jahnke, W., Liebetanz, J., Szyttenholm, A., Fendrich, G., Strauss, A., Zhang, J., Gray, N. S., and Adrian, F. (2010) Inhibitors of the Abl kinase directed at either the ATP-or myristate-binding site, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1804, 454-462.
- [4] Liu, Y., and Gray, N. S. (2006) Rational design of inhibitors that bind to inactive kinase conformations, *Nature chemical biology* 2, 358-364.
- [5] Iacob, R. E., Zhang, J., Gray, N. S., and Engen, J. R. (2011) Allosteric interactions between the myristate-and ATP-site of the Abl kinase, *PLoS One* 6, e15929.
- [6] Zhang, J., Adrián, F. J., Jahnke, W., Cowan-Jacob, S. W., Li, A. G., Iacob, R. E., Sim, T., Powers, J., Dierks, C., and Sun, F. (2010) Targeting Bcr–Abl by combining allosteric with ATP-binding-site inhibitors, *Nature* 463, 501-506.
- [7] Rabuck, J. N., Hyung, S.-J., Ko, K. S., Fox, C. C., Soellner, M. B., and Ruotolo, B. T. (2013) Activation state-selective kinase inhibitor assay based on ion mobility-mass spectrometry, *Analytical chemistry* 85, 6995-7002.
- [8] Ruotolo, B. T., Benesch, J. L., Sandercock, A. M., Hyung, S.-J., and Robinson, C. V. (2008) Ion mobility–mass spectrometry analysis of large protein complexes, *Nature Protocols* 3, 1139-1152.
- [9] Freeke, J., Bush, M. F., Robinson, C. V., and Ruotolo, B. T. (2012) Gas-phase protein assemblies: Unfolding landscapes and preserving native-like structures using noncovalent adducts, *Chemical Physics Letters* 524, 1-9.
- [10] Hyung, S.-J., Robinson, C. V., and Ruotolo, B. T. (2009) Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes, *Chemistry & biology* 16, 382-390.
- [11] Brocca, S., Testa, L., Sobott, F., Šamalikova, M., Natalello, A., Papaleo, E., Lotti, M., De Gioia, L., Doglia, S. M., and Alberghina, L. (2011) Compaction properties of an

- intrinsically disordered protein: Sic1 and its kinase-inhibitor domain, *Biophysical journal* *100*, 2243-2252.
- [12] Canon, F., Ballivian, R., Chirot, F., Antoine, R., Sarni-Manchado, P., Lemoine, J., and Dugourd, P. (2011) Folding of a salivary intrinsically disordered protein upon binding to tannins, *Journal of the American Chemical Society* *133*, 7847-7852.
- [13] Beck, M. W., Oh, S. B., Kerr, R. A., Lee, H. J., Kim, S. H., Kim, S., Jang, M., Ruotolo, B. T., Lee, J.-Y., and Lim, M. H. A Rationally Designed Small Molecule for Identifying an In Vivo Link of Metal–Amyloid- $\beta$  Complexes to the Pathogenesis of Alzheimer’s Disease.
- [14] He, X., Park, H. M., Hyung, S.-J., DeToma, A. S., Kim, C., Ruotolo, B. T., and Lim, M. H. (2012) Exploring the reactivity of flavonoid compounds with metal-associated amyloid- $\beta$  species, *Dalton Transactions* *41*, 6558-6566.
- [15] Hyung, S.-J., DeToma, A. S., Brender, J. R., Lee, S., Vivekanandan, S., Kochi, A., Choi, J.-S., Ramamoorthy, A., Ruotolo, B. T., and Lim, M. H. (2013) Insights into anti-amyloidogenic properties of the green tea extract (–)-epigallocatechin-3-gallate toward metal-associated amyloid- $\beta$  species, *Proceedings of the National Academy of Sciences* *110*, 3743-3748.
- [16] Young, L. M., Saunders, J. C., Mahood, R. A., Revill, C. H., Foster, R. J., Ashcroft, A., and Radford, S. (2015) ESI-IMS–MS: A method for rapid analysis of protein aggregation and its inhibition by small molecules, *Methods*.
- [17] Berkowitz, S. A., Engen, J. R., Mazzeo, J. R., and Jones, G. B. (2012) Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars, *Nature Reviews Drug Discovery* *11*, 527-540.
- [18] Debaene, F. o., Wagner-Rousset, E., Colas, O., Ayoub, D., Corvaia, N., Van Dorsselaer, A., Beck, A., and Cianféroni, S. (2013) Time resolved native ion-mobility mass spectrometry to monitor dynamics of IgG4 Fab arm exchange and “bispecific” monoclonal antibody formation, *Analytical chemistry* *85*, 9785-9792.
- [19] Tian, Y., Han, L., Buckner, A. C., and Ruotolo, B. T. (2015) Collision Induced Unfolding of Intact Antibodies: Rapid Characterization of Disulfide Bonding Patterns, Glycosylation, and Structures, *Analytical chemistry*.

- [20] Seeliger, M. A., Young, M., Henderson, M. N., Pellicena, P., King, D. S., Falick, A. M., and Kuriyan, J. (2005) High yield bacterial expression of active c-Abl and c-Src tyrosine kinases, *Protein Science* 14, 3135-3139.
- [21] Eschweiler, J. D., Rabuck, J. N., Tian, Y., and Ruotolo, B. T. (2015) CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-phase Protein Ions, *Analytical chemistry*.
- [22] Han, S., Zhou, V., Pan, S., Liu, Y., Hornsby, M., McMullan, D., Klock, H. E., Haugen, J., Lesley, S. A., and Gray, N. (2005) Identification of coumarin derivatives as a novel class of allosteric MEK1 inhibitors, *Bioorganic & medicinal chemistry letters* 15, 5467-5473.
- [23] Zhang, J., Yang, P. L., and Gray, N. S. (2009) Targeting cancer with small molecule kinase inhibitors, *Nature Reviews Cancer* 9, 28-39.

## Chapter 5.

### **Ion Mobility-Mass Spectrometry-Based Assay Distinguishes between DFG-out and DFG-in Src Kinase Inhibitors**

Protein tyrosine kinase conformation can be modulated by the binding of small molecule inhibitors that induce a conformational change to the active or inactive kinase structure. However, few technologies are capable of correctly assigning inhibitor-binding modes on changes in protein structure. Here, we have developed an assay based on ion mobility-mass spectrometry to rapidly assess the structure and stoichiometry of kinase-inhibitor complexes using the Src protein tyrosine kinase, a promising therapeutic target due to its role in cancer. We use parameters developed from a broad set of known kinase inhibitors to identify three inhibitors that bind in an unknown mode using micrograms of protein without the need for protein modification or tagging.

#### **5.1 Introduction**

Src is a non-receptor protein tyrosine kinase that plays a key role in many cell signaling processes, including cell adhesion, growth, movement, and differentiation<sup>1</sup> by associating with cell membranes to transduce signals from a number of receptors to internal signaling pathways<sup>2</sup>. The activation of Src by overexpression or by downregulation of C-terminal Src kinase (Csk) has been observed in a number of carcinomas and glioblastomas<sup>3</sup>. However, there are no kinase inhibitors that target Src due to the limited clinical efficacy of known kinase inhibitors<sup>4</sup>.

Most small molecule kinase inhibitors target the catalytic kinase domain. Of these inhibitors, the majority are ATP-competitive and fall into three categories: type II, type I, and  $\alpha$ C helix-out. Type II inhibitors stabilize the inactive DFG-out conformation, in which the conserved Asp-Phe-Gly (DFG) at the start of the activation loop is flipped outward, displacing the catalytic aspartic acid residue and phenylalanine side chain<sup>5</sup> (Figure 5.1A,E,I, red). These movements partially

occlude the ATP-binding site. Type I inhibitors contrastingly stabilize the active DFG-in kinase conformation in which the DFG triad flips in toward the center of the kinase, opening the ATP-binding site and extending the activation loop to serve as a dock for potential substrates<sup>6</sup> (Figure 5.1A,E,I, blue). The DFG-in conformation is also stabilized by  $\alpha$ C helix-out inhibitors, with the addition of the  $\alpha$ C helix swinging outward to disrupt a salt bridge between the catalytic lysine (Lys295) and a conserved glutamic acid (Glu310) that is on the  $\alpha$ C helix<sup>7</sup>.

Local changes in the kinase domain upon inhibitor binding have allosteric consequences on the global kinase conformation. Notably, in recent studies using a construct of Src with the preceding two regulatory domains (SH3-SH2-kinase domain, Src3D, Figure 5.1I), type II

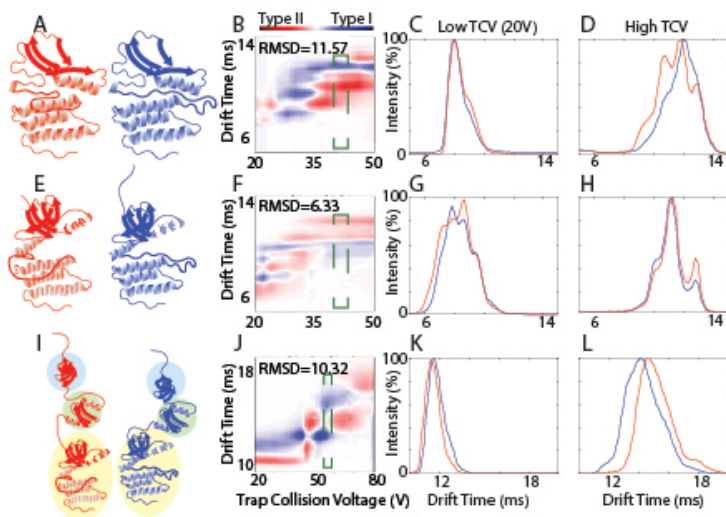


Figure 5.1 Results of CIU fingerprints for nilotinib-bound (red) and dasatinib-bound (blue) Abl and Src. (A) Cartoon representations of the inactive (red) versus active (blue) conformation of the Abl kinase domain. (B) Difference plots of the CIU fingerprints. The two CIU fingerprints of nilotinib and dasatinib bound to Abl have an RMSD value of 11.57%. (C) Drift time chromatograms of nilotinib- and dasatinib-bound Abl at low TCV. (D) Drift time chromatograms at high TCV. (E) Cartoon representations of the Src kinase domain. (F) Difference plots of nilotinib- and dasatinib-bound Src, which have an RMSD value of 6.33%. (G) Drift time chromatograms of Src bound to nilotinib and dasatinib at low TCV. (H) Drift time chromatograms at high TCV. (I) Cartoon representations of Src3D. The kinase domain is colored by a yellow circle, the SH2 domain in a green circle, and the SH3 domain in blue. (J) Difference plots of the CIU fingerprints of nilotinib- and dasatinib-bound Src3D. The RMSD is 10.32%. (K) Drift time chromatograms of Src3D bound to nilotinib and dasatinib at low TCV and (L) high TCV.

of imatinib<sup>8</sup>, a type II inhibitor that is highly selective for its target, the Abelson protein tyrosine kinase (Abl, Figure 5.1A), led to the incorrect belief that while the active forms of kinases are

inhibitors were found to stabilize an elongated, string-like Src3D structure<sup>7, 8</sup> (Figure 5.1I, red). Type I inhibitors were also found to open Src3D, although not to the extent as type II inhibitors (Figure 5.1I, blue)<sup>7</sup>. Conversely,  $\alpha$ C helix-out inhibitors stabilized a conformation similar to the autoinhibited<sup>8</sup>, closed Src3D structure, where the SH3 and SH2 domains pack against the side of the kinase domain<sup>7, 9</sup>.

Discovering highly selective kinase inhibitors is a challenging problem<sup>7</sup> in drug discovery screens, as kinase structure is highly homologous. The discovery



structurally similar, the inactive forms are thought to be conformationally diverse. Inhibitors that target the inactive conformation were erroneously thought to be more specific than type I inhibitors until the discovery of promiscuous type II inhibitors<sup>9, 10</sup>. However, determining how a small molecule inhibitor affects protein kinase conformation at the start of the drug discovery screen is crucial to future discovery efforts, as knowledge of the structure of the kinase:inhibitor complexes can better guide lead compound optimization and provide critical insights into the effects of inhibitor binding on the global kinase conformation. Furthermore, as both specific type I and type II inhibitors exist, and the knowledge of binding type at the front of a drug discovery screen aids in both discovery and optimization.

Previously, we developed an ion mobility-mass spectrometry (IM-MS)-based collision induced unfolding (CIU) assay that is able to differentiate between type I and type II kinase inhibitors using Abl as a model system<sup>11</sup>. IM-MS is able to analyze native-like proteins and protein complexes from small,  $\mu$ l volumes of complex mixtures that would frustrate other biophysical characterization techniques. IM-MS has been used for analyzing biotherapeutics<sup>12-15</sup>, determining the structures of difficult protein targets<sup>16-18</sup>, identifying small molecules that modulate membrane protein stability and conformation<sup>19-21</sup>, and for finding inhibitors for amyloidogenic proteins<sup>22-26</sup>. While IM can separate proteins and protein complexes that are >3% different in global size<sup>27</sup>, more subtle conformational changes can be separated by CIU<sup>28</sup>, a gas-phase unfolding experiment analogous to differential scanning calorimetry or fluorimetry in solution. In a CIU experiment, ions are generated by nano-electrospray ionization (nESI), isolated in the quadrupole, and then accelerated against inert, neutral gas molecules prior to IM separation. Collisions with gas molecules initiate protein complex unfolding, and the sizes of the unfolded ions are measured by IM, followed by detection and mass analysis in the time-of-flight (TOF) mass analyzer. By mapping the size of the ions against collision voltage and normalized intensity, the unfolding pathways of protein complex ions can be mapped and compared.

As the active (DFG-in) and inactive (DFG-out) conformations of kinase domains differ by <1% in global size<sup>11</sup>, which is too small for IM alone to differentiate (Figure 5.1A,E), we previously developed a CIU in order to distinguish between type I and type II Abl inhibitors. This CIU screen distinguished between type I and type II inhibitors with a dynamic range >10 fold more than a previously published fluorescence-based assay<sup>29</sup>. In order to demonstrate the dynamic range of the CIU experiments, the differences between the CIU fingerprints of the 11+ charge

states of Abl bound to the type I inhibitor dasatinib and the type II inhibitor nilotinib were compared. By using the CIU data analysis software package CIUSuite<sup>30</sup>, we were able to compute the differences in the two CIU fingerprints in terms of an RMSD value and a difference plot. We found that the CIU fingerprints of dasatinib- and nilotinib-bound Abl have an overall RMSD value of 11.57% (Figure 5.1B). Upon closer inspection of the CIU fingerprint, we find that at low collision voltages the differences in the drift time profiles of dasatinib-bound and nilotinib-bound Abl are small (Figure 5.1C), but at high TCV (Figure 5.1B, green box, and Figure 5.1D), but at high collision voltages, the differences between the two types of inhibitors become apparent due to both the number conformational families present at high voltages and the drift times at which the peaks appear.

In this study, we develop a similar CIU methodology for the more-challenging Src kinase capable of rapidly distinguishing between type I and type II inhibitors within the parameters of our screening method. Furthermore, we challenged our training data set of known type I and type II inhibitors with three low affinity type II inhibitors and three inhibitors with an unknown binding type. We CIU developed scoring metrics using an extensive with the training data set to correctly identify the low affinity type II inhibitors, and we found that all of the unknown inhibitors align with type II binding signatures.

## 5.2 Experimental

### 5.2.1 General

Imatinib, dasatinib, ponatinib, nilotinib, tozasertib, staurosporine, saracatinib, foretinib, and sorafenib were purchased from LC Laboratories (Woburn, MA); PP2 was purchased from Sigma (St. Louis, MO). DCC-2036 was purchased from SelleckChem (Houston, TX), and bosutinib was purchased from Tocris Bioscience (Bristol, United Kingdom). KB-6149 was made as previously described<sup>31</sup>. PP5 was made as previously described<sup>32</sup>. Unknown inhibitors were donated by the Soellner lab at the University of Michigan. Protein samples were buffer exchanged into 200 mM ammonium acetate at pH 7 (Sigma, St. Louis, MO) using Micro Bio-Spin 40 columns (Bio-Rad, Hercules, CA) for a final concentration of 12  $\mu$ M. Inhibitors were added at a 3:1 molar ratio of Src3D: inhibitor, and the samples were incubated for at least 15 minutes on ice before analysis.

### 5.2.2 Protein Expression and Purification

Src DNA was synthesized by GeneArt (Life Technologies, Grand Island, NY) using *E. coli* modified codons and subcloned into pET28a with a modified TEV-protease cleavable N-terminal 6x-His tag. The plasmid was transformed by electroporation into BL21 DE3 electrochemically competent cells with a YopH in pCDFDuet-1. Cell growth, protein expression, and purification were performed without cleavage of the His-tag as previously described<sup>33</sup>.

### 5.2.3 Ion mobility-Mass Spectrometry

7  $\mu$ l of protein:inhibitor sample were analyzed on a nano-electrospray ionization-quadrupole-ion mobility-time-of-flight mass spectrometer (nESI-Q-IM-TOF) instrument (Synapt G2 HDMS, Waters, Milford, Ma) as previously described<sup>34, 35</sup>. The source was operated in positive mode with a capillary voltage of 1.5-1.7 kV. The sample cone was operated at 50V, the extraction cone was set to 5V, and the source temperature was set to 20°C. The quad profile was set to ramp between 4000 and 5000 m/z, while the traveling wave ion guide was pressurized to  $2.91 \times 10^{-2}$ , the wave height set to 40V, and the wave velocity set to 900 m/s. The TOF was operated over an m/z range of 1000-8000 at a pressure of  $1.42 \times 10^{-6}$  and was externally calibrated using a solution of cesium iodide (20 mg/ml). Protein:inhibitor concentrations were kept sufficiently low so as to not induce artifact complexes. All data conformed to the expected 1:1 inhibitor binding stoichiometry.

Before IM separation, protein:inhibitor ions were activated in the trap traveling wave ion guide to collisionally unfold the complexes. The trap collision voltage was increased in steps of 2V from 20V to 80V after mass selection in the quadrupole. Data were processed using the Python-based CIUSuite<sup>30</sup> software package, which uses a z-score-based scoring metric to classify inhibitors according to user-defined groups.

## 5.3 Results and Discussion

### 5.3.1 IM-MS of Src Kinase Domain vs. Three Domain Constructs

Given our previous work with Abl, we expected that a similar CIU strategy would prove successful for the Src protein tyrosine kinase. Therefore, we began by recording the CIU of the

kinase domain of Src following the experimental protocol we established for our previous Abl kinase domain dataset. An initial comparison of the 11+ charge states of dasatinib-bound and nilotinib-bound Src showed no significant differences between the two CIU fingerprints. We attribute differences in the CIU behavior of the kinase domains of Src and Abl to their previously reported differences in structural flexibility<sup>36, 37</sup>. As Src is less flexible than Abl, the CIU of the 11+charge state lacked the large number of conformational transitions that were responsible for the high dynamic range of our previous Abl CIU assay<sup>11</sup>. As an attempt to impart additional CIU transitions to Src, we sought to enhance the dynamic range of our Src:inhibitor complex IM-MS and CIU datasets through charge manipulation. Previous studies have indicated that, in general, a greater number of CIU transitions are observed for more highly charged proteins<sup>38, 39</sup>. To increase the charge state of the Src:inhibitor complex ions, charge amplification reagents were added in order to change the surface tension of the electrospray droplets produced during nESI and enhance protein charging<sup>40</sup>. Using this method, a 14+charge state was achieved for Src complexes, and the CIU fingerprints of dasatinib-bound and nilotinib-bound Src were collected and analyzed (Figure 5.1F). However, CIU fingerprint RMSD values ranged as low 6.33%. This RMSD is within error of the 3-5% baseline RMSD often seen in replicate CIU experiments for the Src kinase. The drift time plots at low collision voltage (Figure 5.1G) and high collision voltages (Figure 5.1H) also showed very little variation between the two ligand-bound species.

In contrast to Abl, we hypothesized that the three domain version of Src (Src3D, Figure 5.1I) would increase the overall flexibility and range of motion for the kinase and therefore increase the dynamic range of our CIU screen. Comparison of the 13+ charge state of nilotinib-bound and dasatinib-bound Src3D CIU fingerprints produces an RMSD of 10.32% (Figure 5.1J), almost double that of the RMSD of the dasatinib- and nilotinib-bound Src kinase domain and nearly the RMSD of the two Abl CIU fingerprints. The 13+ charge state was chosen because it exhibited the most differences between the screened type I and type II inhibitors. Similar to Src KD and Abl KD, there is very little difference in the drift time chromatograms of dasatinib- and nilotinib-bound Src3D at low collision voltage (Figure 5.1K), but there are differences in the centroids of the drift time chromatograms at high collision voltages. Furthermore, dasatinib exhibits a shoulder at a lower drift time, whereas nilotinib has a shoulder at a higher drift time (Figure 5.1L). Since we were able to increase the dynamic range of our CIU assay through the Src3D construct, we carried on to guild a general Src inhibitor screen using this form of the protein.

### 5.3.2 Identifying a CIU Fingerprint Region that Distinguishes Type I and Type II Src Inhibitors

The dynamic range of the CIU assay is maximized in the area of the CIU fingerprints where type I and type II inhibitors are the most differentiated. By using the analysis module of CIUSuite,<sup>30</sup> a scaled z-score (scaled deviation score, SDS) for all known inhibitors against a type II average was computed (Figure 5.2A). The SDS score is computed by normalizing the z-scores for each inhibitor against a user-specified average. The output of CIUSuite is a plot of SDS vs collision

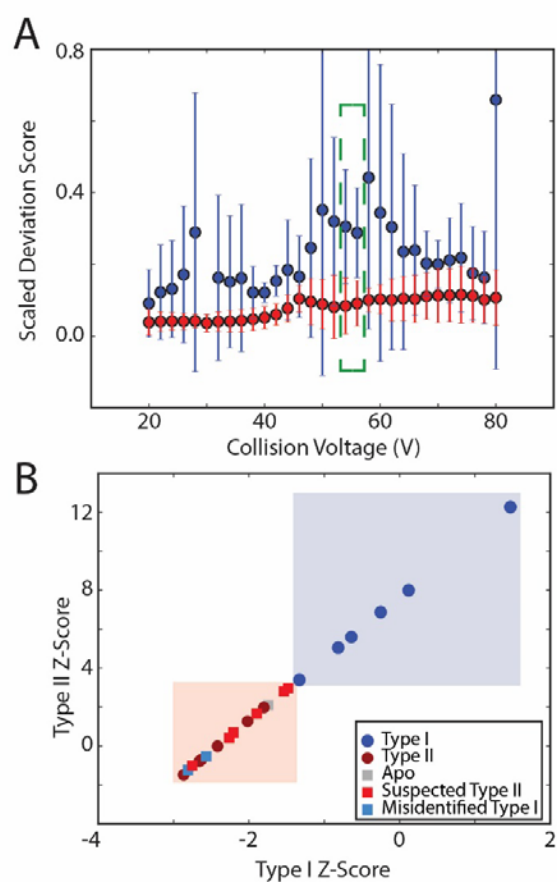


Figure 5.2. Results from CIUSuite\_analysis. (A) SDS vs. collision voltage plot. Type I inhibitors are in blue, and type II inhibitors are in red. The error bars represent the spread of SDS for each inhibitor type. The green box represents the area where type I and type II inhibitors are the most differentiated from each other. (B) Z-score analysis of all inhibitors. Z-scores against a type I average lie along the x-axis, and z-scores against a type II average lie on the y-axis. The scores are the most differentiated along the type II average, and the red and blue boxes represent standard deviations along each axis.

voltage for all inhibitors. The average type I SDS (in blue) and average type II SDS values (in red) are represented by dots, where the error bars represent the range of SDS values for the type I and type II inhibitors at each collision voltage. This analysis revealed an area from 52 to 54V where the type I and type II inhibitors were the most distinguished from one another (Figure 5.2A, green dashed box), as evidenced by minimized overlap between the type I and type II error bars are shown. An area between 42 to 44V is also well separated; however, upon further analysis, several inhibitors from our control group were misclassified in this region.

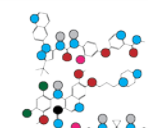
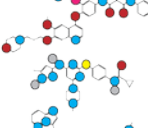
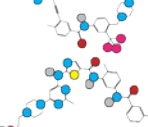
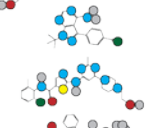
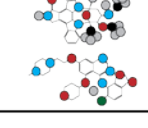
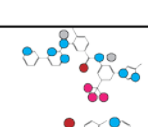
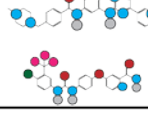
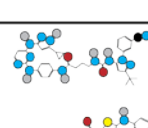
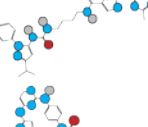
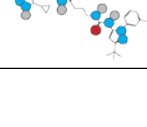
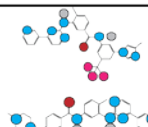
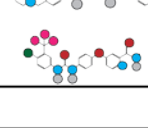
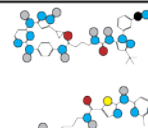
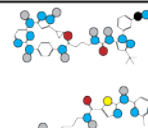
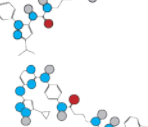
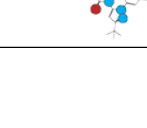
To analyze the spread and range of the type I and type II inhibitors, z-scores for each inhibitor against a type I average or a type II average at 52-54V were plotted, with the z-score against a type I average on the x-axis, and the z-score against a type II average on the y-axis (Figure 5.2B). The light blue and pink boxes represent the standard deviations from the average type I inhibitors or type II inhibitors respectively along each axis. It

is important to note that a z-score against a type I average is not the inverse of a z-score against a type II average. When plotted against a type I average, the maximum z-score of all inhibitors (type I and type II) is 1.53, and the minimum z-score is -2.37, which indicates a maximum of 1.5 standard deviations above the type I average and -2.37 standard deviations below the type I average. The z-score against a type II average provides more dynamic range, as the type I z-score ranges between 5.72 and 13.97 while the type II z-scores range from -1.41 to 1.41. In this analysis, the type II inhibitors are clustered more tightly together than the type I inhibitors, which is indicative of type I inhibitors overall stabilizing a greater breadth of conformational states in Src3D<sup>10</sup>.

### 5.3.3 CIU Accurately Classifies Weakly Bound and Unknown Inhibitors

After identifying the region of CIU fingerprint space discussed above, scores were normalized so that the type II inhibitor PP5, which had the smallest SDS score of any inhibitor within our control group, to a final type II similarity score of 100% (Table 5.1). This resulted in an average type II score is  $47.30 \pm 27.29\%$ , and an average type I score is  $14.07 \pm 3.94\%$ , while the average error associated with our measurements is 4.5% (n=4). We applied a type II cut-off score of 20, which identified nilotinib, imatinib, and sorafenib as type II inhibitors. Although these inhibitors exhibit weak binding with  $K_d$  values in the micromolar range to Src, we were able to observe binding in the mass spectra and to classify these inhibitors as type II. Our ability to properly identify even weak-binding inhibitors provides a dimension to our analysis that is lost by many kinase inhibitor-screening technologies<sup>41</sup>. This type of conformation-sensitive analysis also speaks to the strength of native mass spectrometry and collision induced unfolding to identify novel inhibitors that cause a conformational change in a protein system. Furthermore, by creating a z-score based scoring metric, a cut-off value can be applied to adapt the stringency of the screen. Using this framework, we predicted the results of a theoretical high throughput screen with ~1800 inhibitors using average type I and type II scores with their standard deviations (Figure 5.3). While there is a low percentage of false positives predicted, as indicated by the small overlap of type I and type II inhibitors (~100 out of 1800 inhibitors), the false positive discovery rate can be modulated by applying a higher or lower scoring cut-off.

Table 5.1. Scores resulting from the CIUSuite analysis from Figure 5.2.

Name	MW	Structure	K <sub>d</sub> (nm)	Type	Score	Confirmed type	
Known Type I and Type II							
DCC-2036	553.6		--	II	100.00±5.76	II	✓
Bosutinib	529.2		1	I	81.93	II	✗
Foretinib	632.6		20	II	68.17	II	✓
Tozasertib	464.6		170	I	65.98±4.36	II	✗
Ponatinib	532.4		--	II	40.41	II	✓
KB-6149	641.3		--	II	30.61	II	✓
PP2	301.1		--	I	19.87	I	✓
Dasatinib	487.2		0.21	I	12.65±3.52	I	✓
Staurosporine	466.5		86	I	12.57	I	✓
Saracatinib	541.4		--	I	11.15	I	✓
Weak Type II Inhibitors							
Nilotinib	529.2		1900	II	38.62	II	✓
Imatinib	493.6		--	II	27.45±4.22	II	✓
Sorafenib	637.0		--	II	25.81	II	✓
Unknown Type I and Type II							
KK2-145	658.8		--	UK	68.12	II	✓
KK3-22	647.8		--	UK	48.03	II	✓
KK2-32	646.8		--	UK	30.77	II	✓

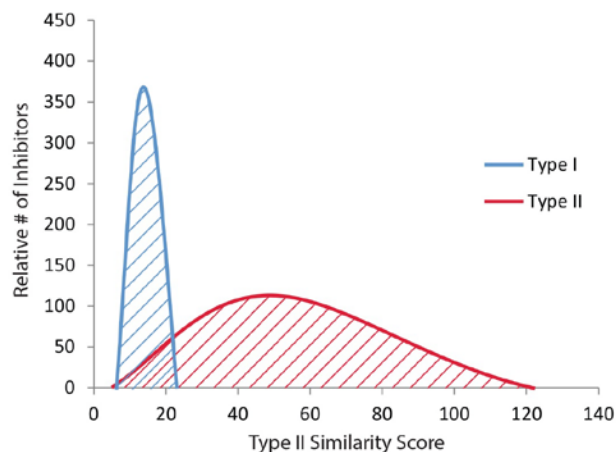


Figure 5.3 An illustration of a theoretical HTS using the scores from Table 1. The averages and standard deviations are used to plot the normal distributions for scores expected for these two classes of inhibitors. (~1800 compounds, 50% type I, 50%).

In our analysis, two inhibitors that are considered to be type I are identified as type II. Tozasertib has been identified as a type I inhibitor in an imatinib-resistant mutant of the Abl kinase<sup>42</sup> and in our previous CIU assay<sup>11</sup>, although in our assay, tozasertib falls along the type I/type II border. However, in a co-complex crystal structure with the protein

tyrosine kinase Aurora A, tozasertib is shown to bind in the inactive conformation<sup>43</sup>. Due to this dichotomy, it is difficult to predict if tozasertib would stabilize a type I- or type II-

like Src3D conformation. Bosutinib is also identified as a type I inhibitor in the co-complex crystal structures of Abl<sup>44</sup> and Src<sup>45</sup>; however, it also falls along the type I/II border in our previous CIU assay<sup>11</sup>. We attribute the assignment of bosutinib as a type II inhibitor to the potential formation of a hydrogen bond between bosutinib and the  $\alpha$ C helix of the Src kinase domain, which would cause a type II-like conformation. We also identified three unknown inhibitors built using a tozasertib scaffold as type II inhibitors. These inhibitors, KK2-145, KK3-22, and KK2-32, had normalized scores of 58.19, 41.03, and 26.28 (Table 5.1). Since tozasertib has a score of 56.26%, it is unsurprising that these inhibitors are also identified as type II. As all of these scores are higher than the cut-off of 20, they were assigned as type II inhibitors. As KK3-22 and KK2-32 cluster more closely to other type II inhibitors, we predict that these inhibitors align more closely to a type II binding signature than to a type I binding signature. Based on the probability curves in Figure 5.3, a cut-off score of 20 would lead to a 3% false positive rate. With using a cut-off of 1 standard deviation, the false positive rate increases to 14%, with a cut-off of 2 standard deviations, the false positive rate is 3%. Therefore, modulating the stringency of the scoring cut-off can decrease the number of false positives, although potential type II hits would likely be missed.



## 5.4 Conclusions

This CIU assay covers a wide range of inhibitor chemical space and binding strength, and to our knowledge, contains the largest number of inhibitors validating a conformationally-selective kinase screen. Using this CIU methodology with Src3D, a differentiating region within the CIU fingerprints where type I and type II inhibitors can be readily distinguished was identified in order to create a normalized score with known type I and type II inhibitors for the purpose of identifying three unknown inhibitors. These inhibitors were classified as type II inhibitors, which is unsurprising given the structure of these inhibitors and our classification of tozasertib as type II. Our approach was further supported by the proper identification of three weak type II inhibitors.

It is important to view our CIU data in light of the local changes that occur within the Src kinase domain upon inhibitor binding, as they have well-known allosteric consequences on the global kinase conformation. Notably, in recent studies using a Src3D, type II inhibitors were found to increase the accessibility of SH3 domain accessibility, which is incompatible with a closed tertiary structure and is most likely related to an elongated, string-like Src3D structure<sup>46, 47</sup> (Figure 5.1I, red). Type I inhibitors, conversely, have been found to stabilize a range of structures from a fully open Src3D structure to a more neutral position between the open/closed Src3D structure (Figure 5.1I, blue)<sup>46</sup>. Recently, another class of ATP-competitive inhibitors have been discovered that close the Src3D structure. These inhibitors are commonly referred to as  $\alpha$ C helix-out<sup>37</sup>. The DFG-in conformation is also stabilized by such  $\alpha$ C helix-out inhibitors through a destabilization of a salt bridge between the catalytic lysine (Lys295) and a conserved glutamic acid (Glu310) that is on the  $\alpha$ C helix<sup>46</sup>.  $\alpha$ C helix-out inhibitors have been found to stabilize a conformation similar to the autoinhibited<sup>47</sup>, closed Src3D structure, where the SH3 and SH2 domains pack against the side of the kinase domain<sup>46, 48</sup>.

Without the inclusion of  $\alpha$ C helix-out inhibitors in the data set, it is difficult to determine if our identification of bosutinib and tozasertib as type II-like inhibitors are true type I outliers. The inhibitors chosen for this screen likely stabilize a range Src3D structures from fully extended to a neutral position between the open and closed Src3D structure. In order to completely classify these inhibitors, future work will seek to add  $\alpha$ C helix-out to our control group. In this future screen, ligands would be scored according to their stabilization of the three domain structure

instead of being ranked purely by the DFG-in/out paradigm, placing all inhibitors on a continuum according to how they stabilized the closed, open, or neutral form of the kinase tertiary structure. Further information about the open/closed paradigm would also benefit from the addition of mutations that stabilize the open or closed conformation<sup>46</sup>. This ranking is especially important as recently it has been discovered that changes in kinase conformation affect downstream signaling<sup>49-55</sup>. Consequently, the potential to direct down-stream signaling in cancerous cells through modulation of a kinase via ligand binding is an important mechanism to realize in drug discovery assays.

For the focused screen presented here, and the narrow range of CIU space that must be probed to assay the kinase structure, we project an approximate throughput of ~3 minutes per sample., given current limitations in nESI sample delivery and ion source brightness. With this throughput, ~480 inhibitors could be screened in a 24 hour time period. We anticipate that as brighter ion sources and faster sample introduction methods are built for high-throughput screens, the time it takes for data collection can be reduced to less than 1 minute/sample. For example, a microfluidics system would be ideal for introducing samples into the instrument<sup>56</sup>. We have also shown that while the structures adopted by protein:ligand complexes during CIU transitions in the gas-phase are unknown, by correlating the differences in CIU fingerprints to known inhibitors types, CIU can be used to accurately build a high-throughput screen that is representative of solution-phase binding mode. By rapidly assessing ligand binding mode in the first step of a screen, rather than relying on only a  $K_d$  or  $IC_{50}$  information, inhibitors that bind in a type II mode can be quickly identified for further high-resolution structural studies and optimization instead of being discarded as non-binders, which will facilitate the creation of a new sets of small molecule kinase inhibitors in the future.

## 5.5 References

- [1] Brown, M. T., and Cooper, J. A. (1996) Regulation, substrates and functions of src, *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* 1287, 121-149.
- [2] Parsons, S. J., and Parsons, J. T. (2004) Src family kinases, key regulators of signal transduction, *Oncogene* 23, 7906-7909.

- [3] Roskoski, R. (2015) Src protein-tyrosine kinase structure, mechanism, and small molecule inhibitors, *Pharmacological Research* 94, 9-25.
- [4] Creedon, H., and Brunton, V. G. (2012) Src kinase inhibitors: promising cancer therapeutics?, *Critical Reviews™ in Oncogenesis* 17.
- [5] Mol, C., Fabbro, D., and Hosfield, D. (2004) Structural insights into the conformational selectivity of STI-571 and related kinase inhibitors, *Current opinion in drug discovery & development* 7, 639-648.
- [6] Traxler, P., and Furet, P. (1999) Strategies toward the design of novel and selective protein tyrosine kinase inhibitors, *Pharmacology & therapeutics* 82, 195-206.
- [7] Knight, Z. A., and Shokat, K. M. (2005) Features of selective kinase inhibitors, *Chemistry & biology* 12, 621-637.
- [8] Druker, B. J., Talpaz, M., Resta, D. J., Peng, B., Buchdunger, E., Ford, J. M., Lydon, N. B., Kantarjian, H., Capdeville, R., and Ohno-Jones, S. (2001) Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia, *New England Journal of Medicine* 344, 1031-1037.
- [9] Zhang, J., Yang, P. L., and Gray, N. S. (2009) Targeting cancer with small molecule kinase inhibitors, *Nature Reviews Cancer* 9, 28-39.
- [10] Liu, Y., and Gray, N. S. (2006) Rational design of inhibitors that bind to inactive kinase conformations, *Nature chemical biology* 2, 358-364.
- [11] Rabuck, J. N., Hyung, S.-J., Ko, K. S., Fox, C. C., Soellner, M. B., and Ruotolo, B. T. (2013) Activation state-selective kinase inhibitor assay based on ion mobility-mass spectrometry, *Analytical chemistry* 85, 6995-7002.
- [12] Tian, Y., Han, L., Buckner, A. C., and Ruotolo, B. T. (2015) Collision Induced Unfolding of Intact Antibodies: Rapid Characterization of Disulfide Bonding Patterns, Glycosylation, and Structures, *Analytical chemistry* 87, 11509-11515.
- [13] Debaene, F. o., Wagner-Rousset, E., Colas, O., Ayoub, D., Corvaia, N., Van Dorselaer, A., Beck, A., and Cianféroni, S. (2013) Time resolved native ion-mobility mass spectrometry to monitor dynamics of IgG4 Fab arm exchange and “bispecific” monoclonal antibody formation, *Analytical chemistry* 85, 9785-9792.

- [14] Terral, G., Beck, A., and Cianf  rani, S. (2016) Insights from native mass spectrometry and ion mobility-mass spectrometry for antibody and antibody-based product characterization, *Journal of Chromatography B*.
- [15] Ferguson, C. N., and Gucinski-Ruth, A. C. (2016) Evaluation of Ion Mobility-Mass Spectrometry for Comparative Analysis of Monoclonal Antibodies, *Journal of The American Society for Mass Spectrometry*, 1-12.
- [16] MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V., and Doudna, J. A. (2008) In vitro reconstitution of the human RISC-loading complex, *Proceedings of the National Academy of Sciences* 105, 512-517.
- [17] Sharon, M., and Robinson, C. V. (2007) The role of mass spectrometry in structure elucidation of dynamic protein complexes, *Annu. Rev. Biochem.* 76, 167-193.
- [18] Shepherd, D. A., Marty, M. T., Giles, K., Baldwin, A. J., and Benesch, J. L. (2015) Combining tandem mass spectrometry with ion mobility separation to determine the architecture of polydisperse proteins, *International Journal of Mass Spectrometry* 377, 663-671.
- [19] Allison, T. M., Reading, E., Liko, I., Baldwin, A. J., Laganowsky, A., and Robinson, C. V. (2015) Quantifying the stabilizing effects of protein-ligand interactions in the gas phase, *Nature communications* 6.
- [20] Laganowsky, A., Reading, E., Allison, T. M., Ulmschneider, M. B., Degiacomi, M. T., Baldwin, A. J., and Robinson, C. V. (2014) Membrane proteins bind lipids selectively to modulate their structure and function, *Nature* 510, 172.
- [21] Landreh, M., Marty, M. T., Gault, J., and Robinson, C. V. (2016) A sliding selectivity scale for lipid binding to membrane proteins, *Current Opinion in Structural Biology* 39, 54-60.
- [22] Hyung, S.-J., DeToma, A. S., Brender, J. R., Lee, S., Vivekanandan, S., Kochi, A., Choi, J.-S., Ramamoorthy, A., Ruotolo, B. T., and Lim, M. H. (2013) Insights into anti-amyloidogenic properties of the green tea extract (–)-epigallocatechin-3-gallate toward metal-associated amyloid- $\beta$  species, *Proceedings of the National Academy of Sciences* 110, 3743-3748.
- [23] Illes-Toth, E., Dalton, C. F., and Smith, D. P. (2013) Binding of dopamine to  $\alpha$ -synuclein is mediated by specific conformational states, *Journal of The American Society for Mass Spectrometry* 24, 1346-1354.

- [24] Susa, A. C., Wu, C., Bernstein, S. L., Dupuis, N. F., Wang, H., Raleigh, D. P., Shea, J.-E., and Bowers, M. T. (2014) Defining the Molecular Basis of Amyloid Inhibitors: Human Islet Amyloid Polypeptide–Insulin Interactions, *Journal of the American Chemical Society* 136, 12912-12919.
- [25] Woods, L. A., Platt, G. W., Hellewell, A. L., Hewitt, E. W., Homans, S. W., Ashcroft, A. E., and Radford, S. E. (2011) Ligand binding to distinct states diverts aggregation of an amyloid-forming protein, *Nature chemical biology* 7, 730-739.
- [26] Young, L. M., Saunders, J. C., Mahood, R. A., Revill, C. H., Foster, R. J., Tu, L.-H., Raleigh, D. P., Radford, S. E., and Ashcroft, A. E. (2015) Screening and classifying small-molecule inhibitors of amyloid formation using ion mobility spectrometry–mass spectrometry, *Nature chemistry* 7, 73-81.
- [27] Benesch, J. L., and Ruotolo, B. T. (2011) Mass spectrometry: come of age for structural and dynamical biology, *Current opinion in structural biology* 21, 641-649.
- [28] Hyung, S.-J., Robinson, C. V., and Ruotolo, B. T. (2009) Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes, *Chemistry & biology* 16, 382-390.
- [29] Simard, J. R., Pawar, V., Aust, B., Wolf, A., Rabiller, M., Wulfert, S., Robubi, A., Klüter, S., Ottmann, C., and Rauh, D. (2009) High-throughput screening to identify inhibitors which stabilize inactive kinase conformations in p38 $\alpha$ , *Journal of the American Chemical Society* 131, 18478-18488.
- [30] Eschweiler, J. D., Rabuck-Gibbons, J. N., Tian, Y., and Ruotolo, B. T. (2015) CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-phase Protein Ions, *Analytical chemistry* 87, 11516-11522.
- [31] Brandvold, K. R., Steffey, M. E., Fox, C. C., and Soellner, M. B. (2012) Development of a highly selective c-Src kinase inhibitor, *ACS chemical biology* 7, 1393-1398.
- [32] Dar, A. C., Lopez, M. S., and Shokat, K. M. (2008) Small molecule recognition of c-Src via the Imatinib-binding conformation, *Chemistry & biology* 15, 1015-1022.
- [33] Seeliger, M. A., Young, M., Henderson, M. N., Pellicena, P., King, D. S., Falick, A. M., and Kuriyan, J. (2005) High yield bacterial expression of active c-Abl and c-Src tyrosine kinases, *Protein Science* 14, 3135-3139.

- [34] Zhong, Y., Hyung, S.-J., and Ruotolo, B. T. (2011) Characterizing the resolution and accuracy of a second-generation traveling-wave ion mobility separator for biomolecular ions, *Analyst* 136, 3534-3541.
- [35] Ruotolo, B. T., Benesch, J. L., Sandercock, A. M., Hyung, S.-J., and Robinson, C. V. (2008) Ion mobility–mass spectrometry analysis of large protein complexes, *Nature Protocols* 3, 1139-1152.
- [36] Lovera, S., Sutto, L., Boubeva, R., Scapozza, L., Dölker, N., and Gervasio, F. L. (2012) The different flexibility of c-Src and c-Abl kinases regulates the accessibility of a druggable inactive conformation, *Journal of the American Chemical Society* 134, 2496-2499.
- [37] Hari, S. B., Perera, B. G. K., Ranjitkar, P., Seeliger, M. A., and Maly, D. J. (2013) Conformation-selective inhibitors reveal differences in the activation and phosphate-binding loops of the tyrosine kinases Abl and Src, *ACS chemical biology* 8, 2734-2743.
- [38] Mao, Y., Ratner, M. A., and Jarrold, M. F. (1999) Molecular dynamics simulations of the charge-induced unfolding and refolding of unsolvated cytochrome c, *The Journal of Physical Chemistry B* 103, 10017-10021.
- [39] Konermann, L., and Douglas, D. (1998) Equilibrium unfolding of proteins monitored by electrospray ionization mass spectrometry: distinguishing two-state from multi-state transitions, *Rapid communications in mass spectrometry* 12, 435-442.
- [40] Yin, S., and Loo, J. A. (2011) Top-down mass spectrometry of supercharged native protein–ligand complexes, *International journal of mass spectrometry* 300, 118-122.
- [41] Davis, M. I., Hunt, J. P., Herrgard, S., Ciceri, P., Wodicka, L. M., Pallares, G., Hocker, M., Treiber, D. K., and Zarrinkar, P. P. (2011) Comprehensive analysis of kinase inhibitor selectivity, *Nature biotechnology* 29, 1046-1051.
- [42] Young, M. A., Shah, N. P., Chao, L. H., Seeliger, M., Milanov, Z. V., Biggs, W. H., Treiber, D. K., Patel, H. K., Zarrinkar, P. P., and Lockhart, D. J. (2006) Structure of the kinase domain of an imatinib-resistant Abl mutant in complex with the Aurora kinase inhibitor VX-680, *Cancer research* 66, 1007-1014.
- [43] Bebbington, D., Binch, H., Charrier, J.-D., Everitt, S., Fraysse, D., Golec, J., Kay, D., Knegt, R., Mak, C., and Mazzei, F. (2009) The discovery of the potent aurora inhibitor MK-0457 (VX-680), *Bioorganic & medicinal chemistry letters* 19, 3586-3592.

- [44] Levinson, N. M., and Boxer, S. G. (2012) Structural and spectroscopic analysis of the kinase inhibitor bosutinib and an isomer of bosutinib binding to the Abl tyrosine kinase domain, *PLoS One* 7, e29828.
- [45] Levinson, N. M., and Boxer, S. G. (2014) A conserved water-mediated hydrogen bond network defines bosutinib's kinase selectivity, *Nature chemical biology* 10, 127-132.
- [46] Krishnamurty, R., Brigham, J. L., Leonard, S. E., Ranjitkar, P., Larson, E. T., Dale, E. J., Merritt, E. A., and Maly, D. J. (2013) Active site profiling reveals coupling between domains in SRC-family kinases, *Nature chemical biology* 9, 43-50.
- [47] Cowan-Jacob, S. W., Fendrich, G., Manley, P. W., Jahnke, W., Fabbro, D., Liebetanz, J., and Meyer, T. (2005) The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation, *Structure* 13, 861-871.
- [48] Leonard, S. E., Register, A., Krishnamurty, R., Brighty, G. J., and Maly, D. J. (2014) Divergent modulation of Src-family kinase regulatory interactions with ATP-competitive inhibitors, *ACS chemical biology* 9, 1894-1905.
- [49] Hatzivassiliou, G., Song, K., Yen, I., Brandhuber, B. J., Anderson, D. J., Alvarado, R., Ludlam, M. J., Stokoe, D., Gloor, S. L., and Vigers, G. (2010) RAF inhibitors prime wild-type RAF to activate the MAPK pathway and enhance growth, *Nature* 464, 431-435.
- [50] Poulikakos, P. I., Zhang, C., Bollag, G., Shokat, K. M., and Rosen, N. (2010) RAF inhibitors transactivate RAF dimers and ERK signalling in cells with wild-type BRAF, *Nature* 464, 427-430.
- [51] Chan, T. O., Zhang, J., Rodeck, U., Pascal, J. M., Armen, R. S., Spring, M., Dumitru, C. D., Myers, V., Li, X., and Cheung, J. Y. (2011) Resistance of Akt kinases to dephosphorylation through ATP-dependent conformational plasticity, *Proceedings of the National Academy of Sciences* 108, E1120-E1127.
- [52] Andraos, R., Qian, Z., Bonenfant, D., Rubert, J., Vangrevelinghe, E., Scheufler, C., Marque, F., Régnier, C. H., De Pover, A., and Ryckelynck, H. (2012) Modulation of activation-loop phosphorylation by JAK inhibitors is binding mode dependent, *Cancer discovery* 2, 512-523.
- [53] Wang, L., Perera, B. G. K., Hari, S. B., Bhatarai, B., Backes, B. J., Seeliger, M. A., Schürer, S. C., Oakes, S. A., Papa, F. R., and Maly, D. J. (2012) Divergent allosteric

- control of the IRE1 $\alpha$  endoribonuclease using kinase inhibitors, *Nature chemical biology* 8, 982-989.
- [54] Koppikar, P., Bhagwat, N., Kilpivaara, O., Manshouri, T., Adli, M., Hricik, T., Liu, F., Saunders, L. M., Mullally, A., and Abdel-Wahab, O. (2012) Heterodimeric JAK-STAT activation as a mechanism of persistence to JAK2 inhibitor therapy, *Nature* 489, 155-159.
- [55] Gilani, R., Phadke, S., Bao, L. W., Lachacz, E., Dziubinski, M., Brandvold, K., Steffey, M., Kwarcinski, F., Graveel, C. R., and Kidwell, K. M. (2016) UM-164: a potent c-Src/p38 kinase inhibitor with in vivo activity against triple negative breast cancer, *Clinical Cancer Research*, clincanres. 2158.2015.
- [56] Sun, S., and Kennedy, R. T. (2014) Droplet electrospray ionization mass spectrometry for high throughput screening for enzyme inhibitors, *Analytical chemistry* 86, 9309-9314.



## Chapter 6

# Collision Induced Unfolding Reveals Unique Fingerprints for Remote Protein Interaction Sites in the KIX Regulatory Domain

The kinase inducible domain (KIX) of the transcriptional activator CREB-binding protein acts as a hub to recruit multiple transcriptional regulators through two allosterically-connected binding sites. However, it is challenging to target these protein:protein interactions therapeutically due to a lack of assays capable of rapidly detecting the binding of potential inhibitor peptides at relatively low affinities within dynamic interaction regions on the protein surface, connected through allostery. Here, we have describe an ion mobility-mass spectrometry-based approach that leverages the gas-phase protein unfolding to produce unique fingerprints for peptide inhibitors bound to each of the two available sites within KIX, as well as a third identifiable fingerprint for doubly-bound KIX:peptide complexes. Furthermore, we can evaluate the analytical utility of the unfolding fingerprints for KIX complexes, as well as assess the structural origins of the conformational families created from KIX:peptide complexes following collisional activation.

### 6.1 Introduction

Transcriptional activators, in general, play a key role in human disease. As modular proteins, they are capable of binding to genetic material through a DNA binding domain (DBD) and contain transcriptional activation domains (TAD) to recruit other transcriptional regulators to stimulate gene transcription. Misregulation of transcription has been implicated in multiple including many cancers and inflammatory disorders<sup>1-3</sup>. Despite their relative importance, transcriptional activators are often difficult to target therapeutically, in part due to the low affinity and dynamic nature of the protein:protein interactions involved within the transcription process<sup>4</sup>. Additionally, the protein:protein interaction (PPI) interfaces are composed of large, hydrophobic surfaces, so designing appropriate inhibitors is challenging and requires extensive efforts in order to collect

high resolution protein structure information in order to identify key regions that may contribute to binding<sup>5</sup>. The TADs within activator proteins, for instance, are intrinsically disordered and form transient secondary structure only when forming a complex with a binding partner<sup>6, 7</sup>.

One of the protein domains that has been the focus of recent drug discovery efforts is the kinase-inducible (KIX) domain of the coactivator CREB-binding protein (CPB). CPB and its homolog

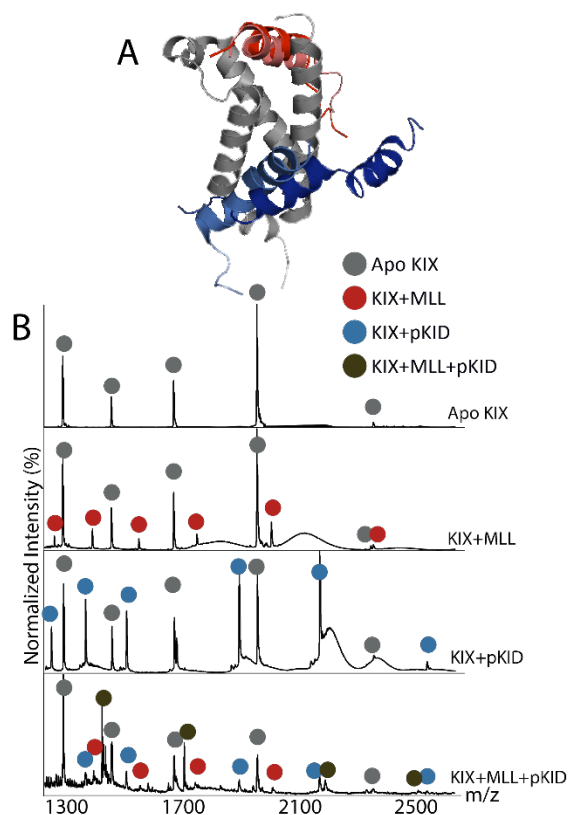


Figure 6.1 Structure of KIX with inhibitors and corresponding mass spectra, (A) KIX (gray) with four inhibitors: MLL (red), E2A (pink), c-Myb (light blue), and pKID (blue). (B) Mass spectra of (from top to bottom) apo KIX, KIX:MLL, KIX:pKID, and KIX:MLL:pKID. The apo peaks are in red, the KIX+MLL peaks in red, KIX+pKID in blue, and the ternary KIX+MLL+pKID in green.

p300 are large, multi-domain proteins that coordinate over a hundred different interacting partners<sup>4</sup>. The KIX domain serves as a hub that can bind over 12 different activators through two binding sites: the phosphorylated kinase-inducible domain (pKID) binding site and the mixed-lineage leukemia (MLL) binding site<sup>2, 3</sup> (Figure 6.1). Although KIX is a conformationally dynamic protein, the NMR data for the binary and ternary KIX complexes reveal that KIX is composed of three alpha-helices and two 3-10 helices<sup>8</sup>. The MLL binding site is formed by the first and third alpha helix (Figure 6.1, red), while the pKID binding site is formed through a groove between the first and third alpha helices, with the pKID peptide adopting a conformation that wraps around the third alpha helix (Figure 6.1, blue).

The transcriptional activator CREB (cAMP response element binding protein) contains a TAD known as pKID, which upon pKID phosphorylation at Ser133 directly participates in binding to KIX with a low,

micromolar affinity and ~150-fold higher affinity than the unphosphorylated form of CREB<sup>9</sup>. As an inducible activator, pKID interacts more strongly than its constitutive counterpart, c-Myb, leading to alterations in gene expression<sup>10, 11</sup>. Both the c-Myb and pKID regions share distinct, but overlapping binding sites, with pKID exhibiting stronger binding when compared to c-Myb, thus leading to a key regulation mechanism for in gene expression. The MLL binding site is also

shared by other activators such as the E-protein activation domain 1 (E2A)<sup>12</sup>, c-Jun<sup>13</sup>, the human t-lymphotrophic virus type I (HTVL-1) Tax<sup>14</sup>, and the human immunodeficiency virus type 1 (HIV-1) Tat<sup>15</sup>. NMR and circular dichroism (CD) studies have revealed that the pKID and MLL peptides are unstructured, and upon binding to KIX, adopt alpha helical structures<sup>16</sup>.

Peptides and peptidomimetics have been described that can inhibit KIX protein:protein interactions that function *in vivo*, and these inhibitors have been extremely valuable in defining the characteristics of KIX activators<sup>17, 18</sup>. Despite these advances, localizing the effects of KIX inhibitors remains challenging, primarily due to the allosteric connectivity between the pKID and MLL binding sites, which can increase KIX binding affinities for these ligands by as much as two-fold<sup>16, 19</sup>. Other peptide binders of the KIX domain have been discovered through disulfide trapping experiments, and exhibit similar levels of allostery<sup>20</sup>. In addition, two natural products have been discovered for the KIX:MLL complex: sekikaic and lobaric acid<sup>21</sup>. These inhibitors also have an allosteric effect on the pKID binding site of KIX.

The plasticity of both KIX and its binding partners, combined with the allosteric connection between its pKID and MLL binding sites, make it challenging to assess potential KIX inhibitors using traditional biophysical techniques<sup>22</sup>. While standard mass spectrometry (MS)-based screens can detect target-ligand binding, the functional and structural consequences of such binding is not often available<sup>20, 23</sup>. Similarly, assays such as competitive fluorescence polarization assays can directly target an interaction between a specific binding partner and a target protein<sup>24</sup>; however, secondary assays typically are required to reveal the full structural consequences of ligand binding of any allosteric sites present in the system. Thus, it is clear that new rapid technologies, sensitive to both ligand binding and protein structure, are needed for PPI inhibitor discovery efforts for KIX, as well as for PPI inhibitor search efforts more broadly.

To this end, we describe here an ion mobility-mass spectrometry (IM-MS) based method capable of detecting peptide binding and assigning the location of that binding uniquely either to either of the two available sites within KIX. In general, IM-MS can measure both protein mass and size within complex mixtures, requiring only small amounts (<1 ng) of unlabeled analyte. In an IM-MS experiment, intact protein:inhibitor complexes are ionized by nano-electrospray ionization (nESI) prior to separation by size and charge by IM and mass analysis by a time-of-flight mass analyzer. While protein complexes that are >2% different in total global size (collision cross

section, CCS) can be readily separated by IM alone, proteins that undergo smaller conformation changes induced by ligand binding can be separated by a gas-phase activation technique called collision induced unfolding (CIU)<sup>25</sup>, which is analogous to a gas-phase calorimetry experiment. Additionally, collision induced dissociation (CID) can be used to measure the ejection of ligands from protein:ligand complexes<sup>26</sup>. CIU has been used to distinguish between modes of ATP-competitive kinase inhibitors<sup>27</sup>, in biosimilar analysis<sup>28, 29</sup>, and to probe the polydispersity of intrinsically disordered proteins<sup>30-34</sup>. In this report, we demonstrate the ability of CIU information to localize the binding of four peptides from MLL, E2A, pKID, and c-Myb, within the KIX domain. Through a detailed quantitative analysis of the CIU data presented, we speculate on the origins of the conformational transitions observed and their broader analytical utility in KIX inhibitor discovery efforts.

## 6.2 Methods and Materials

### 6.2.1 Protein Expression and Purification

The DNA sequence encoding the KIX domain from mouse CBP (residues 586-672) was cloned into the bacterial expression pRSETB vector with an additional hexahistadine tag and a short polar linker fused to the N-terminus of KIX resulting in protein with the sequence (tag and linker residues are shown in lower case)<sup>35</sup>:  
 mrgshhhhhhgm<sup>a</sup>sGVRKGWHEHVTQDLRSHLVHKL<sup>V</sup>QAIFTPDPAALKDRRMENLVAYA  
 KKVEGDMYESANSRDEYYHLLAEKIYKIQKELEEKRRSRL

The pRSETB KIX plasmid was transformed into Rosetta<sup>TM</sup> 2(DE3) pLysS Escherichia coli (Novagen) competent, and a colony was used to inoculate 50 mL LB starter cultures containing 0.1 mg/ml ampicillin and 0.34 µg/ml chloramphenicol and allowed to shake overnight at 250 rpm, 37 C. The overnight culture was added to 1 L of sterile terrific broth (12 g tryptone, 24 g yeast extract, 4% (v/v) glycerol, 900 mL water, 100 mM 1M potassium phosphate buffer) at a ratio of 1 mL overnight to 100 mL media. Cells were grown to an OD<sub>600nm</sub> of approximately 0.8-1.0 (370C, 250 rpm) and induced with 0.25 mM IPTG (25 C, 18 hours). The cells were harvested by centrifugation for 15 minutes at 8,000xg and the cell pellet was stored at -800C. From a 1 L expression the pellet was thawed on ice and resuspended in 20 mL (or approximately 2-5 mL per gram wet weight) of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 7.2 10 mM β-ME). Cells were lysed by sonication and the lysate was separated from cellular

debris by centrifugation at 9,000xg for 30 minutes at 4 C. For nickel affinity purification standard batch purification was followed as previously described using Ni-NTA agarose resin (Qiagen) according to the manufacture's protocol. The protein was eluted with a high concentration of imidazole (50 mM sodium phosphate, 300 mM NaCl, 400 mM Imidazole, pH 7.2, 10 mM  $\beta$ -ME). The KIX protein was further isolated from contaminants by cation-exchange chromatography with Source 15S column (GE Healthcare) using the AKTA FPLC system using buffer A (50 mM sodium phosphate pH 7.2 1 mM DTT) and buffer B (50 mM sodium phosphate, 1 M NaCl, pH 7.2 1 mM DTT). After the protein was applied to the column unbound sample was washed for 40 mL or two column volumes. For the first segment 60% of buffer B was reached over 100 mL or 5 column volumes. For the second segment 100% of buffer B ran through for 40 ml or two column volumes.

Purified protein was buffer then exchanged into 10 mM sodium phosphate, 100 mM NaCl, pH 6.8 using a PD-10 column (GE Healthcare). The protein concentration is determined by ultraviolet light with a wavelength at 280 nm using the extinction coefficient  $\epsilon$  of 12950 cm<sup>-1</sup> M<sup>-1</sup>. Protein purity was assessed by electrophoresis using a Bis-Tris gel in MES running buffer.

### **6.2.2 Peptide Synthesis**

Peptides were synthesized on CLEAR<sup>TM</sup> (Cross-linked Ethoxylate Acrylate Resins) amide resin (Peptides international) by standard HBTU/HOBT/DIEA and N-9 Fluorenylmethoxycarbonyl (Fmoc) solid phase synthesis methods as previously described and purified by HPLC<sup>36</sup>. Dr. Ningkun Wang provided the c-Myb25mer and pKID29mer peptide.

### **6.2.3 Ion Mobility-Mass Spectrometry**

Peptide was added to KIX in a 10:1 molar ratio and incubated at room temperature for 30 minutes. Samples were then buffer exchanged into 100 mM ammonium acetate at pH 7 using Micro Bio-Spin 6 columns (Bio-Rad, Hercules, CA) and prepared to a final concentration of 25  $\mu$ M. After buffer exchange, all samples were moved to ice before analysis. 5-7 $\mu$ L of sample was analyzed on a quadrupole-ion mobility-time-of-flight mass spectrometer (Q-IM-TOF) instrument (Synapt G2 HDMS, Waters, Milford, Ma). Samples were ionized with a nESI source as previously described<sup>37, 38</sup>. Capillary voltages ranged from 1.5-1.7V, with the sample cone operating at 10V and the extraction cone set to 1V in positive ion mode. The quadrupole was operated at a range of 500-

5000 m/z and set to scan automatically. The trap traveling-wave ion guide was pressurized to 2 ml/min with the ion mobility separation cell operating with a wave height of 25 V and a wave velocity of 600 m/s. The TOF was operated at a pressure of  $8.70 \times 10^{-7}$  mbar. The protein:peptide ratios were kept sufficiently low to avoid the formation of artifact complexes. All mass spectra were externally calibrated using a solution of cesium iodide (20 mg/ml) and processed using the Masslynx 4.1 software (Waters, UK).

For collision induced unfolding experiments, ions were activated in the trap traveling-wave ion guide. The 8+ charge state was selected on the basis of its foldedness, high intensity, and the number of intermediate conformations observed during CIU. Each ion was mass-selected in the quadrupole prior to activation by increasing the collision voltage from 4V to 90V in steps of 2V. The corresponding drift times for the 8+ ions were extracted using an in-house data extraction tool and processed using the CIUSuite software package<sup>39</sup>.

## **6.3 Results**

### **6.3.1 Each Peptide Binding Type Has a Unique CIU Fingerprint**

In the mass spectra, the apo KIX along with the KIX:peptide peaks are observed. Furthermore, in the presence of two peptides that target different KIX binding sites, we are able to observe binding of both peptides individually and distinguish these signals from the intact ternary complex. An advantage to MS is that these signals can be individually analyzed, as opposed to most bulk solution-phase studies. However, neither evidence of cooperative binding nor binding site information is not readily apparent from these data. Although peptide binding caused a shift in the CCS in addition to m/z of the KIX:peptide complexes, these changes were related to the overall size of the peptide rather than a shift in the conformation of KIX induced by peptide binding. For example, the measured CCS of the KIX:pKID 8+ charge state is  $1397 \pm 20 \text{ \AA}^2$ , compared to the  $1382 \pm 19 \text{ \AA}^2$  CCS measured for the KIX:MLL 8+ charge state. This is a less than 2% difference between the CCS of these ions, which within the error of the CCS measurement. Therefore, in order to separate KIX complexes according to their resulting structures, we implemented a CIU strategy that aimed to identify differences in the unfolding patterns of 8+ KIX complexes

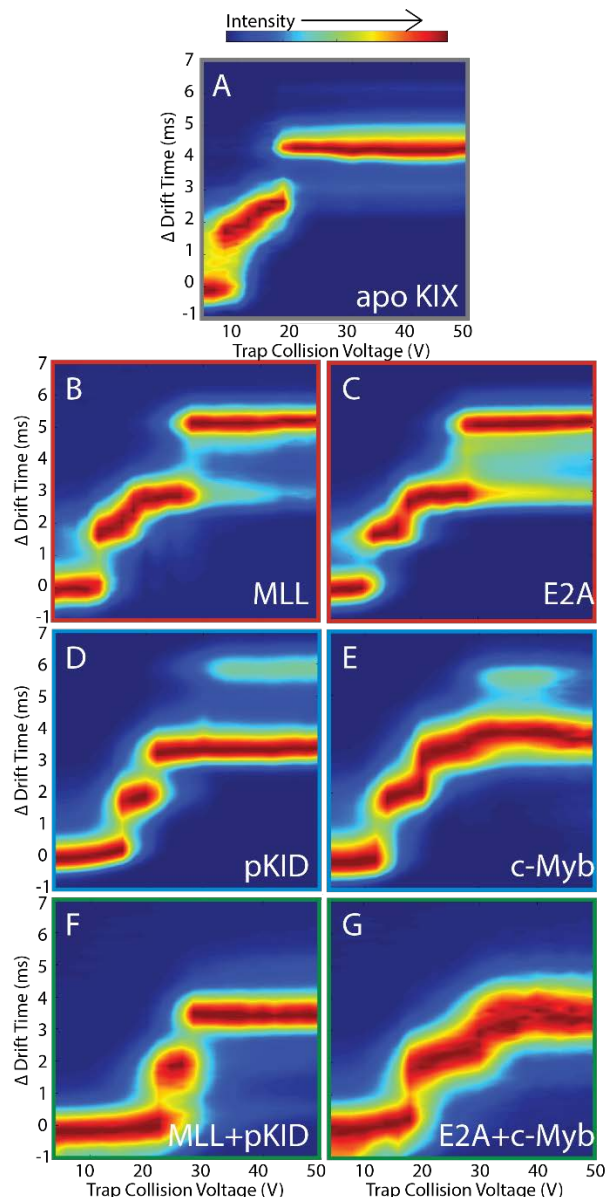


Figure 6.2 CIU fingerprints for KIX:peptide complexes. (A) CIU fingerprint for the apo KIX 8+ ions (B) KIX:MLL CIU fingerprint, which has for CIU features. (C) KIX:E2A also has four CIU features that occur at similar areas in the  $\Delta$ drift time axis. (D) KIX:pKID also has four CIU features, although the fourth has less than 50% of the normalized intensity. (E) KIX:c-Myb also has four unfolding features, and the fourth feature has less intensity than the third. (F) CIU fingerprint of the ternary KIX:MLL:pKID complex, which lacks the fourth feature. (G) The ternary complex of KIX:E2A:c-Myb also lacks the fourth unfolding feature.

diagnostic for the occupied binding locations within the protein. Here, CIU fingerprints were plotted as a function of  $\Delta$ drift time in order to normalize data for the increase in ion CCS caused by peptide binding (Figure 6.2). For the binary KIX:peptide complexes, we observe four CIU features having a  $\Delta$ drift time centroid values of -0.01, 2.26, 4.75, and 5.06 (Figure 6.2B-E), while the ternary KIX:peptide complexes have  $\Delta$ drift time centroid values of 0.21 ms, 2.30 ms and 3.93 ms. A cursory examination of data shown in Figure 6.2 indicates that MLL and E2A bound KIX ions appear to follow similar unfolding pathways in the gas-phase (Figure 6.2A,B), whereas pKID and c-Myb exhibit similar CIU data (Figure 6.2C, D). In the CIU fingerprints of the ternary complexes formed by KIX:MLL:pKID and KIX:E2A:c-Myb, the fourth feature is not observed (Figure 6.2E,F), and the first feature is highly stabilized (24V versus 4V for apo KIX) (Figure 6.2E,F). It is also apparent that the fourth feature observed in Figure 6.2E and F is composed of ~50% normalized intensity, as opposed to this

feature maintaining 100% intensity in the MLL-like binders (Figure 6.2B,C). The first three features are also stabilized relative to the apo KIX CIU fingerprint (Figure 6.2A). While there are

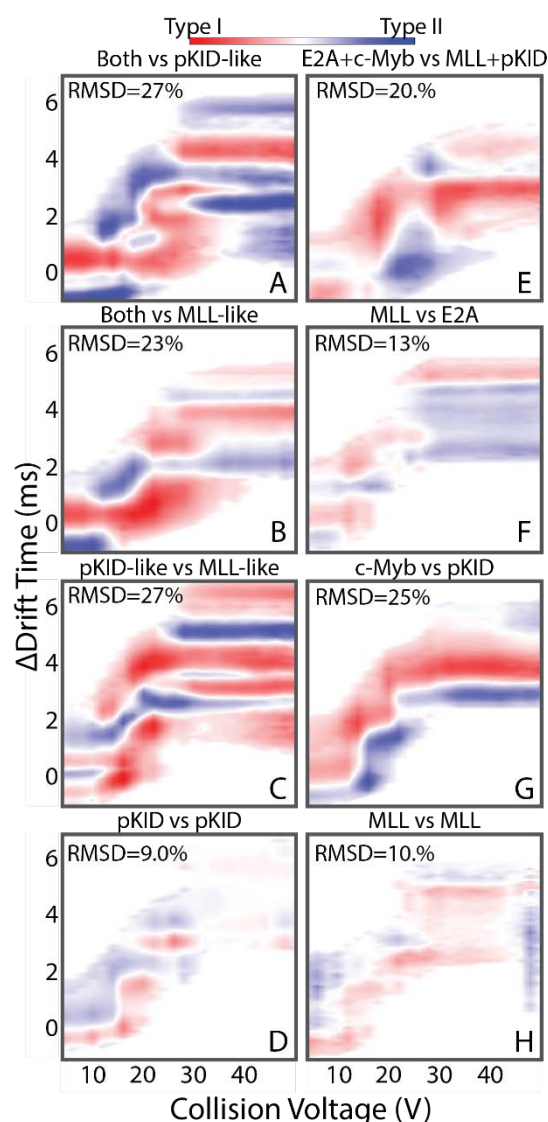


Figure 6.3 RMSD values computed from CIUSuite. (A) RMSD plot the average ternary complex against the average pKID-like CIU fingerprint. (B) RMSD for the ternary complex and average MLL-like CIU fingerprints (C) RMSD values for the average pKID-like vs MLL-like CIU fingerprints. (D) RMSD values for two replicates of KIX:pKID CIU fingerprints. (E) E2A+c-Myb vs ML+pKID. (F) MLL vs E2A. (G) c-Myb vs pKID. (H) MLL versus a replicate of MLL.

KIX:pKID CIU fingerprint has a stability of  $19 \pm 0$ , whereas KIX:c-Myb fourth feature has a stability of  $19 \pm 1$ . However, KIX complexes follow the trend of low RMSD values (9-10%) for replicates, medium RMSD values within the same class of binders (13% to 25%), and high RMSD values when cross-comparing classes (23% to 27%). Furthermore, when comparing the RMSD

slight differences in the Δdrift time centroid values, the main force between the differences of the CIU fingerprints are the number and stability of the CIU features.

Although the differences between the MLL-like, pKID-like, and ternary-like CIU fingerprints that are discernible by eye, we used CIUSuite<sup>39</sup> to calculate RMSD values to determine precise RMSD values for the similarities and differences between the CIU fingerprints of each type of binder (Figure 6.3).

The RMSD values between the three different classes (Figure 6.3A,B,C) range between 23 and 27 %. This is 2-3x higher than the RMSD values between the replicates of pKID and MLL (Figure 6.3D,H), which are 9% and 10 %, respectively. The RMSD values between the average CIU fingerprints of within each class (Figure 6.3 E,F,G) range from 13 to 25%. Interestingly, c-Myb and pKID have the greatest RMSD value between similar binding modes, with an RMSD value of 25%. We attribute this to the slightly different binding locations of these two peptides<sup>9, 11</sup> (Figure 6.1A) and the differences in the stability of the fourth feature. The fourth feature in the



values to each peptide type (Figure 6.4A), it is clear that the lowest RMSD values are for similar binders. For example, when comparing all RMSD value for all peptide against E2A, MLL has the lowest RMSD value.

### 6.3.2 Determining Differences in CIU Stability for Ligand Screening

The stabilities of each CIU feature were also calculated (Figure 6.4A) with CIUSuite, which uses

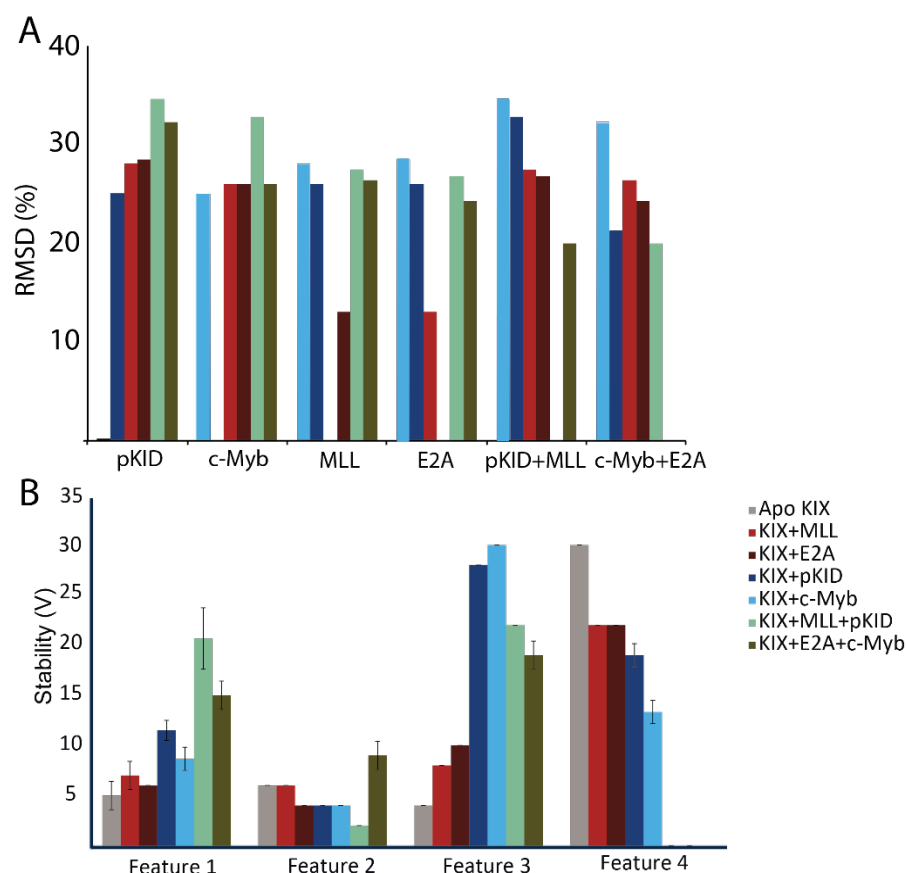


Figure 6.4. Histograms of the RMSD values and stabilities of the CIU features observed. (A) RMSD values from Figure 6.3 in histogram form. (B) Histogram of the stabilities of each of the CIU features observed in the gas-phase unfolding process.

a first derivative test to identify local maxima in the data to define the features of the CIU fingerprints in terms of centroid drift time and stability (Figure 6.4B). In this analysis, it was found that the first feature in apo KIX had a stability of  $5 \pm 1$  V, compared to stability values of  $7 \pm 1$  V and  $6 \pm 0$  V for KIX:MLL and KIX:E2A, stability values of  $12 \pm 1$  V and  $9 \pm 1$  V for KIX:pKID and KIX:c-Myb, and  $21 \pm 3$  V and  $15 \pm 1$  V for KIX:MLL:pKID and KIX:E2A:c-Myb. While

there is not much stability imparted by peptide binding on the second feature, the third ternary feature is greatly stabilized by pKID-like binders, with stabilities of  $28 \pm 0$  V and  $22 \pm 0$  V for KIX:pKID and KIX:c-Myb respectively, as compared to  $4 \pm 0$  V,  $8 \pm 0$  V,  $10 \pm 0$  V,  $22 \pm 0$  V and  $19 \pm 1$  V for apo KIX, KIX:MLL, KIX:E2A, KIX:MLL:pKID, and KIX:MLL:E2A. Notably, the fourth feature is missing in the ternary complexes, and the stabilities increase from pKID-like binders

( $19\pm 1V$  and  $13\pm 1V$  for KIX:pKID and KIX;MLL) to the MLL-like binders ( $22\pm 0V$  for both KIX:MLL and KIX:E2A), to apo KIX with a stability of  $30\pm 0V$ .

These differences in the stabilities of these features can be used to hypothesize about their origins. Recently, we discovered a strong correlation between the number of CIU transitions and the number of domains for multi-protein domain systems<sup>40</sup>. While KIX is a single domain, there are three  $\alpha$ -helices, which correlates well to our observations of four CIU transitions for our apo and binary complex CIU fingerprints. Furthermore, we note that the fourth CIU feature appears to dramatically decrease in intensity upon pKID and c-Myb binding and is completely obliterated upon ternary complex formation. These observations lead us to speculate that the third to fourth CIU transition is caused by the unfolding of the third  $\alpha$ -helix, as all four peptides contact this helix, and binding in both sites abolishes the fourth unfolding feature. As the second unfolding feature is the shortest-lived for all KIX complexes, we hypothesized that the transition from the second to

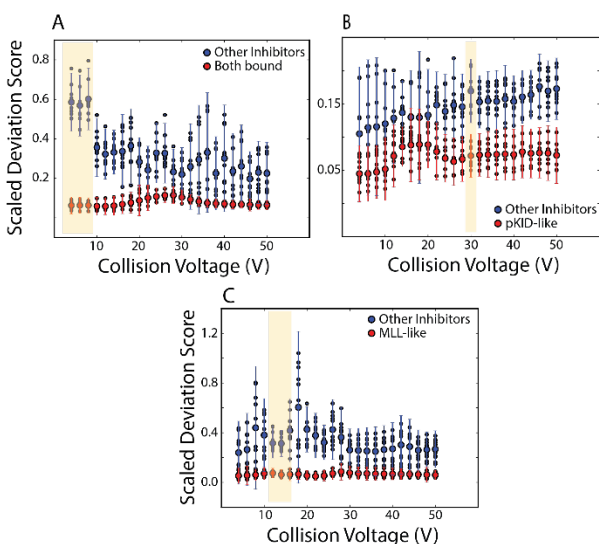


Figure 6.5 SDS vs collision voltage for peptides compared to an (A) ternary complex average, (B) a pKID-like average, or a (C) MLL-like average. The blue circles represent the average and standard deviations of the other peptides, and the red circles represent the average and standard deviation of the binding complex that was used for the comparison.

third CIU feature is caused by the unfolding of the second  $\alpha$ -helix. Finally, we hypothesize that the transition from the first to second unfolding feature is caused by the unfolding of the first  $\alpha$ -helix, as the first CIU feature is stabilized by ligand binding.

### 6.3.3 Using CIU to Localize Ligand Binding

Previously, we have demonstrated that the best way to exploit CIU results to screen for ligands is to use a region in the CIU fingerprints where the differences between different classes of ligands are maximized<sup>27</sup>. In order to achieve this goal, we utilized the analysis module of CIUSuite. The output of the analysis module is

a graph of the average and standard deviation scaled deviation score (scaled z-score, SDS) for two peptide types collision voltage against collision voltage. The SDS is computed from the z-score of all peptides against either a “type I” or “type II” average. The CIU fingerprints of all KIX:peptide

complexes were compared against the averages of the MLL-like, pKID-like, or ternary-complex-like average CIU fingerprint (Figure 6.5). Whenever the KIX:peptide complexes were compared to a ternary (Figure 6.5A) or MLL-like (Figure 6.5B) average, a clear separation was achieved (green box). However, there was no clear separation between the complexes when compared to a pKID-like average (Figure 6.5B). Furthermore, while the MLL-like (Figure 6.5C) and ternary-like (Figure 6.5A) complexes had SDS scores near zero in comparison to the other peptides, the pKID-like peptides exhibited a wide variation (Figure 6.5B). The area between 4 and 8V, 30V, and 12-16V was selected for further analysis for comparison against both bound, pKID-like bound, and

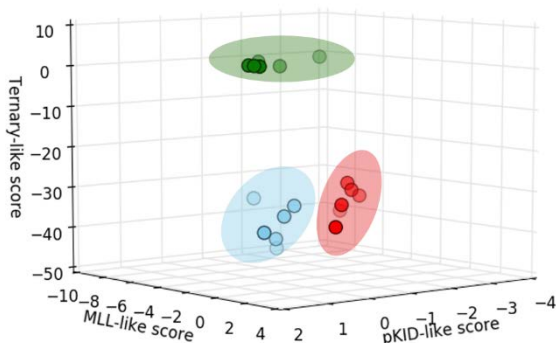


Figure 6.6. Z-scores for each KIX:peptide complex against the MLL-like average, the pKID-like average, and the ternary complex. The pKID-like peptides are represented by the blue circles the MLL-like by the red circles, and the ternary-like in the green circles. The ellipses around the points are to illustrate the clustering of the peptides.

MLL-like bound respectively, as these regions exhibited the best area of differentiation according to the SDS graphs.

In order to better represent the clustering of each peptide type., the z-scores for each peptide against an MLL-like, pKID-like, or ternary-like complex were computed (Figure 6.6). The z-score indicates how far away a given point is from the average. Against the average MLL-like score, MLL-like binders had an average score of  $0.1 \pm 1.1$ , while pKID-like peptides had a score of  $-2.1 \pm 3.1$  and the ternary complexes had a score of  $-5.3 \pm 2.4$ . Against

the pKID-like score, MLL-like binders had a score of  $-1.1 \pm 0.6$ , pKID-like binders had a score of  $1.7 \times 10^{-10} \pm 1.1$ , and ternary binders a score of  $-0.4 \pm 1.5$ . Against a ternary average, pKID-, MLL-, and ternary-like peptides had scores of  $-4.0 \pm 1.9$ ,  $-5.7 \pm 1.9$ , and  $0.2 \pm 1.1$ , respectively. These scores indicate that each complex type clusters within one standard deviation of the average score within its type. Overall, this clustering analysis provides a way forward for a screen to rapidly localize the binding site of peptides to KIX.

## 6.4 Conclusions

Here, we have presented an assay based on IM-MS that can rapidly identify the binding site of a peptide to KIX by using CIU. This is the first example to use CIU to identify separate, unique fingerprints for remote binding sites on a single protein. The detailed comparisons of the differences in the CIU fingerprints, quantified by RMSD values, reveal that the differences in the observed CIU features arise from differences in the stabilization of the CIU features imparted by peptide binding, which can be used in combination with the knowledge of where a peptide binds to infer the gas-phase unfolding pathway of the KIX:peptide complexes. We hypothesize that KIX unfolds in a linear  $\alpha$ -helix 1-3 fashion in the gas-phase. Further, we identified a region between 12-16V that provided the best separation between complex types, which could be used in a potential screen to rapidly identify peptide localization.

However, neither MS nor CIU data were able to reveal information about the allostery between the pKID and MLL binding sites. Careful  $K_d$  measurements using MS<sup>41</sup> by titrating in one peptide and holding the other constant, could reveal the allosteric mechanism between the two binding sites. Careful monitoring and controls should be used with this approach, as the ionization efficiencies of the peptides in this study are different. These  $K_d$  measurements could later be built into a full CIU-MS screen where both information about allostery and ligand localization could be combined for a high-throughput screen.

## 6.5 References

- [1] Haura, E. B., Turkson, J., and Jove, R. (2005) Mechanisms of disease: Insights into the emerging role of signal transducers and activators of transcription in cancer, *Nature clinical practice Oncology* 2, 315-324.
- [2] Lee, L. W., and Mapp, A. K. (2010) Transcriptional switches: chemical approaches to gene regulation, *Journal of Biological Chemistry* 285, 11033-11038.
- [3] Mapp, A. K. (2003) Regulating transcription: a chemical perspective, *Organic & biomolecular chemistry* 1, 2217-2220.
- [4] Thompson, A. D., Dugan, A., Gestwicki, J. E., and Mapp, A. K. (2012) Fine-tuning multiprotein complexes using small molecules, *ACS chemical biology* 7, 1311-1320.
- [5] Wells, J. A., and McClendon, C. L. (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces, *Nature* 450, 1001-1009.

- [6] Babu, M. M., van der Lee, R., de Groot, N. S., and Gsponer, J. (2011) Intrinsically disordered proteins: regulation and disease, *Current opinion in structural biology* 21, 432-440.
- [7] Sugase, K., Dyson, H. J., and Wright, P. E. (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein, *Nature* 447, 1021-1025.
- [8] Radhakrishnan, I., Pérez-Alvarado, G. C., Parker, D., Dyson, H. J., Montminy, M. R., and Wright, P. E. (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator: coactivator interactions, *Cell* 91, 741-752.
- [9] Zor, T., De Guzman, R. N., Dyson, H. J., and Wright, P. E. (2004) Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb, *Journal of molecular biology* 337, 521-534.
- [10] Mayr, B., and Montminy, M. (2001) Transcriptional regulation by the phosphorylation-dependent factor CREB, *Nature reviews Molecular cell biology* 2, 599-609.
- [11] Radhakrishnan, I., Pérez-Alvarado, G. C., Dyson, H. J., and Wright, P. E. (1998) Conformational preferences in the Ser 133-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB, *FEBS letters* 430, 317-322.
- [12] Denis, C. M., Chitayat, S., Plevin, M. J., Wang, F., Thompson, P., Liu, S., Spencer, H. L., Ikura, M., LeBrun, D. P., and Smith, S. P. (2012) Structural basis of CBP/p300 recruitment in leukemia induction by E2A-PBX1, *Blood* 120, 3968-3977.
- [13] Bannister, A. J., Oehler, T., Wilhelm, D., Angel, P., and Kouzarides, T. (1995) Stimulation of c-Jun activity by CBP: c-Jun residues Ser63/73 are required for CBP induced stimulation in vivo and CBP binding in vitro, *Oncogene* 11, 2509-2514.
- [14] Vendel, A. C., McBryant, S. J., and Lumb, K. J. (2003) KIX-mediated assembly of the CBP-CREB-HTLV-1 tax coactivator-activator complex, *Biochemistry* 42, 12481-12487.
- [15] Vendel, A. C., and Lumb, K. J. (2004) NMR mapping of the HIV-1 Tat interaction surface of the KIX domain of the human coactivator CBP, *Biochemistry* 43, 904-908.
- [16] Brüschweiler, S., Konrat, R., and Tollinger, M. (2013) Allosteric communication in the KIX domain proceeds through dynamic repacking of the hydrophobic core, *ACS chemical biology* 8, 1600-1610.
- [17] Rowe, S. P., and Mapp, A. K. (2008) Assessing the permissiveness of transcriptional activator binding sites, *Biopolymers* 89, 578-581.

- [18] Gee, C. T., Koleski, E. J., and Pomerantz, W. C. (2015) Fragment Screening and Druggability Assessment for the CBP/p300 KIX Domain through Protein-Observed  $^{19}\text{F}$  NMR Spectroscopy, *Angewandte Chemie International Edition* 54, 3735-3739.
- [19] Palazzesi, F., Barducci, A., Tollinger, M., and Parrinello, M. (2013) The allosteric communication pathways in KIX domain of CBP, *Proceedings of the National Academy of Sciences* 110, 14237-14242.
- [20] Wang, N., Majmudar, C. Y., Pomerantz, W. C., Gagnon, J. K., Sadowsky, J. D., Meagher, J. L., Johnson, T. K., Stuckey, J. A., Brooks III, C. L., and Wells, J. A. (2013) Ordering a dynamic protein via a small-molecule stabilizer, *Journal of the American Chemical Society* 135, 3363-3366.
- [21] Majmudar, C. Y., Højfeldt, J. W., Arevang, C. J., Pomerantz, W. C., Gagnon, J. K., Schultz, P. J., Cesa, L. C., Doss, C. H., Rowe, S. P., and Vásquez, V. (2012) Sekikaic acid and lobaric acid target a dynamic interface of the coactivator CBP/p300, *Angewandte Chemie International Edition* 51, 11258-11262.
- [22] Wilson, C. G., and Arkin, M. R. (2013) Probing structural adaptivity at PPI interfaces with small molecules, *Drug Discovery Today: Technologies* 10, e501-e508.
- [23] Erlanson, D. A., Wells, J. A., and Braisted, A. C. (2004) Tethering: fragment-based drug discovery, *Annu. Rev. Biophys. Biomol. Struct.* 33, 199-223.
- [24] Owicki, J. C. (2000) Fluorescence polarization and anisotropy in high throughput screening: perspectives and primer, *Journal of Biomolecular Screening* 5, 297-306.
- [25] Hyung, S.-J., Robinson, C. V., and Ruotolo, B. T. (2009) Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes, *Chemistry & biology* 16, 382-390.
- [26] Niu, S., Rabuck, J. N., and Ruotolo, B. T. (2013) Ion mobility-mass spectrometry of intact protein–ligand complexes for pharmaceutical drug discovery and development, *Current opinion in chemical biology* 17, 809-817.
- [27] Rabuck, J. N., Hyung, S.-J., Ko, K. S., Fox, C. C., Soellner, M. B., and Ruotolo, B. T. (2013) Activation state-selective kinase inhibitor assay based on ion mobility-mass spectrometry, *Analytical chemistry* 85, 6995-7002.

- [28] Tian, Y., Han, L., Buckner, A. C., and Ruotolo, B. T. (2015) Collision induced unfolding of intact antibodies: rapid characterization of disulfide bonding patterns, glycosylation, and structures, *Analytical chemistry* 87, 11509-11515.
- [29] Beck, A., Debaene, F., Diemer, H., Wagner-Rousset, E., Colas, O., Dorsselaer, A. V., and Cianférani, S. (2015) Cutting-edge mass spectrometry characterization of originator, biosimilar and biobetter antibodies, *Journal of Mass Spectrometry* 50, 285-297.
- [30] Saikusa, K., Kuwabara, N., Kokabu, Y., Inoue, Y., Sato, M., Iwasaki, H., Shimizu, T., Ikeguchi, M., and Akashi, S. (2013) Characterisation of an intrinsically disordered protein complex of Swi5–Sfr1 by ion mobility mass spectrometry and small-angle X-ray scattering, *Analyst* 138, 1441-1449.
- [31] Pagel, K., Natan, E., Hall, Z., Fersht, A. R., and Robinson, C. V. (2013) Intrinsically disordered p53 and its complexes populate compact conformations in the gas phase, *Angewandte Chemie International Edition* 52, 361-365.
- [32] Vahidi, S., Stocks, B. B., and Konermann, L. (2013) Partially disordered proteins studied by ion mobility-mass spectrometry: implications for the preservation of solution phase structure in the gas phase, *Analytical chemistry* 85, 10471-10478.
- [33] Dickinson, E. R., Jurneczko, E., Nicholson, J., Hupp, T. R., Zawacka-Pankau, J., Selivanova, G., and Barran, P. E. (2015) The use of ion mobility mass spectrometry to probe modulation of the structure of p53 and of MDM2 by small molecule inhibitors, *Frontiers in molecular biosciences* 2.
- [34] D'Urzo, A., Konijnenberg, A., Rossetti, G., Habchi, J., Li, J., Carloni, P., Sobott, F., Longhi, S., and Grandori, R. (2015) Molecular basis for structural heterogeneity of an intrinsically disordered protein bound to a partner by combined ESI-IM-MS and modeling, *Journal of The American Society for Mass Spectrometry* 26, 472-481.
- [35] Buhrlage, S. J., Bates, C. A., Rowe, S. P., Minter, A. R., Brennan, B. B., Majmudar, C. Y., Wemmer, D. E., Al-Hashimi, H., and Mapp, A. K. (2009) Amphipathic small molecules mimic the binding mode and function of endogenous transcription factors, *ACS chemical biology* 4, 335-344.
- [36] Pomerantz, W. C., Wang, N., Lipinski, A. K., Wang, R., Cierpicki, T., and Mapp, A. K. (2012) Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization, *ACS chemical biology* 7, 1345-1350.

- [37] Zhong, Y., Hyung, S.-J., and Ruotolo, B. T. (2011) Characterizing the resolution and accuracy of a second-generation traveling-wave ion mobility separator for biomolecular ions, *Analyst* 136, 3534-3541.
- [38] Hernández, H., and Robinson, C. V. (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry, *Nature protocols* 2, 715-726.
- [39] Eschweiler, J. D., Rabuck-Gibbons, J. N., Tian, Y., and Ruotolo, B. T. (2015) CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions, *Analytical chemistry* 87, 11516-11522.
- [40] Zhong, Y., Han, L., and Ruotolo, B. T. (2014) Collisional and Coulombic Unfolding of Gas-Phase Proteins: High Correlation to Their Domain Structures in Solution, *Angewandte Chemie* 126, 9363-9366.
- [41] Jecklin, M. C., Touboul, D., Bovet, C., Wortmann, A., and Zenobi, R. (2008) Which electrospray-based ionization method best reflects protein-ligand interactions found in solution? A comparison of ESI, nanoESI, and ESSI for the determination of dissociation constants with mass spectrometry, *Journal of the American Society for Mass Spectrometry* 19, 332-343.



## **Chapter 7.**

### **Native Mass Spectrometry Exercises for the Undergraduate Laboratory**

Native mass spectrometry involves the transfer of intact macromolecular structures from solution to the gas-phase, offering a unique molecular viewpoint on biological complexes such as proteins, nucleic acids, and lipids. The ability to interrogate samples that are difficult or impossible by more traditional biophysical techniques, has led to the rapid growth of native mass spectrometry as a key analytical skill set in academic and industrial careers alike. By contrast, the chemical education literature is markedly lacking in experimental protocols that introduce undergraduate students to this important technology. Here, we present two laboratory protocols that were specifically developed for undergraduate education in a biomedical analytical chemistry laboratory using systems that are well-studied and which expose students to some of the most crucial aspects of this approach.

#### **7.1 Introduction**

Native mass spectrometry (native MS) is an important analytical technique employed to determine structural information about macromolecular complexes that frustrate traditional biophysical techniques. In this approach, native-like biological complexes can be transferred intact from the solution to the gas-phase by applying “soft” ionization techniques like electrospray ionization (ESI) or nano-electrospray ionization (nESI), along with and carefully controlled instrument parameters<sup>1, 2</sup> to maintain native-like complexes in the gas-phase. For more than 20 years, native MS has provided a novel molecular viewpoint on a vast range of difficult targets from small aggregating systems<sup>3-5</sup> to measurements of intact virus capsids<sup>6-9</sup>.

While the use of native MS has grown exponentially over the past this time<sup>10</sup>, with a little under 150 papers published in 1997 and over 500 papers published in 2015, there has been few concomitant efforts to introduce such topics into undergraduate curricula. For most students, their

first introduction to applied native MS, if any, occurs in graduate school. One of the reasons for this lack of instruction is the aforementioned dearth of experimental protocols available to undergraduate instructors. This results in a workforce that is ill-prepared for the challenges and opportunities of using this technology.

To address this discrepancy, we designed two laboratory experiments to equip undergraduate students with skills in applied native MS technologies using two well-studied protein systems: the lysozyme:N-acetyl-chitose protein:ligand complex<sup>11</sup>, and the concanavalin A (ConA) tetramer<sup>12</sup>. These experiments were created as a part of a Future Faculty Graduate Student Instructor (FFGSI) position developed by the Chemical Sciences at the Interface of Education|University of Michigan (CSIE|UM) program<sup>13</sup>. During this experiment, students learned the principles behind native MS experiments, such as the importance of MS-compatible buffers, the importance of tuning an instrument parameter file to maintain protein complexes, and the general principles behind native MS. Here, we describe the development and deployment of two separate, two-laboratory period experiments designed to introduce students to native MS techniques.

### **7.1.1 Curriculum Interview**

The University of Michigan introduced the Biomedical Analytical Chemistry Laboratory in the winter semester of 2011 in order to impart the fundamental principles of applied analytical chemistry in a clinical chemistry/biomedical measurement setting. The course consisted of eleven laboratory experiments, nine of which are completed in a single four-hour session. These experiments are executed in two parts, with the students introduced to the equipment via the generation of a calibration curve followed by measurement of a concentration of an unknown sample. The last two experiments, involving high performance liquid chromatography (HPLC) and a Special Project, each comprise two four-hour laboratory sessions. For the Special Projects, students work in groups of two or three to develop their own experiments, which are then vetted by the faculty instructor and GSIs. This year, three groups were selected to participate in the native MS experiments instead of their proposed special projects. Two groups performed the lysozyme experiments, and one group performed the ConA experiments.

As part of this pilot program, the experiments were performed on a Waters Synapt HDMS (Waters, Milford MA) or an LCT premier (Waters, Milford, MA) housed in the Ruotolo research group.

We plan for future protocols to be implemented in the Biomedical Analytical Chemistry Laboratory as a part of the regular experiment schedule. The Synapt was chosen for the ConA experiments because it has the ability to break apart complexes using collision induced dissociation (CID). The LCT was selected for the lysozyme experiments due to its sensitivity, simplicity, and relative ease of use for the students. While both of these platforms are commonly available, it is important to note that other MS platforms are also available that will suffice for implementing the exercises described here<sup>1</sup>. Samples are introduced via nESI using gold coated glass capillaries which are available commercially, but in this case were prepared in-house<sup>1, 14</sup> by the FFGSI and were loaded by the students themselves.

## **7.2 Mass Spectrometry of Multiprotein Complexes for the Undergraduate Laboratory**

ConA (Figure 7.1A) was selected as a model system for this course due to its importance in structural biology and the mass spectrometry community. It is the first carbohydrate-binding protein for which a three-dimensional structure was determined<sup>15</sup>, and it is among the first proteins to be studied by native MS<sup>12</sup>. In 1994, the Smith group demonstrated that specific solution structures of multimeric protein complexes could be observed via native MS<sup>16</sup>, including the dimeric structure of ConA at pH 5.7 and the tetrameric form at pH 8.4. The demonstration that oligomeric species in the gas-phase follow well-known trends in the solution-phase highlighted the potential of mass spectrometry for protein complexes. As such, this system was employed to acclimate undergraduates to native MS software and technology. This experiment was broken into two laboratory periods

The instructor began with a brief tutorial about native MS technologies and demonstrated data collection on the Synapt. Before sample analysis, students sprayed a solution of cesium iodide (CsI) in order to determine how well the mass spectrometer was calibrated using an instrument parameter file developed by the instructor. Students were then given two samples of ConA in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) which is a common buffer in structural biology experiments. One sample was kept on ice as a standard, and the other sample was buffer exchanged into 200 mM ammonium acetate at pH 7, which is a native mass spectrometry-compatible buffer, following the protocols used by the Ruotolo lab<sup>2</sup> and others<sup>1, 14</sup>. Students took

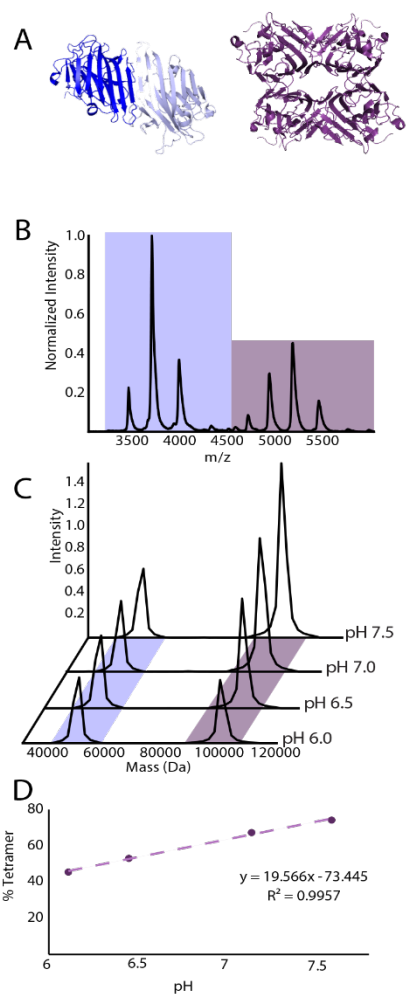


Figure 7.1. MS analysis of ConA (A) Structures of the ConA dimer (blue) and tetramer (purple) from PDB ID 4P9W. (B) Mass spectrum at pH ~6.0. The dimer population is highlighted in blue, and the tetramer population of purple. (C) Deconvoluted spectra showing the mass of the two species. The tetramer species, ~50 kDa, decreases in intensity compared to the tetramer (purple) species as the pH increases. (D) Example calibration curve of percent tetramer versus pH. The  $R^2$  value is 0.9957.

UV-Vis measurements at 280 nm before and after buffer exchange to report how much sample was lost during the buffer exchange process.

After collecting a spectrum of CsI, students sprayed the buffer exchanged and non-buffer exchanged ConA samples. They observed that while no protein peaks were detected for ConA in the non-buffer exchanged sample, due to the number of HEPES salt adducts on the protein and high salt background, they could see peaks corresponding to the dimer and tetrameric forms of ConA in the buffer exchanged samples. Students then initiated CID to collisionally clean the protein sample of residual salt molecules in order to calculate an accurate mass using the Masslynx 4.1 software<sup>1</sup> (Waters, UK). Samples were collected in triplicate for replicate analysis.

During the second lab period, students were then given 200 mM ammonium acetate buffer at pH 6.11, 6.45, 7.14, 7.59, and one ammonium acetate buffer of an unknown pH. The students buffer exchanged the ConA samples into each of the five buffers. And loaded the samples into the capillaries to acquire a three-minute spectra for each of the buffer-exchanged samples (Figure 7.1B). Students observed an increase in the tetramer as a function of pH (Figure 7.1C), demonstrating the use of Cona as a pH probe.

Using the Masslynx software, students extracted the intensities of the tetramer and plotted the percent tetramer as a function of pH (Figure 7.1D). The calibration was then used

to find the pH of an unknown sample, and they were asked to perform 95% confidence intervals and a LINEST analysis using Excel for statistical analysis. All experiments were done in triplicate to determine the error associated with MS measurements in comparison to the other techniques

they had used over the course of the class, such as spectrophotometry, ion-selective electrodes, and fluorimeters. In order to assess these results from a biomedical perspective, students were asked if the unknown sample had a pH in the proper biological range of ~7.4 and to justify their answer.

### 7.3 Protein:Ligand Complex Analysis Using Mass Spectrometry for the Undergraduate Laboratory

Lysozyme is an enzymatic antibacterial agent in eggs that catalyzes the hydrolysis of the 1,4-beta-linkages in bacterial peptidoglycans. Ganem *et al.* reported the kinetics of lysozyme with the hexasaccharide of N-acetylglucosamine in 1991<sup>11, 17</sup>. These studies were among the first to observe intact noncovalent enzyme:substrate and enzyme:product complexes through the use of ESI-MS. More recently, MS was used to determine the  $K_d$  values of lysozyme with various sugars, which were then compared to  $K_d$  values measured using other biophysical techniques<sup>18</sup>. We developed protocols following this work for students to explore the binding kinetics of lysozyme with penta-N-acetylchitopentaose (NAG5).

Similarly for ConA, the lysozyme project started with a brief tutorial of native MS and a demonstration on how to use the LCT instrument. Again, students used sodium iodide to check the calibration of the LCT before testing the samples of lysozyme in a biological buffer versus a buffer exchanged lysozyme sample. For sample analysis, students were given an aliquot of lysozyme (Figure 7.2A) and an aliquot of NAG5 (Figure 7.2A). They then calculated how much NAG5

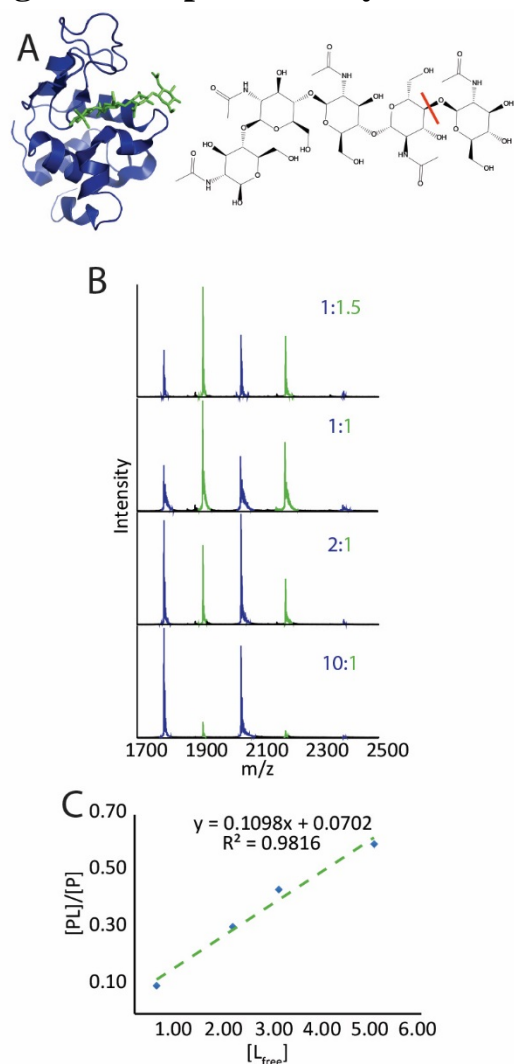


Figure 7.2. Overview of the lysozyme experiment. (A) Structure of lysozyme with NAG4 (left) and the structure of NAG5 (right). The lysozyme cleavage site is indicated by the red line. (B) Example spectra of lysozyme:NAG5 complexes as a function of increasing NAG5 concentration. The free lysozyme peaks are in blue, and the lysozyme:NAG5 peaks are in green. (C) Example plot of  $[PL]/[P]$  vs. free ligand concentration. The  $R^2$  value obtained from these experiments was 0.9816. The slope of the line gives the  $K_a$  value.

was needed to make 0.1, 0.5, 1, and 1.5 molar ratios of NAG5:lysozyme sample. Here, the two groups of students followed two methods: the first group incubated the NAG5:lysozyme complex for five minutes on ice before analysis, and the second group collected spectra immediately after mixing the NAG5:lysozyme samples together. All samples were collected in triplicate for replicate analysis. The intensities of the NAG5, lysozyme, and lysozyme:NAG5-bound peaks (Figure 7.2B) were calculated by the students using the Masslynx 4.1 software to create a plot of protein:ligand/protein ( $[PL]/[P]$ ) versus molar ratio of free ligand ( $[L_{\text{free}}]$ ) (Figure 7.2C). This curve was used to find the concentration of the unknown ligand. The slope of the curve reported the  $K_a$  value, which were compared to the values found by previous MS and biophysical technique measurements. Students reported their average  $K_a$  values and used a student t-test to determine how close their value was to the reported MS literature value. Students also reported on other complexes that were observed in their mass spectra; for example, students were able to observe lysozyme NAG4:complexes as well as free NAG (data not shown).

## 7.4 Potential Modifications

There are several variations that would be suitable for these laboratory exercises. First, instead of using capillaries pulled in-house by the instructor, the purchase of pre-cut needles would greatly facilitate the laboratory experience, as the pre-cut needles had a wide variety of tip sizes, which affected the needle spray. The authors also note that HEPES salts are very difficult to remove from protein samples<sup>1</sup>; therefore, phosphate and Tris buffers should be used as the native buffers. The students also noted that having a lecture dedicated to native MS would also be helpful, as this topic has not been covered in-depth in the corresponding lecture series in this course. The lysozyme MS protocols could be easily modified to measure enzyme turnover to teach students about enzyme kinetics. MS is particularly advantageous because the concentrations of each ligand-bound and digest sugar species can be measured individually as a function of ligand concentration and/or time.

## 7.5 Discussion

After the experiments were completed, students wrote a several page report on their findings that included an introduction explaining the basic principles behind the experiment, materials and methods citing the procedure that was used during the experiment and any deviations that occurred,

discussion of the mass calculation and buffer exchange, analysis of the calibration curves, unknown calculations, and  $K_a$  values (if applicable), a summary and comment of the results of their experiments, including sources of error, and a survey written by the instructor to help determine the success of the project. Overall, students gave positive feedback on the experiment. They enjoyed using an instrument that they had been briefly introduced to in previous classes, but had not encountered in their laboratory or research experiences. All students ( $n=6$ ) agreed that native MS should be included in the experiment rotation in the Biomedical Analytical Chemistry laboratory. However, the students did note that native MS was more complex than the other methods and instruments that they had used up to this point, which made it a more challenging experiment.

By designing these laboratory protocols, we have developed two native MS laboratory experiments that explore different aspects of the strengths of native MS: probing noncovalent protein:protein interactions, and exploring the kinetics of an enzyme:ligand system. These experiments can be performed on any MS instruments that are designed to perform native MS with chemicals that are cheap and accessible. We envision that these experiments would also translate well to any analytical chemistry or biochemistry laboratory and will be helpful for the introduction of native MS to these courses at all levels.

## 7.6 References

- [1] Hernández, H., and Robinson, C. V. (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry, *Nature protocols* 2, 715-726.
- [2] Zhong, Y., Hyung, S.-J., and Ruotolo, B. T. (2011) Characterizing the resolution and accuracy of a second-generation traveling-wave ion mobility separator for biomolecular ions, *Analyst* 136, 3534-3541.
- [3] Hyung, S.-J., DeToma, A. S., Brender, J. R., Lee, S., Vivekanandan, S., Kochi, A., Choi, J.-S., Ramamoorthy, A., Ruotolo, B. T., and Lim, M. H. (2013) Insights into antiamyloidogenic properties of the green tea extract (–)-epigallocatechin-3-gallate toward metal-associated amyloid- $\beta$  species, *Proceedings of the National Academy of Sciences* 110, 3743-3748.
- [4] Young, L. M., Saunders, J. C., Mahood, R. A., Revill, C. H., Foster, R. J., Tu, L.-H., Raleigh, D. P., Radford, S. E., and Ashcroft, A. E. (2015) Screening and classifying small-molecule

- inhibitors of amyloid formation using ion mobility spectrometry–mass spectrometry, *Nature chemistry* 7, 73-81.
- [5] Pacholarz, K. J., Garlish, R. A., Taylor, R. J., and Barran, P. E. (2012) Mass spectrometry based tools to investigate protein–ligand interactions for drug discovery, *Chemical Society Reviews* 41, 4335-4355.
- [6] Fuerstenau, S. D., Benner, W. H., Thomas, J. J., Brugidou, C., Bothner, B., and Siuzdak, G. (2001) Mass spectrometry of an intact virus, *Angewandte Chemie International Edition* 40, 541-544.
- [7] Snijder, J., Rose, R. J., Veesler, D., Johnson, J. E., and Heck, A. J. (2013) Studying 18 MDa virus assemblies with native mass spectrometry, *Angewandte Chemie International Edition* 52, 4020-4023.
- [8] Uetrecht, C., Barbu, I. M., Shoemaker, G. K., van Duijn, E., and Heck, A. J. (2011) Interrogating viral capsid assembly with ion mobility–mass spectrometry, *Nature chemistry* 3, 126-132.
- [9] Uetrecht, C., Versluis, C., Watts, N. R., Roos, W. H., Wuite, G. J., Wingfield, P. T., Steven, A. C., and Heck, A. J. (2008) High-resolution mass spectrometry of viral assemblies: molecular composition and stability of dimorphic hepatitis B virus capsids, *Proceedings of the National Academy of Sciences* 105, 9216-9220.
- [10] Marcoux, J., and Robinson, C. V. (2013) Twenty years of gas phase structural biology, *Structure* 21, 1541-1550.
- [11] Ganem, B., Li, Y. T., and Henion, J. D. (1991) Observation of noncovalent enzyme-substrate and enzyme-product complexes by ion-spray mass spectrometry, *Journal of the American Chemical Society* 113, 7818-7819.
- [12] Light-Wahl, K., Winger, B., and Smith, R. D. (1993) Observation of the multimeric forms of concanavalin A by electrospray ionization mass spectrometry, *Journal of the American Chemical Society* 115, 5869-5870.
- [13] Coppola, B. P. (2016) Broad & Capacious: A New Norm for Instructional Development in a Research Setting, *Change: The Magazine of Higher Learning* 48, 34-43.
- [14] Ruotolo, B. T., Benesch, J. L., Sandercock, A. M., Hyung, S.-J., and Robinson, C. V. (2008) Ion mobility–mass spectrometry analysis of large protein complexes, *Nature Protocols* 3, 1139-1152.



- [15] Hardman, K. D., and Ainsworth, C. F. (1972) Structure of concanavalin A at 2.4-Å resolution, *Biochemistry* 11, 4910-4919.
- [16] Light-Wahl, K. J., Schwartz, B. L., and Smith, R. D. (1994) Observation of the noncovalent quaternary associations of proteins by electrospray ionization mass spectrometry, *Journal of the American Chemical Society* 116, 5271-5278.
- [17] Ganem, B., Li, Y. T., and Henion, J. D. (1991) Detection of noncovalent receptor-ligand complexes by mass spectrometry, *Journal of the American Chemical Society* 113, 6294-6296.
- [18] Veros, C. T., and Oldham, N. J. (2007) Quantitative determination of lysozyme-ligand binding in the solution and gas phases by electrospray ionisation mass spectrometry, *Rapid Communications in Mass Spectrometry* 21, 3505-3510.

## **Chapter 8.**

### **Conclusions and Future Directions**

#### **8.1 Conclusions**

The majority of drug discovery efforts have primarily been focused on discovering potent small molecule inhibitors using screening methodologies that output a  $K_i$  or  $IC_{50}$  value using some sort of enzymatic assay. However, not all targets can be accessed by enzymatic assays; furthermore, these simple assays do not always provide all of the information that are advantageous for drug discovery. The reason for this change is twofold. First, protein structure is intimately related to function, and changes in protein structure can have consequences on downstream pathways. An example of this is recent work in the kinase inhibitor community. It has been discovered that altering the 3D form of the kinase can have consequences in the downstream signaling pathways<sup>1-3</sup>, which could potentially be used to better develop inhibitors for kinases that are implicated in disease. Second, there are many potential targets that are not enzymes or easily coupled with enzymatic assays. Many proteins in this target class are PPIs, such as the KIX domain of CPB with its binding partners. It has been difficult for traditional biophysical techniques such as NMR and X-ray crystallography to assess consequences of ligand binding on the structure of the protein target due to the size of the protein system, conformational heterogeneity, or difficulty in crystallizing the protein targets.

A key technology to overcoming these challenges has been the combination of ion mobility with mass spectrometry. As IM-MS can separate protein conformers that are >2% different in global size, CIU has been a useful strategy for the differentiation of smaller protein conformational changes that have been induced by ligand binding<sup>4</sup>. However, a challenge for the application of CIU for protein:ligand HTS screening protocols has been the lack of software to analyze the dense datasets that are created with our CIU strategy. The use of CIU for this study has been made possible by the development of a set of Python-based modules for the analysis of CIU data (Chapter 2). More importantly, by building our own tools for data analysis, we determined what data is useful and how this data should be implemented into workflows

determine information about CIU experiments, for example, for protein structure, biosimilar analysis, and protein:ligand screening<sup>5</sup>. We then created a CIU strategy to differentiate between type I and type II kinase inhibitors, which bind to the active and inactive kinase conformation respectively, using the Abl kinase domain as a model system (Chapter 3). We expanded our CIU technique to include allosteric Abl inhibitors that bind to the myristate binding site that is remote from the ATP-binding site (Chapter 4) and to include type I and type II kinase inhibitors using the Src3D protein construct (Chapter 5). Through these experiments, we have identified regions the CIU fingerprints that are ideal for separating each class of inhibitor, cutting the screening time from >1 hr for each inhibitor to <3 minutes, enabling the potential analysis of ~480 inhibitors in 24 hours. Ultimately, given sufficient S/N, we predict that these experiments could be further reduced to 1 second/inhibitor, which would enable the screening of 86,400 inhibitors in a 24 hour time period.

We then turned to inhibitors of PPIs using the KIX domain of CPB with peptides that bind to the MLL or pKID binding site (Chapter 6). We explored six unique complexes: KIX:MLL, KIX:pKID, KIX:E2A: KIX:c-Myb, KIX:MLL:pKID, and KIX:E2A:c-Myb. Through these CIU experiments, we find each class of peptide and the ternary complex has a unique CIU fingerprint. We then used the stability of each CIU feature to suggest a mechanism of gas-phase complex unfolding based on the known solution-phase interactions. We also used the CIU data to determine regions within the fingerprints that could be used for screening unknown peptides, and rationalized the CIU fingerprints based on detailed quantitative comparisons.

Finally, we introduced native MS to an analytical chemistry laboratory geared towards students in the biomedical sciences.. Despite the growing use of IM-MS in academic and industrial careers in the life sciences, students are not exposed to this technology unless they use the instrumentation in graduate school. This leads to a workforce that is ill-prepared to deal with the challenges of native MS experiments. The chemical education literature also reflects this trend, as most protein MS experiments are geared towards proteomics experiments, with only two providing protocols for protein:ligand or structural biology MS experiments. To this end, we developed two sets of experiments exploring the pH dependence of the ConA tetramer<sup>6, 7</sup> and the binding of N-acetylchitopentaose to hen egg white lysozyme<sup>8, 9</sup>. These systems are well-studied within the mass spectrometry community and easily facilitated in the undergraduate laboratory settings.

## **8.2 Immediate Future Directions**

As CIU is a newer technology, its potential to determine information about proteins and protein complexes has not yet been achieved. For example, mechanisms of gas-phase protein unfolding remains unknown. Through CIU experiments with smaller proteins, such as KIX, with ligands that are known to stabilize a certain part of the protein, more structural information about the structure of the gas-phase structures along the CIU unfolding pathway can be achieved.

### **8.2.1 Screening Type I, Type II, and Allosteric Inhibitors for the Abl Kinase**

We have developed a CIU screening protocol to identify type I, type II, and allosteric Abl kinase inhibitors. The next step for this work would be to screen a library, such as one of the Cambridge small molecule fragment libraries, to find new type II and allosteric kinase inhibitors. As most assays focus on binding strength, many of these potential inhibitors might have been discarded due to a lack of binding strength. Thus, by identifying fragments that induce a conformational change and locating the binding site to the ATP or allosteric binding sites, new classes of inhibitors could be discovered. The challenge to this step is the introduction of the kinase:ligand samples to the mass spectrometer. Faster sample manipulation systems such as microfluidics systems represent an exciting development to make HTS a reality for CIU.

### **8.2.2 Exploring Src3D Kinase Inhibitors that Induce a Closed Conformation**

While the type I/type II classification is appropriate for describing how an inhibitor interacts with the kinase domain, a different convention is needed for the Src3D structure. It has been found that type II inhibitors will completely open the 3D conformation into an extended structure, whereas type I inhibitors will stabilize structures ranging from fully open to a neutral position. Recently,  $\alpha$ C-helix out inhibitors have been discovered. These inhibitors induce a conformational change to the closed kinase conformation. Furthermore, it is now believed that kinase inhibitors stabilize a range of conformations ranging from fully open to fully closed. By including  $\alpha$ C-helix out in a screen, along with Src3D mutants that induce an open or closed conformation, the full extent of CIU can be explored. Furthermore, there are many kinase inhibitors that are known to stabilize the DFG-in or DFG-out kinase structure; however, it is unclear how these inhibitors stabilize the 3D structure. For example, tozasertib stabilizes a DFG-in conformation in Abl, but it is unknown how tozasertib affects the 3D structure of Src. As the open/closed form can affect downstream pathways, this CIU screen could provide information that can lead to better knowledge of how an inhibitor will affect its downstream signals.

### 8.2.3 Exploring KIX Allostery

While our CIU strategy can be used to localize peptide binding, any cooperative effects between the MLL and pKID binding site are difficult to detect because of weak binding. In order to observe any allosteric effects on the pKID and MLL binding sites, KIX should be loaded with either MLL or pKID, and then titrated with the other inhibitor, and vice versa. MS data can be used to measure  $K_d$  as a function of inhibitor binding. Additionally, these  $K_d$  measurements can be taken with the I660V mutant of KIX, which has been found to destroy the allosteric connection between the pKID and MLL binding sites, in order to explore the allosteric effects caused by peptide binding. These experiments to study allostery can be incorporated in a screen that includes CIU to determine information about peptide localization and allostery in one screen.

### 8.2.4 Interactions between Abl and Dscam

Immunoprecipitation (IP) pull-down assays have identified an interaction between the Abl kinase domain and the cytoplasmic domain of Dscam, and the addition of the kinase inhibitor nilotinib has been found to rescue presynaptic growth to wild-type levels in a DS *Drosophila* model<sup>10</sup>. However it is unclear if Abl and Dscam form a direct interaction, or if this interaction is mediated by other binding partners. Work moving forward can use the native MS and CIU protocols established in this thesis to determine if there is an interaction between Dscam and Abl, and how this interaction can change as a function of nilotinib and other small molecule kinase inhibitor binding. Dscam protein expression and purification need to be optimized prior to further analysis. Alternatively, complexes from IP pull down assays can be cleaned for MS analysis, although the pull-down assay would need to be optimized for the highest concentration and purity of Abl:Dscam complex prior to MS analysis.

## 8.3 Future Directions—Looking Ahead

This thesis has extensively focused on the use of CIU and IM-MS in the context of drug discovery for small conformational changes. In order to target difficult protein targets that do not align with the typical enzyme:inhibitor interactions, such structural information needs to be incorporated at the beginning of the drug discovery to operate beyond the typical binding strength information. CIU and IM-MS is an ideal technology to probe structural information that can affect the mechanism of function and the etiology of disease and merge the two worlds together. Additionally, the technology can be applied to other systems that are difficult to probe

by other mechanisms. For example, CIU can be applied to nucleic acid:protein interactions, such as the intrinsically disordered fragile X mental retardation protein, which forms an ordered structure when bound to target G quadruplex RNA. Many proteins that target nucleic acids are disordered, and the order-to-disorder transition upon nucleic acid binding is difficult to study by traditional biophysical assays. However, a better understanding of the physics behind protein and nucleic acid unfolding in the gas-phase is needed to obtain biophysical thermodynamic characteristics and structural information that can be used to relate the gas-phase structures to solution-phase behaviors. As the gas-phase unfolding behavior of more protein and protein:ligand complexes are explored in the gas-phase, the realization of the physics behind gas-phase unfolding will lead to *de novo* protein structural characterization without necessitating previous solution-phase structural knowledge.

### 8.3 References

- [1] Wang, L., Perera, B. G. K., Hari, S. B., Bhattacharai, B., Backes, B. J., Seeliger, M. A., Schürer, S. C., Oakes, S. A., Papa, F. R., and Maly, D. J. (2012) Divergent allosteric control of the IRE1 $\alpha$  endoribonuclease using kinase inhibitors, *Nature chemical biology* 8, 982-989.
- [2] Koppikar, P., Bhagwat, N., Kilpivaara, O., Manshour, T., Adli, M., Hricik, T., Liu, F., Saunders, L. M., Mullally, A., and Abdel-Wahab, O. (2012) Heterodimeric JAK-STAT activation as a mechanism of persistence to JAK2 inhibitor therapy, *Nature* 489, 155-159.
- [3] Andraos, R., Qian, Z., Bonenfant, D., Rubert, J., Vangrevelinghe, E., Scheufler, C., Marque, F., Régnier, C. H., De Pover, A., and Ryckelynck, H. (2012) Modulation of activation-loop phosphorylation by JAK inhibitors is binding mode dependent, *Cancer discovery* 2, 512-523.
- [4] Rabuck, J. N., Hyung, S.-J., Ko, K. S., Fox, C. C., Soellner, M. B., and Ruotolo, B. T. (2013) Activation state-selective kinase inhibitor assay based on ion mobility-mass spectrometry, *Analytical chemistry* 85, 6995-7002.
- [5] Eschweiler, J. D., Rabuck-Gibbons, J. N., Tian, Y., and Ruotolo, B. T. (2015) CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions, *Analytical chemistry* 87, 11516-11522.

- [6] Light-Wahl, K., Winger, B., and Smith, R. D. (1993) Observation of the multimeric forms of concanavalin A by electrospray ionization mass spectrometry, *Journal of the American Chemical Society* 115, 5869-5870.
- [7] Light-Wahl, K. J., Schwartz, B. L., and Smith, R. D. (1994) Observation of the noncovalent quaternary associations of proteins by electrospray ionization mass spectrometry, *Journal of the American Chemical Society* 116, 5271-5278.
- [8] Ganem, B., Li, Y. T., and Henion, J. D. (1991) Observation of noncovalent enzyme-substrate and enzyme-product complexes by ion-spray mass spectrometry, *Journal of the American Chemical Society* 113, 7818-7819.
- [9] Veros, C. T., and Oldham, N. J. (2007) Quantitative determination of lysozyme-ligand binding in the solution and gas phases by electrospray ionisation mass spectrometry, *Rapid Communications in Mass Spectrometry* 21, 3505-3510.
- [10] Sterne, G. R., Kim, J. H., and Ye, B. (2015) Dysregulated Dscam levels act through Abelson tyrosine kinase to enlarge presynaptic arbors, *Elife* 4, e05196.

## Appendices

### Appendix I. Chapter 3 Supporting Information

#### I.I On the Origin of CIU Fingerprint bands

We state that our CIU fingerprints contain bands that originate from multiple processes that occur

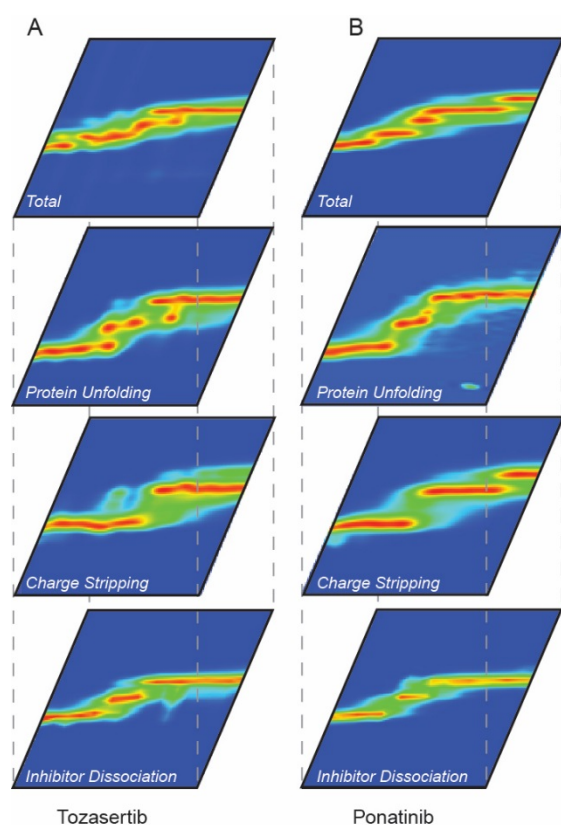


Figure I.1 Individual mass-resolved components of the CIU fingerprints for Abl-inhibitor complex. The CIU data is shown for  $m/z$  values corresponding to inhibitor dissociation (bottom), charge stripping/inhibitor dissociation (bottom-middle) and protein unfolding (top-middle) for the intact protein-inhibitor complexes of Abl bound to (A) Tozasertib (type I) and (B) Ponatinib (type II).

upon the coalitional activation of Abl-inhibitor complexes. All of these transitions generate mass-resolvable features that we can track independently to determine their relative influence on the composite fingerprints we use to differentiate type I and II inhibitor complexes (Figure I.1). Both composite and individual fingerprint contributions for tozasertib (type I) and ponatinib (type II) indicate that while inhibitor stripping and charge reduction (which predominately occurs with concomitant inhibitor loss) both occur at relatively low voltage values to some extent, both populations of ions also undergo unfolding (as indicated by the larger drift times achieved for each individual ion population shown). In general, inhibitor and charge stripped species tend to give rise to many of the CIU features recorded at high collision voltages. Thus their inclusion within our composite screening approach is critical to its success, as we identify this region as the most differentiating between different inhibitor classes.



## I.II Average Drift Time Spectra for Three CIU Regions

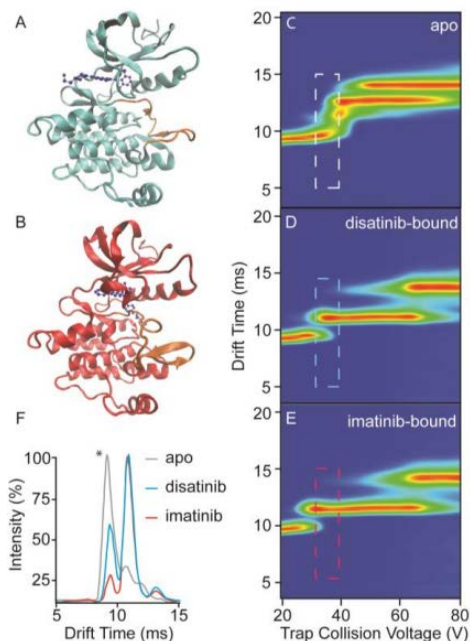


Figure I. 2 X-ray structures (1IEP) for Abl bound to dasatinib (A) shown in blue, and imatinib (B) shown in red. The activation loop (orange) changes between the two structures. CIU fingerprints derived from 10+ ions corresponding to Apo-Abl kinase (C), Abl bound to dasatinib (D) and Abl bound to imatinib (E). The dashed boxes highlight a potential region of the plot useful for high-throughput screening, and are color coded to match the other sections of the figure. Drift time plot from the three dashed-box regions in C-E (F), a factor of 2.5 fold difference in compact protein conformer intensity (\*) is observed between the two inhibitor data-sets (colors as indicated).

In determining the area within CIU fingerprints that provide the maximum level of discrimination between type I and type II kinase inhibitors, we evaluated three principle voltage ranges: 20-28V, 32-36V, and 40-44V. These three regions were selected for detailed analysis based on a visual inspection of CIU fingerprints for Abl-inhibitor complexes, and preliminary data for 10<sup>+</sup> protein-inhibitor complexes shown in Figure I.2. As such, we computed the average IM drift time spectra for all three regions and compared both the number of unique features detected, the resolving power achieved, and the total number of peaks observed (Figure I.2). The IM data collected between 20V and 28V, while useful for differentiating imatinib and dasatinib complexes in Figure I.2, contains only two features on average, with only 1 feature having a unique centroid value for each type I and II averaged IM spectrum. This, combined with the relatively poor peak resolution within this region makes it a non-optimized choice for differentiating type I and I inhibitors in a broad dataset. Intermediate voltage regions show two well-

resolved peaks for the average type II CIU response, and three poorly-resolved features for type I-bound complexes, with both sets being mutually exclusive to their respective datasets. High voltage data provides the largest number of peaks for type II average data (5 resolved features and one shoulder), all with unique centroid drift times when compared with the two main features recorded for the average type I CIU response. The superior discriminating power of this final higher-voltage region of our CIU data is supported by the similarity scores calculated in Tables I.1 and I.2. In Table I.2 specifically, we use other methodologies to calculate a score value useful for

screening. For example, we attempt to normalize our scores to the average type I and type II

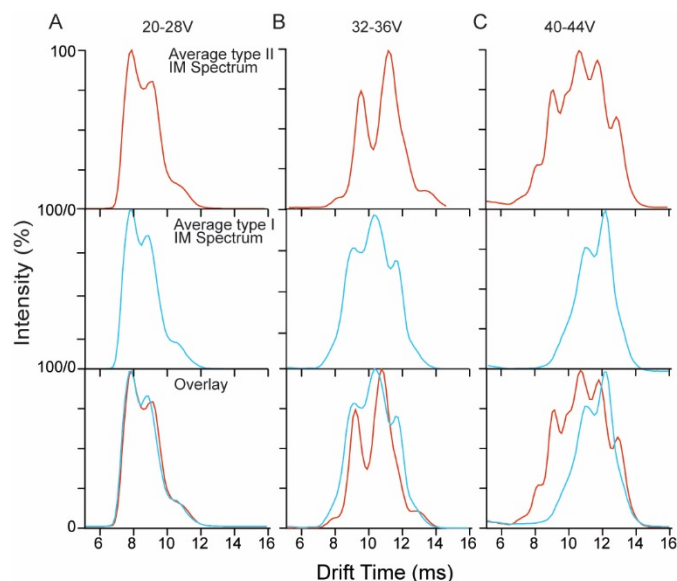


Figure I. 3 Averaged IM Spectra generated from all type I (blue) or type II (red) datasets are shown for selected regions of the CIU fingerprints recorded. Type II data is shown across the top of the diagram, with type I below and the two datasets are overlaid across the bottom. The integrated areas include (A) 20-28V, (B) 32-36V, and (C) 40-44V. The data shown in (C) is evaluated to have the highest discriminating power of those compared here and most useful for inhibitor screening. The features observed are discussed in more detail in the text.

Table I.1

Drugs	type	Region1: Trap Collision energy 40-44V				Region2: Trap Collision energy 32-36V				Region1: Trap Collision energy 20-28V			
		$\chi^2$ Value for drug and average type I	$\chi^2$ Value for drug and average type II	Normalized type I similarity score	Normalized type II similarity score	$\chi^2$ Value for drug and average type I	$\chi^2$ Value for drug and average type II	Normalized type I similarity score	Normalized type II similarity score	$\chi^2$ Value for drug and average type I	$\chi^2$ Value for drug and average type II	Normalized type I similarity score	Normalized type II similarity score
Imatinib	II	2311.921	155.467	15.81	100	505.61	145.82	71.63	100	253.79	127.05	13.28	100
Nilotinib	II	563.26	210.37	64.91	73.90	481.20	143.83	75.26	101.38	100.58	142.55	33.50	89.13
Ponatinib	II	610.72	251.19	67.62	61.89	159.83	379.37	226.59	38.44	106.65	106.65	31.60	119.13
Sorafenib	II	254.97	319.98	143.42	48.59	344.37	10340.1	105.08	1.41	131.71	1279.12	25.60	9.93
DCC-2036	II	247.09	358.28	147.98	43.40	197.86	455.63	183.04	32.00	241.45	234.82	13.96	54.11
Bosutinib	I	150.01	365.51	243.75	42.53	246.68	18241.89	146.82	0.80	222.08	229.33	15.18	55.4
Tozasertib	I	113.80	368.26	321.30	42.21	75.19	1264.73	481.66	11.53	48.28	31.84	69.70	399.03
PP2	I	246.62	389.67	148.25	38.89	523.36	1483.46	69.20	9.83	315.17	418.12	10.69	30.39
Saracatinib	I	486.57	442.21	75.15	35.16	485.38	634.82	74.61	22.97	328.32	642.12	10.26	19.79
Staurosporine	I	218.46	634.61	167.37	24.50	139.73	2367.54	259.18	6.16	126.92	1377.02	26.56	9.22
Dasatinib	I	365.64	939.87	100	21.75	362.163	156.24	100	93.33	33.70	22.79	100	557.50

signatures, rather than to a maximum type II response. While we find some of these methods to

Table I.2

Region1: Trap Collision energy 40-44V						Region2: Trap Collision energy 32-36V						Region1: Trap Collision energy 20-28V					
Drugs	type	Normalized type I score	Normalized type II score	Ratio type II/type I		Drugs	type	Normalized type I score	Normalized type II score	Ratio type II/type I		Drugs	type	Normalized type I score	Normalized type II score	Ratio type II/type I	
Imatinib	II	8.99	152.77	16.99		Imatinib	II	37.98	182.99	4.42		Imatinib	II	34.27	134.30	3.92	
Nilotinib	II	36.90	112.76	3.06		Nilotinib	II	39.91	185.52	4.65		Nilotinib	II	86.48	119.70	1.38	
Ponatinib	II	38.47	94.34	2.45		Ponatinib	II	120.16	70.34	0.59		Ponatinib	II	81.55	159.99	1.96	
Sorafenib	II	81.79	74.00	0.90		Sorafenib	II	55.72	2.58	0.05		Sorafenib	II	66.04	13.34	0.20	
DCC-2036	II	84.10	66.14	0.79		DCC-2036	II	97.06	58.56	0.60		DCC-2036	II	36.02	72.67	2.02	
Bosutinib	I	138.5	64.70	0.47		Bosutinib	I	77.85	1.46	0.02		Bosutinib	I	39.16	74.41	1.90	
Staurosporine	I	95.30	37.35	0.39		Staurosporine	I	137.44	11.27	0.08		Staurosporine	I	68.53	12.39	0.18	
PP2	I	84.45	48.42	0.57		PP2	I	36.69	17.99	0.49		PP2	I	27.60	40.81	1.48	
Saracatinib	I	42.75	53.57	1.25		Saracatinib	I	39.57	42.03	1.06		Saracatinib	I	26.49	26.57	1.00	
Tozasertib	I	182.24	64.34	0.35		Tozasertib	I	255.42	21.10	0.08		Tozasertib	I	180.14	535.91	2.98	
Dasatinib	I	56.76	25.19	0.44		Dasatinib	I	53.03	170.79	3.22		Dasatinib	I	258.08	748.73	2.90	
Normalized type I equation: $\frac{A1}{\chi^2_{A1}} \times 100$						Normalized type II equation: $\frac{C1}{\chi^2_{C1}} \times 100$						Normalized type I equation: $\frac{C1}{\chi^2_{C1}} \times 100$					
A1 is the average type I score for type I drugs $\chi^2_{A1}$ is the type I $\chi^2$ value B is averaged A1 value for type I drugs						B1 is the average type I score for type I drugs $\chi^2_{B1}$ is the type II $\chi^2$ value D is averaged C1 value for type II drugs						B1 is the average type I score for type I drugs $\chi^2_{B1}$ is the type II $\chi^2$ value D is averaged C1 value for type II drugs					

provide strong discrimination power, none seem substantially superior to the approach indicated in the main text (Table 3.1).

### I.III Simple IM Separations are Insufficient to Differentiate between Type I and Type II Binders in Complex with the Abl Kinase Domain.

Our initial approach to differentiate type I and II kinase inhibitors using IM separation included a simple separation strategy that relies exclusively on the orientationally-averaged collision cross-section (CCS) difference between the two bound states of the protein to generate differentiated responses.

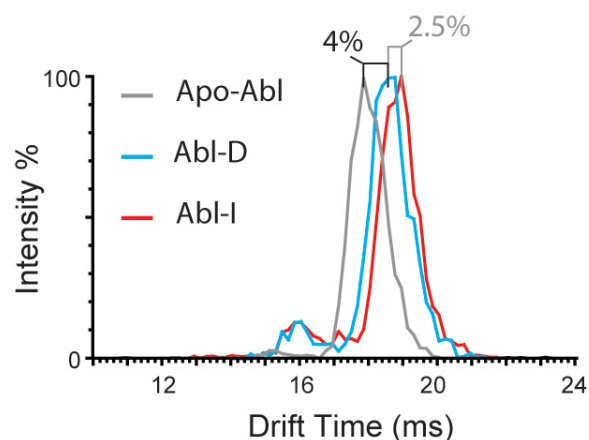


Figure I.4 Ion mobility spectra for three samples. Apo-Abl kinase (grey), Abl bound to dasatinib (type I inhibitor, blue) and Abl bound to imatinib (type II inhibitor, red). There is a 4-6% difference in drift time between apo and holo forms of Abl, and a much smaller, but measurable drift time difference (2.5%) between the two bound forms.

We found that the two bound states differ, at most, by IM drift time values of 2.5%, which correlates to a ~1.5% difference in CCS (Figure I.3). This correlates well to the difference in the projected areas (PA) calculated for the two X-ray structures of the proteins bound to the same inhibitors (PDB ID: 1IEP for imatinib-bound Abl, and PDB 2GQG for dasatinib-bound Abl), which predicts a CCS difference of 1% for the gas-phase ions<sup>1</sup>. We also note that both of these bound states are, interestingly, somewhat larger (~4-6%) than the apo form of the protein in the gas-phase.

#### I.IV Extrapolating Type II similarity scores to a larger chemical library screen.

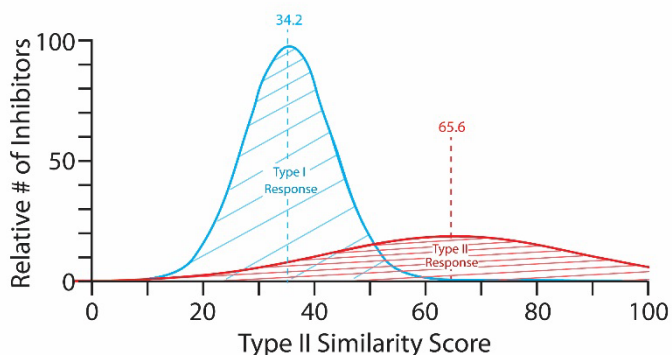


Figure I 5 An illustration using the scores from Table 1, this graph uses the average and standard deviations of the recorded similarity scores to plot the normal distributions for the scores expected for these two classes of inhibitors if the observations made here hold for a larger screen (~1800 total compounds, 50% type I, 50% type II simulated here). Average values are shown on the plot for Type I (blue) and Type II (red) inhibitors.

The 11 compounds used to train our IM-MS approach represent all of the commercially obtainable inhibitors available where there is strong structural evidence that identifies the kinase binding mode involved. While this dataset is small in comparison to a typical library screen, we felt it instructive to use the mean and standard deviations of our current dataset to graphically project the type I and II responses expected for an 1800

compound inhibitor library comprised equally of type I and II inhibitors. In this scenario, depicted in Figure I.4, while overlap exists between the two scoring ranges identified in this report for the two inhibitor classes, type II compounds having strong similarity scores are easily distinguished from type I inhibitors, which tend to cluster more strongly in our measurement. Depending upon the stringency of the screen employed, false positive values from the CIU-based approach can be tuned to nearly zero, if desired.

#### I.IV Phosphorylation of c-Abl.

Abl was phosphorylated following a previously described procedure<sup>1, 2</sup>. Briefly, since c-Abl is inefficient at autophosphorylation, it was incubated with hematopoietic cell kinase (Hck) at room temperature for 1 hour in Buffer D (50 mM Tris (pH 8.0), 100 mM NaCl, 5% glycerol, and 1 mM DTT), 30 mM MgCl<sub>2</sub>, 10  $\mu$ M c-Abl, 5 mM ATP, 100  $\mu$ M Na<sub>3</sub>VO<sub>4</sub>, and 50 nM Hck and passed through a Sephadex G25 column eluting with Buffer D.

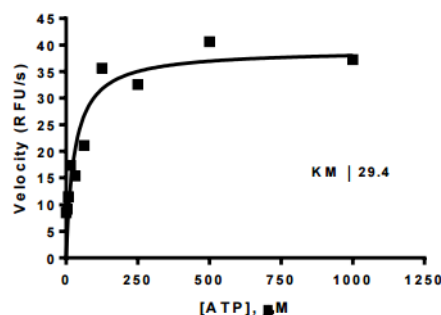
#### I.V General procedure for ATP K<sub>m</sub> determination.

The previously described fluorescence assay<sup>3</sup> was used to determine K<sub>m</sub> values. Reaction volumes of 100  $\mu$ L were used in 96-well plates. 85  $\mu$ L of enzyme in buffer was added to each well. 2.5  $\mu$ L of DMSO was then added followed by 2.5  $\mu$ L of substrate peptide (AEXIYAAPF-OH, where X is

2,3-diaminopropionic acid) solution (1.8 mM in DMSO). The reaction was initiated with 10  $\mu$ L of the appropriate ATP dilution (typically 1000, 500, 250, 125, 62.5, 31.3, 15.6, 7.8, 3.9, 2.0  $\mu$ M in H<sub>2</sub>O) and reaction progress was immediately monitored at 405 nm (ex. 340 nm) for 10 minutes. Reactions had final concentrations of 100 nM enzyme, 45  $\mu$ M peptide substrate, 100  $\mu$ M Na<sub>3</sub>VO<sub>4</sub>, 100 mM Tris buffer (pH 8), 10 mM MgCl<sub>2</sub>, 0.01% Triton X-100. The initial rate data collected was used for determination of  $K_m$  values. For  $K_m$  determination, the kinetic values were obtained directly from nonlinear regression of substrate-velocity curves in the presence of varying concentrations of ATP. The equation  $Y = (V_{max} * X)/(K_m + X)$ ,  $X$  = substrate concentration ( $\mu$ M) and  $Y$  = enzyme velocity (RFU/s); was used in the nonlinear regression. Each ATP  $K_m$  value was determined using at least three independent experiments; a representative  $K_m$  curve is shown for both phosphorylated and non-phosphorylated Abl (Figure I.5 and I.6).

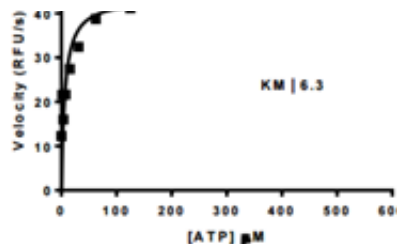
#### I.VI General Procedure for Determination of Inhibitor $K_i$ for Phosphorylated and Non-Phosphorylated Abl:

Abl inhibition assay was performed using a continuous, fluorimetric assay as previously described.<sup>3</sup> Reaction volumes of 100  $\mu$ L were used in 96-well plates. To each well was added 85  $\mu$ L of buffer + enzyme. 2.5  $\mu$ L of varying concentrations of inhibitor was then added (typically 10000, 2500, 625, 156, 39, 10, 2.4, 0.61, 0.15, 0  $\mu$ M in DMSO). 2.5  $\mu$ L of peptide substrate (AEXIYAAPF-OH, where X is 2,3-diaminopropionic acid) solution (1.8 mM in DMSO) solution was added. 10  $\mu$ L of ATP (50 mM in water) was added to initiate the reaction and was immediately monitored at 405 nm (ex. 340 nm) for 10 minutes. Final concentrations in the reaction are 100 nM enzyme, 45  $\mu$ M peptide substrate, 5 mM ATP, 100  $\mu$ M Na<sub>3</sub>VO<sub>4</sub>, 100 mM Tris buffer (pH 8),



$K_m = 22 \pm 7.0 \mu\text{M}$   
rel.  $V_{max} = 45 \pm 4.2 \text{ RFU/s}$

Figure I.6 ATP  $K_m$  Curve of Non-Phosphorylated c-Abl



$K_m = 5.9 \pm 1.5 \mu\text{M}$   
rel.  $V_{max} = 38 \pm 4.7 \text{ RFU/s}$

Figure I.7 ATP  $K_m$  Curve of Phosphorylated c-Abl

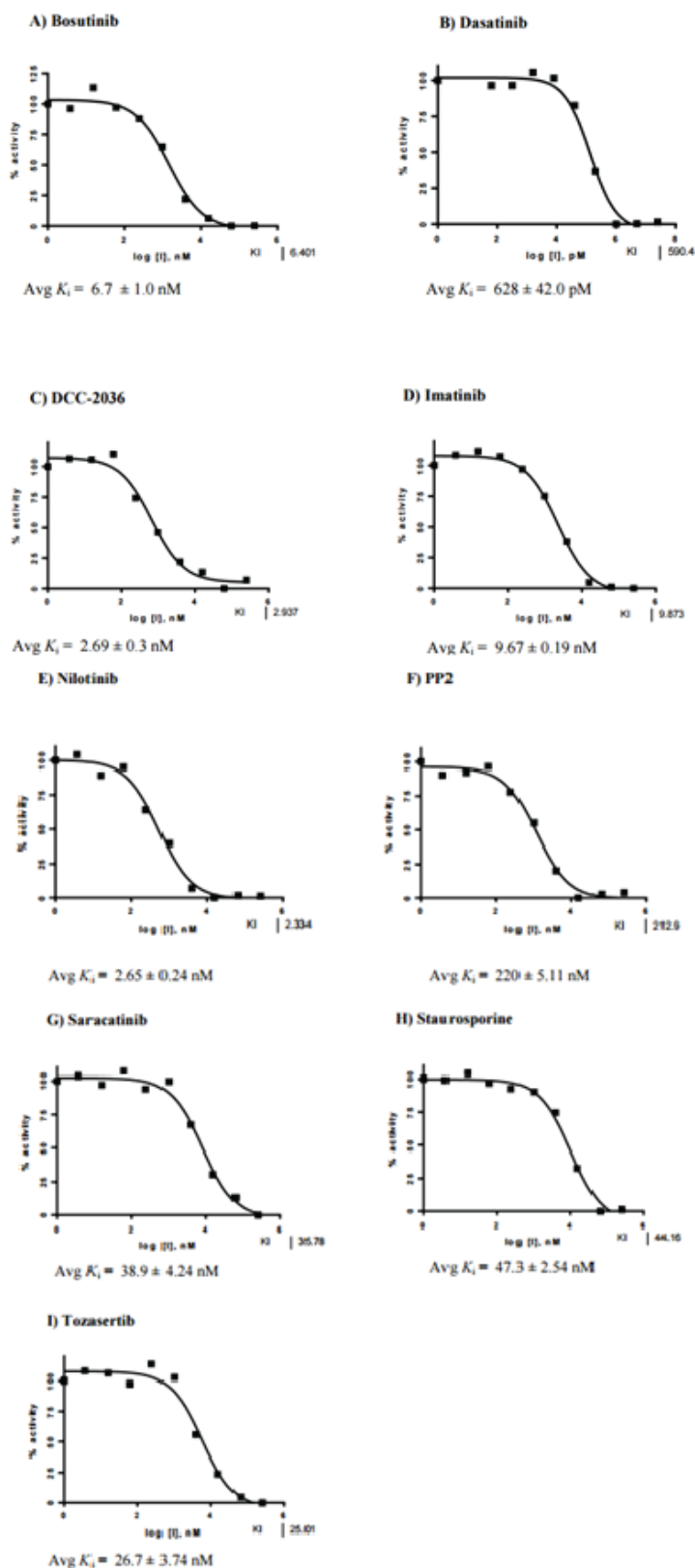


Figure I.8 Non-Phosphorylated Abl data

10 mM  $\text{MgCl}_2$ , 0.01% Triton X-100. The initial rate of the reaction was used to determine  $K_i$  values. For  $K_i$  determination, the kinetic values were obtained directly from nonlinear regression of substrate-velocity curves in the presence of various concentrations of the inhibitor. The equation  $Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10^{X - \text{LogEC}_{50}})$ ,  $X = \log(\text{concentration})$  and  $Y = \text{binding}$ ; was used in the nonlinear regression.

## I.VII Analytical data for $K_i$ determination.

Each inhibitor  $K_i$  value was determined using at least 3 independent measurements. An example curve is provided for each inhibitor.



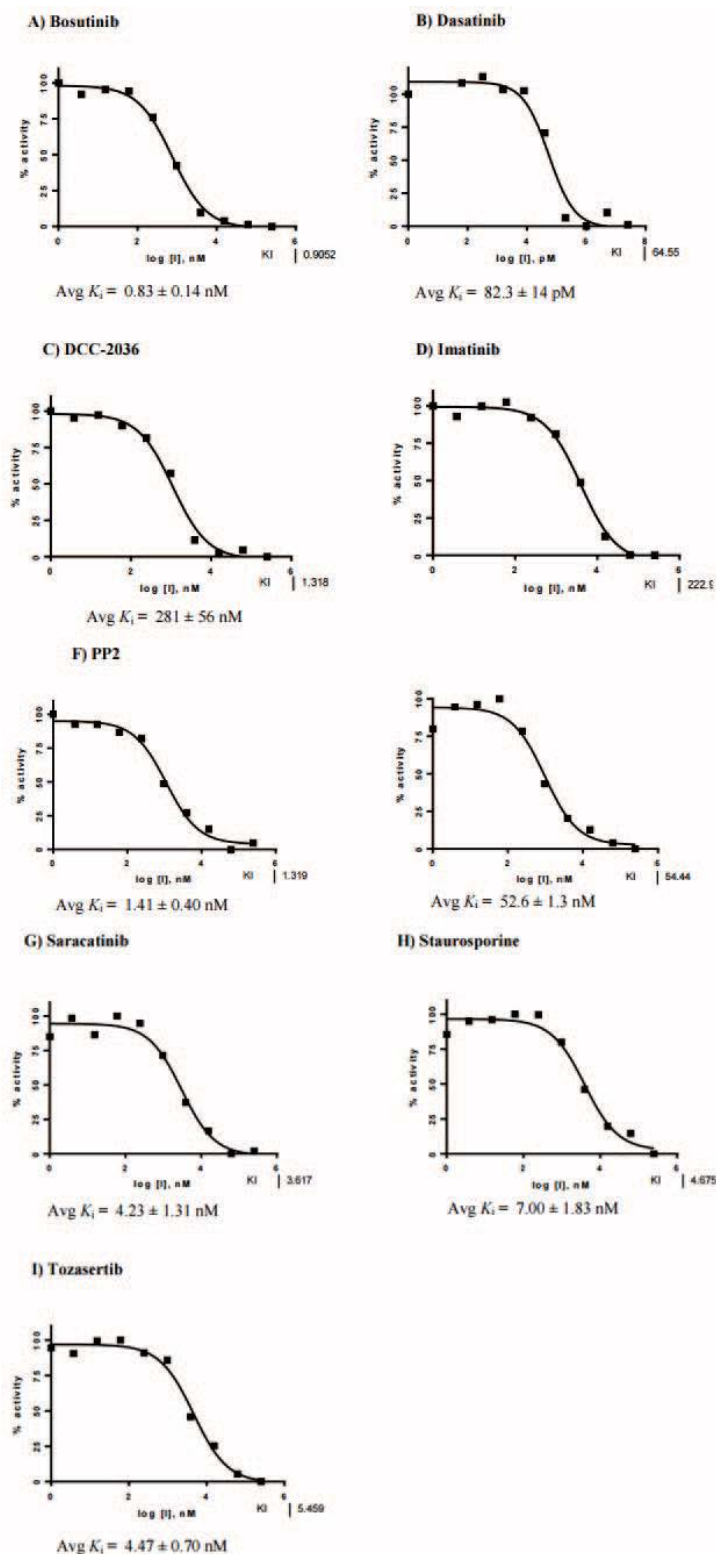


Figure I.9 Phosphorylated c-Abl Data

## I.VIII Mass spectrometry of Apo Abl and P-Abl.

Apo Abl and P-Abl samples were buffer exchanged into 200 mM ammonium acetate using Micro Bio-Spin 6 columns (Bio-Rad, Hercules, CA) and prepared into a final concentration of 12  $\mu$ M (Apo Abl) or XX12  $\mu$ M (P-Abl). Sample aliquots ( $\sim 5\mu$ l) were analyzed on a Q-IM-TOF MS instrument (Synapt G2 HDMS, Waters, Milford, Ma) and ionized using a nESI source, as described previously<sup>4, 5</sup>. The capillary voltage was held at 1.2 kV, with the source operating in positive mode, the sample cone operating at 65V, the extraction cone operating at 6V, and the TCE operating at 20V to help with desolvation for an accurate mass. The other instrumental conditions were set as previously described in the methods section. The mass of ApoAbl was calculated to be  $33180.70 \pm 0.88$ Da. The mass of P-Abl was calculated to be

33262.70 $\pm$ 0.18. The difference between the two masses is 82.00 Da, which agrees well with the protein being fully phosphorylated. There was a small amount of Abl that was not phosphorylated (less than 10%).

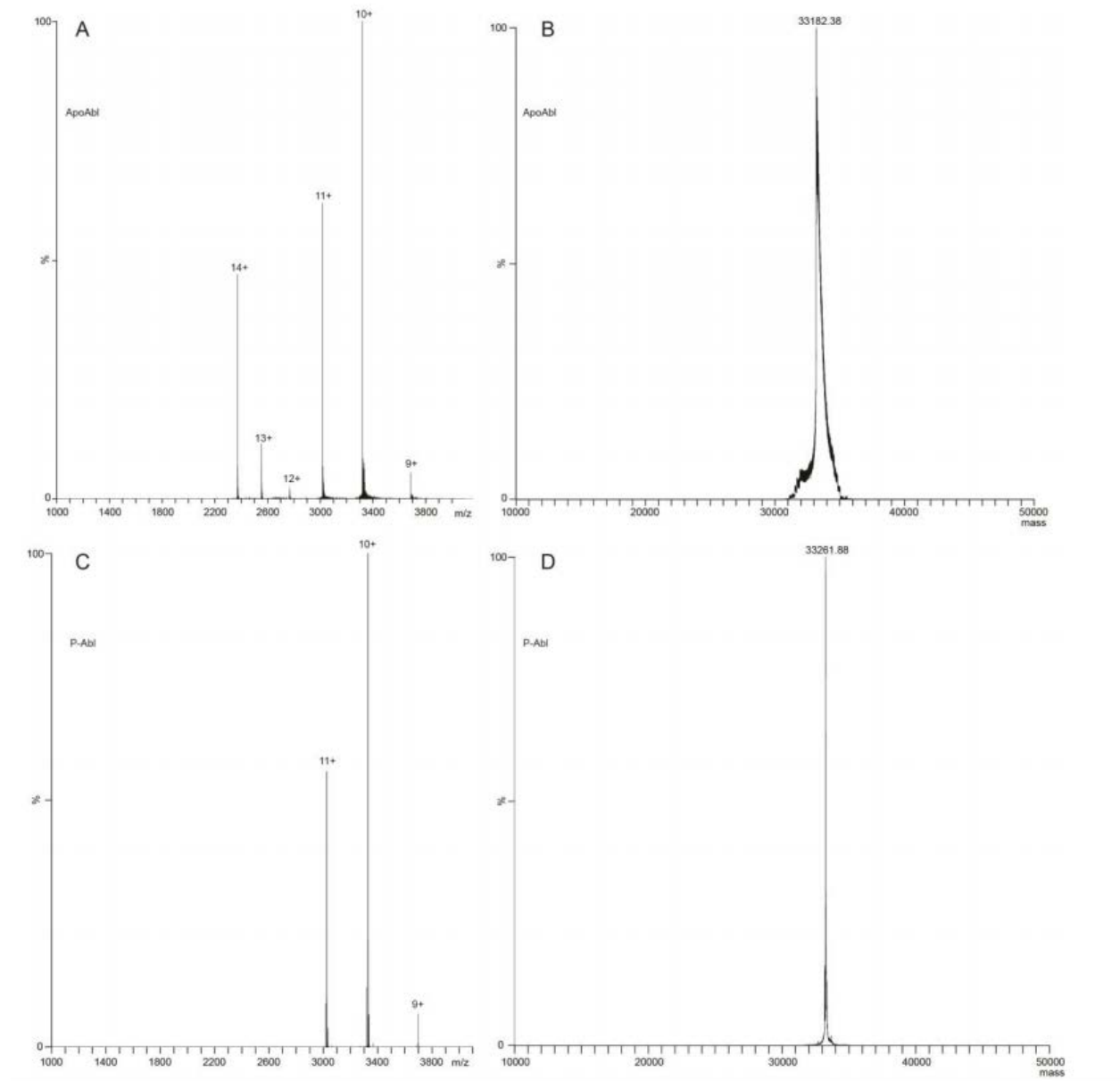


Figure I 10 Mass Spectra of ApoAbl (A, B) and phosphorylated Abl (P-Abl, C,D). A) Mass spectrum of ApoAbl. B) Deconvoluted mass of ApoAbl. C) Mass spectrum of P-Abl. D) Deconvoluted mass of P-Abl. The recorded average mass difference between the two forms is 79.52  $\pm$  0.90Da, indicating complete mono-phosphorylation of the P-Abl protein sample.

## I.IX References

- [1] Nagar, B., Bornmann, W. G., Pellicena, P., Schindler, T., Veach, D. R., Miller, W. T., Clarkson, B., and Kuriyan, J. (2002) Crystal Structures of the Kinase Domain of c-Abl in Complex with the Small Molecule Inhibitors PD173955 and Imatinib (STI-571), *Cancer Research* 62, 4236-4243.
- [2] Schindler, T., Bornmann, W., Pellicena, P., Miller, W. T., Clarkson, B., and Kuriyan, J. (2000) Structural mechanism for STI-571 inhibition of abelson tyrosine kinase, *Science* 289, 1938-1942.
- [3] Wang, Q., Cahill, S. M., Blumenstein, M., and Lawrence, D. S. (2006) Self-reporting fluorescent substrates of protein tyrosine kinases, *Journal of the American Chemical Society* 128, 1808-1809.
- [4] Zhong, Y., Hyung, S.-J., and Ruotolo, B. T. (2011) Characterizing the resolution and accuracy of a second-generation traveling-wave ion mobility separator for biomolecular ions, *Analyst* 136, 3534-3541.
- [5] Ruotolo, B. T., Benesch, J. L., Sandercock, A. M., Hyung, S.-J., and Robinson, C. V. (2008) Ion mobility–mass spectrometry analysis of large protein complexes, *Nature Protocols* 3, 1139-1152.

## Appendix II. Supporting Information for Chapter 7

### II.I Chemicals

Chemical Name, amount	Company	Catalog Number
chicken egg white lysozyme, 1g	Sigma	L6876-1G
Concanavalin A, 25MG	Sigma	C2010-25MG
Ammonium acetate, 100g	Sigma	09689-100G
Penta-N-acetylchitopentaose, 5 mg	Cayman Chemicals	36467-68-2

### II.II Protein Ligand Interaction Student Handout

# Special Project-HEWL v2

## 1. Introduction

The aim of this experiment is to introduce you to the concepts and applications of native mass spectrometry. Specifically, you will use a Micromass LCT Premier, check the instrumental calibration, determine the recovery of protein after a buffer exchange step, determine how solution conditions affect the charge state distribution and therefore the structure of a protein, and determine the concentration of an unknown amount of ligand using a calibration curve constructed from mass spectrometry data.

## 2. Background

Mass spectrometry (MS) is an extremely important analytical technique that measures the mass-to-charge ( $m/z$ ) ratio and abundance of gas-phase ions. MS has been used widely to determine the elemental or isotopic signature of a sample, measure the mass of molecules, and determine the chemical structures of measured molecules. For example, MS has been used in the field of proteomics (to identify protein-protein interaction networks, and to quantify protein abundance as a function of time or some stimuli), to discover biomarkers for cancers and other diseases, and in drug discovery. In a typical MS workflow, a sample (which can be a solid, liquid, or a gas) is ionized. The ions are then introduced into the instrument and separated according to  $m/z$ . Over the past 20 years, a technique called native mass spectrometry has been increasingly used in the analysis of large biomolecules. In a native MS experiment, gentle instrument and buffer conditions are used to maintain the solution structure of a protein or protein complex. For example, native mass spectrometry has been used to calculate the drug to antibody ratio of antibody:drug conjugates, which are used in cancer therapeutics.

One of the most important contributions to mass spectrometry, and especially native mass spectrometry has been the development of electrospray ionization (ESI). ESI is a gentle technique that allows for the introduction of whole, native proteins into a mass spectrometer for analysis, including protein:ligand complexes that are held together by weak, noncovalent interactions. ESI additionally is advantageous because it results in multiple charges on an analyte, and each  $m/z$  from an analyte gives an individual measure of its mass. The charge state distribution (i.e. the number of  $m/z$  peaks there are for a given analyte) is proportional to the surface area of the analyte. Thus, for a protein system, more extended, unfolded proteins will exhibit a broader charge state distribution with a greater number of charges than smaller, more compact proteins.

In this lab, you will be using a variant of ESI called nano-ESI. nanoESI is achieved by pulling glass capillaries to tips of a few micrometers. nanoESI has lower flow rates than ESI (nl/min as compared  $\mu$ l/min flow), and therefore is more tolerant for salts and other contaminants in buffers and has higher sensitivity (which means less sample is required, both in volume and concentration). For a native protein (i.e. not denatured by heat, chemicals, or pH), the average charge on the protein generated by the electrospray process is related to molecular weight based on equation 1:

$$\text{Average charge state} = 0.067\sqrt{\text{molecular weight}} \quad (1)$$

Proteins are typically buffered *in vitro* using salts like Tris, phosphate, and NaCl or KCl, among others. At the concentrations used in these biological buffers, the salt will mask the protein signal in the mass spectrometer. Therefore, the sample must be exchanged into a buffer more compatible with mass spectrometry. You will be using buffer exchange columns to quickly remove excess salt and switch your buffer to 200 mM ammonium acetate.

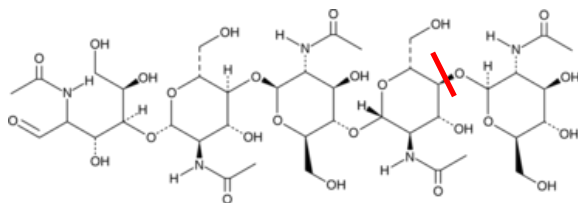


Figure 1. Penta-N-acetylchitopentaose (NAG5). The cleavage site is shown in red.

You will be using native mass spectrometry to evaluate the kinetics of hen egg white lysozyme (HEWL) with penta-N-acetylchitopentaose (NAG5). Lysozyme is an enzyme that is an antibacterial agent in eggs. HEWL damages bacterial cell walls by catalyzing the hydrolysis of the 1,4-beta-linkages in peptidoglycans. Lysozyme is also found in a number of human secretions. In fact, reduced lysozyme levels are

associated with bronchopulmonary dysplasia in newborns and diarrheal disease. Conversely, high levels of lysozyme caused by certain cancers can lead to low blood potassium and kidney failure. It is also an important to structural biology, as it was the second protein structure and the first enzyme structure to be solved by X-ray diffraction.

This experiment has four goals: (1) you will determine the recovery of protein after a buffer exchange step, (2) use native mass spectrometry to measure the charge state distribution of the protein in several native and denaturing buffers, and (3) determine the concentration of NAG5 in an unknown sample.

### 3. Procedure

#### Required Apparatus:

- 1) UV-Vis spectrophotometer
- 2) Micromass LCT Premier
- 3) 6 kDa molecular weight cut-off (MWCO) buffer exchange columns
- 4) Microcentrifuge
- 5) UVettes (cuvettes for the UV-Vis)

#### Reagents:

- 1) NaI
- 2) Unknown NAG5 sample
- 3) HEWL samples, starting concentration of 18.4  $\mu\text{M}$  in 200 mM ammonium acetate, pH 8.16
- 4) NAG5, 25  $\mu\text{M}$
- 5) HEWL in biological buffer
- 6) Biological buffer
- 7) 200 mM ammonium acetate buffers with varying pH and methanol concentrations

## Day One Procedure

### Part 1: Buffer Exchange

1) Obtain a HEWL sample in ammonium acetate (marked “L” on the tube) and a HEWL sample in a biological buffer (marked “BL” on the tube). Thaw your HEWL aliquots on ice. Measure the concentration of 50  $\square$  the HEWL sample in the biological buffer ( $\epsilon = 36000 \text{ M}^{-1}\text{cm}^{-1}$ )

2) Obtain one buffer exchange column and one balance. The buffer exchange column is stored in a Tris buffer, which is incompatible with electrospray, so the columns will have to be cleaned before buffer exchanging the protein. Snap off the bottom part of the buffer exchange column and place it in a 1.5 ml collection tube (the one without a lid), and remove the green cap of the buffer exchange column. Follow this sequence to prepare the buffer exchange columns and exchange your sample:

a) Place the buffer exchange column AND BALANCE in the microcentrifuge. **\*Caution: Make sure the centrifuge is properly balanced!!!!!!\*** Put the lid on the centrifuge, and shut the top. Set the speed to 1.0xg (not  $\text{min}^{-1}$ ) and the time to 2 minutes (the centrifuge is set to read minutes, except for the final minute, where the centrifuge will count down in seconds). If you are using one of the temperature-controlled microcentrifuges, make sure the centrifuge is set to 25°C. Hit the play button to start the first cycle. This step will drain the Tris buffer from the column. When the centrifuge is finished spinning (after it beeps), pour the Tris buffer that is now in your collection tube into the waste.

b) Place the buffer exchange column back in the collection tube (you can re-use the collection tube until you collect your protein). Carefully pipet 500  $\mu\text{l}$  of 200 mM ammonium acetate pH 8.16 onto the top of the column. **Do not do this with too much force—the goal is to not disturb the column bed!** Place the buffer exchange column and the collection tube back into the microcentrifuge. Keep the speed on 1.0xg, but change the time to 1 minute. Hit the play button to start the next cycle. When the centrifuge is finished spinning (after it beeps), pour the ammonium acetate buffer that is now in your collection tube into the waste. Repeat the step 3 more times, so you have washed the columns with ammonium acetate buffer for a total of 4 times. This will make sure that all of the Tris buffer has been replaced with ammonium acetate, and your protein will be effectively exchanged into your buffer of choice (in this case, ammonium acetate).

c) *Without adding in any buffer or protein*, place the buffer exchange column and the collection tube back into the microcentrifuge. Keep the speed on 1.0xg, and keep the time set to 1 minute. Hit the play button to start the next cycle. This will remove any excess buffer from your column that can dilute your protein during the buffer exchange process.

d) Label a 1.5 ml Eppendorf tube for your sample. Place the buffer exchange column into this new tube and throw away the collection tube. Carefully pipet 50  $\mu$ l of your HEWL protein in the biological buffer onto the top of the column bed. Be very careful not to introduce bubbles into your sample! Place the buffer exchange column and the Eppendorf tube back into the microcentrifuge. Keep the speed on 1.0xg, but change the time to 4 minutes. Hit the play button to start the final cycle. This is the buffer exchange step. After this, your protein sample will be in your ammonium acetate buffer of choice.

3) Measure the absorbance at 280 nm of 50  $\mu$ l of your cleaned HEWL sample. Make a note of the absorbance and also the volume of your sample, being careful not to introduce bubbles into your protein. You may need to dilute your sample if the absorbance is above 1

4) Obtain 6 more buffer exchange columns. Buffer exchange 50  $\mu$ l of the HEWL sample in 200 mM ammonium acetate into the 9 different ammonium acetate buffers. (In the end, you will have 9 HEWL samples in 9 different buffers). After buffer exchange, dilute your samples by adding in 50  $\mu$ l of buffer.

## **Part 2: Mass Spectrometry**

### *Part 2a: Checking the calibration of the instrument*

1) Carefully insert a nESI needle into the sample holder. Using a gel-loader pipet, load ~5  $\mu$ l of sodium iodide (NaI) into the needle. Screw the sample holder into the stage. Start the voltages and gases by pulling on the black screw on the side of the stage, and push the stage into the instrument.. Adjust the needle position to be close proximity to the cone (see Needle Position).

2) Open the .ipr file named “FFGSI\_NAI.IPR”. If you do not see any signals like the example in the MS guide, change the capillary voltage up or down by 0.5 kV increment (keep the capillary voltage in the 1000-3500 V range). If you still do not observe any signal, increase the nanoflow gas to 3 psi. Once you see your sample spraying, decrease the nanoflow gas to 0 bar. Make a note of the final capillary voltage you use to obtain your spectrum. When your spray is steady, hit the play button. Collect a mass spectrum from 200-8000 m/z for 2 minutes. Use a scan speed of 1 second.

3) Disengage the stage from the instrument. Carefully remove the needle from the sample holder and insert a new needle. Load ~5  $\mu$ l of the HEWL sample in the biological sample. Open the .ipr file named “FFGSI\_HEWL\_MS.IPR”. Play around with parameters given in MS guide-Instrument Parameters to obtain the best signal intensity. When your spray is steady, hit the play button and collect a mass spectrum from 200-8000 m/z for 2 minutes using a scan speed of 1 second. \*Note: if your spray stops in the middle of the data collection, or if you decide that you’re not happy with the parameters, you can collect multiple spectra. Just make sure you keep track of the spectrum you’re going to use for your data analysis.

4) Repeat step 3 with the HEWL sample from the biological buffer that has been buffer exchanged. Use new needles for each replicate.



6) Repeat step 3 with each of the six other buffer exchanged HEWL samples. Collect replicates of each sample.

7) Save all of your data to the Ruotolo Lab network drive under the FFGSI folder. Make a new folder for your project with your last names and your section, and copy and paste your data files to this folder.

### **Day 1 Data Analysis:**

1) Use the “Automatic Peak Detection” feature for your NaI spectrum. Copy and paste the chromatogram into MassDiff and note the mass accuracy. If any peaks did not agree well with the expected peaks, make a note of the error.

2) Use Masslynx to calculate the molecular weight of the charge state distributions you see.

3) Smooth the data with a smooth window of 15, the number of smooths 2, and the smoothing method Savitzky-Golay.

4) Calculate the weighted average charge state for each spectra by using the percent intensity of each charge state (TIC of the charge state/TIC total). Report the average and standard deviation of each charge state for each buffer condition.

### **Day 2- unknown ligand concentration:**

**\*Note: You will make the sample, and immediately spray the sample!**

1) Obtain one HEWL aliquot and one NAG5 aliquot. Dilute the HEWL sample to 5  $\mu\text{M}$  in a total volume of 800  $\mu\text{l}$ . Calculate how much of the NAG5 you need to make a 0.1, 0.5, 1, and 1.5 molar ratio sample of NAG5:lysozyme with 50  $\mu\text{l}$  of HEWL.

2) Pipette 50  $\mu\text{l}$  of HEWL into a clean eppendorf tube. Create your 0.1 molar ratio sample by pipetting the appropriate amount of NAG5 into the HEWL. Mix by gently flicking the eppendorf tube.

3) Immediately after mixing, collect a two minute spectrum of your 0.1 molar ratio sample.

4) Repeat steps 2 and 3 for your 0.5, 1, and 1.5 molar ratio samples, and then collect triplicates at each molar ratio (you should have a total of 12 spectra). This reaction happens quickly, so make a new sample for each trial.

5) Using Masslynx, extract the intensity of your cleavage product (this should have a 1+ charge state at  $\sim 830$  m/z). Divide the TIC of the cleavage product by the TIC of your whole spectrum for each sample.

6) Obtain a sample of the unknown NAG5. Add 13  $\mu\text{l}$  of the unknown sample to 50  $\mu\text{l}$  of 5  $\mu\text{M}$  HEWL. Mix by gently flicking the eppendorf tube.

- 7) Immediately after mixing, collect a two minute spectrum of your unknown sample.
- 8) Repeat steps 6 and 7 two more times for your unknown sample replicates.

## **Day 2 Data Analysis**

- 1) Using Masslynx, extract the intensity of your cleavage product (this should have a 1+ charge state at ~830 m/z). Divide the TIC of the cleavage product by the TIC of your whole spectrum for each sample to find the % cleavage product. Create a calibration curve by plotting the % cleavage product vs. molar ratio.
- 2) Using your calibration curve, find the concentration of the unknown ligand. Make sure you account for your dilution.

## **Lab Report**

- 1) Make a note of the error from MassDiff. Is the instrument properly calibrated? Explain your thinking.
- 2) Create a table of buffer conditions (pH and % methanol) and the average charge states. Compare the average charge state of HEWL calculated from equation (1) to each buffer condition. What does this tell you about how folded the protein is in solution (given that the lower charge state generally can be equated to a more folded protein structure).
- 3) Show an example spectra of buffer exchanged HEWL and HEWL in the native buffer using the smoothed data. What are the advantages and disadvantages of buffer exchanging your sample? How much protein, if any, did you lose during the buffer exchange process?
- 4) Show your calibration curve with error bars and the least squares best fit linear regression line for this plot.
- 5) Show an example spectrum of the cleavage product and the 1+ ligand peak at ~1034 m/z.
- 6) Make a note of the other species you see. (For example, do you see any cleavage products? Each reaction cleaves a sugar with a molecular weight of 203 Da. Do you see any other protein:ligand species forming?)
- 5) Report the average value you obtained for the n=3 measurements of the unknown ligand. Make sure you keep in mind the dilution that was done to make the sample. Provide the concentration for the sample with the 95% confidence interval.

## **II.III K<sub>a</sub> Measurement Student Handout**

# Special Project-HEWL-K<sub>a</sub> Amendment

## 1. Additional background

The binding strength of a ligand to a protein target is a key piece of data for inhibitor screening and drug discovery. Most often the binding strength is measured in terms of a K<sub>a</sub> value (dissociation constant), or it's inverse, the K<sub>d</sub> value. For a simple system, K<sub>a</sub> is measured in equation (1)

$$K_a = \frac{[PL]}{[P][L]} \quad (1)$$

Where [P] is the concentration of the free protein, [L] is the concentration of the free ligand, and [PL] is the concentration of the protein:ligand complex. In this experiment, [P] was kept the same, and [L] was titrated into the solution. A plot of [PL]/[P] vs [L] will give a straight line with a slope of K<sub>a</sub>. We can calculate the concentration of free ligand in solution from [L]<sub>total</sub> - [PL].

There are a number of assumptions that one needs to make in order to calculate K<sub>a</sub> values from mass spectrometry data:

- (1) Protein:ligand complexes are transferred from solution to the gas phase without disrupting the interactions between the protein and the ligand.
- (2) The intensities of the ions recorded by the mass spectrometer are proportional to the concentrations of each species observed (i.e. each species ionizes with the same efficiency)

**\*Note: you don't need to collect any more data—we'll use the data you already have!**

## 2. Data analysis

\*Note: you will have three of these curves—one for each “set” of concentrations you collected (minus the unknown)

*For each mass spectra...*

- 1) There is a lot of “noise” in the lower part of the mass spectrum due to peptides associated with lysozyme. Display from 266 to 480 m/z, and combine that region to find the TIC of the lysozyme noise.
- 2) Record the TIC for each individual charge state for the apo protein and the lysozyme:NAG5 complex.
- 3) Sum up the signals for the lysozyme noise and each apo charge state. This will be [P].
- 4) Sum up the signals for the lysozyme:NAG5 complex. This will be [PL].
- 5) For your y-values, divide [PL] by [P].

6) To determine  $[L]_{\text{free}}$  (for your  $x$  values), find the TIC for each spectra. The percent of your free ligand will be  $(1-[PL]/TIC)$ . Your values should be above  $\sim 0.6$ . Find the exact concentration of ligand in your solution by using  $M_1 V_1 = M_2 V_2$ . Multiply your exact ligand concentration by the percent free ligand to find your  $x$  value.

7) Make a plot of  $[PL]/[P]$  versus  $[L]_{\text{free}}$ . The slope of the line of your data will be the  $K_a$  in  $\mu M^{-1}$ . Report your  $K_a$  in  $M^{-1}$ .

### 3. Lab Report

1) Report your average  $K_a$  with standard deviations. Compare your findings with a 95% confidence interval and a student t-test to the mass spectrometry “true value” of  $1.2 \times 10^5 M^{-1}$ . Discuss how close your value is to the true value.

2) Show a graph with all three of your data curves. Make sure it's clear which  $R^2$  values and equations belong to which set of data.

### II.IV Protein Complex Analysis Student Handout

## Special Project-ConA v1

### 1. Introduction

The aim of this experiment is to introduce you to the concepts and applications of native mass spectrometry. Specifically, you will use a Waters Synapt HDMS, check the instrumental calibration, determine the recovery of protein after a buffer exchange step, determine the pH dependence of the tetrameric form of a protein, and use the relationship between dimer/tetramers to determine the pH of an unknown solution.

### 2. Background

Mass spectrometry (MS) is an extremely important analytical technique that measures the mass-to-charge ( $m/z$ ) ratio and abundance of gas-phase ions. MS has been used widely to determine the elemental or isotopic signature of a sample, measure the mass of molecules, and determine the chemical structures of measured molecules. For example, MS has been used in the field of

proteomics (to identify protein-protein interaction networks, and to quantify protein abundance as a function of time or some stimuli), to discover biomarkers for cancers and other diseases, and in drug discovery. In a typical MS workflow, a sample (which can be a solid, liquid, or a gas) is ionized. The ions are then introduced into the instrument and separated according to  $m/z$ . Over the past 20 years, a technique called native mass spectrometry has been increasingly used in the analysis of large biomolecules. For example, native mass spectrometry has been used to calculate the drug to antibody ratio of antibody:drug conjugates, which are used in cancer therapeutics.

One of the most important contributions to mass spectrometry, and especially native mass spectrometry has been the development of electrospray ionization (ESI). ESI is a gentle technique that allows for the introduction of whole, native proteins into a mass spectrometer for analysis, including oligomeric proteins that are held together by weak, noncovalent interactions. ESI additionally is advantageous because it results in multiple charges on an analyte, and each  $m/z$  from an analyte gives an individual measure of its mass. The charge state distribution (i.e. the number of  $m/z$  peaks that are present for a given analyte) is proportional to the surface area of the analyte. Thus, for a protein system, more extended, unfolded proteins will exhibit a broader charge state distribution with a greater number of charges than smaller, more compact proteins.

In this lab, you will be using a variant of ESI called nano-ESI. nanoESI is achieved by pulling glass capillaries to tips of a few micrometers. nanoESI has lower flow rates than ESI (nl/min as compared  $\mu$ l/min flow), and therefore is more tolerant for salts and other contaminants in buffers and has higher sensitivity (which means less sample is required, both in volume and concentration).

However, despite the tolerance of nanoESI to salts, proteins are typically buffered *in vitro* using salts like Tris, phosphate, and NaCl or KCl, among others. Which are typical for biological buffers. At the concentrations used in these biological buffers, the salt will mask the protein signal in the mass spectrometer. Therefore, the sample must be exchanged into a buffer more compatible with mass spectrometry. You will be using buffer exchange columns to quickly remove excess salt and switch your buffer to 200 mM ammonium acetate.

You will be using native mass spectrometry to evaluate the pH dependence of the tetrameric form of concanavalin A (abbreviated ConA). ConA is a carbohydrate-binding protein that binds to specific structures found in various sugars, glycoproteins, and glycolipids. It was one of the first carbohydrate-binding proteins for which a three-dimensional structure was determined. Crystallography experiments found a tetramer of nearly identical, 237 amino acid ( $M_r$  25,500) monomers. However, in solution, a dimer-tetramer equilibrium has been reported that is both temperature and pH dependent.

This experiment has four goals: (1) to determine recovery of protein after a buffer exchange step, (2) use native mass spectrometry to measure the dimer/tetramer equilibrium at various pHs and determine how close it is to the “ideal” charge state at each pH using equation 1, (3) confirm the oligomeric assignment using MS/MS, and (4) use this relationship to determine the pH of a ConA sample at an unknown pH.

### **3. Procedure**

#### **Required Apparatus:**

- 1) UV-Vis spectrophotometer
- 2) Waters Synapt HDMS
- 3) 6 kDa molecular weight cut-off (MWCO) buffer exchange columns
- 4) Microcentrifuge
- 5) UVettes (cuvettes for the UV-Vis)

#### **Reagents:**

- 1) 200 mM ammonium acetate at pH 6.11, 6.45, 7.14, and 7.59
- 2) Unknown ammonium acetate buffer
- 3) ConA samples, starting concentration of 50  $\mu$ M, made up in 200 mM ammonium acetate
- 4) CsI
- 5) Biological buffer

## Day One Procedure

### Part 1: Buffer Exchange

1) Thaw two ConA aliquot on ice. One will be your “biological buffer” sample, and one will be the “buffer exchanged” sample. Measure the absorbance of 50  $\mu$ l of the biological buffer sample ConA at 280nm using the provided cuvettes. Use the biological sample as the blank.

2) Obtain one buffer exchange column and one balance column. The buffer exchange column is stored in a Tris buffer, which is incompatible with electrospray, so the columns will have to be cleaned before buffer exchanging the protein. Snap off the bottom part of the buffer exchange column and place it in a 1.5 ml collection tube (the one without a lid), and remove the green cap of the buffer exchange column. Follow this sequence to prepare the buffer exchange columns and exchange your sample:

a) Place the buffer exchange column AND BALANCE in the microcentrifuge. **\*Caution: Make sure the centrifuge is properly balanced!!!!!!\*** Put the lid on the centrifuge, and shut the top. Set the speed to 1.0xg (not  $\text{min}^{-1}$ ) and the time to 2 minutes (the centrifuge is set to read minutes, except for the final minute, where the centrifuge will count down in seconds). If you are using one of the temperature-controlled microcentrifuges, make sure the centrifuge is set to 25°C. Hit the play button to start the first cycle. This step will drain the Tris buffer from the column. When the centrifuge is finished spinning (after it beeps), pour the Tris buffer that is now in your collection tube into the waste.

b) Place the buffer exchange column back in the collection tube (you can re-use the collection tube until you collect your protein). Carefully pipet 500  $\mu$ l of 100 mM ammonium acetate pH 7.0 onto the top of the column. **Do not do this with too much force—the goal is to not disturb the column bed!** Place the buffer exchange column and the collection tube back into the microcentrifuge. Keep the speed on 1.0xg, but change the time to 1 minute. Hit the play button to start the next cycle. When the centrifuge is finished spinning (after it beeps), pour the ammonium acetate buffer that is now in your collection

tube into the waste. Repeat the step 3 more times, so you have washed the columns with ammonium acetate buffer for a total of 4 times. This will make sure that all of the Tris buffer has been replaced with ammonium acetate, and your protein will be effectively exchanged into your buffer of choice (in this case, ammonium acetate).

c) *Without adding in any buffer or protein*, place the buffer exchange column and the collection tube back into the microcentrifuge. Keep the speed on 1.0xg, and keep the time set to 1 minute. Hit the play button to start the next cycle. This will remove any excess buffer from your column that can dilute your protein during the buffer exchange process.

d) Label a 1.5 ml Eppendorf tube for your sample. Place the buffer exchange column into this new tube and throw away the collection tube. Carefully pipet 100  $\mu\text{l}$  of your ConA protein onto the top of the column bed. Be very careful not to introduce bubbles into your sample! Place the buffer exchange column and the Eppendorf tube back into the microcentrifuge. Keep the speed on 1.0xg, but change the time to 4 minutes. Hit the play button to start the final cycle. This is the buffer exchange step. After this, your protein sample will be in your ammonium acetate buffer of choice.

3) Measure the absorbance at 280 nm of 50  $\mu\text{l}$  of your buffer exchanged ConA sample, using the 200 mM ammonium acetate as a blank. Make a note of the absorbance and also the volume of your sample, being careful not to introduce bubbles into your protein. Calculate the concentration of both your “buffer exchanged” and “biological” samples ( $\epsilon=33920 \text{ M}^{-1}\text{cm}^{-1}$ ). You may need to dilute your sample if the absorbance is  $>1.0$

## **Part 2: Mass Spectrometry**

### *Part 2a: Checking the calibration of the instrument*

1) Carefully insert a nESI needle into the sample holder. Using a gel-loader pipet, load  $\sim 5 \text{ ul}$  of cesium iodide (CsI) into the needle. Carefully screw the sample holder into the stage. Start the voltages and gases by pulling on the black screw on the side of the stage, and push the stage into



the instrument. The box in the corner of the MS tune page will turn from yellow to green when the stage is fully pushed in. Adjust the needle position to be close proximity to the cone (see MS guide-Synapt help-Needle Position).

2) Open the .ipr file named “FFGSI\_CSI.IPR”. If you do not see any signals like the example in the MS guide, change the capillary voltage up or down by 0.5 kV increment (keep the capillary voltage in the 1.00-2.00 kV range). If you still do not observe any signal, increase the nanoflow gas to 0.1. Once you see your sample spraying, decrease the nanoflow gas to 0 bar. \*Note: it will take a while for the nanoflow gas to fully decrease to 0. Make a note of the final capillary voltage you use to obtain your spectrum. When your spray is steady, hit the play button. Collect a mass spectrum from 1000-10000 m/z for 3 minutes. Use a scan speed of 1 second.

3) Disengage the stage from the instrument. Carefully remove the needle from the sample holder and insert a new needle. Load ~5 ul of “non-buffer exchanged” ConA sample. Open the .ipr file named “FFGSI\_CONA\_MS.IPR”. Play around with parameters given in MS guide-Instrument Parameters to obtain the best signal intensity. When your spray is steady, hit the play button and collect a mass spectrum from 1000-10000 m/z for 3 minutes using a scan speed of 1 second. \*Note: if your spray stops in the middle of the data collection, or if you decide that you’re not happy with the parameters, you can collect multiple spectra. Just make sure you keep track of the spectrum you’re going to use for your data analysis.

4) Increase the trap collision voltage to 75V and take a mass spectrum of your “buffer exchanged” ConA sample (there’s no reason to change the needle for this). Increasing the trap collision voltage increases energy with which the protein ions collide with neutral gas molecules. This will help to knock off excess salt ions so you can have an accurate mass.

5) Change the trap collision voltage back to 8V. Using new needles, collect three more mass spectra of your “buffer exchanged” ConA samples for a total of 3 spectra at 8V and three spectra at 50V.

6) Repeat steps 4 and 5 with your “non-buffer exchanged” ConA sample. Use new needles for each replicate.

7) Save all of your data to the Ruotolo Lab network drive under the FFGSI folder. Make a new folder for your project with your last names and your section, and copy and paste your data files to this folder.

### **Day 1 Data Analysis:**

1) Use the “Automatic Peak Detection” feature for your CsI spectrum. Copy and paste the chromatogram into MassDiff and note the mass accuracy. If any peaks did not agree well with the expected peaks, make a note of the error.

2) Obtain a molecular weight for each oligomeric species in each spectrum using Masslynx (see Masslynx help guide) for the samples sprayed with the trap collision voltage at 8V and the trap collision voltage of 50V. Make a note of how the two compare.

3) Smooth the data with a smooth window of 15, the number of smooths 2, and the smoothing method Savitzky-Golay.

4 ) Fill in the ConA\_MW worksheet in order to see the full set of  $m/z$  for each oligomeric species.

### **Day 2-Measuring the pH dependence of the ConA tetramer:**

1) Obtain six buffer exchange columns. Clean the columns according to the Day 1 procedure using the 200 mM ammonium acetate 6.11, 6.45, 7.14, 7.59, and unknown pH buffers. Buffer exchange ConA samples into each buffer.

2) Insert a needle into the sample holder. Load the pH 4 sample into the needle and spray. Use the FFGSI\_CONA\_MS.ipr file.

3) Collect one 3 minute mass spectrum. Use Masslynx to calculate the masses and charge states of the peaks.

4) Confirm the peak assignments for overlapping dimer/tetramer charge states (see your ConA\_MW worksheet) by using MS/MS (there is no need for a new needle at this point). Increase the trap collision voltage to 75V, and acquire a MS/MS spectrum for 3 minutes (see Help-MSMS). Inspect the mass spectrum and identify the peaks you see using the MSMS chart (see Example-peak assignment check).

5) Using new needles each time, obtain two replicates for the pH 4 buffer sample for a total of 3 replicates.

6) Repeat this process with the other buffer exchanged samples to construct your calibration curve.

7) Obtain a sample of ConA with an unknown pH. Repeat steps 2-5 to find the dimer/tetramer ratio, and repeat the experiment in triplicate.

## **Day 2 Data Analysis**

1) Export example spectra for each pH. Make a note of the assignments of each of the charge states

2) After confirming the identification of the peaks in the mass spectrum, find the intensity of the tetramer TIC. Find the ratio of percent of tetramer in each sample (tetramer TIC/total TIC). Plot the percent tetramer against pH for your calibration curve.

3) Using your calibration curve, find the pH of the unknown sample. Include a table that includes 95% confidence intervals and LINEST analysis.

## **Lab Report**

- 1) Make a note of the error from MassDiff. Is the instrument properly calibrated? Explain your thinking.
- 2) Create a table with the average molecular weights of the monomer, dimer, and tetramer. Given that the molecular weight derived from the protein sequence is 25539.40 Da, how accurate is your molecular weight? Comment on which sample and conditions gave you the most accurate molecular weight, and comment on the precision of your measurements among similar samples and conditions. Explain what could contribute to more or less molecular weight.
- 3) Show an example spectra of buffer exchanged ConA and non-buffer exchanged ConA using the smoothed data. Comment as to the resolution and centroid  $m/z$  of the peaks between the samples at 8V. What are the advantages and disadvantages of buffer exchanging your sample? Also comment as to how the resolution of the peaks and the centroid  $m/z$  changes when the trap collision voltage is increased. Use a student t-test to calculate how close the high collision voltage masses are to the “ideal” masses of the monomer, dimer, and tetramer based on the sequence mass.
- 4) Show one example MSMS dimer identification and one example MSMS tetramer identification.
- 5) Show an example mass spectrum for each pH value.
- 6) Show your calibration curve with error bars and the least squares best fit linear regression line for this plot.
- 7) Report the average value you obtained for the  $n=3$  measurements of the unknown 200 mM ammonium acetate buffer. Provide the pH range for the sample with the 95% confidence interval. Assuming that human blood has a pH of  $\sim 7.4$ , is your pH measurement in the proper biological range? Explain your reasoning.

## **II.V Instrument Parameters for Special Projects**

### LCT Instrument Parameters

Sample aliquots of ~5  $\mu$ l were analyzed in positive ion mode. The capillary voltages ranged from 2.0-2.5 kV with the sample cone set to 0.0 V. The hexapole and ion guide energies were minimized to reduce protein:ligand dissociation. The TOF mass analyzer was operated over the range of 200-8000 m/z at a pressure of  $1.25 \times 10^{-6}$  mbar.

#### Synapt Instrument Parameters

Sample aliquots of ~5  $\mu$ l were analyzed in positive ion mode. The capillary voltages ranged from 1.5-1.8 kV, with the sampling cone set to 15.0V and the extraction cone set to 0.0V. Trap collision voltage was increased from 8.0V to 70V for collisional induced dissociation of mass selected oligomers. The TOF mass analyzer was operated over 1000-10000 m/z at a pressure of  $1.77 \times 10^{-6}$  mbar, with the quad profile set to auto.

## II.VI Protein Ligand Interaction Grading Rubrics

### FFGSI Project - Lysozyme

Part	Description	Points Possible	Points Obtained	Comments
Introduction and Presentation	<ul style="list-style-type: none"> <li>- Clear and concise writing style</li> <li>- Demonstration of knowledge of subject and purpose of experiment</li> </ul>	20		
Materials and Methods	<ul style="list-style-type: none"> <li>- Citing the procedure that was used and make note of any deviations from the procedure provided.</li> </ul>	10		
Day One Data Analysis	<ul style="list-style-type: none"> <li>-Discussion of MassDiff error</li> <li>-Table of ammonium acetate concentrations and weighted average charge state.</li> <li>-Comparison of measured MW to theoretical MW with a 95% confidence interval.</li> </ul>	15		
Day Two Data Analysis	<ul style="list-style-type: none"> <li>-Calibration curve of protein:ligand ratios including the <math>R^2</math> value and linear regression analysis.</li> <li>-Report the average concentration of the unknown NAG5 sample.               <ul style="list-style-type: none"> <li>– 95% confidence interval</li> <li>– Student t-test</li> </ul> </li> <li>-Report <math>K_a</math> and compare to mass spec “true value”.               <ul style="list-style-type: none"> <li>– 95% confidence interval</li> <li>– Student t-test</li> </ul> </li> </ul>	20		
Figures and Tables	<ul style="list-style-type: none"> <li>-Representative raw data shown (if applicable) and calibration curves.</li> <li>-Correct formatting.</li> <li>-Figures and tables mentioned in results including proper labeling, title, and captions .</li> </ul>	10		
Conclusions	<ul style="list-style-type: none"> <li>- Summary and comments on the results of the experiments.</li> <li>-Discuss sources of error.</li> </ul>	15		
Survey	<ul style="list-style-type: none"> <li>-Answered all survey questions.</li> <li>-Provide detailed feedback</li> </ul>	10		
		100	Total	

## II.VII Protein Complex Analysis Grading Rubric

### FFGSI Project - Lysozyme

Part	Description	Points Possible	Points Obtained	Comments
Introduction and Presentation	<ul style="list-style-type: none"> <li>- Clear and concise writing style</li> <li>- Demonstration of knowledge of subject and purpose of experiment</li> </ul>	20		
Materials and Methods	<ul style="list-style-type: none"> <li>- Citing the procedure that was used and make note of any deviations from the procedure provided.</li> </ul>	10		
Day One Data Analysis	<ul style="list-style-type: none"> <li>-Discussion of MassDiff error</li> <li>-Table of ammonium acetate concentrations and weighted average charge state.</li> <li>-Comparison of measured MW to theoretical MW with a 95% confidence interval.</li> </ul>	15		
Day Two Data Analysis	<ul style="list-style-type: none"> <li>-Calibration curve of protein:ligand ratios including the <math>R^2</math> value and linear regression analysis.</li> <li>-Report the average concentration of the unknown NAG5 sample.               <ul style="list-style-type: none"> <li>– 95% confidence interval</li> <li>– Student t-test</li> </ul> </li> <li>-Report <math>K_a</math> and compare to mass spec "true value".               <ul style="list-style-type: none"> <li>– 95% confidence interval</li> <li>– Student t-test</li> </ul> </li> </ul>	20		
Figures and Tables	<ul style="list-style-type: none"> <li>-Representative raw data shown (if applicable) and calibration curves.</li> <li>-Correct formatting.</li> <li>-Figures and tables mentioned in results including proper labeling, title, and captions .</li> </ul>	10		
Conclusion s	<ul style="list-style-type: none"> <li>- Summary and comments on the results of the experiments.</li> <li>-Discuss sources of error.</li> </ul>	15		
Survey	<ul style="list-style-type: none"> <li>-Answered all survey questions.</li> <li>-Provide detailed feedback</li> </ul>	10		
		100	Total	