Comparing Solution-Phase and Gas-Phase Protein Stability Using Ion Mobility and Differential Mobility Mass Spectrometry

LUCIENNE NOUCHIKIAN

THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULLFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN CHEMISTRY
YORK UNIVERSITY
TORONTO, ONTARIO

April 2019

Abstract

A critical quality attribute for many proteins concerning biopharmaceutical products is the protein's stability. Current stability assays are lengthy but some quick stability assays to determine the protein's melting temperature include thermal ramps of a protein using a fluorescent dye to monitor unfolding. Protein unfolding has also been studies with mass spectrometry through collision induced unfolding (CIU) whereby the ion temperature in the trap collision cell is increased and unfolding changes are tracked in the ion mobility cell. However, whether a protein retains its structural properties in the gas phase has been a topic of wide discussion. Here, we present that CIU and solution thermal melts provide the same information on protein stability using a homo-tetrameric model protein, pyruvate kinase, and four of its point mutants. Efforts to employ a second ion mobility technique – differential mobility spectrometry, to determine gas phase stability will also be discussed.

Acknowledgements

I would first of all like to thank Dr. Derek Wilson for all his support during my masters. Thank you for not worrying about me despite me not being able to purify my WT for about a year. I would also like to thank Dr. Philip Johnson, Dr. Ryan Hili, and Dr. Mark Bayfield for taking the time to be on my committee and for their encouragement and guidance in my research work. A special thank you to Dr. Brendon Seale and Dr. Yves LeBlanc for not letting me drown in the complicated sea of DMS, and helping me learn so much. A huge thank you to all my lab members for helping me from the beginning as I was a newbie in the world of MS. Cristina Lento, thanks for helping me and all the other student as they were having a mid-experiment crisis, always being there even if it was only for moral support, and fueling our lab's Rubik's cube obsession. Xiaojing Huang, thanks for being my lab buddy during our purification days and being just as crazy (maybe more) as me and helping me start the Wilson Lab rock climbing club and signing up for silks with me. I would not have done those things alone. Irina Oganesyan and Kerene Brown thanks for all the help that I got from you about MS and almost anything else that I asked and for bringing such positive and hard-working energy to the lab every day. Thank you to Peter Liuni, Shaolong Zhu, Ruth Knox, Lisa Szymkowicz, Shenbaga Moorthy, Esther Wolf, John Van Nostrand, Banafsheh Mahrazma and Alex Poirier for their added support to everyone who needed it. Finally, to my family, my dad –Armen Nouchikian, my mom – Valentina Nouchikian, my little sister –Irina Nouchikian, and my best friend –Tamara Babic (and her cat) thank you for being my biggest supporters in all the decisions that I made. You helped me destress during long days and celebrated my victories. Thank you all very much! This work would not have been such a success if I didn't receive all this help along the way.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Ampicillin (Amp)

Collision Energy (CE)

Collision Voltage (CV)

Collisional Induced Unfolding (CIU)

Compensation Voltage (CoV)

Differential Mobility Spectrometry (DMS)

Differential Scanning Calorimetry (DSC)

Differential Scanning Fluorimetry (DSF)

Electrospray Ionization (ESI)

Ethylenediaminetetraacetic acid (EDTA)

Flow Through (FT)

High Definition Mass Spectrometer (HDMS)

High Fidelity (HF)

Hydrogen Deuterium Exchange (HDX)

Ion Mobility (IM)

Ion Mobility Spectrometry (IMS)

Isopropyl β-D-1-Thiogalactopyranoside (IPTG)

Kanamycin (Kan)

Luria Bertani (LB)

Mass Spectrometry (MS)

Melting Temperature (T_m)

Overnight (O/N)

Pyruvate Kinase (PK)

Pyruvate Kinase gene code (*pykF*)

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Time Resolved (TR)

Time-of-Flight (TOF)

Total Lysate (TL)

Travelling Wave (TW)

Tris-Acetate EDTA (TAE)

List of Publications

1. Nouchikian, L., Roque, C., Song, J.Y., Rahman, N., Ausar, S.F.: An intrinsic fluorescence method for the determination of protein concentration in vaccines containing aluminum salt adjuvants. *Vaccine*. **36**, 5738–5746 (2018).

Chapter 1: Introduction

1.1 Electrospray Ionization for Native Protein Analysis by Mass Spectrometry

The modern mass spectrometer was reported in 1918 by a Canadian born scientist, Arthur Jeffrey Dempster[1]. However, for over 60 years, mass spectrometry (MS) has been used to solely analyze the mass to charge (m/z) of small molecules[2]. Only until the invention of "soft" ionization methods such as electrospray ionization (ESI) did the use of proteins in mass spectrometry take-off. ESI is considered a soft ionization method as it produces minimal fragmentation thus allowing for the study of non-covalent interactions such as quaternary protein structure, protein-ligand, or protein-protein interactions[2]. Some unspecific interactions may also be captured using ESI which requires some caution to be undertaken when interpreting data. However, these non-specific interactions can typically be ruled out based on prior protein structure knowledge or by increasing capillary voltages or other in-source parameters which may result in dissociation of weakly held non-specific interactions and preserve slightly stronger specific non-covalent formations.

1.1.1. Mechanism of Positive Electrospray Ionization

Depending on what the analyte of interest is, ESI can either be used in positive or negative mode. Analytes that are easily protonated and thus carry a positive charge are analyzed in positive mode. Positive mode analytes include proteins as certain amino acids can be protonated such as histidine and to a certain extent aspartic acid and glutamic acid[3]. Analytes that are prone to deprotonation such as lipids, are typically analyzed in negative mode. Here, the process of positive ESI mode is further explained with the mechanistic features shown in **Figure 1**.

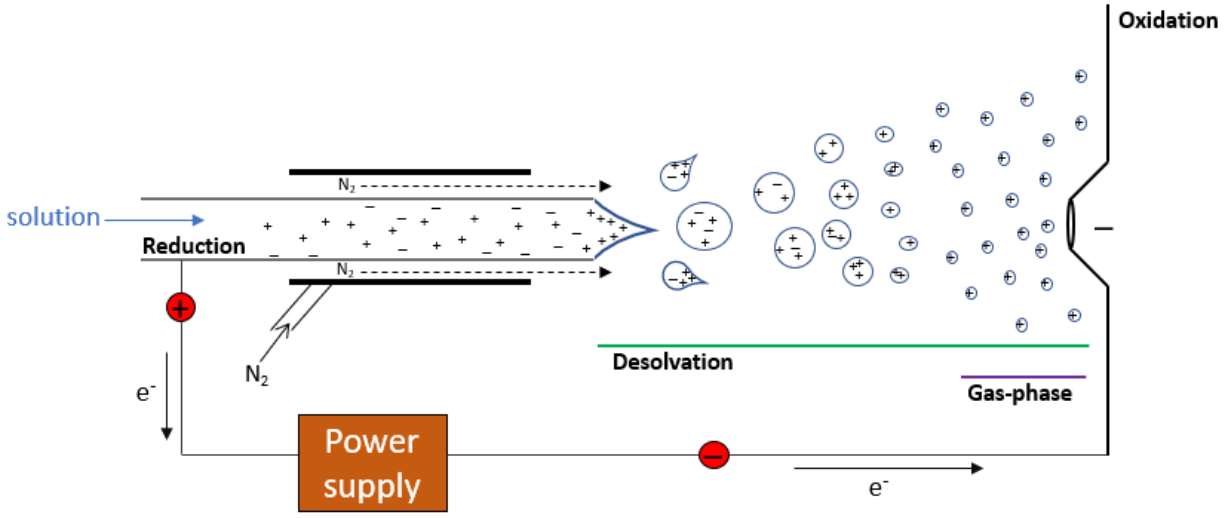

Figure 1. Mechanism of Positive Electrospray Ionization. A voltage is applied to the metal capillary resulting in an accumulation of positive charges at the Taylor cone. Droplets are further desolvated until only a single gas-phase ion (indicated as a positive charge) is left.

For positive ion mode ESI, liquid is flowed through a metal capillary where a high voltage is applied[4]. For native MS, voltages in the lower end of the spectrum (~1.5 -2 kV) are used to minimize any dissociation of weakly held interaction and premature unfolding of the protein that may occur. However, too weak of a voltage may cause insufficient desolvation of the protein droplets[3] and thus a balance must be maintained. This voltage allows for reduction to occur on the inside of the metal capillary and electrons get carried away onto the cone where oxidation occurs which helps charge the protein in solution[3]. Once the proteins are charged, they are sprayed into the orifice of the mass spectrometer through droplet formation. There are three main stages to ESI droplet formation[5]. Initially, small charged droplets form at the tip of the capillary as a result of the electric field applied. These droplets contain a large population of positively charged ions which creates an elongated droplet shape which is known as a Taylor cone. Dry N_2 gas is applied on the outside of the capillary to assist in desolvation and formation of a spray cone. As solvent evaporates, the droplets decrease in size until a certain limit, known

as the Rayleigh limit, is reached. At this point, the coulombic forces overpower surface tension and resulting in droplet fission resulting in small and highly charged droplets[5]. This process continues until finally gaseous ions are formed and detected by the mass spectrometer.

Depending on the analyte type, there are two main mechanisms that have been proposed as to how the final gas stage phase is reached (**Figure 2**). For smaller ions such as organic or inorganic compounds, the Ion Evaporation mechanism (IEM) is dominant. Here, the gas phase ion is capable of being emitted from the droplet due to the electric field caused by the charges on the surface of the droplet[5, 6]. The second model is called the charged residue model (CRM) where the last stages of the droplet completely evaporate releasing a single ion. CRM has been shown to be effective when analyzing folded proteins[7, 8]. For unfolded proteins, a chain ejection model (CEM) has been proposed which combines the two previous models where the ion is ejected before full desolvation[8, 9].

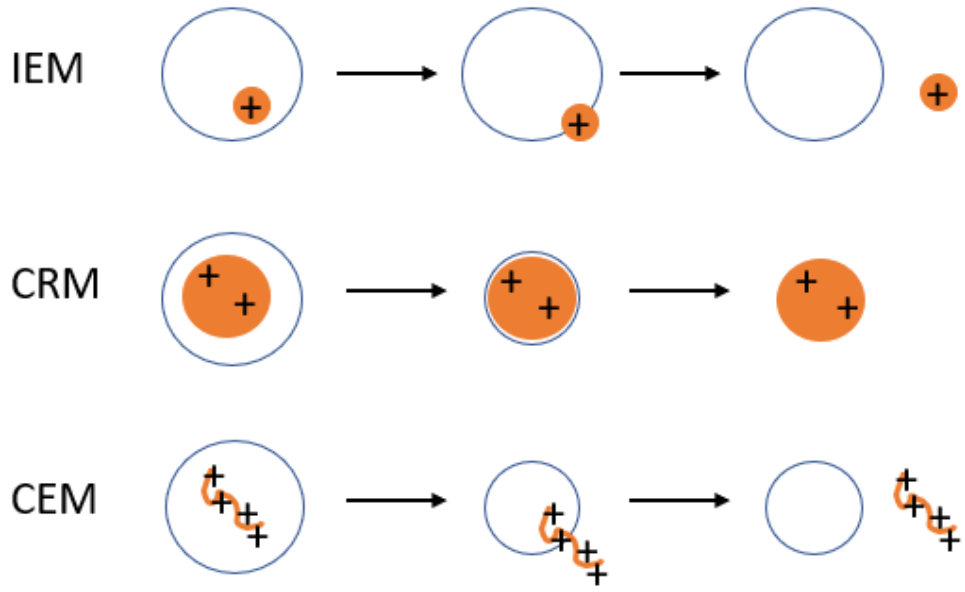

Figure 2. Desolvation Mechanisms. (top) Ion Evaporation Mechanism (IEM) for small molecules. (middle) Charged Residue Model (CRM) for structured proteins or peptides. (bottom) Chain Ejection Model (CEM) for unfolded proteins

1.1.2. What is "Native" Mass Spectrometry?

The term "native" mass spectrometry has been scrutinized for years. A large question arises when proteins are transferred over to the gas-phase which is: whether their biological structure is preserved in the gas-phase and if this information can be biologically relevant. Without solvent, many important interactions such as the hydrophobic effect are lost. Many have reported that upon transition to the gas phase, charges collapse onto the backbone of the protein due to water loss, resulting in a compact ion[10] (**Figure 3**). Others have reported that due to high Coulombic interactions, there is a repulsion of charges which results in an expanded protein ion structure[11, 12].

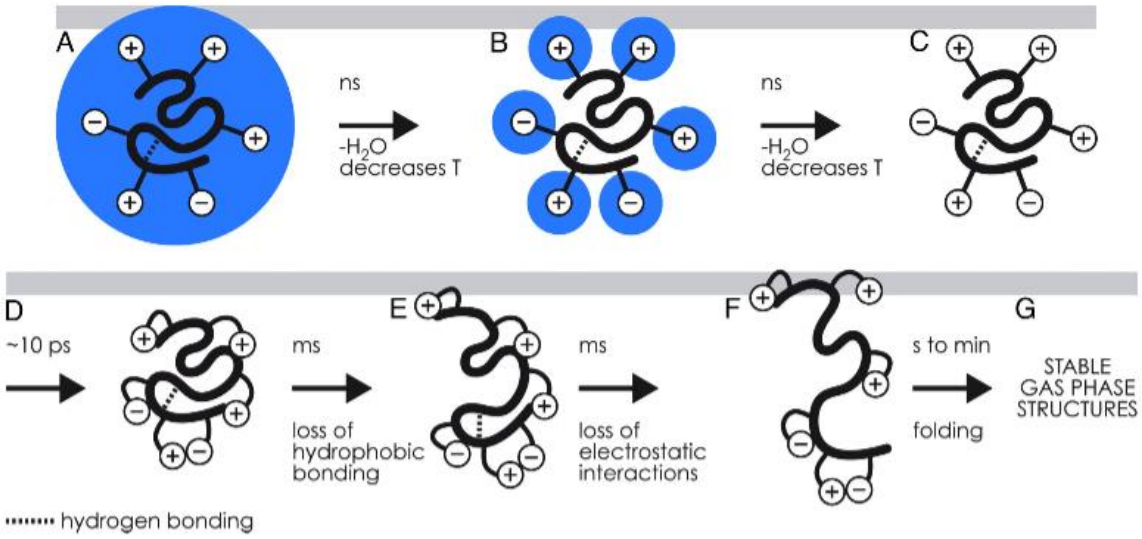

Figure 3. Progression of a structured protein after ESI in the gas phase. (a) native protein surrounded by water. (b) Protein undergoes water loss. (c-e) Collapse of protein side-chains and slight unfolding from picoseconds to milliseconds. (f-g) more stable gas-phase equilibrium is reached. Adapted from Breuker and Mclafferty (2018)[10].

Many studies have also shown that proteins maintain their native solution phase conformations. A "VIP" paper published by Seo *et al*. (2016)[13] compared the charge state envelope of native Myoglobin and β-Lactoglobulin in non-denaturing (ammonium acetate) and

denaturing solutions[13]. Proteins in non-denaturing solutions showed a charge state envelope that had a predominant charge of 8+ while in denaturing solutions the charge state was predominately 10+ for Myoglobin and 14+ for β-Lactoglobulin (**Figure 4**).

Not only was the protein more charged in denaturing solutions indicating protein unfolding as more surface area was available to take up charge is solution, but with the use of Ion Mobility Spectrometry (IMS) (described in more detail in section 1.2), Seo and colleagues managed to show that the average collisional cross section (CCS) of the proteins increased in denaturing solution compared to ammonium acetate further confirming unfolding (**Figure 4 c,d**). These results showed that the solution phase protein conformation was maintained and reflected in the gas-phase.

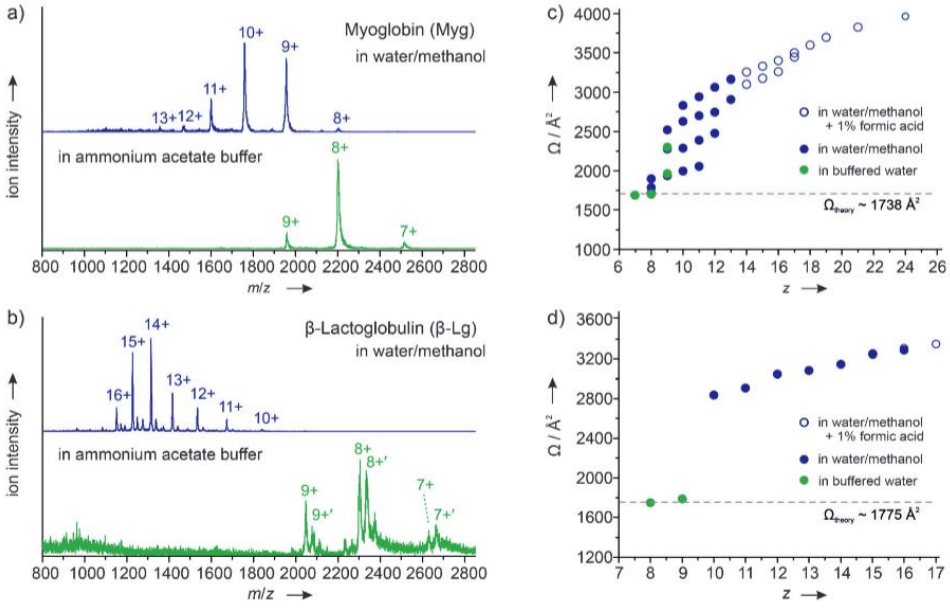

Figure 4. Mass spectra and collisional cross section of Myoglobin and β-Lactoglobulin. Mass spectra for Myoglobin (a) and β-Lactoglobulin in non-denaturing (bottom spectrum) and denaturing solutions (top spectrum). Collisional cross section of charge states for Myoglobin (c) and β-Lactoglobulin (d) in non-denaturing (green circle) and denaturing (solid and open purple circles). Adapted from Seo *et al.* (2016)[13].

Others have shown that folded protein maintain their solution-like state even after 100 ms (**Figure 5**) and only begin to adapt unfolded intermediates after ~1 s and gas-phase equilibrium

structures after 1 min or more[14]. Additionally, some experiments showed using gas-phase hydrogen deuterium exchange (HDX) that a gas-phase equilibrium is not reached even after one hour[15]. However, for unstructured and highly charged proteins, this gas-phase equilibrium is reached much faster, even before 100 ms[14]. Despite some studies showing differences in gas-phase protein ion structure, there is an increasing amount of evidence that "native" solution-like protein structure is captured in the gas-phase especially with the use of more "gentle" settings. Due to its high mass-accuracy and its ability to preserve structural information, native ESI-MS has been used in many biological applications including the redefining the model of RNA polymerase III[2].

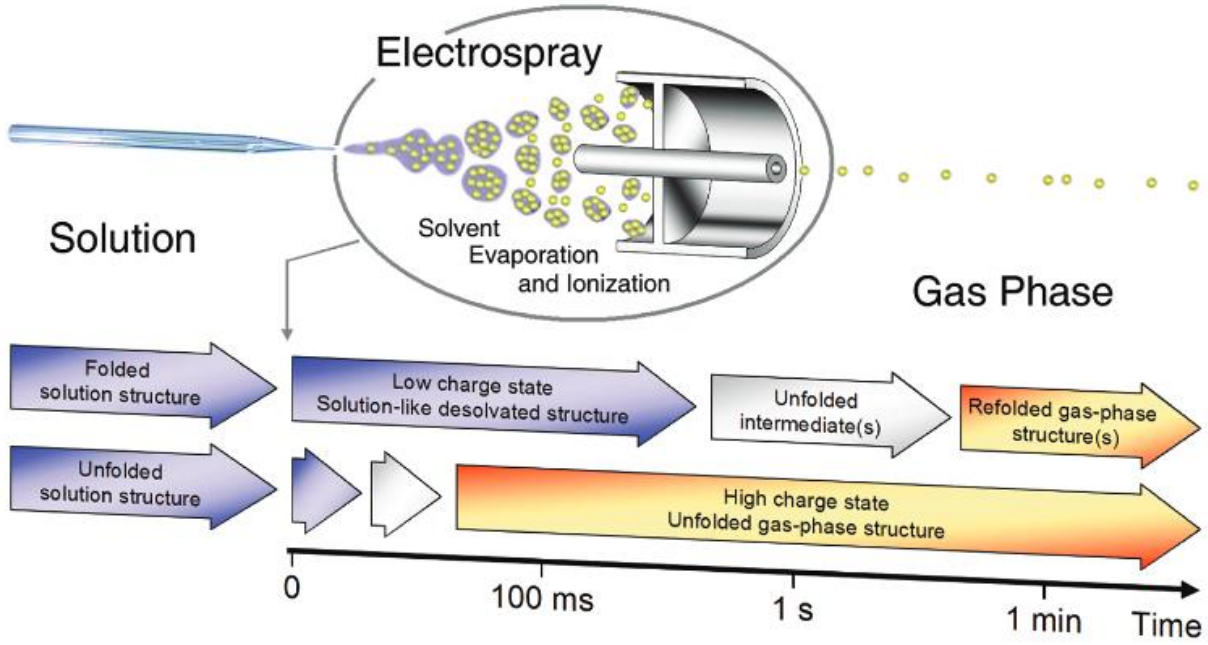

Figure 5. A time-line representation of the transition of folded and unfolded proteins to their gas-phase conformations. Folded proteins retain their solvent like structure long enough for analysis to take place. Unfolded proteins lose their solution phase structure much quicker. Adapted from Wyttenbach and Bowers (2011)[14].

1.2 Ion Mobility Spectrometry - a powerful technique for studying protein structure in the gas phase

Ion mobility spectrometry (IMS) is a technique that characterizes a swarm of gaseous ions as they enter a gas filled separation tube with an applied electric field[16]. Ions in this swarm are then separated based on their charge and cross-section[17] which can be thought of as concept that is similar to that of SDS-PAGE where smaller, more highly charged ions move through the separation chamber faster than larger, less charged ions. IMS was coupled to MS in 1962 however it took some time for it to gain popularity[17]. With the introduction of ESI-MS and the ability to do native MS work on proteins, IMS has been a huge resource to determine protein structure in the gas-phase. IMS has been used in a wide variety of applications, from airport drug screening, to quality control in the pharmaceutical and food industry, to medical diagnosis[16]. Although there are various versions of IMS, three of the most commonly used types are: drift

tube ion mobility spectrometry (DTIMS), travelling wave ion mobility spectrometry (TWIMS), and field-asymmetric ion mobility spectrometry (FAIMS) which is also commonly known as differential mobility spectrometry (DMS). TWIMS and DMS will further be discussed in detail in sections 1.2.1 and 1.2.2 respectively.

Despite being different techniques, these three methods rely on common principles. As ions move through an electric field or voltage gradient (*E*, expressed as V/cm), their drift velocities (v_d) depend on the time needed for a swarm of ions to travel the distance (*d*, cm) of the drift tube (drift time, t_d)[16] (Equation 1).

$$v_d = \frac{d}{t_d} \quad \text{(Equation 1)}$$

Drift velocity is then normalized to the electric field by the following equation (2), where K is the mobility coefficient usually reported as cm^2/Vs[16].

$$K = \frac{v_d}{E} \quad \text{(Equation 2)}$$

K provides information on the collision frequency of the ions with the buffer gas in the drift region and is therefore dependent on the density of the gas in the drift regions (*N*), which is related to pressure (*P*) and temperature (*T*)[18]. As pressure and temperature may vary *K*, a reduced mobility measurement (K_o) can be reported for inter-laboratory comparisons[18] (Equation 3).

$$K_o = K\frac{N}{N_o} = K\left(\frac{P}{P_o}\right)\left(\frac{T_o}{T}\right) \quad \text{(Equation 3)}$$

N_o is Loschmidt's number (number of particles of an ideal gas, Avagadros number, per given volume) and P_o and T_o are values at STP (100 kPa and 0 °C respectively)[18].

Finally, K_o also depends on the ratio of *E/N* and can be written as the following:

$$K_o\left(\frac{E}{N}\right) = K_o(0)\left[1 + \alpha_2\left(\frac{E}{N}\right)^2 + \alpha_4\left(\frac{E}{N}\right)^4 + \cdots\right] \quad \text{(Equation 4)}$$

The alpha function characterizes the electric field dependence of ion mobility and has been shown to be a unique ion feature that can be used for separation[19]. In linear methods, such as DTIMS and TWIMS, it can be assumed that *E/N* is low enough to not influence mobility and thus Equation 4 is simply $K_o=K_o(0)$, however, in non-linear methods, such as DMS, Equation 4 is used[18].

1.2.1. Travelling Wave

The introduction of TWIMS coupled to Time-of-Flight (TOF) MS has made a significant impact on the analysis of large biological molecules[16]. **Figure 6** shows a schematic of the commercially available TWIMS mass spectrometer, Synapt^{TM} G2-S HDMS (high definition mass spectrometer), that were released by Waters Corporation. Here, the ions enter the mass spectrometer by ESI and make their way to the ion guide which helps maximize ion transmission to the mass analyzer. Once past the ion guide, ions enter the quadrupole where certain ions based on their m/z may be selected. After the quadrupole, ions enter the TRI-WAVE (highlighted as a red box), which is composed of three travelling wave cells including the IMS cell for separation.

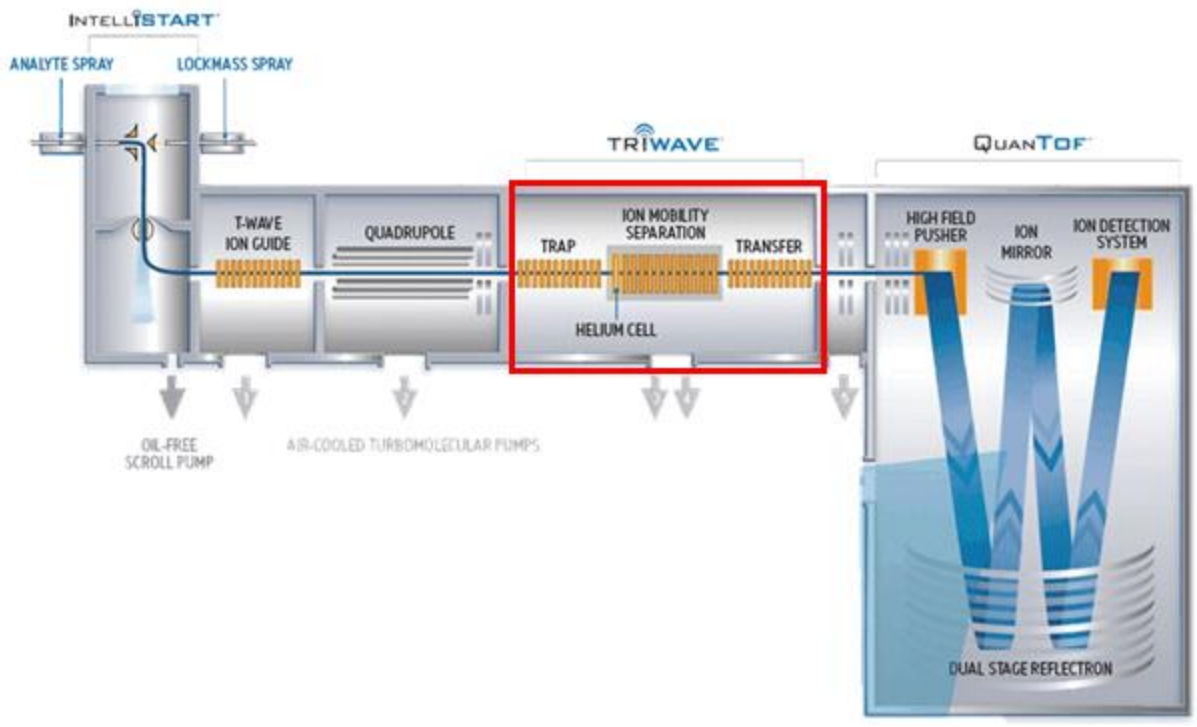

Figure 6. Schematic of Waters G2-S HDMS TOF-MS. The TRI-WAVE containing the three travelling wave cells (Trap, IMS, and Transfer) are highlighted in a red box. Schematic is taken from Waters (Milford, MA, USA).

The first travelling wave cell is called the trap cell and can either be used as an ion transfer cell or a collision cell[16]. Wave height, wave velocity, and gas pressure (typically consisting of an inert gas such as Argon) can play a role in ion transmission depending on ion size. The trap cell can also be used as a pre-IMS collision cell. Here, the collision voltage (CV), also known as collision energy (CE), in the trap cell can be increased which effectively increases the ion temperature and increases the energy that an ion collides with a gas molecule[20]. These collisions can result in ion unfolding, termed collisional induced unfolding (CIU), or dissociation, termed collisional induced dissociation (CID).

After the ions leave the trap cell, they move onto the IMS cell. If the IMS cell is off, this cell acts as transfer cell[16]. However, if the IMS cell is on, ions are separated based on charge (z)

and CCS[17]. **Figure 7** shows how separation in the TWIMS cell works[17]. The IMS cell is made from stacked ring ion guides (SRIG) and operating at low pressure with a drift gas (typically N_2). Initially, ions enter the IMS cell in a swarm (the "swarm" of ions is depicted as the red and blue circles). A wave is generated by a voltage pulse on two rings in the Synapt G2 (one ring on the Synapt G1) then the pulse is moved to neighbouring rings (5 spaces away) and relaxed[16, 20].

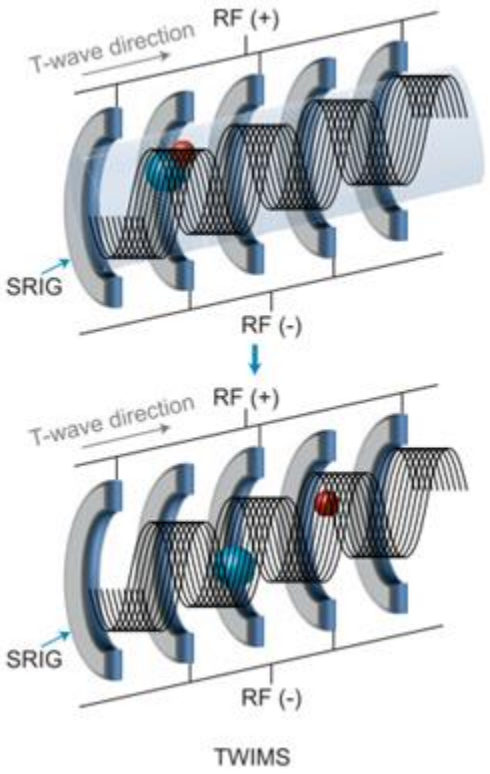

Figure 7. Separation of ions in a Travelling Wave IMS cell. (top) Ions enter the TWIMS cell in a swarm (bottom) ions then surf the wave and separate based on size and charge. Adapted from Lanucara *et al.* (2014)[17].

The ions then "surf" the wave as ions with lower mobility slip over the wave while ions with higher mobility move forward with the wave[16]. Higher mobility is determined by higher charge and a lower cross sectional area. With a higher cross sectional area, ions bump into the drift gas more than those with a lower cross sectional area and thus their mobility is impeded. It

is important to note that CCS is not the exact size of the ion but rather represents a "momentum transfer cross section (Ω)" and is the cross section based on the size of the colliding spheres[18] and their average rotational sphere. **Figure 8** shows two molecules of the same m/z which will not be distinguished based on their mass spectra alone. However, the molecule in **Figure 8a** is much more elongated which causes it to have a much larger average rotational collisional cross section than the more compact ion in **Figure 8b**. By implementing IMS, these two ions, despite their same m/z can be separated based on their drift times. The ion in **Figure 8b** will travel much faster though the IMS cell as it will be impeded by less gas molecules than the ion in **Figure 8a**.

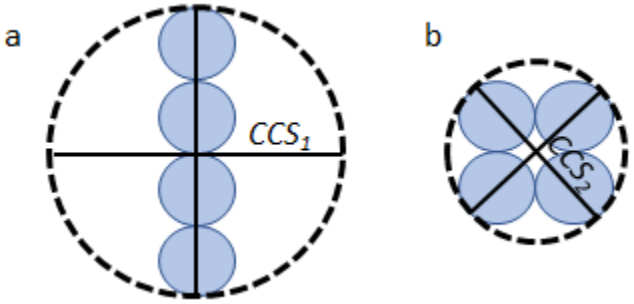

Figure 8. Diagram of average rotational collisional cross section (CCS). Ions of the same m/z will not be distinguished by their mass spectra alone. Ion (a) is more elongated and thus has a larger average rotational collisional cross section than ion (b) which is more compact and can be distinguished by their drift times in IMS.

For DTIMS CCS can be directly calculated using the Mason–Schamp equation[17, 18]:

$$CCS = \frac{3}{16}\sqrt{\frac{2\pi}{\mu k_B T}}\frac{ze}{N_o K_o} \quad \text{(Equation 5)}$$

Where e is the elementary charge, μ is the reduced mass of the ion–neutral drift gas pair, and k_B is the Boltzmann constant, and z is the charge state of the ion[17]. Equation 5 only works for low-field IMS measurements such as DTIMS but as TWIMS operated below this limit, calibrations must be used using standards of known CCS derived from DTIMS measurements[18] and

preferably using analytes of similar physical and chemical properties[17] as this can generate different calibration curves as can be seen in **Figure 9**. Detailed calibration protocols for TWIMS have been written[21, 22].

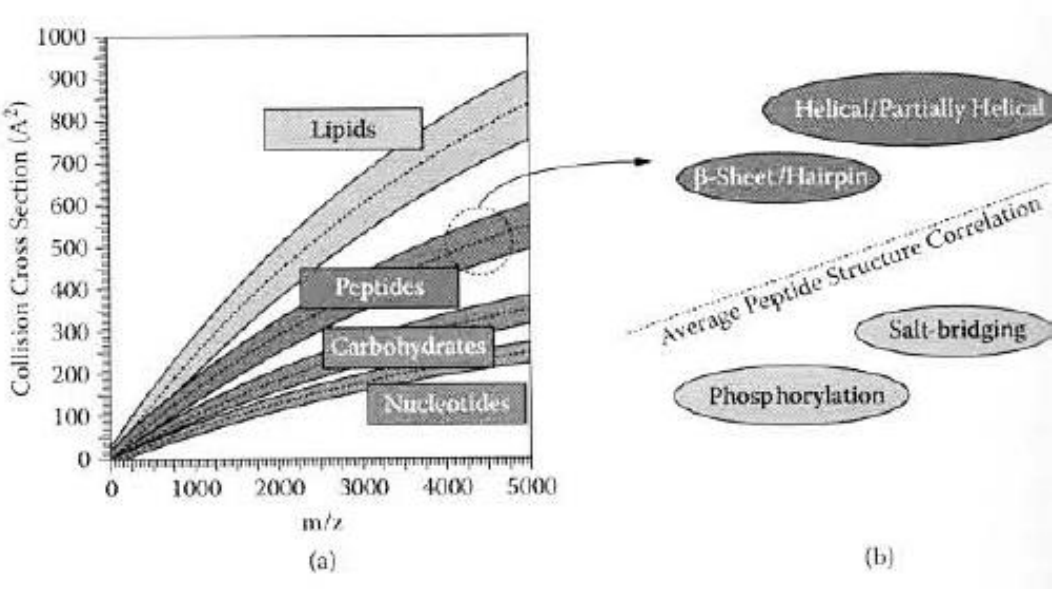

Figure 9. Collisional cross section calibration curve of varying ion species. (a) Calibration curve depicting varying ion species such as nucleotides, carbohydrates, peptides and lipids, (b) closer look into the peptide calibration curve where even variation in peptide structure cause deviations. Adapted from Ion Mobility Spectrometry, CRC press (2016)[16].

Collisional cross sections can be obtained from TWIMS however this requires a more laborious procedure. Initially, a calibration curve of standard analytes as the one in **Figure 9a** must be obtained. From here, experimental constants *m* and *x* are obtained to be used in the following formula[23]:

$$\Omega_{corrected} = m{t_d}^x \quad \text{(Equation 6)}$$

Where t_d is the drift time to which a correction factor can be applied to correct for the non-linearity effect of TWIMS[23]:

$$t_d' = t_d - \frac{EDC \times \sqrt{\frac{m}{z}}}{1000} \quad \text{(Equation 7)}$$

Where EDC is the Enhanced Duty Cycle constant that can be found in experimental parameters. Once t_d, m, and x are known, $\Omega_{corrected}$ can be calculated[23] in $Å^2Da^{-1}C^{-1}$:

$$\Omega_{corrected} = \frac{\Omega}{ez\sqrt{\frac{1}{m_{ion}} + \frac{1}{m_{buffer}}}} \quad \text{(Equation 7)}$$

Where m_{ion} and m_{buffer} is the mass of the ion and buffer gas (usually N_2) in the IMS cell. Equation 7 can then be rearranged to find Ω.

Finally, once the ions pass through the IMS cell, they move onto the third travelling wave cell called the transfer cell where they can undergo post-IMS fragmentation or be directly transferred and focused into the TOF for mass analysis[16].

1.2.2. Differential Mobility Spectrometry

DMS and FAIMS are terms that can sometimes be used interchangeably, and both employ the same method of ion separation with the only difference being that DMS is a planar cell consisting of two flat parallel electrodes, while FAIMS is a cylindrical cell. In 2010, a Toronto based company, AB SCIEX, introduced a DMS cell attachment to their mass spectrometers which they named SelexION Technology[16]. As the DMS cell attaches to the outside of the mass spectrometer, before the orifice, this means that unlike in TWIMS where mobility is done under low pressure, ions in the DMS cell are separated at ambient pressure conditions.

Figure 10 shows how ions are separated in the DMS cell. Ions from the ESI source enter the DMS cell. In the DMS cell ions undergo separation based on an applied waveform that contains a quick high-field pulse followed by a longer period of a low-field pulse which is also known as the separation voltage. Based on the ions alpha function, they will either make their way to the top or bottom electrode. To select for certain ions to get through to the mass spectrometer, a compensation voltage (CoV) is applied to the DMS cell to counteract an ions trajectory to an electrode. By being selective as to which ions can enter the mass spectrometer, DMS acts as a pre-filter device which is analogous to using liquid chromatography before MS analysis with a 15x time saving and reducing chemical background noise[24].

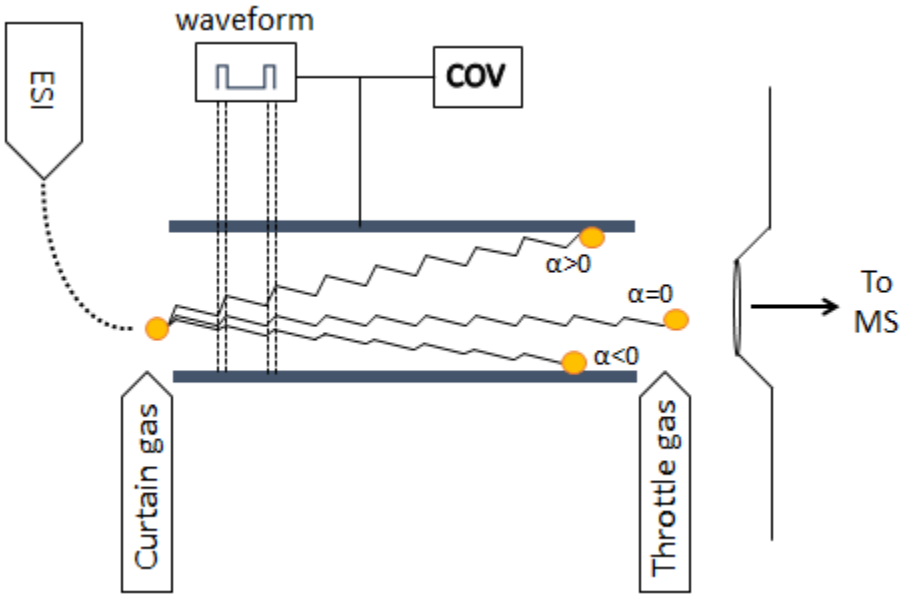

Figure 10. Ion Separation in a Differential Mobility cell. An ion enters the differential mobility cell that has a waveform applied to it with a short period of high voltage followed by a longer period of low voltage. Without any corrections ions will make their way to the top or bottom of the DMS cell. However a compensation voltage (COV) is a DC voltage that is applied to direct certain ions through into the mass spectrometer. Modifier gasses can be introduced in the curtain gas and throttle gas. Typically, N_2 gas is used when no polar modifier gases are used.

1.2.2.1 Chemical Gas Modifiers in DMS

An added benefit of the DMS cell is that different gasses can be introduced in the front of the DMS cell though the curtain gas, or the back of the DMS cell though the throttle gas. This allows for an additional separation parameter as, unlike in TWIMS, ions in the DMS cell are also separated based on their chemical properties and their ability to interact with modifier gasses. Modifier gases with different polarities such as isopropanol, methanol, and acetonitrile can be introduced alone or in combination and based on the interaction with the ions can produce high-levels of separation. **Figure 11** shows the effectiveness of adding a modifier gas. Without modifier gas, all ions were appearing to be clustering around the same CoV and showing no apparent separation even at higher separation voltages. However, as modifier gas was introduced, 5 distinct groupings appear allowing for high-levels of separation by changing CoV transmissions of ions.

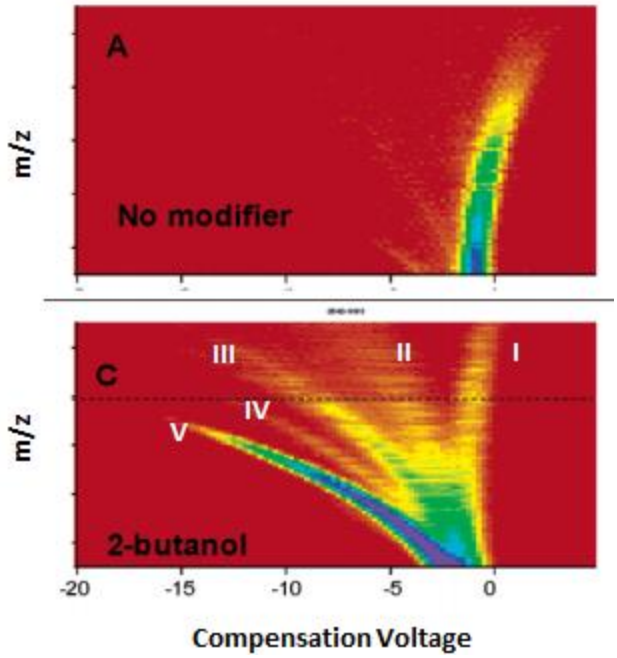

Figure 11. DMS separation of angiotensin and neurotensin. (top) DMS profile of the two compounds with no modifier present. Both compound appear at the same CoV and are therefore inseparable. (bottom) DMS profile with 2-butanol as a modifier. The spectrum separates into five distinct groups (labeled I-V). The additional modifier allowed for the separation of the two monomeric compounds with additional aggregation and clustering seen. Adapted from Levin *et al.* (2006)[25].

1.2.2.2. Deuterium as a Chemical Modifier

Deuterated gas can also be added to the DMS cell. This allows for the use of gas-phase hydrogen deuterium exchange which can also be used to help determine gas-phase protein structure information and dynamics. A recent paper published by Zhu *et al*. (2016), showed using gas-phase HDX that protein with certain structural profiles, show a distinct deuterium uptake as a function of CoV. By adding deuterated ammonium into the DMS cell from the curtain gas, they were able to show that folded proteins exhibit a linear deuterium uptake as a function of their CoV (**Figure 12a**), while unfolded protein do not show any change in deuterium uptake (**Figure 12c**)[26]. Partially folded proteins showed a slight increase in uptake, followed by an inflection point where there was no further change in uptake[26].

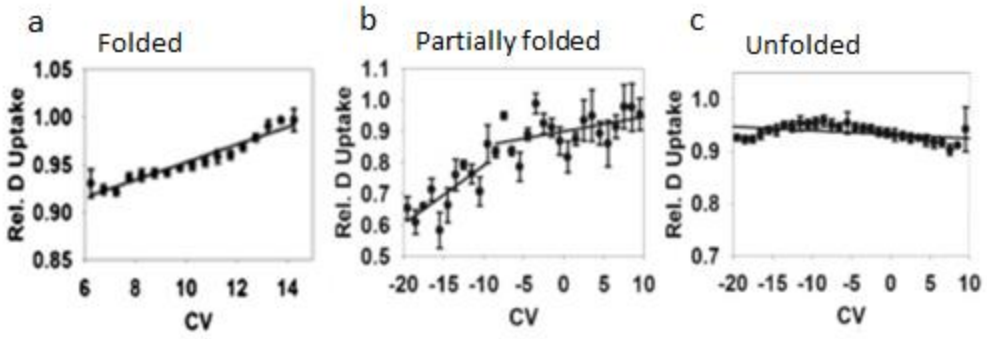

Figure 12. Deuterium uptake in the gas phase as a function of compensation voltage (CV). Gas-phase HDX-CoV profiles for folded (cytochrome C) (a), partially folded (phosphorylated tau) (b), and unfolded (tau) (c) protein ions. Adapted from Zhu *et al* (2016)[26].

A quick pulse-labeling of deuterium can also be done if deuterium is introduced through the throttle gas. This should no result in a CoV change of the ion as it has already travelled through the DMS cell unaltered by modifier gases from the throttle, and just before it enters the mass spectrometer, it can be quickly labeled with deuterium which can lead to the detection of highly dynamic or highly exposed regions of the peptide ion. This quick pulse-label has the potential to provide more sensitive information about structural protein changes compared to adding deuterium from the curtain gas, as the curtain gas is present though the entire DMS cell which can result in detecting an average structural change rather than quick changes in conformational dynamics.

1.3 Protein Stability

Stability is an important critical quality attribute of pharmaceutical drugs. It determines a drugs shelf-life and how changes in temperature such as during transportation or improper storage may affect a drugs expiry date. Drugs that contain protein as their main component need to be more cautious in determining the stability. Proteins are biological and can be very susceptible to aggregation or decomposition. Therefore, monitoring the stability of the protein in almost all aspects of the formulation process and further storage is a major aspect in bio-pharmaceuticals.

1.3.1. Long Term Solution Phase Stability Assays

Currently, there are many solution-phase techniques that can be performed over-time to determine a protein's stability. Methods such as Dynamic Light Scattering, Size Exclusion Chromatography, UV-Visible fluorescence are all methods to determine particle size[27]. Furthermore techniques such as Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) can monitor protein degradation in denaturing and reducing conditions. Implementing all these techniques, aggregation and decomposition of a protein can be measured over a period of time in different formulations conditions (such as different buffers or excipients) and in different storage temperatures (freeze-thawing, 4 °C, 25°C, 37 °C and upwards of 65 °C for accelerated stability testing)[28]. However, this real-time stability testing may take upwards of years and requires lots of sample, especially if multiple formulation conditions are tested. Thus the need for quick and high-throughput methods for protein stability determination is desired.

1.3.2. Differential Scanning Fluorimetry – A Short Term High-Through Put Solution Phase Stability Assay

In order to gain a fair sense of how certain formulation components may affect the stability of the protein, rapid solution phase heating assays to monitor protein unfolding are commonly used. Two of the most widely used methods are Differential Scanning Calorimetry (DSC) and Differential Scanning Fluorimetry (DSF). Both these techniques use a heat ramp typically ranging from room temperature to ~100 °C and measure the melting temperature (T_m) of the protein which indicates the temperature at which half the protein population is unfolded. Occasionally, multiple T_ms of a protein may be observed as the protein undergoes multiple stages of unfolding. DSC has been known to be the industry standard in determining T_m as it has been reported to provide better data resolution, and does not rely on external tags or fluorescent dyes

and only relies of the difference of heat flow as the protein unfolds[29]. However, the use of DSF gained increasing popularity. In DSF, a dye is used to measure fluorescence increase as the protein unfolds. Many dyes can be used but SYPRO Orange has become a common dye in many experiments due to its chemical stability, high fluorescence quantum yield, and its excitation and emission spectra are compatible with many RT-PCR instruments. To obtain T_ms from DSF, the negative derivative of the florescence output, which is typically sigmoidal, is taken and the minimum peaks are determined as T_ms of the protein as this reflects an inflection point in the fluorescence curve. **Figure 13** shows a typical workflow of a DSF experiment.

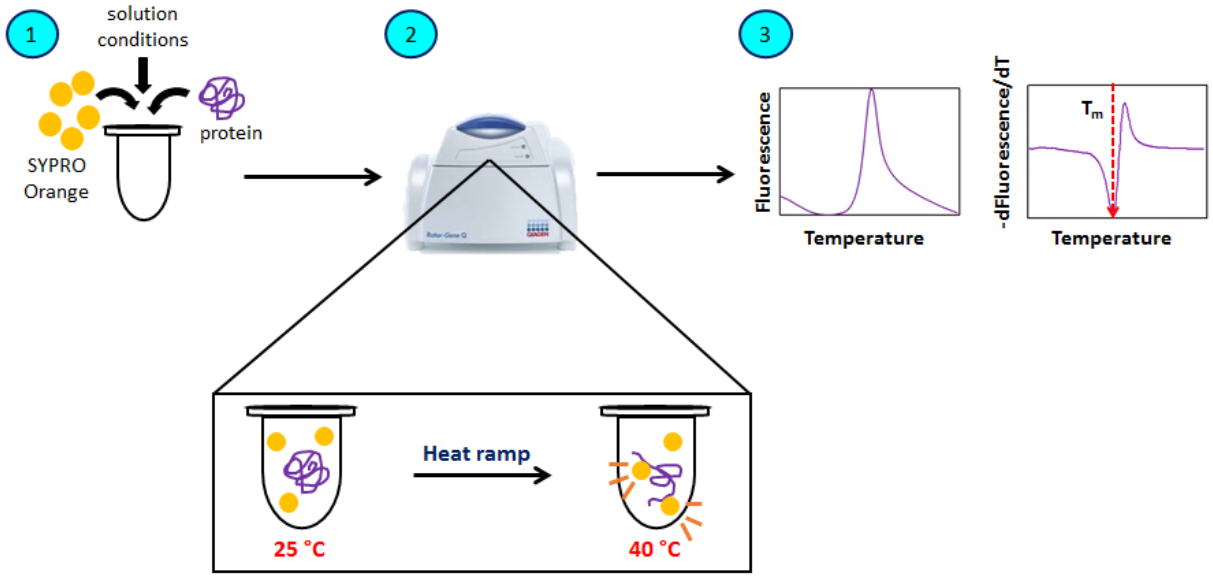

Figure 13. Typical DSF workflow schematic. (1) Protein along with the conditions being tested in mixed with Sypro Orange. (2) The solution is heated in an RT-PCT and as the protein unfolds, the dye bind to the hydrophobic areas of the protein which allows it to fluoresce. (3) Fluorescence is collected as a function of temperature and the negative derivative is taken to find the T_m. The lowest point on the graph indicated the T_m which is indicated by a dashed red arrow. At this point 50 % of the protein is unfolded.

Many experiments have shown that DSF and DSC data correlate. IgG in six different buffer conditions was tested by He and colleagues (2009), and they showed that DSF and DSC data agree as they display the same relative trends and very similar T_m values, with DSF T_m

values being only slightly higher than DSC[30]. Another study compared micro-DSC, conventional DSC, and DSF and showed that all gave similar trends with micro-DSC and DSF agreeing in T_m data while conventional DSC gave higher T_m values in this study[31]. Although DSF and DSC may not correlate to the exact degree Celsius, their trends remain the same and T_m values obtained are similar and quite close to each other, showing that DSC and DSF are comparable. DSF provides major advantages over DSC in that DSF requires less sample, less time, and can be implemented in a high-though put fashion analyzing 96-well plates. Thus, DSF is ideal for optimization studies to screen multiple conditions, including ligand binding and buffer composition, and can be used to get fast and reliable solution phase stability information in varying conditions.

1.3.3. Collisional Induced Unfolding – A Gas Phase Stability Assay

The use of the trap cell as a collision cell before IMS analysis has gained much popularity for analyzing protein structure and relative gas-phase stability. Effectively, protein ions are being heated in the trap cell which leads to unfolding or disassociation[20]. This unfolding of a protein ion can be monitored in the IMS cell (**Figure 14**) with the folded ion species having lower drift times and as the ion unfolds longer drift times are recorded. Finally, the drift time of the protein ion is then plotted against the collision voltage in the trap cell and a CIU fingerprint plot is created. Transition states in the fingerprint plot, where the drift time is seen to jump, indicate the change of the protein ion going from one conformational family to the next and shows the presence of gas-phase intermediate which are stable on a millisecond time scale[32]. The stability of the ion can be determined in different conditions by monitoring the CV at which transitions from one intermediate to the next occur. If the transition occurs at a higher CV for ion A than for ion B, this indicates that more energy was needed in ion A to continue to unfold the protein than in ion B, thus the protein ion A is more stable.

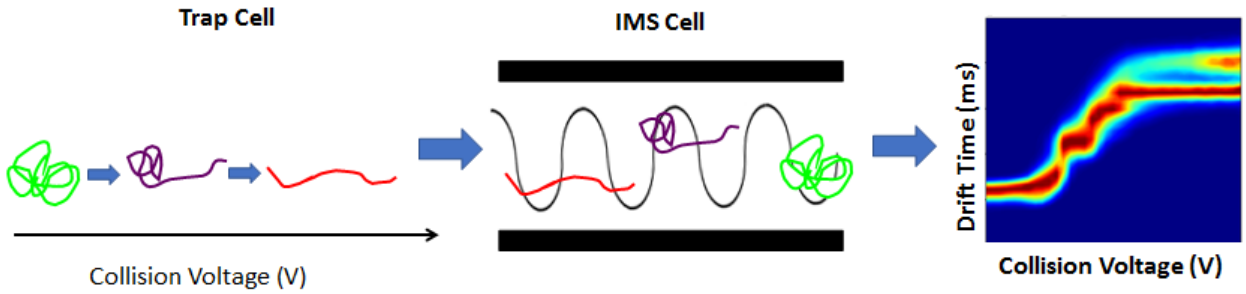

Figure 14. Experimental Schematic for Collision Induced Unfolding of a protein. (left) a protein is unfolded in the trap cell by increasing collisional energy. The folded protein is represented in green, intermediate stages of protein unfolding is represented in purple and completely unfolded protein is represented in red. (middle) drift times of the ion is collected in the IMS cell with the folded protein ion having lower drift times and as it unfolds, obtains increasingly larger drift times. These drift times are then plotted as a function of the collisional energy in the trap cell and a CIU fingerprint plot is obtained.

Many studies have already been done comparing the gas-phase stabilities of proteins and highlighting the effectiveness and usefulness of CIU as a relative gas-phase stability technique. CIU has been used to compare how protein-ligand interactions affect stability[33], how using different salt additives can affect gas-phase stability[34, 35], and how different disulphide and post-translational modifications can effect gas-phase stability of antibodies[36]. Although not yet developed, a high-throughput approach may be a viable future for CIU[37]. By selecting a specific CV range that indicates stability changes between different protein conditions, a CV ramp can be minimized to a specific transition change thus minimizing the time needed to analyze each protein and condition. The use of auto-samplers and automated methods that currently exist can help develop this high-throughput screening method for CIU.

1.3.4. DMS-HDX for Probing Gas-Phase Stability

As described in section 1.2.2.2, gas-phase HDX using DMS led to different deuterium uptake profiles as a function of CoV for different protein conformations. This raises a question as to whether DMS-HDX can be used to test stability of proteins. Folded proteins show a linear trend

or a “sloped” profile of deuterium with CoV, and unfolded proteins show no correlation or a “flat” profile[26]. The question arises if proteins of different stabilities will show these same trends but at varying degrees. For example, more stable proteins might behave more similarly to the folded proteins and exhibit a “sloped” profile, while less stable proteins might behave more similarly to the unfolded proteins and exhibit a more “flat” profile. Less stable proteins may also behave as partially-folded proteins do and exhibit a “sloped” profile in the beginning followed by an inflection point where a “flat” profile is reached. In this case, perhaps differences in the point of inflection can lead to determining the relative gas-phase stability of proteins.

The DMS cell can also be operated at varying temperatures. By increasing the temperature in the DMS cell, a “thermal ramping” experiment such as DSF can be performed and protein unfolding can be determined by gas-phase HDX uptake. This heat ramping method will also be analogous to the collision energy ramp in IMS as ions are heated with the increasing input of energy. As proteins unfold due to higher temperatures, and increase in HDX uptake can be expected. More stable proteins should unfold at slower rates than unstable ones and therefore might show a slower rate of deuterium uptake.

1.4 Protein Dynamics

1.4.1. Hydrogen Deuterium Exchange Mass Spectrometry to Probe Solution Phase Protein Dynamics

Proteins are non-static biological molecules and are constantly moving and flexing which helps their biological function[38, 39]. Biochemical techniques such as X-ray crystallography can give atomic-level resolution but they present proteins as stagnant while other biochemical techniques such as nuclear magnetic resonance (NMR), can be powerful but require high concentrations and are difficult to perform on large protein or protein complexes[38]. To circumvent such problems

that other techniques present, a relatively recent development of hydrogen deuterium exchange coupled to mass spectrometry has revolutionized how protein dynamics can be characterized. The fundamental basis of HDX comes from the exchange of the hydrogen (H) atom on chemically labile groups such as N-H, O-H, and S-H with the surrounding water[38]. By introducing deuterated water (D_2O), this exchange of H with D can now be detected on the mass spectrometer as a mass shift. Typically, for HDX-MS workflows, only the N-H exchange on the amide backbone hydrogen is of concern as this is present on every amino acid residue and therefore changes throughout the entire protein structure can be monitored[38]. To obtain HDX data for backbone amide hydrogens only, a rapid acid-quench to a pH of 2.5 at which the rate of H to D conversion is at a minimum[40]. Here, all the hydrogens on the backbone that exchanges with deuterium become locked, and no information is obtained from the labile side-chain hydrogens because of high back-exchange rates[38, 40] and therefore any hydrogens that were exchanged to deuterium, revert back to hydrogen. To minimize any back-exchange of the amide backbone hydrogens before analysis, quenching is usually performed at 0 °C[40].

Deuteration of a protein measures protein dynamics by revealing the more "flexible" or more "solvent exposed" part of the protein. The labile hydrogens that are more solvent exposed have a higher chance of deuterium exchange than those that are buried within the protein core. As the protein undergoes its regular "breathing" motions in solution by opening and closing, some of the inner parts of the protein become solvent exposed. If the protein spends more time in an "open" configuration rather than a "closed" one, or the rate at which it opens and closes is high, then more areas of the protein become exposed to solvent which results in more deuteration[38]. In comparison, proteins that are typically more compact and spend more time in a closed configuration, do not experience as much deuterium uptake and thus will show less

deuteration and can be shown to be less dynamic or more structured in nature. HDX can either be analyzed by native mass spectrometry to obtain "global" HDX information, as to how much deuterium the protein picked up on average. Or, the protein can then be digested using enzymes such as pepsin to obtain more local, peptide-level, HDX information to pin-point more precise locations on the protein that experienced deuterium exchange.

1.4.1.1. Time Resolved Hydrogen Deuterium Exchange

Typical HDX-MS workflows have a time scale of minutes to hours and can last for days. Early work by Dr. Lars Konnermann and then further development by Dr. Derek Wilson, gave rise to a technique called time-resolved-ESI (TRESI)[41] HDX where HDX can be monitored on a millisecond (ms) timescale. Using TRESI-HDX useful information about protein folding intermediates[42], dynamics of unstructured proteins[43], and allostery[44–46] can be obtained as these motions occur quickly and can be missed using conventional time-scales. A typical TRESI set-up can be seen in **Figure 15**. Here, protein and deuterium are flown through two glass capillaries at certain flow rates. Flow rates can vary depending on the protein to deuterium ratio and the final mixing time scales desired. The two lines then meet at a tri-union, where the protein line continues through the tri-union and into an outer metal ESI capillary and the deuterium flows from the glass capillary, into the metal outer capillary (highlighted as dotted oval). Mixing of protein with deuterium occurs in the metal capillary (highlighted as solid oval). The protein from the glass capillary, which is sealed at the end, flows into the metal capillary though a notch typically cut 4 mm from the end with a laser. The protein is then mixed with the deuterium solution flowing from the outside and the mixing time scale is determined by the inner diameter of the metal capillary, the outer diameter of the glass capillary, notch cut position, distance of notch from the end of the metal capillary, and flow rates. Longer time points can be obtained by

pulling back the glass capillary thus increasing the time the protein interacts with the deuterium. The metal capillary can then be used for ESI for global HDX analysis by MS.

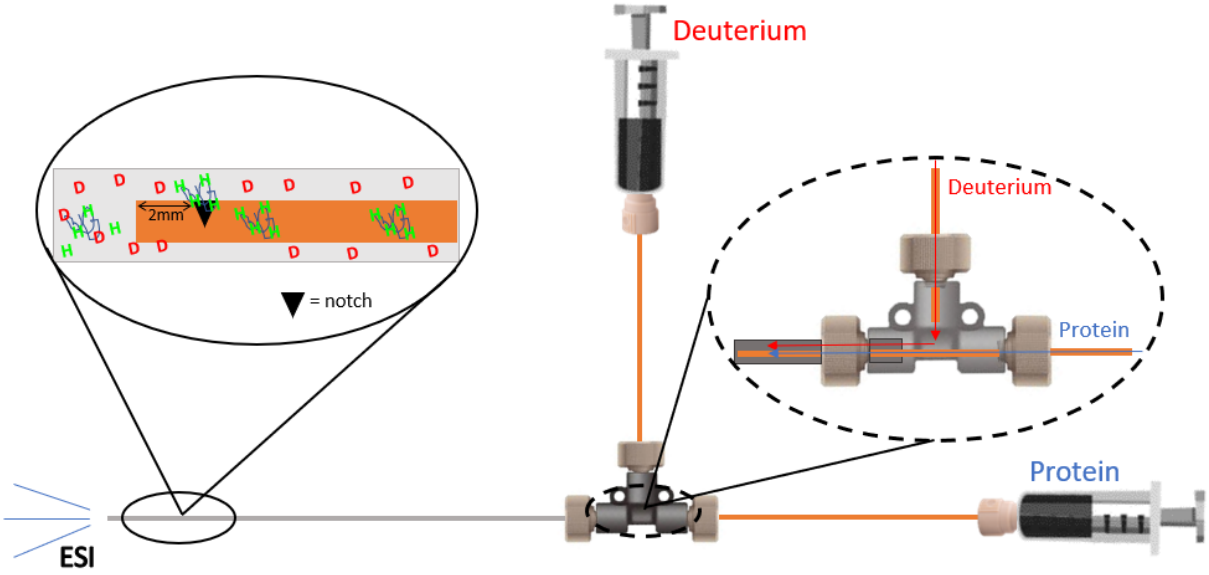

Figure 15. TRESI experimental set-up. Protein and deuterium flow into a tri-union (dotted oval inlet) where the protein line feeds into a metal capillary while the deuterium flows into the metal capillary, outside the protein line. The end of the protein line is sealed with a laser (solid oval inlet) with a notch cut 4 mm from the end (although this distance may be altered). The proteins flow from inside the protein line, through the notch, and out into the metal capillary where they mix with deuterium. The reaction time is dictated by the distance the notch is from the end of the metal capillary, inner diameter of the metal ESI capillary, the outer diameter of the protein line, and flow rates. The metal capillary acts as an ESI line where the deuterated protein mixture is sprayed into the mass spectrometry for HDX analysis.

Figure 15 will result in global non-quenched HDX analysis, where side-chain exchange will also be observed. However, simple modification can be done as seen in **Figure 16** to capture only back-bone amide hydrogens. Adding a T-mixer and an acid line to the end of the metal capillary will quench the reaction in the mixing chamber and a second metal line can then be used as an ESI source to spray the quenched solution into the mass spectrometer. A second modification (highlighted as a black dotted box) can be added after the quenching chamber. Here, a small pepsin reaction chamber can be added to digest the intact protein into peptides and

local HDX analysis can be performed. In this thesis, a non-quenched global set up as seen in **Figure 15** is performed.

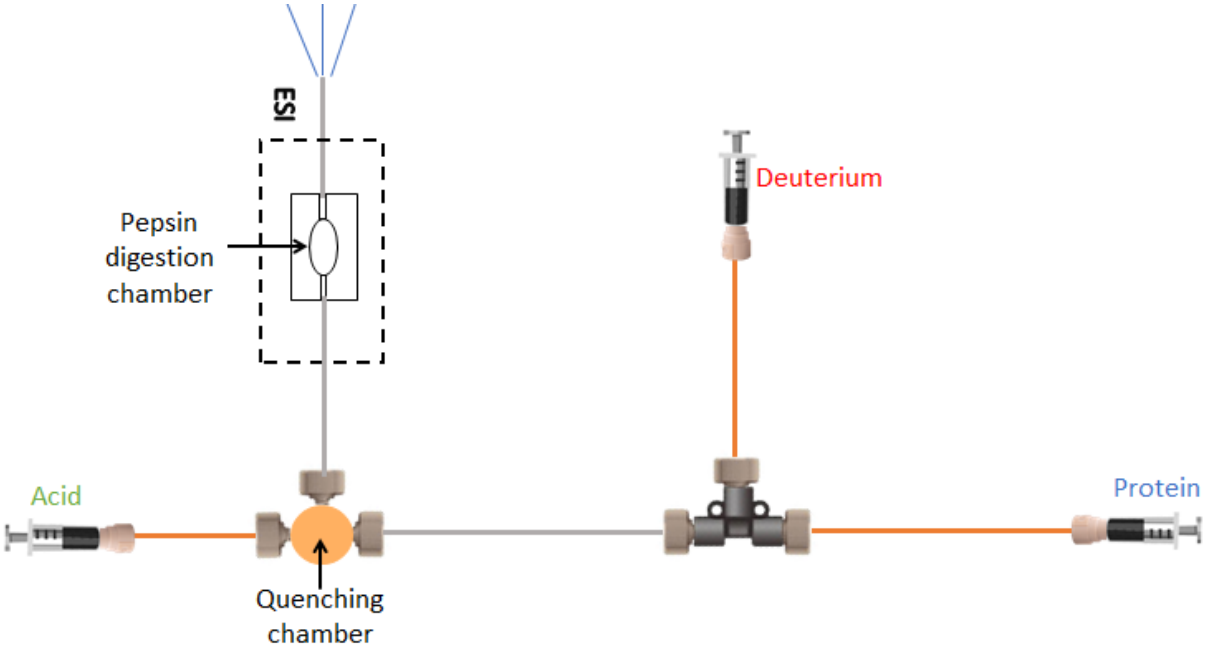

Figure 16. TRESI set-up with acid quenching and a protein digestion chamber. The TRESI set-up from Figure 15 can be connected to a T-mixer to which an acid line is attached that results in quenching of the HDX reaction. This allows for the detection of back-bone amide hyrgoen exchange only. The quenched solution may be sprayed directly into the mass spectrometer for intact global protein analysis or, an additional protein digestion chamber may be added (dotted rectangle) to allow for digestion of the protein for peptide-level HDX analysis.

1.4.2. Hydrogen Deuterium Exchange to Probe Gas-Phase Protein Dynamics

Although not a widely employed technique due to its relatively new emergence, the use of gas-phase HDX is being studied to probe protein conformations. A little about gas-phase HDX has been described in section 1.2.2.2, using a DMS cell, however, other methods employ drift tubes and ion traps. Relatively recent studies by Dr. Kasper Rand and Dr. John R. Engen's group showed that gas-phase can be done using a tri-wave[47–49]. ND_3 gas can be added into the trap, instead of the typical Argon gas, for HDX labeling with the added bonus of using IMS as a second conformational probing technique[49]. Through gas-phase HDX, multiple gas-phase conformers can be detected that are sometimes difficult to detect by IMS alone[49]. Furthermore,

probing these gas-phase conformers can give insight into protein misfolding pathways or determining intermediates[49]. Rand *et al.* (2009), have shown that using ND_3 gas over $D_2O_{(g)}$ of $CH_3OD_{(g)}$ is more representative of probing surface accessible labile hydrogens and protein conformations due to its labeling mechanism[49]. Sub-millisecond time scales can also be achieved though pressure control and thus can also lead to time-resolved conformational probing[49]. In the gas-phase, there is no quenching and so all labile hydrogens can be subject to exchange. At the shorter time scales of gas-phase HDX, deuterium uptake data consist mainly of the exchange of side-chain hydrogens as the back-bone amide hydrogens exchange at a much slower rate[49]. However, due to steric hindrance, some side-chain amino acids will be more easily exchanged than others, which will give rise to the deuteration differences[49]. Although much more testing will be needed for gas-phase HDX to become more accepted and more wide-spread, experiments have shown its potential.

1.5 Research Objectives

Although many studies have focused on showing that solution phase structure is maintained in the gas phase when analyzed by mass spectrometry, not many studies have focused on showing that a proteins solution phase stability is retained in the gas phase. Despite CIU being a widely used technique to compare relative gas phase stability between proteins such as those that were post-translationally modified or bound to a ligand, not many studies have correlated their results to solution phase data. In this report, gas and solution phase stability will be compared using a model system protein, wild type (WT) pyruvate kinase (PK) from *Escherichia coli* type 1, and four of its point mutants: Lys404Ala (K404A), Thr405Ala (T405A), Gln408Ala (Q408A), and Leu411Glu (L411E).

Pyruvate kinase is an enzyme that catalyzes a highly exergonic reaction of PEP to pyruvate in glycolysis and thus needs to be regulated through phosphorylation and cooperativity mechanisms[50]. This regulation requires modifications on different sites of the protein which is why the protein can only be function in a tetrameric form. The tetramer of pyruvate kinase consists of four identical monomers consisting of three domains: the allosteric C domain, the TIM barrel A domain, and the Lid B domain (**Figure 17a**)[45]. All of the point mutations occur in the Cα4'-helix that are involved in hydrogen bonding between the A and C domains of the monomer. The tetramer itself is held together by ionic and salt bridge interaction at the A/A' and C/C' interfaces and has been shown to be highly stable (**Figure 17b**)[45]. As the tetrameric complex is highly stable, we focus on monitoring changes in unfolding of the tetramer species as a result of the different monomer stabilities within the protein complex. Two ion mobility methods will be used to try and assess the gas-phase stability of the different protein variants and results will be compared to their solution phase stability data obtained using DSF. The first ion mobility method will be employing CIU and the second will be employing DMS and gas-phase HDX. Furthermore, dynamics of these protein variants will be compared to their stability to see if there is correlation. In both the gas and solution phases, heat ramps of the protein will be performed and the degree of unfolding will act as the stability indicator of all the protein variants. Thus stability in our work is dependent on the temperature at which the protein unfolds. If being able to show that a proteins stability data is maintained in the gas-phase, then mass spectrometry can be used a fast and effective stability indicating assay.

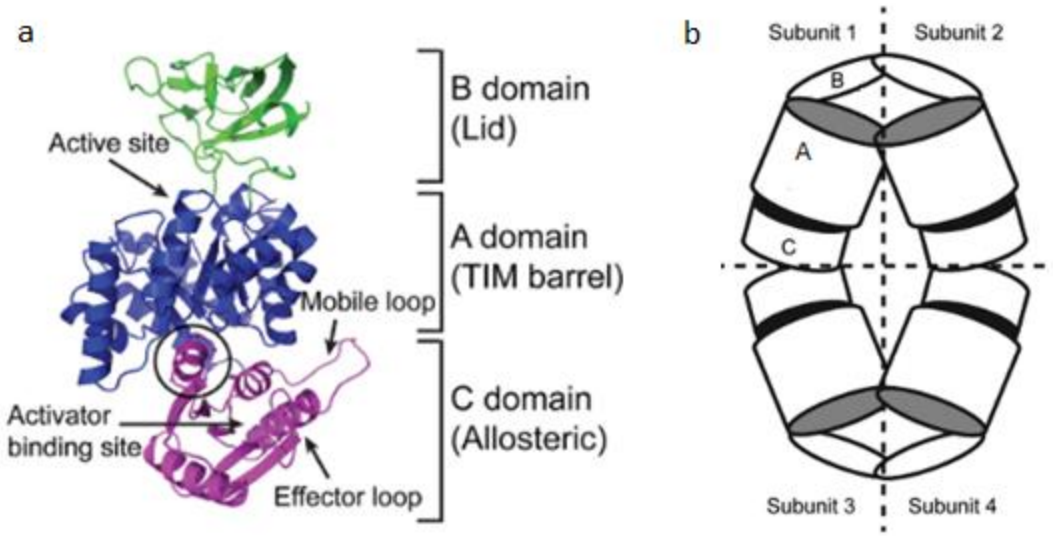

Figure 17. *E. coli* Pyruvate Kinase monomer subunit and assembly. (a) Three subunits comprising the Pyruvate Kinase monomer. (b) Assembly of the monomeric units into a tetramer. Adapted from Dobson *et al.* (2016)[45].

Chapter 2: Experimental Methods

2.1 Materials

Chemicals were purchased from Sigma-Aldrich unless otherwise specified. Buffers were made in-house using ultrapure water made in-house using a Millipore Milli-Q Advantage A10 system. Empty pET30ΔSE expression vector, wild type *pykF* gene in a pBluescript II KS+ vector, and plasmids of the recombinant point mutants of *pykF* in pET30ΔSE expression vectors were donated by Professor Renwick Dobson (University of Canterbury).

2.2 Cloning of Wild Type *pykF* gene

2.2.1. Preparation of Competent Cells and Bacterial Transformation BL21/DH5α cells

Both DH5α and BL21(DE3) cells were transformed using the same protocol. Competent cells were made by inoculating 0.5 mL of an overnight (O/N) culture of *E. coli* Bl21 cells into 50 mL of Luria-Bertani (LB) medium and letting cells grow at 37 ^{o}C at 200 rpm until an OD_{600} of 0.2-0.4 was obtained. Cells were pelleted and suspended in 5 mL of cold 100 mM $CaCl_2$ and incubated on ice for 30 min. Cells were pelleted again and re-suspended in 0.5 mL of 100 mM $CaCl_2$ and incubated on ice for an additional 20 min. 0.5 mL of cold 50% glycerol was added and 200 µL aliquots were made and stored at -80 ^{o}C. To transform the cells, a ratio of 1 ng: 1 µL of plasmid to competent cell stock was used. Plasmids of interest were incubated with competent cells on ice for 20 min, then heat shocked for 90 s at 42 ^{o}C and quickly placed back on ice for 5 min. 800 µL of fresh LB broth was added and the cells were allowed to grow for 1 hr at 37 ^{o}C shaking. 100 µL of cells were then plated on LB plates as a viability control and remaining cells were pelleted, re-suspended in 100 µL of LB, and plated on LB plated containing the appropriate antibiotic, 50 µg/mL Kanamycin (Kan) for plasmids containing pET30ΔSE expression vectors

and 100 µg/mL ampicillin (Amp) for plasmids containing pBluescript II KS+ vector. An overnight culture was then done by selecting a single colony from the antibiotic plates, and inoculated in 5mL of LB with appropriate antibiotic. Cells were pelleted the next day and stored in 25% glycerol with appropriate antibiotic at -80 °C. DH5α cells were used for long term storage and mini-prep of the gene for future use and BL21(DE3) cells were used for expression purposes.

2.2.2. Plasmid Purification Using Mini-Prep

Empty pET30ΔSE expression vector and pBluescript II KS+ expression vector containing the WT *pykF* gene were transformed into DH5α cells. Plasmids were purified using GeneJET Plasmid Miniprep Kit (ThermoFisher). A detailed step-wise procedure is available in a PDF version through their website. All centrifugation steps were done at > 12,000g. 5mL of O/N culture was spun down for 2 min and cell pellet was re-suspended in 250 µL of Resuspension Solution containing RNase. 250 µL of Lysis buffer was added and mixed followed quickly by 350 µL of Neutralization Solution and mixed to ensure neutralization reaction was complete. Percipitate was removed by centrifugation for 5 min and supernatant was removed and placed into a Thermo Scientific GeneJET Spin Column. The column was centrifuged for 1 min and flow though was discarded. 500 µL of Wash Solution containing 96% ethanol was added and centrifuged again for 1 min. The wash step was repeated once more and an additional spin of 1 min was done to remove any remaining wash buffer. Finally, 50 µL of Elution Buffer was added and incubated for 2 min at room temperature, then centrifuged for 1 min and the final solution was collected. DNA concentration was determined using NanoDrop2000 UV-Vis Spectrometer.

2.2.3. Restriction Digests

In order to extract the WT *pykF* gene from the pBluescript II KS+ vector, SacI and HindIII were used to cut the gene out. To insert the gene into the empty pET30ΔSE vector, the same restriction enzymes were used to cut the multiple cloning site. For 1 ug of both the empty pET30ΔSE vector and the pBluescript II KS+ vector containing *pykF*, 0.5 µL of Hind III (20, 000 units/mL) and 0.5 µL of Sac I (20, 000 units/mL) were added to 5 µL of 10X Buffer 2.1 and Milli-Q water to top-up to 50 µL. The reaction was mixed and incubated at 37 °C for 1 hr. The restriction enzymes were then heat inactivated at 80 °C for 20 min.

2.2.2. Gel Electrophoresis

After restriction digestion, products of the reaction were run on a 1.2 % Agarose gel was run to determine if the restriction digestion was successful and for subsequent DNA extraction. 1.2 g of Agarose was added to 100 mL of 1X Tris-Acetate EDTA (TAE) buffer. Once the agarose was dissolved by bringing the solution to a gentle boil, 2 µL of 10 mg/mL of Ethidium Bromide was added for DNA visualization and stirred. The gel was cast and once solidified 0.2 volumes of loading dye was added to 80 µL of sample and a DNA ladder was run to visualize DNA length. The gel was run for 65 min at 50V.

2.2.3. DNA Gel Extraction

Once the gel was run, the products were first visualized with a gel imager using UV light to identify if the WT gene was successfully cut out and if the empty pET30ΔSE vector was also cut open to allow for gene insertion. Once the WT gene was identified, gel DNA gene extraction was done using Thermo Scientific GeneJet Gel Extraction Kit. A detailed procedure is available online of their website. The DNA band was identified through visualization with a UV lamp. Visualization time was minimized to avoid damage to the DNA. The band was then excised using a clean razor

blade and placed in a pre-weight test tube. A 1:1 volume of Binding buffer was added to dissolve the gel and promote DNA binding to a silica membrane. The mixture was incubated at 55 °C for 10 min to allow for the gel to completely dissolve. Up to 800 µL of the dissolved mixture was transferred over to a GeneJET purification column and centrifuged for 1 min (>12, 000 g). 700 µL of wash buffer containing ethanol was added to the column and centrifuged for 1 min (>12, 000 g) and the flow through was discarded. To ensure complete removal of the wash buffer, the column was centrifuged again for 1 min. 50 µL of elution buffer was added to column and spun for 1 min. DNA concentration was determined using NanoDrop2000 UV-Vis Spectrometer.

2.2.4. Ligation

Once the WT *pykF* gene was extracted, it was added to the cut pET30ΔSE vector and ligated. A total of 100 ng of DNA (a 3:1, 5:1, and 7:1 ratio of insert:vector was done to test efficiency) was added to 2 µL of 10X T4 ligase buffer and topped to 20 µl (if needed) of Milli-Q water. Lastly, 1 µL of T4 DNA Ligase was added to the mixture. The reaction was mixed and incubated at 25 °C for 30 min. The T4 DNA ligase was then heat inactivated for 10 min at 65 °C and products were stored in 4 °C.

2.3 Expression and Purification

2.3.1. Protein Expression

0.8 mL of overnight culture of transformed *E. coli* Bl21(DE3) cells was inoculated into 0.8 L of LB broth with 50 µg/mL of Kan and incubated at 37 °C and shaking at 150 rpm until an OD_{600} of ~0.4 was reached. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added and to a final concentration of 1 mM (0.8 mL of 1 M IPTG) and induction continued for 3-4 hrs at 37 °C shaking at 150 rpm. Cells were then centrifuged at 5,000 g for 10 min at 4 °C and cell pellets

were stored at -20 °C until further use. Due to the instability of the protein variant L411E, induction was carried out at 16 °C O/N and the cells pellets were harvested the next day.

2.3.2. Purification using Ni^{2+}affinity chromatography

The induced pellet were re-suspended in 30-35 mL Nickel-binding buffer (20 mM Tris, 200 mM KCl, 30 mM imidazole, pH 8) and protease inhibitor coctail containing broad-spectrum protease inhibitors was added 1:100 (100x). Cells were sonicated for a total of 3 min (15 seconds on, 15 seconds off) at 30% amplitude. The supernatant was collected by pelleting the cells at 20,000 g for 30 – 40 min at 4 °C. The supernatant was loaded onto an immobilized Ni^{2+}-affinity gravity chromatography column (GE Healthcare) which was pre-equilibrated with the binding buffer and allowed to incubate for 30 min gently shaking at 4 °C. Unbound proteins were initially washed with ~100-150 mL of binding buffer. A second wash step using 25 – 50 mL of binding buffer containing 50 mM imidazole was done to eliminate any tightly bound unspecific proteins. Finally, a step gradient using 15mL of elution buffers (20mM Tris, 200mM KCl, 50, 75, 150, 250, and 500mM imidazole, pH 8) were performed to elute the protein. Samples of all wash and elution fraction were collected for purity determination using Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) gels. Elution fractions were collected and dialyzed overnight into storage buffer (10 mM Tris-HCl, 100 mM KCl, pH 8). Fractions that were determined to be pure (see section 2.3.4) were concentrated using a 10,000 Da Millipore centrifugal filter concentrator spinning at 2,500 g the following day to a final volume of ~10 mL. 20 µL of sample was stored for conentration determination and the rest were stored at -20 °C in 500 µL aliquots to minimize freeze-thawing.

2.3.3. Determining Protein Purity with SDS-PAGE

To determine protein purity of the elution fraction, samples at each step of the purification process were collected and run on a 12% SDS-PAGE gel along with an un-stained protein ladder. The gel was run at 40 min at a 220 V then stained for 20 min with Coomassie Brilliant Blue R-250, 50% methanol and 10% glacial acetic acid gently shaking at room temperature. The gel was de-stained using de-staining solution (40 % methanol, 10 % acetic acid) for 10 – 20 min and continued de-staining in Milli-Q water overnight.

2.3.4. Determining Protein Concentration

Protein concentration was determined using Bradford Assay. Standards of BSA in water at 0, 0.0625, 0.125, 0.25, 0.5, and 1 mg/mL were made to create the calibration curve. Protein samples were prepared using 2 or 3 dilutions within the middle of the calibration curve range. 5 µL of standard or protein sample was added in duplicate to a 96-well plate followed by 250 µL of Bradford reagent (no mixing). The plate was incubated for 5-10 min in the dark at room temperature. Sample readings were taken using a plate reader set at a wavelength of 595 nm. Protein concentrations were determined using the linear standard curve and the average reading of the protein dilutions was taken.

2.4 Native Ion Mobility Mass Spectrometry

2.4.1. Sample Preparation

1 – 2mL of sample diluted to ~10 – 15 µM 200 mM ammonium acetate using a Slide-A-Lyzer™ MINI Dialysis Device with a 10 kDa cutoff in a 50 mL falcon tube (ThermoFisher). A mini stir rod was used to allow for continuous mixing of the buffer. 200 mM ammonium acetate buffer was changed every 1.5 – 2 hrs and the buffer was replaced 3 times. A final buffer change was

done for a fourth time and left overnight. Protein concentration was measured and stored in 100 – 200 µL aliquots for future use. Samples were diluted to a final concentration of 2 – 2.5 µM in 200mM ammonium acetate for experimental use.

2.4.2. Native Ion Mobility Mass Spectrometry and Collision Induced Unfolding

2 µM protein in 200mM ammonium acetate was analyzed using a Quadrupole-Ion mobility-Time of Flight (TOF) mass spectrometer (Synapt G2-S HDMS, Waters, Milford, MA, USA). Ions were analyzing in positive ESI mode. Capillary voltage was set to 2 kV, source temperature at 120 °C, and sampling and extraction cone were both set to 50 V to reduce any in-source unfolding or fragmentation. The trap cell containing Argon gas was kept at a pressure of 3.5 × 10-2 mbar. Nitrogen gas was used in the ion mobility cell and operated at a pressure 3.2 mbar with the travelling wave height set to 40 V and wave velocity set to 800 m/s. The TOF pressure was operated at 1.6×10^{-6} mbar. For CIU experiments, collision energy in the trap cell was ramped from 20 – 200 V in 5 V increments with a 1 min scan time. Enhanced duty cycling at an m/z of 7000 was used to select for the tetrameric ions in the quadrupole for higher sensitivity during CIU experiments.

2.4.3. Data Analysis

Data was acquired using MasyLynx v4.1. CIU data was plotted and analyzed using CIUSuite[51] and CIUSuite2[52]. Ion mobility data for each charge state was extracted using TWIMExtract[53].

2.5 Differential Scanning Fluorimetry

Solution phase stability was determined using differential scanning fluorimetry (DSF). Samples were prepared to a final volume of 50 µL containing 2 µM protein in 200mM ammonium acetate and 10X SYPRO Orange. SYPOR Orange was diluted to a 500x stock from the original 5000x

stock solution using 200mM ammonium acetate. Samples were then heat ramped from 25 – 99 °C at a rate of 1 °C/min using a Rotor Gene Q RT-PCR (QIAGEN).

2.6 Global Solution-Phase Hydrogen Deuterium Exchange

2.6.1. Sample Preparation and Set-up

A TRESI mixer with an inner glass capillary, with an inner diameter of 178 µm, and a 28-gauge outer metal capillary, with an inner diameter of 151 µm was used for global HDX analysis (**Figure 15**). A VersaLaser™ was used to seal the end of the glass capillary and cut a notch at 2 mm from the end. 5 µM of protein in 200 mM ammonium was mixed with D_2O at a flow rate of 5 µL/min resulting in a 1:1 protein to deuterium ratio and a final protein concentration of 2.5 µM. Data was acquired using Quadrupole-Ion mobility-TOF mass spectrometer (Synapt G1 HDMS, Waters, Milford, MA, USA) instrument in positive ion mode and reaction times of 180 ms to 15.1 s were analyzed.

2.6.2. Data Analysis

Peak smoothing using Savitzky-Golay method, baseline correction, and centroid determination was performed using MasyLynx v4.1 software. Deuterium uptake rate curves were analyzed with GraphPad Prism v5 using a two-phase association curve fit function.

2.7 Differential Mobility Spectrometry

2.7.1. Sample Preparation

Proteins were dialysed same as in section 2.4.1. However, 50 mM ammonium acetate was used for dialysis and proteins were diluted to a final concentration of 15 µM before being aliquoted in 100 – 200 µL samples and stored in -20°C. Samples were transported in dry ice to SCIEX where they were stored in -20°C. For experimental use, samples were diluted to 0.75 µM in 5 mM

ammonium acetate (100 µL of 15 µM protein, 400 µL HPLC water, 500 µL of 5 mM ammonium acetate HPLC grade).

2.7.2. Instrument Parameters

Samples were run on a Triple TOF 6600 using a DuoSpray Ion source. Source temperature was set to 225 °C, DMS offset set at -10, capillary voltage at 4.8 kV, and initial collision energy set at 10 V. For better high-mass transmission a GeneJET injector quadrupole (Q1) was added. Instrument parameters were adjusted if necessary.

2.7.3. Global Gas-Phase Hydrogen Deuterium Exchange

2.7.3.1. CoV Ramps

To investigate stability of proteins as a relationship of HDX uptake as a function of CoV, CH_3OD was inserted through the curtain gas at 120 µL/min and spectra were collected for 30 s from CoV -12 to -4 at 0.5 V intervals.

2.7.3.2 Temperature Ramps

To investigate the difference of HDX uptake on temperature, ND_3OD was introduced into the DMS through the curtain gas at 10 µL/min and ACN (a carrier gas) was introduced at 138.1 µL/min. The DMS temperature was ramped from 150 °C to 350 °C at 50 °C intervals with a 10 min equilibration point between each temperature. Spectra were collected in triplicate for 30 s at each point at a set CV of -6 V to select for the tetramer. A quick ramp using two temperatures (150 °C and 350 °C) was done introducing ND_3OD (10 psi) and ACN from the throttle gas to see if a quick pulse-label HDX would show differences.

2.7.4. Data Analysis

Data was analyzed using PeakView® Software V 1.2.2.0. Gaussian smoothing was done with smoothing widths of 10 points and peak centroids were obtained using centroid function with intensity sum above 50%.

Chapter 3: Results and Discussion

3.1 Cloning of WT *pykF* gene into pET30ΔSE Expression Vector

Four point-mutants for *pykF* gene were provided in pET30ΔSE expression vectors while the WT *pykF* gene was provided in a pBluescript II KS+ expression vector. This meant that the mutants would be expressed with a His_6-tag while the WT protein would not have a tag. As this study is focusing on the effect of a single amino acid change on the stability of the protein while keeping the rest of the structure the same, we wanted to ensure that either all mutants had or did not have a tag. Initially, many attempts were done to purify the untagged WT protein using anion exchange chromatography followed by hydrophobic interaction chromatography and size exclusion, yet none resulted in a successful purification of the protein. Furthermore, an effort to cut the his-tag off the mutants was not very successful and the protein itself was seen to un-specifically bind to the Ni^{2+}affinity column which made separation of the tagged and un-tagged proteins somewhat difficult and a large loss in protein was observed. After many optimization attempts, it was decided that the WT gene should be cloned into the pET30ΔSE expression vector and all protein variants should contain a his-tag for consistency.

3.1.1. Optimization of Restriction Enzyme Digestions

A restriction enzyme digest using SacI and Hind III as in section 2.2.3 was done to excise the WT gene from the pBluescript II KS+ expression vector and insert it into the pET30ΔSE expression vector containing the his-tag. To test the efficiency of the restriction digest, high-fidelity (HF) restriction enzymes and different buffer conditions were tested to see which conditions are the most optimal for DNA digestion. **Figure 18** shows five different digestion conditions in two different buffers. Single digests of SacI, HindIII, and HindIII-HF were set up to see if all the restriction enzymes worked individually (lanes 2-4, and 7-9). Double digests of

SacI/Hind-III and SacI/HindIII-HF were also run to see how efficiently the gene would be cut out in each case (lanes 5-6, and 10-11). Buffer 2.1 was used for the single and double digestion in lanes 2-6 and Cutsmart buffer was used in the single and double digestion in lanes 7-11 to see which buffer conditions the restriction enzymes were most efficient. It was clearly seen that in all the single digests, in both buffer conditions, all the enzymes were able to digest the plasmid leaving a single band just above 4,000 bp as the weight of the vector and the WT DNA insert is 4,374 bp. HindIII in Cutsmart buffer (lane 8) was the only one to show only slight amounts of uncut DNA left (red box). The double digestion worked in all cases with the WT gene showing just below 1500 bp as the WT gene alone is 1413 bp alone. However, the presence of the *pykF* gene was stronger when buffer 2.1 was used compared to Cutsmart (green boxes). The high-fidelity and regular HindIII restriction enzymes performed equally showing no undigested DNA. From these results, it was established that regular HindIII, SacI and buffer 2.1 will be used for subsequent cloning purposes.

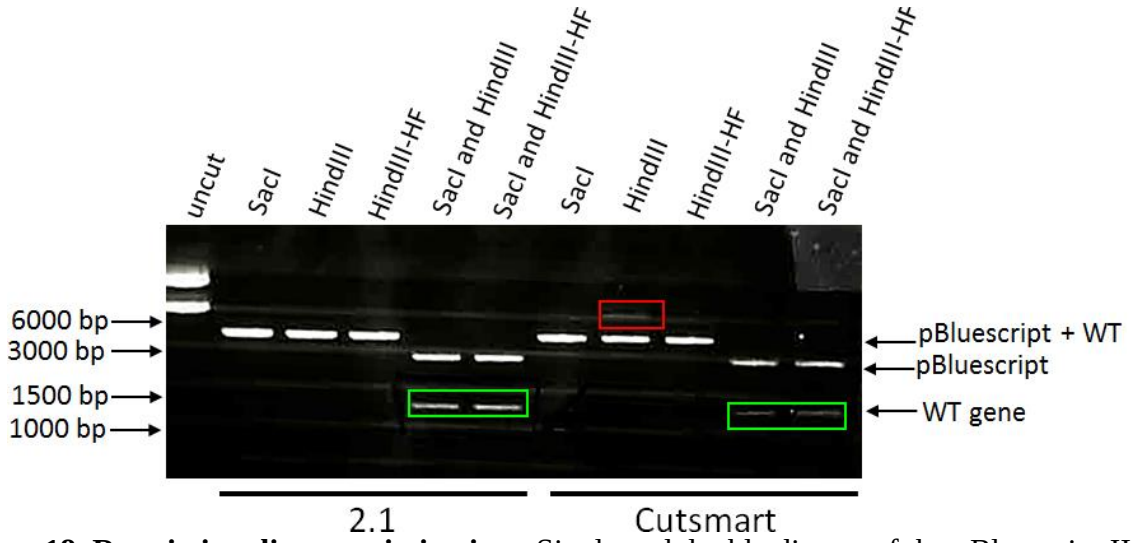

Figure 18. Restriction digest optimization. Single and double digests of the pBluescript II KS+ vector containing the WT *pykF* gene in different buffers. A control uncut plasmid was run in lane 1 and followed by single digests of SacI (lane 2 and 7), HindIII (lanes 3 and 8), HindIII-HF (lanes 4 and 9) and double digests of SacI/HindIII (lanes 5 and 10) and SacI/HindIII (lanes 6 and 11). Buffer 2.1 was used for digestion conditions in lanes 2-6 and Cutsmart buffer was used in lanes 7-11.

3.1.2. Double Digestion to Ensure Successful Gene Cloning

Using the conditions specified in the last section, the WT *pykF* gene in pBluescript II KS+ plasmid and the empty pET30ΔSE plasmid were double digested. The double digest of the pBluescript plasmid was run on an agarose gel and the wild type gene was gel extracted, purified, and ligated into the pET30ΔSE expression vector using a 3:1 and 5:1 insert to vector ratio. The ligated plasmid was transformed into DH5α cells for plasmid storage and future plasmid purification as more plasmid would be needed to transform BL21 (DE3) cells for protein expression. Mini-prep of the plasmid from DH5α that were transformed with the ligation mixtures of 3:1 and 5:1 was performed and a final double digestion was done to ensure the WT *pykF* gene was inserted into the pET30ΔSE expression vector, as the selected colony could have been a false positive (**Figure 19**). Bands for supercoiled, nicked, and linear DNA can be seen in the uncut lanes. The double digestion lanes showed two distinct bands, with one band appearing at the exact mass of the WT *pykF* gene while the other band appearing at the mass for pET30ΔSE. Some undigested plasmid can be seen just above the band for pET30ΔSE along with some nicked DNA above in some double digested wells however, digest efficiency is not important in this case. This additional digestion showed that the WT gene was successfully cloned into the pET30ΔSE vector and the selected colonies from the transformation did not containing an empty pET30ΔSE plasmid with both the 3:1 and 5:1 ligation conditions being successful. Stocks from 3:1-2 and 5:1-1 were kept for future experiments as 5:1-2 showed very faint DNA results that cannot be seen in the figure. Once the successful cloning of the gene was confirmed, mini-prep was done to extract plasmid for BL21 transformations.

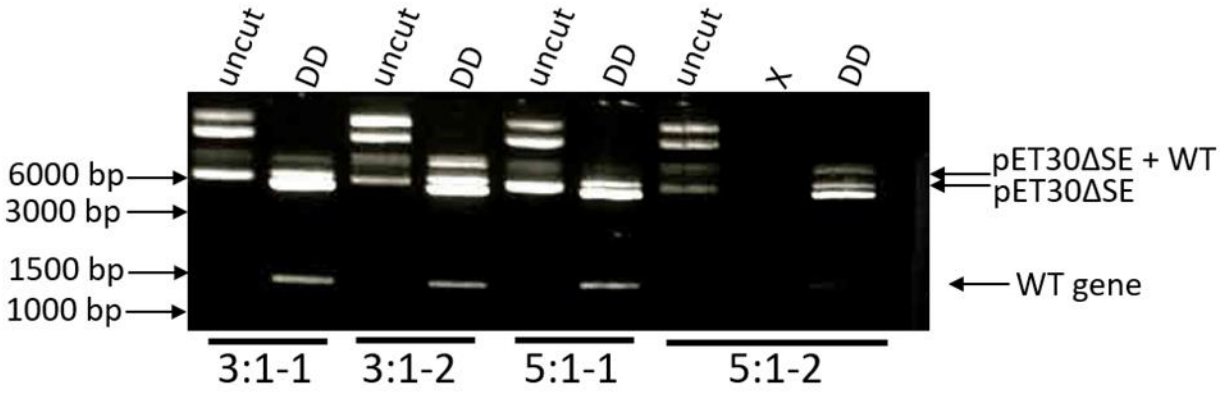

Figure 19. Double digestion of DH5α gene transformed with pET30ΔSE containing wild-type *pykF*. Uncut (lanes 1, 3, 5, and 7) and double digested (DD) (lane 2, 4, 6, and 9) pET30ΔSE products from two different colonies selected from the 3:1 and 5:1 ligation ratios. 3:1-1 (lane 1, 2) 3:1-2 (lane 3, 4) 5:1-1 (lane 5, 6) 5:1-2 (lane 7, 9). Lane 8 is empty.

3.2 Expression and Purification of Wild Type Pyruvate Kinase and Point Mutants

3.2.1. Expression Trials for Maximal Protein Yields

Before beginning purification, mini expression trials (50 mL) of the protein were done to not only obtain maximum protein production from expression but to also ensure that the plasmid was successfully transformed into BL21(DE3) cells. For the induction time trials, T405A was used as a model. Two different expression methods were used to test protein production (**Figure 20**). After the cells reached an optical density of OD_{600} ~0.4 (lane 1 and 4), they were induced with IPTG (1 mM). The first induction method was a 3-hour induction at 37 °C after which a sample was taken (lane 2), and induction was continued overnight at 20 °C to see if protein yield would increase (lane 3). The second method was just to induce the protein overnight at 20 °C (lane 5). To ensure equal cell volume loading, OD_{600} of each sample was taken and appropriate dilutions were done to ensure all wells were loaded with approximately the same amount of cell mass. Relative protein yield was calculated using ImageJ v1.5, where the intensity ratio of the

corresponding PK protein band and a control band present in all samples was taken and compared for all conditions (**Figure 20**).

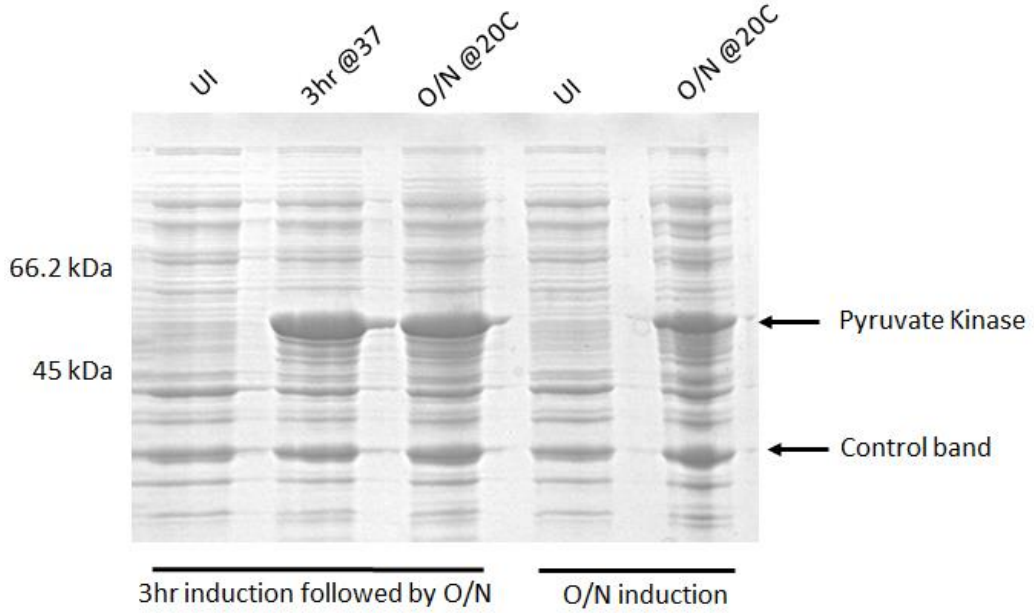

Figure 20. Optimizing time of induction for maximal protein expression. Un-induced (UI) cells at OD_{600} ~0.4 (lane 1 and 4), induction after 3 hrs at 37 °C (lane 2), followed by overnight at 20 °C (lane 3) and an overnight induction at 20 °C only.

A ratio with a control band was done to correct for any dilution errors or gel loading errors to ensure equal comparisons. The relative intensities were 1, 0.7, and 0.6 for lanes 2, 3, and 5 respectively. These results showed that 3 hours yielded the most protein and that perhaps some degradation of the protein was happening for the overnight samples. A quick test for all mutants using a 3-hour expression revealed that all were well over-expressed and gave high protein yield (**Figure 21**). Overall, these results showed that a 3-hour induction was enough to produce maximum protein yields and this condition was chosen for large scale expression for protein purification.

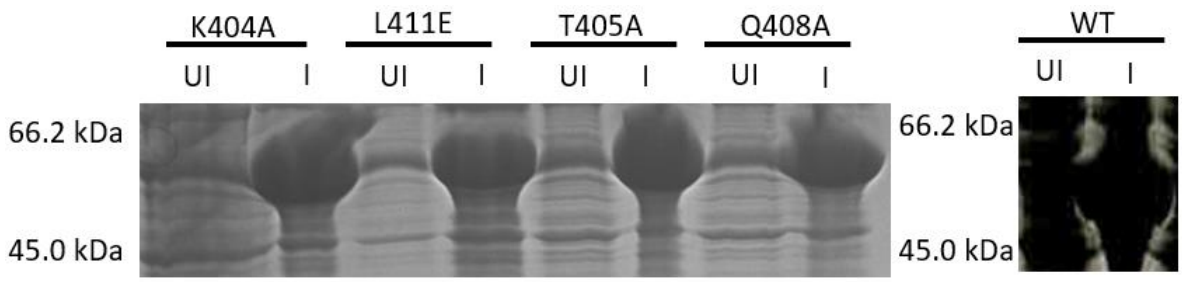

Figure 21. Induction trials of all PK protein variants. Un-induced (UI) sample at OD_{600} and induced (I) samples after 3 hr at 37 °C of all the PK variants.

3.2.2. His-tag Purification of Pyruvate Kinase Variants

Each mutant was purified using a $Ni2^{+-}$affinity gravity column. **Figure *22*** shows the results from the purification for each variant on a 12 % SDS-PAGE gel. Before the purification, an elution fraction (500mM imidazole) was collected to ensure the column was clean (lane 1). No protein can be seen in any of the gels in lane 1 indicating there was no residual protein left over from other purifications and therefore no cross contamination between proteins from other purifications occurred. There is clearly a lot of protein seen in the total lysate (TL), and quite a bit of protein seen in the flow through (FT). This could be due to overloading the column with protein or perhaps the beads where losing their binding affinity as they have been previously used for multiple purifications. Despite a relatively large loss of protein in the flow-through, this was not a major concern as lots of protein was retained and collected in the elution fractions. The 30 mM imidazole wash fraction showed that all the impurities have been washed away. Some imidazole in the wash buffer is necessary to remove any proteins that are non-specifically bound to the column and do not elute with the desired protein during the purification. Further elution fractions containing higher imidazole concentrations show no more impurities present and only the elution of pyruvate kinase. In all cases the protein is seen to be mostly eluting at 150mM imidazole. The final protein concentration after purification ranged from 4 – 11 mg/mL (18 – 60

μM) in a total of 10 mL, which is an abundant amount of protein for use in mass spectrometry despite the loss of protein in flow-through fractions.

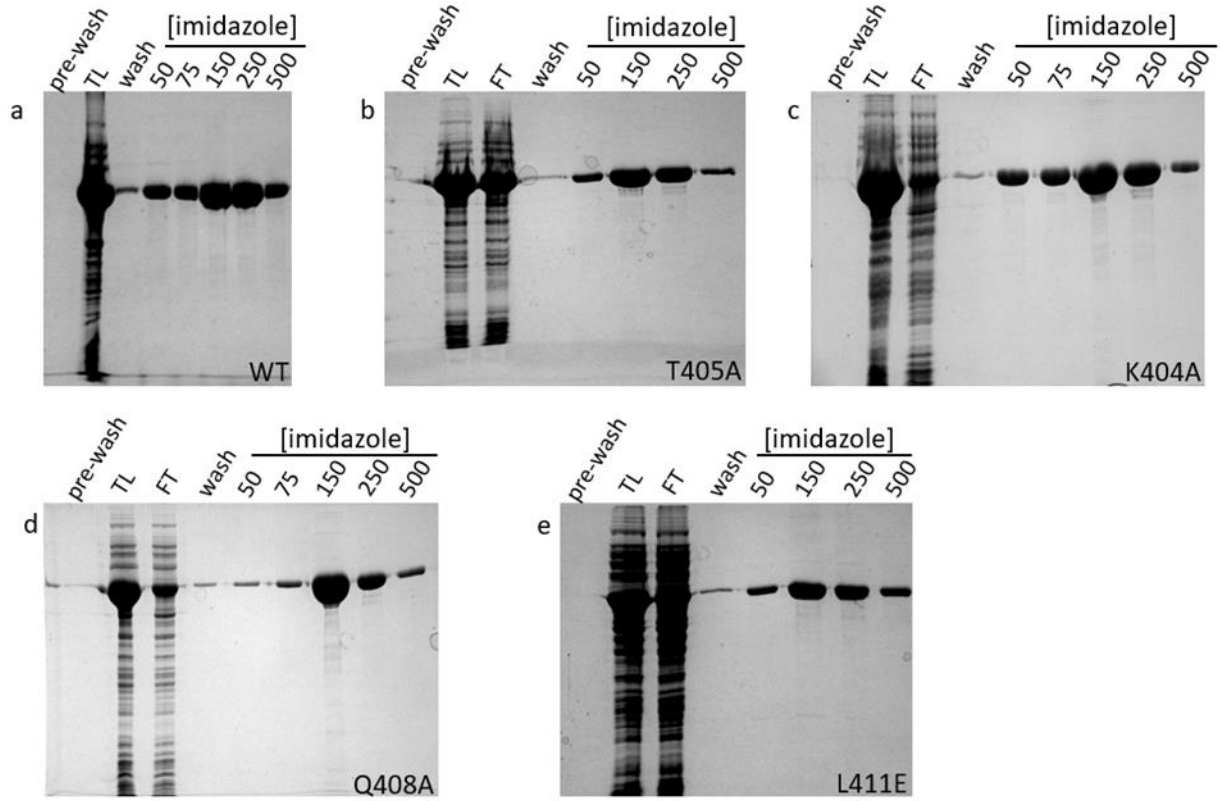

Figure 22. SDS-PAGE results of Pyruvate Kinase protein variants purification. (a) WT (b) T405A (c) K404A (d) Q408A (e) L411E. Elution from column prior to protein purification (500mM imidazole) (lane 1), followed by total lysate, flow through (this sample is not present in the WT gel), 30 mM imidazole wash, and increasing imidazole elution fractions.

3.3 Comparing Solution and Gas Phase Stability Using Native Traveling Wave Ion Mobility Mass Spectrometry (TWIMS-MS)

3.3.1 Native MS of Pyruvate Kinase Variants

To obtain initial information for the pyruvate kinase protein variants, native mass spectra using gentle ESI and non-activating trap voltage conditions (20 V) was obtained (**Figure 23**). Differences in the native mass spectra between the protein variants can already be seen. The WT variant is highly tetrameric, with the tetramer appearing between 6,000 – 8,000 m/z. Significant

amount of highly charged octamer is also visible although it has previously been reported that the octamer is non-specific[54] as the enzyme operates in a tetrameric state. T405A and K404A mutants both show similar mass spectra. They are highly tetrameric and show very slight amount of monomer peaks that appear at 3,000 – 4,000 m/z. A very small amount of dimer can also be seen at 5,000 m/z. Q408A shows the highest monomer to tetramer ratio with an appreciable dimer population also visible. L411E appears mostly as a monomer with slight amounts of dimer and forms no visible tetramer in the gas phase. The disruption of the tetramer could have been due to the fact that a highly hydrophilic amino acid was introduced into a hydrophobic area of the protein causes significant changes in the structure. This change also led to a high population of unfolded monomer as well (data not shown).

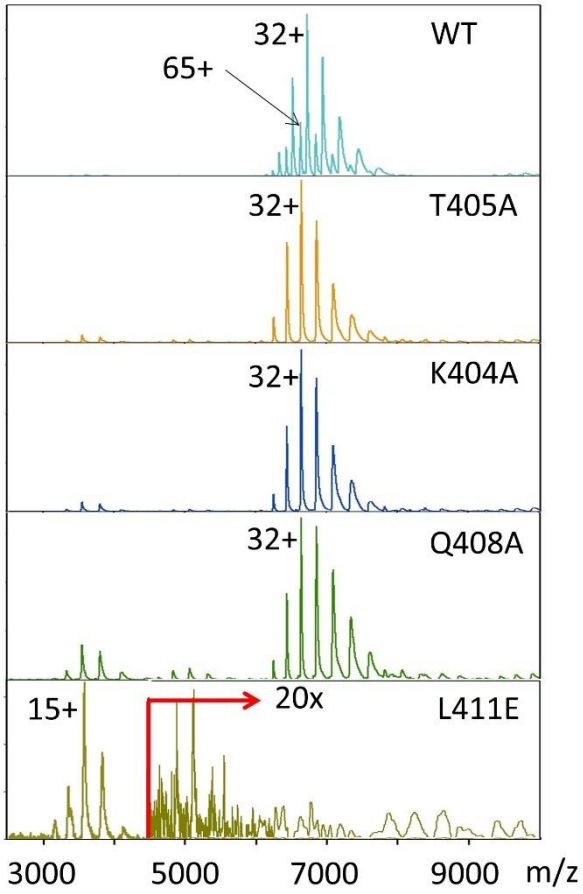

Figure 23. Native Mass Spectra of Pyruvate Kinase protein variants. Native mass spectra of the wild type (top) followed by T405A, K404A, Q408A, and L411E (bottom). The intensity scale for L411E was increase 20x from 4,500 m/z onwards to highlight the absence of the tetramer in the gas phase.

As L411E was previously shown to be inactive[45], this may explain its inability to form a tetrameric complex. A stability assay previous study done, showed that the WT was the most stable in solution, then T405A and K404A with similar stabilities, followed by Q408A, and lastly L411E[45]. Preliminary results from the native mass spectra of all five PK variants show that the most stable mutants have the lowest monomer to tetramer ratios and as solution phase stability decrease, the ratio of monomer to tetramer increases.

3.3.2 Solution Phase Stability Using Differential Scanning Fluorimetry

To investigate the relationship of solution to gas phase stability in more detail, the differential scanning fluorimetry (DSF) assay previously used to characterize the protein variants stabilities was repeated using 200 mM ammonium acetate. This was done to ensure consistency between the solutions used to generate the native spectra for gas phase analysis when comparing to DSF solution phase results. This removes any effects that different salts may have on protein stability[34, 35, 55, 56] as previous results were done in 10 mM Tris, pH 7.5. Melting curves in 200 mM ammonium acetate showed the same trend that was previously reported (**Figure 24a**). T_m results were generated by taking the negative derivative of the melt curves and taking the lowest points which signify the point of inflection (**Figure 24b**). WT was the most stable with a T_m = 60.5 ± 0.3 °C followed by T405A and K404A with T_ms of 58.1 ± 0.7 °C and 57.8 ± 0.4 °C respectively. Q408A mutant had a bi-phasic melting curve with the first T_m at 43.5 ± 2.3 °C and the second at 57.5 ± 0.0 °C. The first melting phase of Q408A being at a temperature of approximately 15 °C lower than T405A, and K404A, shows this mutant has a much lower solution phase. L411E had the lowest solution phase stability with a weak T_m signal at 42.7 ± 0.5 °C which may be due to unfolding of the folded monomer population which can be seen in the native MS spectrum. T405A and K404A have the same solution phase stability as there was no significant difference between their T_ms. WT was at least 2 °C more stable than T405A and K404A.

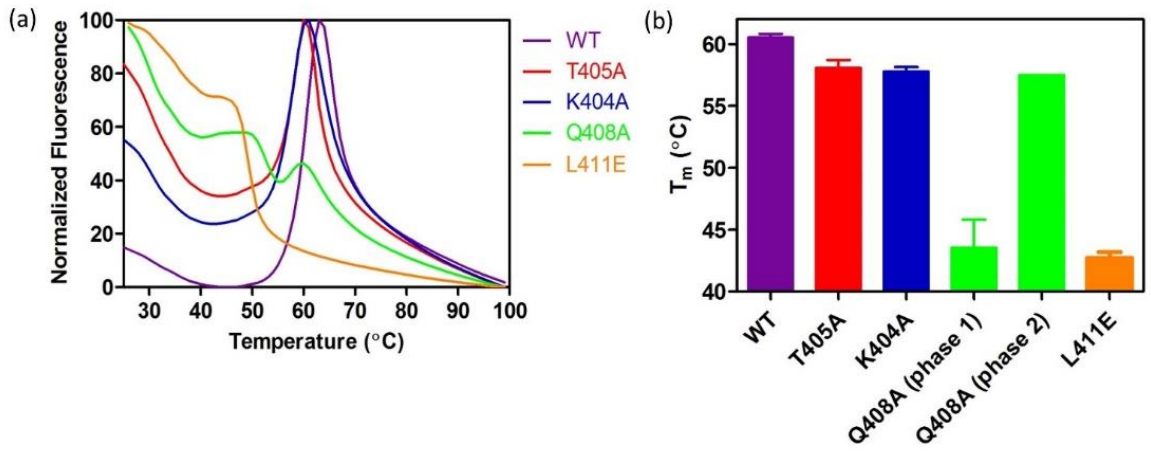

Figure 24. Solution phase stabilities for pyruvate kinase protein variants. (**a**) Melt curves in 200 mM Ammonium Acetate (**b**) Melting temperatures (T_m) taken from the negative derivate of the melt curves. Error bars represent the standard deviation of triplicates for three different days (n=9).

Although it cannot be exactly determined by DSF how the proteins as unfolding as only fluorescence emission is monitored, other studies have used other biophysical methods to characterize multimeric protein unfolding. A review paper has summarized the unfolding results for legume lectins family of proteins[57]. These proteins exhibited similar characteristics to the PK protein mutants as only slight changes in the amino acid sequences led to different types of monomer associations therefore resulting in dimers with varying strengths of subunit interactions[57]. Furthermore, just like PK, these monomer subunits contained all necessary components for ligand interactions and only formed higher order non-covalently bound oligomers for their biological functions[57]. It was shown that monomers that were highly associated with each other in the dimeric state, unfolded and dissociated at the same time at the T_m however, they acted as one unit[57]. Meaning that the monomeric species in the dimer, despite being non-covalently associated, just like in pyruvate kinase, unfolded as though they were a singular entity. It was also shown that as the interactions of the monomer decreased in the dimer, not only did the T_ms decrease, but the monomer and the dimer species unfolded and dissociated independently giving rise to either asymmetrical melt curves or bi-phasic melt curves[57]. It has

also been reported, that distinct bi-phasic melt curves is indicative of intermediates in the unfolding pathway[58].

As pyruvate kinase is highly tetrameric and the A/A' and C/C's monomeric interfaces are tightly held together by salt bridges and hydrogen bonds, it can be seen from the melt curves that the WT and even T405A and K404A unfold and perhaps dissociate as a single subunit as the melt curves seem to be quite symmetric. With the higher T_m of the WT, it can be seen that the tetramer is more stable and is more strongly associated compared to T405A or K404A. This is expected as a hydrogen bond in the monomeric A/C interface has been altered which may have impacted the tetrameric association of the monomers. L411E also shows only one distinct unfolding curve which is most likely due to only unfolding of the monomer, which alone is significantly less than the unfolding of the monomer that was associated as a tetramer. Q408A shows a distinct bi-phasic melt curve which can be indicative of intermediates present in its unfolding pathway. Despite it showing clear properties of tetrameric unfolding, as the second T_m clearly aligns with that of T405A and K404A, Q408A's first T_m is rather asymmetrical and wide, showing distinct dissociation and unfolding events before finally completely unfolding.

3.3.3. Gas Phase Stability: Collision Induced Unfolding

3.3.3.1. Collision Energy Ramps Using Only Native MS

As the initial native spectra and the solution phase stability data showed a correlation, a more in-depth gas phase stability analysis was pursued. To mimic a solution phase heat ramp, unfolding of the tetrameric protein ion in the gas phase was done by increasing the collision energy in the trap cell which results in an increase of the ion temperature [37, 59] leading to unfolding. Initially, a collision energy ramp only monitoring changes in the tetramer and unfolded monomer intensities was performed to see if changes in the folded and unfolded states can be an indicator of stability.

The highest intensity charge states in both the tetramer and the unfolded monomer envelopes (32+ and 22+ respectively) were normalized to one another and plotted as a function of collision voltage (**Figure 25**).

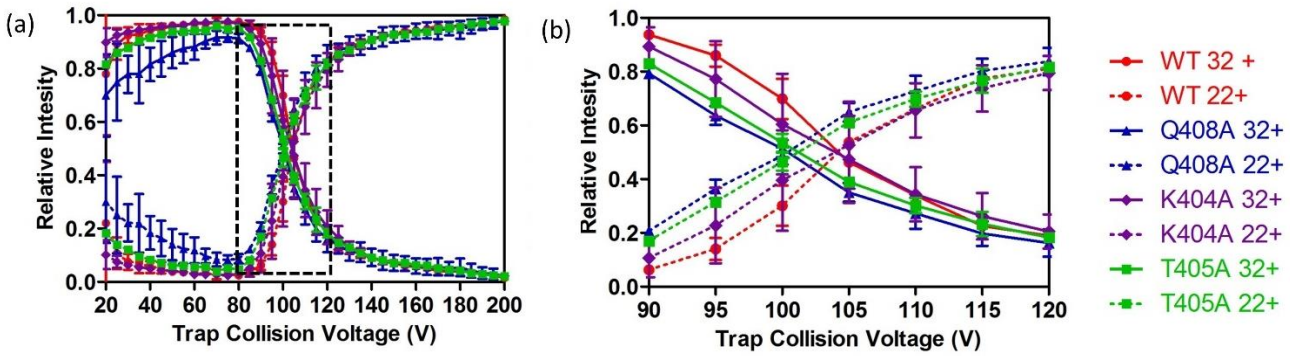

Figure 25. Survivability curves of the folded tetramer and unfolded monomer. (**a**) Full collision voltage ramp (**b**) Enlarged section from panel (**a**) (black dotted box). Solid lines represent the folded tetramer species (32+ charge state) and the dotted lines represent the unfolded monomer species (22+ charge state). Error bars represent standard deviation of duplicates.

The rate at which the tetramer species disappears and the unfolded monomer species appears, seem to be the same for all the mutants. A careful look at the point of intersection between 100 – 110 V shows that initially Q408A has the lowest tetramer and highest unfolded monomer composition between all three mutants and the point of intersection occurs at a slightly lower energy level, while the WT shows the highest folded and lowest unfolded compositions. However, the differences between all the protein mutants are non-significant and based on intensities alone, results may be uninformative.

3.3.3.2 Collision Energy Ramps with IMS

To obtain more sensitive unfolding information of the tetramer ion in the gas phase, the drift time of tetramer ions were monitored using TWIMS. An average CIU fingerprint plot was generated from three replicates for each protein variant. Four charge states were analyzed (34+ to 31+) however, the highest charge state in the protein envelope (34+) showed CIU fingerprint

plots that were not well resolved and did not show differences between the protein mutants. This can be due to premature unfolding of the protein ion, as higher charge states exhibit high coulombic repulsion which cause unfolding[60, 61]. Furthermore, this charge state had the lowest intensity which could contribute to higher noise and un-resolved plots. The 33+, 32+, and 31+ charge states of the protein variants all showed resolved fingerprint plots with the same gas-phase stability trend however, as previously reported[62], lower charge states required more energy for unfolding. The 32+ peak was selected as it showed the clearest differences between transition states and it had the highest intensity ion peak for all mutants (**Figure 26**).

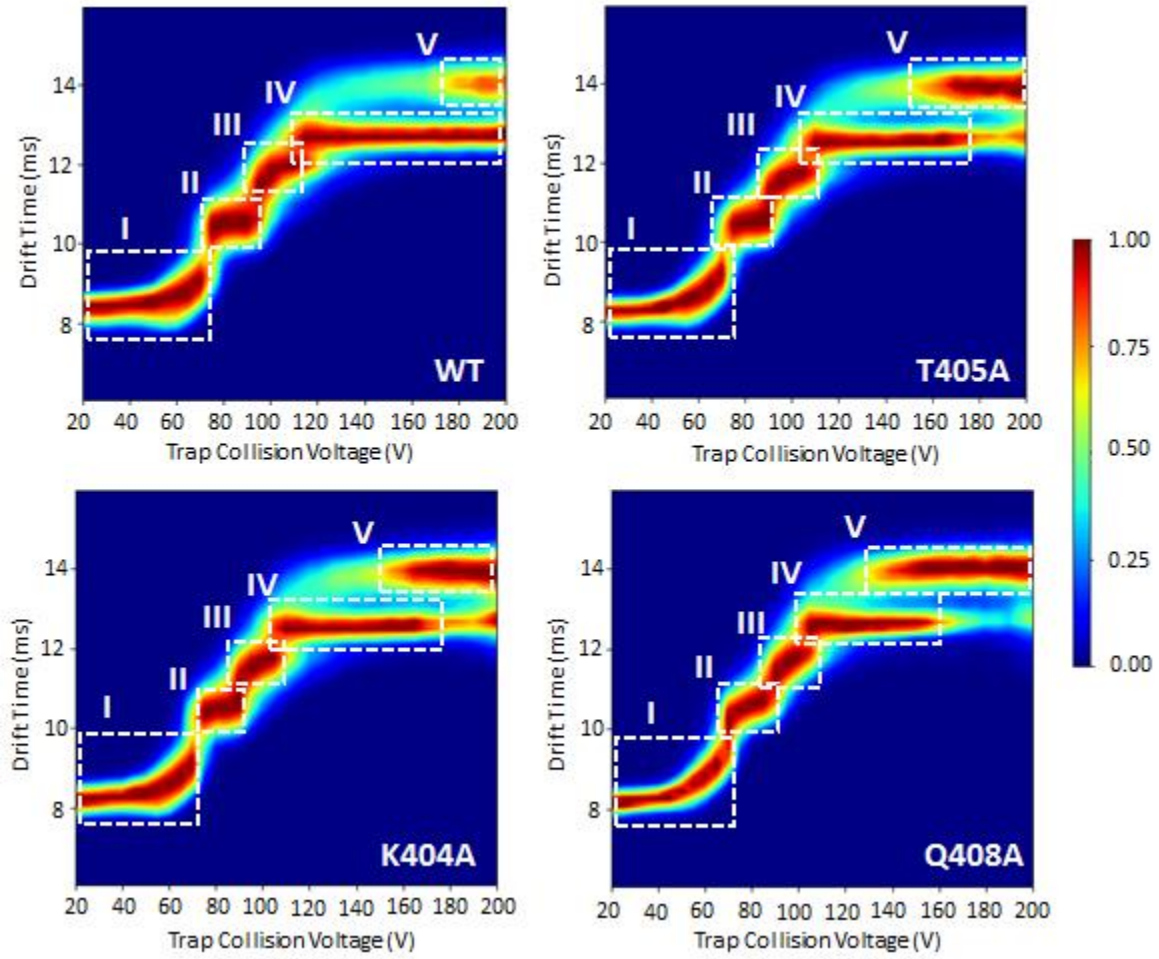

Figure 26. Averaged CIU fingerprint plots for pyruvate kinase variants. CIU plots are an average of three replicates. Conformation families in each plot are grouped into five categories (I, II, III, IV, and V) and highlighted (white dotted box).

In the CIU fingerprint plots for the 32+ charge state ion, transition states were grouped together into five categories. The initial transition phase (I) is between 20 – 70 V with two smaller transitional states (II and III) between around 70 – 100 V. Immediately, differences between the next two transitions (IV and V) can be seen for all variants. Transition state IV in the WT can be seen to last from approximately 115 V to 200 V with high intensity. For T405A and K404A this transition state appears at a slightly lower CV (~110V) and begins to fade after 170 V. Transition state IV for Q408A begins around 105 V and is the shortest lived as it completely fades after 160 V. The last transition state (V) only begins to appear around 180 V for WT while it is already strongly present around 170 V for T405A and K404A. For Q408A the last transition state is strongly present at 150 V. Since transition state IV and V is present at higher CEs for WT, and present at lower energies for Q408A, this indicates that the WT has the highest gas phase stability while Q408A has the lowest. T405A and K404A have similar CIU fingerprint plots and have transition states that appear at CEs that are between WT and Q408A. CIU plots were not done for L411E as it does not exhibit any tetramer species and thus is considered to be highly unstable in the gas phase.

To determine more accurately where the transition phases occur and subtle differences between the fingerprint plots of the protein variants, the CIU_{50} values for the averaged plots at each transition phase were obtained (**Figure 27**) and the CIU fingerprint plots were subtracted between all the protein variants (**Figure 28**). CIU_{50} values are similar to T_m values as they indicate at which collision voltage the protein ion shifts from one conformational family to the next, a point at which the protein ion unfolds from one stable gas phase intermediate state to the next more unfolded state. As **Figure 27** shows, the CIU_{50} values for WT are consistently higher for all transition states compared to all other mutants. A fourth transition phase was not identified

for the WT variant as the fifth state is only beginning to appear and thus it can be assumed that the fourth transition phase for the WT occurs at a collisional voltage above 200 V. T405A and K404A are nearly identical for the first three transition phases, with some difference seen in the last transition phase, with the fourth transition occurring at a higher voltage (167.1 V) than for K404A (162.5 V). However, the difference in the fourth transition phase for T405A and K404A is not nearly as markedly different than that for WT and T405A or K404A. Therefore, T405A and K404A show similar CIU_{50} values with only slight differences. Q408A, on the other hand, shows consistently lower CIU_{50} values for each transition phase compared to the other variants, clearly indicating it requires the least amount of energy to transition from one conformational family to the next.

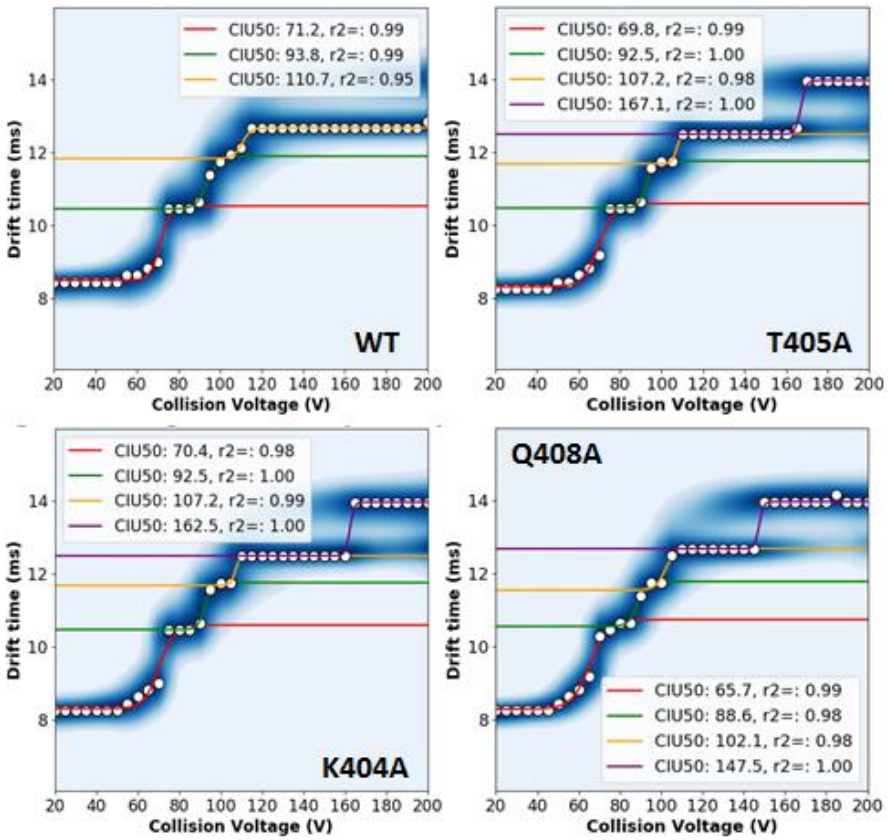

Figure 27. CIU_{50} values for the averaged CIU fingerprint plots for all tetrameric pyruvate kinase variants.

Subtraction plots provide information between the average difference across the entire fingerprint and where the most variation occurs. Subtraction plots for all the mutants can be seen in **Figure 28**. RMSD values are indicated on all the plots which highlight the average difference between the two PK variants being analyzed. It is clear that T405A and K404A have no significant differences between each other as the plots has few areas of variance with a very low RMSD of 3.43. This, along with the CIU_{50} values, reinforces that the gas-phase stability difference between T405A and K404A is non-significant, just as it is in solution.

Subtracting T405A and K404A from WT, results in similar RMSD values of ~13. All transition phases for the WT (blue) transition are seen to occur at higher CEs than those of T405A or K404A (red) which is also confirmed with CIU_{50} values. Despite transition state IV being more stable for the WT, it appears at slightly higher drift times compared to T405A or K404A which might indicated that this intermediate for the WT adopts a slightly more extended conformation in the gas-phase. The large red area from around 120 – 200 V highlights the presence of the fifth transition phase between T405A and K404A compared to its near absence in the WT. T405A and K404A plots were also subtracted from Q408A showing similar RMSDs of ~15. Here, Q408A's fingerprint (red) appear exclusively at lower CEs compared to T405A and K404A (blue) indicating that across the entire range, T405A and K404A are more stable than Q408A. When WT and Q408A were subtracted, they had the highest RMSD difference of 19.56. Interestingly, initial drift times for WT are higher than those of Q408A, however, at activating CE conditions (~50 V) and above, WT shows lower drift times, and therefore higher stability. The initial higher drift times for the WT can be due to the number of adducts initially present at low CE for each variant. Differences in m/z shift at an initial CVs (10 V) and at a higher CVs (80 V) were measured between Q408A and WT. It was seen that WT had higher m/z shifts than

Q408A (data not shown) which can be a result of more adducts present on the WT resulting in higher initial drift times. Overall, the CIU_{50} values and subtraction plots show that, just as in solution, T405A and K404A are closest in stability with no significant differences between them. With the highest RMSD difference, Q408A and WT have the highest difference in stability. The subtraction plots also showed that T405A and K404A are closer in stability to WT than to Q408A. This was interesting as we were able to show that not only the order of stability in the gas phase was representative of solution phase, but the mutants that were closer in stability in solution were also closest in stability in the gas phase.

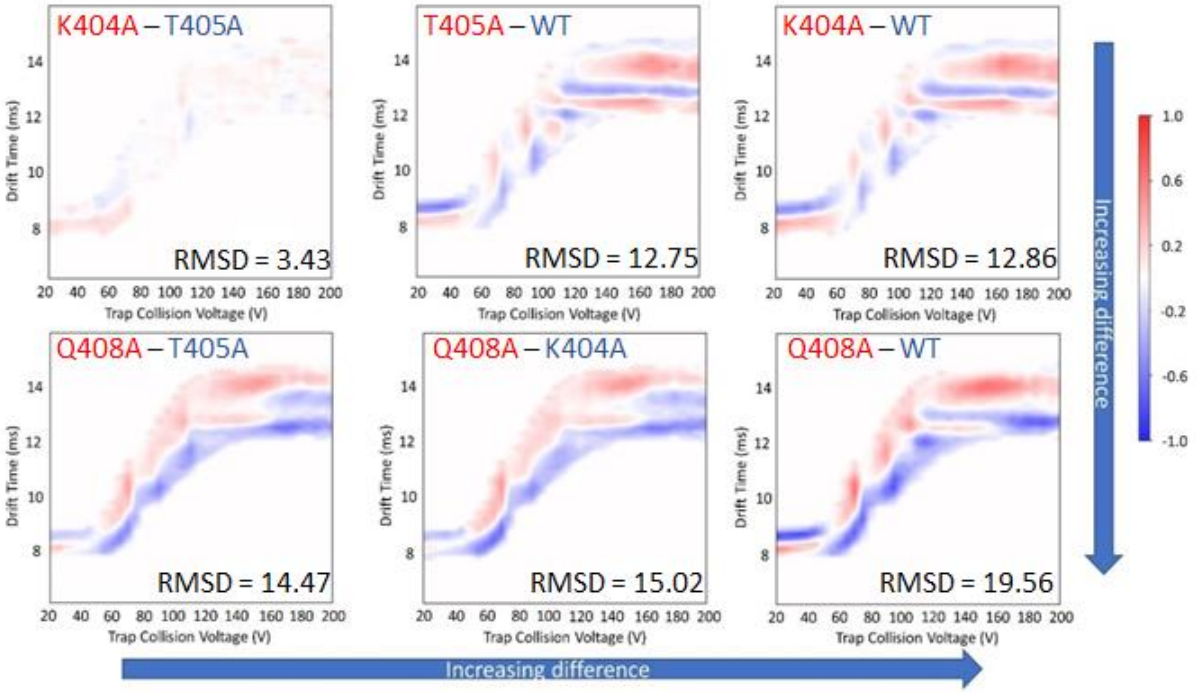

Figure 28. Subtraction of the average CIU plots between all pyruvate kinase variants. (**a**) K404A – T405A (**b**) T405A – WT (**c**) K404A – WT (**d**) T405A – Q408A (**e**) K404A – Q408A (**f**) Q408A – WT. Red colour represent the mutant being subtracted from while blue represent the mutant being subtracted. RMSD values are indicated on the bottom right.

Despite many CIU studies comparing relative gas-phase stability, almost none have done a direct solution-phase comparison. A paper comparing how ligand binding affects gas-phase stability did compare their results to solution phase data gathered from isothermal titration

calorimetry, however it was only shown that both gas and solution phase ligand binding lead to "cooperative stabilization"[62]. Although this does reinforce similarities between gas and solution phase stabilization mechanisms, their results did not show how various ligands affected the proteins solution stability compared to the gas-phase. Another paper showing that mutations of tetrameric transthyretin (TRR) protein, led to decreased gas-phase stability which was in agreements with solution-phase data[63]. The solution phase data that was used as a reference, was monitoring the dissociation of the tetramer in pH 4.0 and 7.0 using mass spectrometry as a function of increasing cone voltage[64]. Although the protein was exposed to varying solution conditions, this is not a true solution phase assay.

A recent paper published in 2016 was one of the only papers that did a direct comparison of solution and gas-phase stabilities using DSF and CIU as well[65]. Here, they looked at a WT and two mutant versions of cAMP-dependent protein kinase (PKA) that affected the proteins ability to autophosphorylate. They showed that the WT version was in fact more stable in the gas phase though its CIU profile and it had a higher T_m than the other two mutants[65]. The R133A mutant, whose autophosphorylation was not affected, showed a bi-phasic melting curve with the second T_m being similar to that of WT (~43-44 °C) while the fist T_m was around 37 °C which was comparable to that of the K72H mutant which was shown to have no phosphorylation[65]. The transition phase for the CIU profiles for R113A and K72H seemed to occur at very similar CVs despite there being some differences in their T_m results. However, the CIU profiles for all three only focused on a single transition phase that occurred below 40 V. It was mentioned that at higher voltages covalent bond dissociation may happen, yet it was unclear if the reason why their ramp was short was indeed because they saw fragmentation or higher voltages were not attempted. Based on their CIU profiles, there seems to be an intense protein ion species still

present at 40 V which may be an indication that higher CVs should be probed to get a better understanding of the gas-phase stability. As it can be seen from the CIU fingerprint plots generated for the PK variants, most of the dramatic differences occurred at CVs above 100 V. It is also common practice for many of the CIU experiments to probe voltages up to 200 V with no dissociation of the protein seen.

3.4 Comparing Protein Stability and Dynamics using Global Time Resolved HDX

To see if any correlation can be drawn between protein stabilities and dynamics, global hydrogen deuterium exchange (HDX) using time-resolved electrospray ionization (TRESI) was performed on all the tetrameric protein variants (**Figure 29**). Surprisingly, the global HDX results revealed a different trend in dynamics for the variants compared to their stability trend. T405A was shown to be the most dynamic with the highest deuterium uptake rate and a higher plateau compared to all other mutants, while WT showed to be the least dynamic. Q408A and K404A were in the middle with Q408A showing it was slightly more dynamic that K404A. If dynamics correlated to stability, it would have been expected that Q408A would have shown to be the most dynamic or unfolded by having the most solvent accessible areas, yet this is not the case. The global HDX results also showed that T405A and K404A, despite being similar when it comes to stability, have major differences in dynamics. These results being surprising at first, can be explained by looking at the catalytic properties in the absence of their allosteric regulator[45]. T405A has a point mutation that tunes it for activation which results in high allosteric signaling, even without the presence of an allosteric regulator, from the C domain to the top of the A domain where the protein undergoes “rocking” movements which creates a flexible TIM barrel that transfers the signal to the active site and further destabilized the active site [45]. This mutation not only results

in the proteins ability to be more dynamic and catalytically active than the WT variant which is evident from the faster initial deuterium uptake rate and final plateaus between the two. K404A has a catalytic activity that is similar to that of the WT while Q408A has been destabilized such that its catalytic rates are much lower that WT in the absence of the allosteric regulator[45]. Higher uptake for Q408A compared to K404A can be a result of the higher instability of Q408A between the two mutants. As K404A has a similar catalytic rate to WT, it might be expected that their HDX rates would be the same, however, with K404A being more unstable, this slight instability can lead to its slightly "unfolded" or "less compact" state giving it more solvent exposed regions than the WT, resulting in a higher uptake of deuterium. The lower deuterium uptake from the WT variant indicates that the WT is more compact than the rest of the mutants and supports the results that the difference in initial drift times between Q408A and WT are due to the higher number of adducts present on the WT variant rather than the initial conformation states of the proteins.

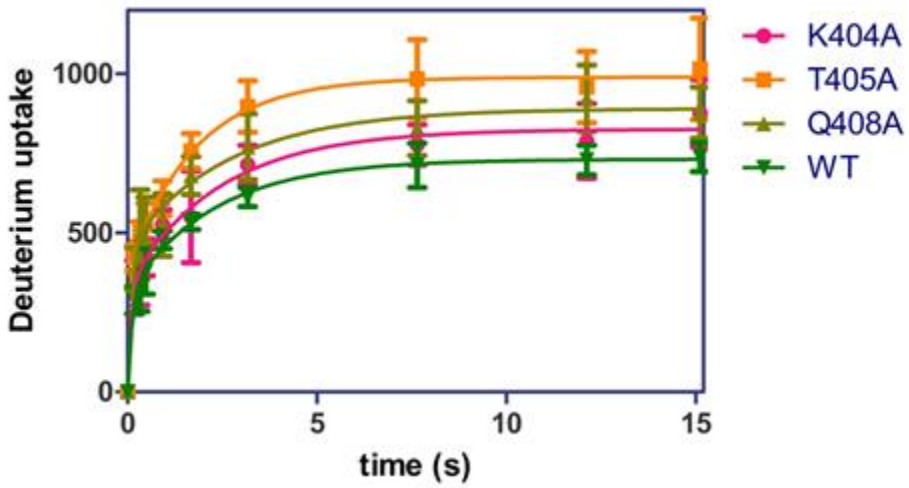

Figure 29. Global HDX on pyruvate kinase variants. Data was fitted as with a two-phase association curve. Error bars represent the standard deviation of triplicate measures on three different days (n=9).

Many studies have been done comparing how mutations affect stability. Many studies have shown that generally, mutations affect a protein's stability or "fitness" with some mutations being tolerated as long as they do not surpass a certain stability threshold after which protein fitness is highly affected[66]. Another paper showing how stability and protein flexibility relate used HDX and measured conformational-motion temperature and compared 30 different sequences of hydrophobic and other amino acids of chain length from 16 to 20 and concluded that protein flexibility is inversely proportional to stability[67]. However, other works have summarized that depending on where the mutation occurs, the a less functional protein may be more stable[68]. It has been shown that in certain "flexible hotspots" such as sites required for allostery, can be mutated to make the protein more stable, but less functional[68]. This global HDX data is in exact correlation with these results. T405A has a mutation in which it assists in positively regulates the allosteric signaling in the PK protein which results in slight destabilization, which may be within the enzymes tolerance range. However, despite K404A having the same stability, its mutation did not affect the allosteric mechanisms and therefore may not be important in signaling. The mutation that occurred in Q408A did affect the protein's fitness as it perhaps surpassed the threshold of stability that the enzyme could tolerate which in turn negatively affected its function. L411E's mutation was highly deleterious as it did not even allow the protein to form into a stable tetramer.

3.5 Implementing DMS as a Gas-Phase Stability Technique.

As described in section 1.3.4, our group was successful in showing that DMS can be used to probe solution phase conformations. From these results we wondered if DMS could be used to probe protein stability in the gas-phase. To be able to compare results to TWIMS, and our solution phase T_m results, the same mutants were used in the DMS study. Lots of optimization

was done to determine which protein concentration, buffer conditions, DMS and mass spectrometer settings would be best to run these large protein samples. High ammonium acetate concentrations, such as those used in TWIMS, were not feasible as this caused arcing due to the high ion concentration and therefore an increase in conductivity. It was finally established 5 mM acetate did not cause such problems and still gave well resolved mass spectra.

3.5.1. Establishing CoV profiles

One of the initial experiments that were performed was to see if a good CoV profile could be established and if this was reproducible with all mutants. **Figure 30a** shows the CoV profiles of all the mutants obtained using DMS with MeOH as a modifier gas. It can be seen that for WT, K404A, and T405 a two distinct peaks are seen around a maximum of -8 V and 4 V for CoV. A closer look using T405A as an example (**Figure 30b**), shows that the peak from -12 V to -2 V contains only the tetrameric form of the protein while the peak from 0 V to 8 V contains both the monomer and the tetramer species. Initially, from the CoV profiles, it may seem as though the mutant Q408A does not exhibit any tetrameric form, however when CoV of -8 V was selected and acquired over 2 min the tetrameric peaks for Q408A were detected (data not shown). This was due to the fact that these results were collected before the installation of the GeneJET Q1 which resulted in extremely low signal as it was difficult to select high-mass ions. As Q408A does not have as much tetramer as the other mutants, the signal was too low to see in a single scan which is why the first tetrameric peak is not very visible. Furthermore, L411E only shown the peak for the monomer and unlike Q408A, when more scans at a CoV of -8 V was done, it did not show a tetrameric peak, indicating that even under atmospheric conditions, L411E fails to show a tetramer form. What is interesting is that all the mutants show the tetramer and monomer peaks to appear at the same CoV which indicates that they are not distinguishably different in

their chemical properties. Q408A's monomeric peak seems slightly shifted to lower CoV values, however this can be due to the low intensity signal problems and potential aggregation.

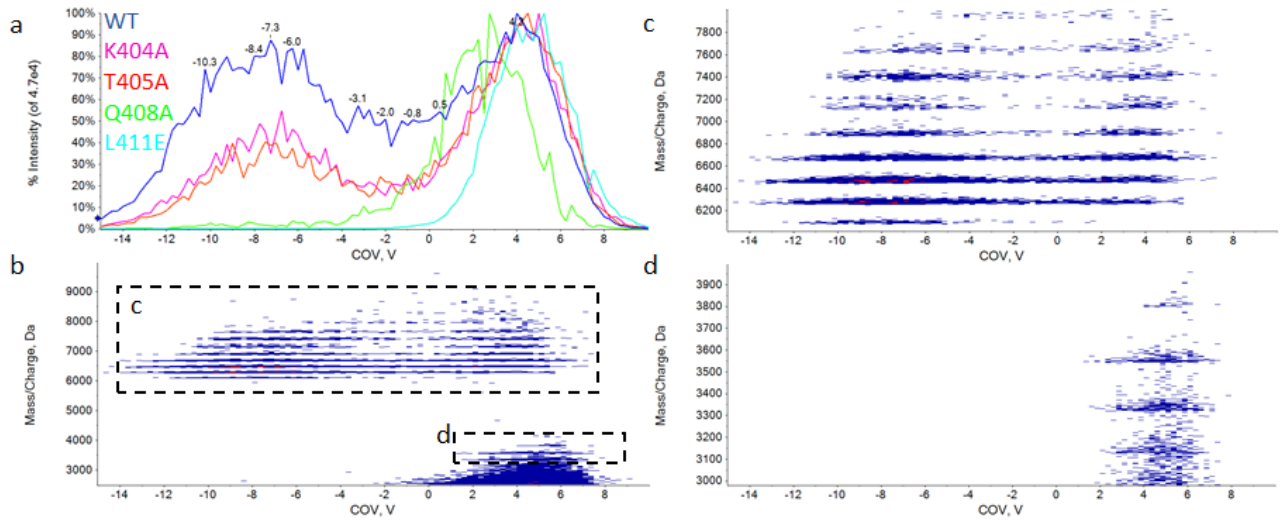

Figure 30. CoV profiles of protein variants. (a) CoV profiles of all mutants from -15 to 10 V (b) example contour map of CoV and m/z of T405A with a closer look of (c) the tetramer (d) the monomer.

The CoV profile also showed that the tetramer was able to be separated from the monomer at negative CoVs and therefore can be analyzed on its own without monomer interference. The monomer on the other hand can never be separated from the tetramer as the massive size of the tetramer makes it have a large CoV spread, appearing in both the positive and negative range, although majority present around -8 V. It is important to note that these CoV profiles were taken with MeOH as a chemical modifier and thus the maximum for each peak does change depending on the type of modifier gas used. The modifier gas will not dramatically affect the peaks; thus, the tetramer can always be separated from the monomer, however, a quick CoV ramp should be done to determine the new peak maximum when using a different modifier gas.

3.5.2. Collision Energy ramps

Once CoV profiles were established, a simple collision energy ramp, like the one in section 3.3.3.1, was done to see if any changes in the survivability of the tetramer can be detected. To maintain consistency between TIMS and DMS data, the survivability of the 32+ charge state is shown in **Figure 31** both using DMS selection at a set CoV of -7 V (**Figure 31a**) and no DMS selection (**Figure 31b**). When the tetramer is specifically selected in the DMS at a CoV of -7 V, there seems to be no noticeable difference in the survivability rates of the tetramer between T405A and K404A compared to WT. These results parallel those seen when a simple CV ramp was done using the Synapt. Yet, when no DMS selection was used, the intensity of the tetramer for the WT seemed to decrease at a faster rate than that of K404A and T405A, although with a small difference. Using DMS selection at a specific CoV, means that a tetramer with a specific conformation is being analyzed, while tetrameric ions with slightly different conformations are being filtered out. This might give rise to the slight difference seen between the two graphs as without DMS selection represents a larger average of the 32+ ion conformations. This trend was not unique to this charge state as three other charge states showed similar results.
Conformational selection may explain the slight difference seen between the two graphs, but, it does not explain the reason as to why the WT would show a slightly faster increase in decay. As these results are only a single replicate, this faster rate of decay for the WT may be statistically insignificant. Q408A was not done as there were aggregation issues. Overall, the initial results of a CV ramp with and without DMS selection did not give any useful results, similar to that of TWIMS. Thus, a simple CV ramp proved not to be a good stability indicating assay using two different methods.

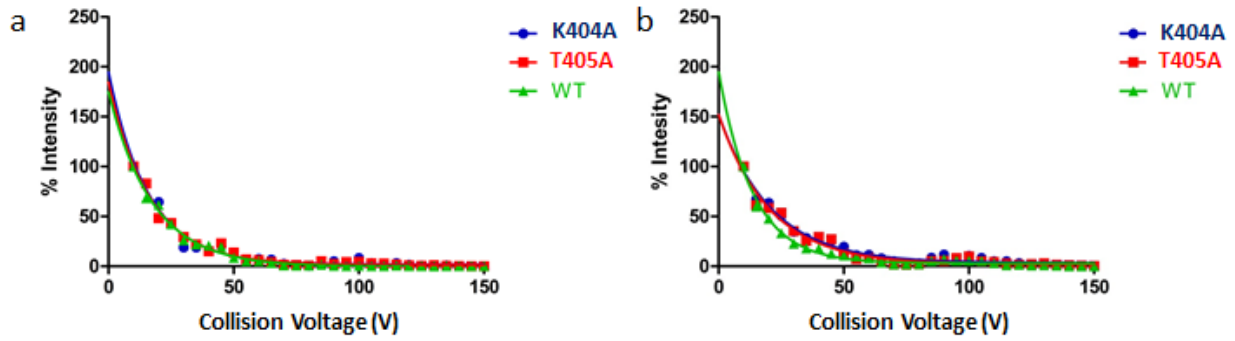

Figure 31. Collision energy ramps of the tetramer. The 32+ charge state intensity as a function of collision voltage both (a) with DMS selection at -7 V and (b) without DMS selection using MeOH as a modifier gas. Data was fit to a decay function using GraphPad.

3.5.3. Stability Monitoring Using DMS-HDX as a Function of CoV

With a previous study showing that solution phase conformations can be retained in the gas phase by monitoring HDX uptake as a function of CoV (Sections 1.2.2.2 and 1.3.4), here we try and replicate the results using the PK variants to try and see if this conformational correlation can translate into a stability correlation. MeOD was introduced into the curtain gas at 120 µL/min and spectra at CoV values of -12 V to -4 V, were the tetramer is most predominant and separated from the monomer, and collected for 30 s at 0.5 V intervals to improve signal intensity. T405A was used to see if any difference in deuterium uptake can be detected with increasing CoV. **Figure 32** shows the results of the gas-phase deuterium uptake with increasing CoV. The royal blue spectrum shows non-deuterated T405A in MeOH. When MeOD is added, the spectra are seen to shift to the right showing a mass increasing corresponding to deuterium uptake. All the CoV spectra are overlaid on the same figure and show no difference in deuterium uptake form -12 V to -4 V. These results are unexpected as all the PK variants, despite their stability differences, are structured proteins. As previously shown, structured proteins would have an increase in deuterium uptake as CoV increases, and un-structured proteins show no increase in deuterium uptake as a results of CoV increases[26]. The DMS-HDX profile obtained with T405A

is similar to that described for unstructured proteins as all CoV values showed the same deuterium uptake.

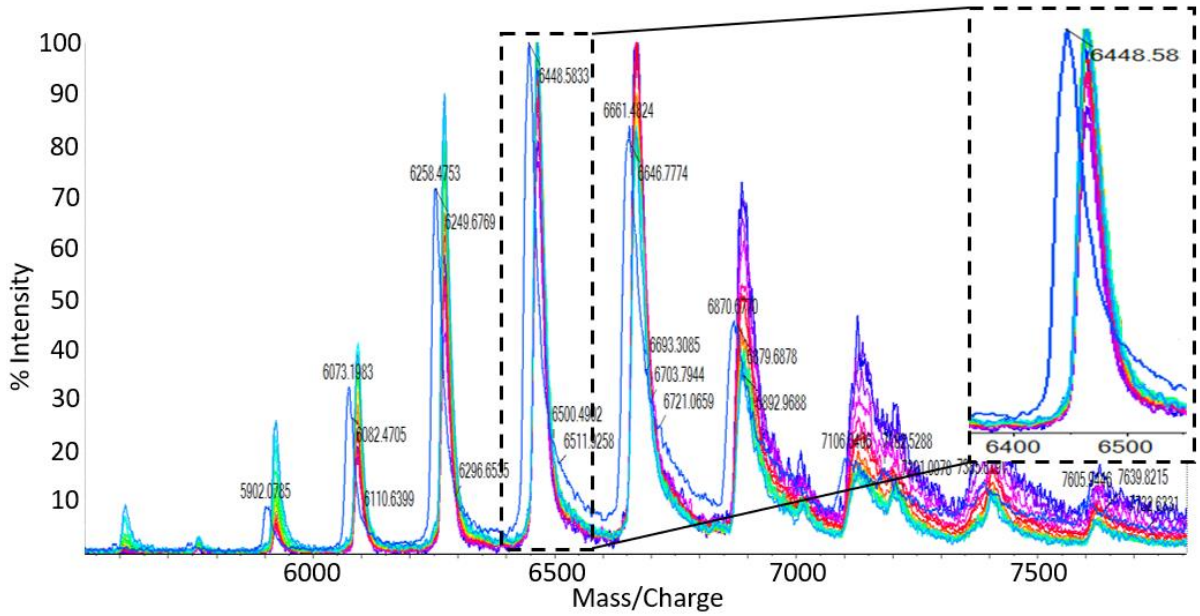

Figure 32. DMS-HDX of T405A with increasing CoV. Gas-phase deuterium uptake of T405A using a DMS cell containing MeOD as a modifier. Dark blue spectrum indicates T405A in MeOH as a control showing no deuterium uptake while all other spectra are in MeOD at varying CoV values from -12 to -4 V at 0.5 V intervals. A close-up of a single charge state is shown in the black dotted inlet box.

One reason that no change was seen between CoVs was perhaps that PK is a very large protein and may have too many ion conformations present at each CoV for there to be any noticeable differences between each interval. The average amount of conformations may have resulted in low sensitivity to changes in deuterium uptake. The previous results were done with Cyt-C which is a much smaller protein. The aim of this study was to see if the variants with different stabilities would show a difference in their HDX uptake rates that then can be correlated back to their stabilities. It was hypothesized that the more stable variants would act more like the “folded” proteins and would show a more “sloped” DMS-HDX profile, or in other words, a larger HDX uptake as CoV increases. However, even with T405A, one of the more

stable variants, there was no increase in HDX seen and therefore this assay was perhaps not sensitive enough to detect stability differences in large protein molecules.

3.5.4. Stability Monitoring Using DMS-HDX as a Function of Temperature

As no difference was seen between the gas-phase HDX rates with changing CoV values, an experiment equivalent to the solution phase DSF assay and the TWIMS CIU gas phase stability assay was attempted using DMS-HDX. In theory, it was hypothesized that a more stable ion would unfold at a slower rate in the gas phase compared to a less stable ion. This can be detected using HDX as ions that are more unfolded should have a larger HDX uptake than those that unfold slower. The ion temperature was ramped in the DMS cell from 150 °C to 350 °C at 50 °C intervals and the ion unfolding was measured in the gas-phase through HDX. ND_4OD was introduced into the curtain gas at 10 μL/min along with ACN as a carrier gas at 1.5% and spectra of the tetramer at CoV -6 V were collected in triplicate after a 10 min equilibration time at each temperature. The deuterium uptake of the 32+ charge state was monitored as the DMS temperature was increased and calculated by using baseline spectra in ACN alone (**Figure 33**).

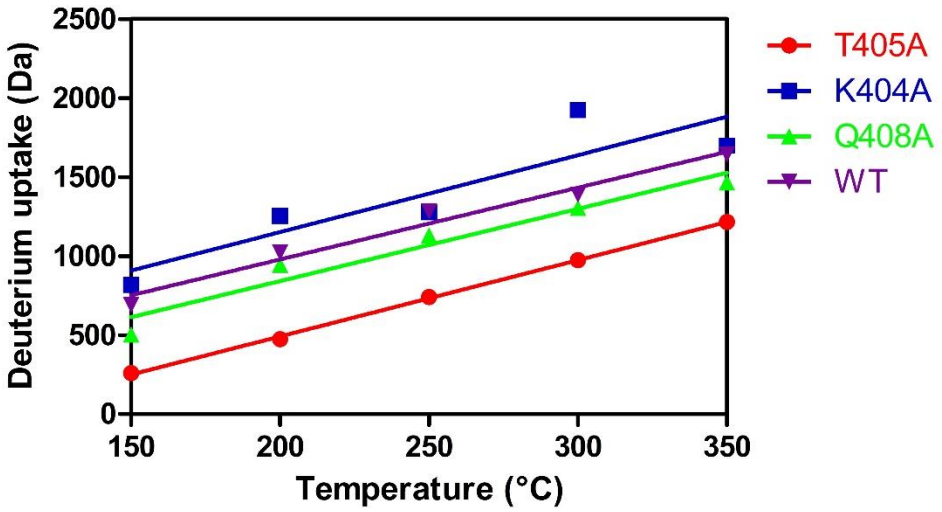

Figure 33. DMS-HDX heat ramp using ND_4OD in the curtain gas. Deuterium uptake rates of the 32+ charge state tetrameric ion at CoV -6 V. Data was acquired in triplicate. Error bars too small to be seen.

Statistical analysis using GraphPad was done to see if there was any statistical significance between the slopes. None of the slopes were significantly different from one another ($P = 0.9089$) which indicates that there was no difference in the rate of deuterium uptake between the PK variants. Again, just as in the last section DMS may not be sensitive enough to capture the changes between the proteins. Another reason as to why no changes may have been seen is that the temperature intervals are too large. As in the DSF assay, data is collected at a heating rate of 0.5 °C/min and from the solution phase data it was shown that the largest difference between the tetrameric protein was about 17 °C. If this correlates over to the gas phase, then the 50 °C temperature intervals were too large to capture a change. If the method was automated, testing smaller temperature intervals may be feasible assuming there will be no aggregation or other issue to disrupt the continuous flow of modifier gas into the DMS cell as this would results in needing more re-equilibration time. Currently, this method required close to 2 hrs to collect 5 data points for all four PK variants and even only doubling the data points to 10 might require a whole day and still not provide low enough resolution to detect changes in deuterium uptake, if any. Therefore, this assay is not a quick way to determine protein stability, however, it would be interesting to see if it may still work using a smaller model protein or peptide variants, which might provide better sensitivity in data.

A second quick heat ramp using only a low temperature (150 °C) and a high temperature (350 °C) was done as a test to see if a quick pulse-label of HDX from the throttle gas might give lower data resolution in terms of detecting slight stability differences between the PK variants. The throttle gas is added from the back of the DMS cell (**Figure 10**) which restricts the HDX reagent to a much smaller area. By adding the HDX in the curtain gas, a more “averaged” data set is collected. By using the throttle gas as a quick pulse-label, only the sites that are highly

labile will have enough time to react with the deuterium and thus showing deuterium uptake changes between highly sensitive sites between all the variants.

Figure 34 shows the deuterium uptake results for three charge states (34+, 33+, and 32+ respectively) when ND_4OD was intruded in the throttle gas at 10 psi. Again, no correlation can be made between the solution phase stability or conformational dynamics of the PK variants as T405A is showing to be the most stable while K404A is showing to be the least stable and most dynamic. No correlation can be made to the CIU results obtained with TWIMS as well. The results do not parallel those that were obtained when HDX was added through the curtain gas. Unlike the other gas phase assays where all charge states exhibited similar behaviours, here the HDX behaviour varies from charge state to charge state and is somewhat difficult to understand which behaviour is the "most correct", if any.

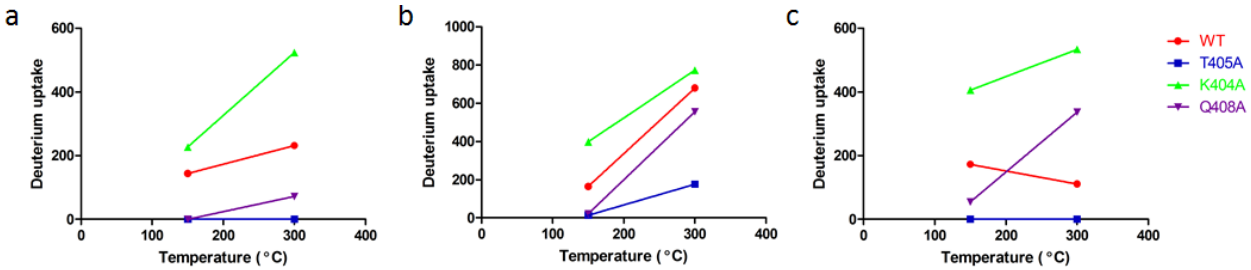

Figure 34. DMS-HDX heat ramp using ND_4OD in the throttle gas. (a) 34+ (b) 33+ (c) 32+ Deuterium uptake rates with no DMS selection. Data was acquired in triplicate.

HDX pulse labeling from the back of the DMS cell did not result in any consistent HDX correlation that could be compared to stability. One interesting result from the gas-phase HDX experiments was that gas-phase conformers just like those reported by Rand *et al.*(2009)[49] were seen. This made it at times, especially for the pulse-label HDX results, to determine the peak centroid as there were a few peaks within one peak. It would be interesting to know if the number of gas phase conformers seen may also somehow represent the dynamics of the protein

ion as perhaps more conformers can show a more dynamic protein ion compared to less conformers. Unfortunately, the data obtained in these experiments with the PK variants was not resolved enough to make a conclusion. Overall, more work needs to be done with these assays to see if gas-phase stability can be detected using DMS and if it correlates to solution stabilities.

Chapter 4: Conclusions and Future Work

4.1. Conclusions

The purpose of this study was to compare solution and gas phase stabilities in hopes that mass spectrometry can become a valuable tool to performing quick stability assays. Overall, despite efforts of others trying to combine solution and gas phase data to show a correlation between stability, here, we present comprehensive analysis that demonstrates that solution phase stability using DSF and gas phase stability using CIU provide identical stability with CIU being sensitive enough to detect T_m changes as low as 2 °C. Comparing the results to the protein dynamics, clear differences between stability and dynamics are seen with no exact correlation between the two. Due to the complex nature of the effect that even single point mutations may have on the protein, dynamics is not a good stability indicator, yet it can provide additional useful information of the protein's function.

With the gas-phase stability results obtained using CIU, it was shown that CIU has the potential of being used as a fast and effective stability indicator requiring low sample concentrations. Information obtained with mass spectrometry provides added benefits of being able to monitor native MS spectra at every CE value which can generate useful information such as fragmentation or disassociation for weakly bound protein-protein or protein-ligand complexes. By implementing IMS, or by selecting ion of interest in the quadrupole, drift times of a singular charge state can be monitored allowing this method to be performed on complex mixtures.

DMS was used as a second ion mobility gas phase technique to try and correlate gas-phase stability with solution phase. Results here were not highly successful as they were with CIU as no distinct correlation could be found between the two techniques despite multiple methods of investigation. Not only was there no distinct correlation found, but the methods

themselves were more laborious which does not make these techniques suited for a quick stability assay.

4.2. Future Work

With CIU showing that both gas and solution phase stability show identical results with the model PK protein variants, it would be interesting to test industry antigens with known solution phase stabilities. Short term DSF and long-term stability results of the industry antigens can be compared to their CIU profiles to see if similar results can be seen with other proteins.

With respect to DMS, much more work can be done to try and see if gas phase stability can still be captured and compared to solution phase. In this study, a very large protein complex was used as we had initial solution phase stability results indicating that there were differences between the PK variants and thus were good candidates to continue with the study. However, DMS is best known to work with small molecules and peptides or if native work on proteins was to be done, small proteins, perhaps less than 10 kDa should be considered to get better signal intensity, resolution, and sensitivity. Peptides with only a single change in the amino acid sequence may be used as a model for such a study and can perhaps give result similar to those found by Zhu *et al* (2016)[26]. Using a smaller model protein or peptide may also lead to being better able to characterize potential gas-phase conformers that were seen with the PK variants. These gas-phase conformers, if well resolved, might be separable using DMS and may also give insight into gas-phase dynamics.

References

1. Dempster, A.J.: A new Method of Positive Ray Analysis. *Phys. Rev.* **11**, 316–325 (1918).
2. Heck, A.J.R.: Native mass spectrometry: A bridge between interactomics and structural biology. *Nat. Methods*. **5**, 927–933 (2008).
3. Konermann, L.: Addressing a Common Misconception: Ammonium Acetate as Neutral pH "Buffer" for Native Electrospray Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **28**, 1827–1835 (2017).
4. Leney, A.C., Heck, A.J.R.: Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass Spectrom.* **28**, 5–13 (2017).
5. Felitsyn, N., Peschke, M., Kebarle, P.: Origin and number of charges observed on multiply-protonated native proteins produced by ESI. (2002).
6. Iribarne, J. V, Thomson, B.A.: Charge and fission of droplets in electrostatic sprays. *II J. Chem. Phys.* **64**, 4977 (1976).
7. Dole, M., Mack, L.L., Hines, R.L., Mobley, R.C., Ferguson, L.D., Alice, M.B.: Molecular Beams of Macroions. *II J. Chem. Phys.* **49**, 2240–2249 (1968).
8. Susa, A.C., Xia, Z., H Tang, H.Y., Tainer, J.A., Williams, E.R.: Charging of Proteins in Native Mass Spectrometry ESI droplet Gas-phase protein ion What limits protein ion charging in native MS? *J. Am. Soc. Mass Spectrom.*
9. Ahadi, E., Konermann, L.: Modeling the Behavior of Coarse-Grained Polymer Chains in Charged Water Droplets: Implications for the Mechanism of Electrospray Ionization. *J. Phys. Chem. B.* **116**, 104–112 (2012).
10. Breuker, K., Mclafferty, F.W.: Stepwise evolution of protein native structure with electrospray into the gas phase, 10 12 to 10 2 s. (2008).
11. Hogan, C.J., Ruotolo, B.T., Robinson, C. V., Fernandez de la Mora, J.: Tandem Differential Mobility Analysis-Mass Spectrometry Reveals Partial Gas-Phase Collapse of the GroEL Complex. *J. Phys. Chem. B.* **115**, 3614–3621 (2011).
12. Czar, M.F., Zosel, F., König, I., Nettels, D., Wunderlich, B., Schuler, B., Zarrine-Afsar, A., Jockusch, R.A.: Gas-Phase FRET Efficiency Measurements To Probe the Conformation of Mass-Selected Proteins. *Anal. Chem.* **87**, 7559–7565 (2015).
13. Seo, J., Hoffmann, W., Warnke, S., Bowers, M.T., Pagel, K., von Helden, G.: Retention of Native Protein Structures in the Absence of Solvent: A Coupled Ion Mobility and Spectroscopic Study. *Angew. Chemie - Int. Ed.* **55**, 14173–14176 (2016).
14. Wyttenbach, T., Bowers, M.T.: Structural stability from solution to the gas phase: Native solution structure of ubiquitin survives analysis in a solvent-free ion mobility-mass spectrometry environment. *J. Phys. Chem. B.* **115**, 12266–12275 (2011).
15. Freitas, M.A., Hendrickson, C.L., Emmett, M.R., Marshall, A.G.: Gas-phase bovine ubiquitin cation conformations resolved by gas-phase hydrogen/deuterium exchange rate and extent. (1999).
16. Eiceman, G.A., Karpas, Z., Hill, H.H.: Ion mobility spectrometry. CRC Press (2016).
17. Lanucara, F., Holman, S.W., Gray, C.J., Eyers, C.E.: The power of ion mobility-mass

spectrometry for structural characterization and the study of conformational dynamics. *Nat. Chem.* **6**, 281–294 (2014).

18. Gabelica, V., Alfonso, C., Barran, P.E., Benesch, J.L.P., Bleiholder, C., Bowers, M.T.: Recommendations for Reporting Ion Mobility Mass Spectrometry Measurements. *ChemRxiv*. **Preprint**, 1–71 (2018).

19. Krylov, E. V, Nazarov, E.G.: Electric field dependence of the ion mobility Non-radioactive Ion Source for Operation in Ambient Pressure View project Differential Mobility Spectrometry View project Electric field dependence of the ion mobility. *Artic. Int. J. Mass Spectrom.* **285**, 149–156 (2009).

20. Shvartsburg, A.A., Smith, R.D.: Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Anal. Chem.* **80**, 9689–9699 (2008).

21. Michaelevski, I., Kirshenbaum, N., Sharon, M.: T-wave ion mobility-mass spectrometry: basic experimental procedures for protein complex analysis. *J. Vis. Exp.* (2010).

22. Ruotolo, B.T., Benesch, J.L.P., Sandercock, A.M., Hyung, S., Robinson, C. V: Ion mobility – mass spectrometry analysis of large protein complexes. *Nat. Protoc.* **3**, 1139–1152 (2008).

23. Smith, D.P., Knapman, T.W., Campuzano, I., Malham, R.W., Berryman, J.T., Radford, S.E., Ashcroft, A.E.: Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. *Eur. J. Mass Spectrom.* **15**, 113–130 (2009).

24. Larry, J., Jc, C., Le Blanc, Y., Kibbey, R.G.: Differential mobility spectrometry: a valuable technology for analyzing challenging biological samples HHS Public Access. *Bioanalysis*. **7**, 853–856 (2015).

25. Levin, D.S., Miller, R.A., Nazarov, E.G., Vouros, P.: Rapid separation and quantitative analysis of peptides using a new nanoelectrospray-differential mobility spectrometer-mass spectrometer system. *Anal. Chem.* **78**, 5443–5452 (2006).

26. Zhu, S., Campbell, J.L., Chernushevich, I., Le Blanc, J.C.Y., Wilson, D.J.: Differential Mobility Spectrometry-Hydrogen Deuterium Exchange (DMS-HDX) as a Probe of Protein Conformation in Solution. *J. Am. Soc. Mass Spectrom.* **27**, 991–999 (2016).

27. Raynal, B., Lenormand, P., Baron, B., Hoos, S., England, P.: Quality assessment and optimization of purified protein samples: why and how? *Microb. Cell Fact.* **13**, 180 (2014).

28. van de Merbel, N., Savoie, N., Yadav, M., Ohtsu, Y., White, J., Riccio, M.F., Dong, K., de Vries, R., Diancin, J.: Stability: recommendation for best practices and harmonization from the Global Bioanalysis Consortium Harmonization Team. *AAPS J.* **16**, 392–9 (2014).

29. Shi, S., Semple, A., Cheung, J., Shameem, M.: DSF Method Optimization and Its Application in Predicting Protein Thermal Aggregation Kinetics. (2013).

30. He, F., Hogan, S., Latypov, R.F., Narhi, L.O., Razinkov, V.I.: High Throughput Thermostability Screening of Monoclonal Antibody Formulations. (2009).

31. Malik, K., Matejtschuk, P., Thelwell, C., Burns, C.J.: Differential scanning fluorimetry: Rapid screening of formulations that promote the stability of reference preparations. *J. Pharm. Biomed. Anal.* **77**, 163–166 (2013).

32. Ruotolo, B.T., Hyung, S.-J., Robinson, P.M., Giles, K., Bateman, R.H., Robinson, C. V, Ruotolo, B.T., Hyung, S., Robinson, P.M., Robinson, C. V, Giles, K., Bateman, R.H.: Protein Structures

Ion Mobility-Mass Spectrometry Reveals Long-Lived, Unfolded Intermediates in the Dissociation of Protein Complexes**. *Angew. Chem. Int. Ed. Engl.* **119**, 8147–8150 (2007).

33. Allison, T.M., Reading, E., Liko, I., Baldwin, A.J., Laganowsky, A., Robinson, C. V.: Quantifying the stabilizing effects of protein–ligand interactions in the gas phase. *Nat. Commun.* **6**, 8551 (2015).

34. Han, L., Hyung, S.-J., Ruotolo, B.T.: Bound cations significantly stabilize the structure of multiprotein complexes in the gas phase. *Angew. Chem. Int. Ed. Engl.* **51**, 5692–5 (2012).

35. Han, L., Ruotolo, B.T.: Traveling-wave Ion Mobility-Mass Spectrometry Reveals Additional Mechanistic Details in the Stabilization of Protein Complex Ions through Tuned Salt Additives. *Int. J. Ion Mobil. Spectrom.* **16**, 41–50 (2013).

36. Tian, Y., Han, L., Buckner, A.C., Ruotolo, B.T.: Collision Induced Unfolding of Intact Antibodies: Rapid Characterization of Disulfide Bonding Patterns, Glycosylation, and Structures. *Anal. Chem.* **87**, 11509–11515 (2015).

37. Dixit, S.M., Polasky, D.A., Ruotolo, B.T.: Collision induced unfolding of isolated proteins in the gas phase: past, present, and future. *Curr. Opin. Chem. Biol.* **42**, 93–100 (2018).

38. Konermann, L., Pan, J., Liu, Y.-H.: Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *This J. is Cite this Chem. Soc. Rev.* **40**, 1224–1234 (2011).

39. Wales, T.E., Engen, J.R.: HYDROGEN EXCHANGE MASS SPECTROMETRY FOR THE ANALYSIS OF PROTEIN DYNAMICS. (2005).

40. Oganesyan, I., Lento, C., Wilson, D.J.: Contemporary hydrogen deuterium exchange mass spectrometry. *Methods*. **144**, 27–42 (2018).

41. Wilson, D.J., Konermann, L.: A Capillary Mixer with Adjustable Reaction Chamber Volume for Millisecond Time-Resolved Studies by Electrospray Mass Spectrometry. *Rapid Commun. Mass Spectrom*. **59**, 6408–6414 (1987).

42. Konermann, L., Pan, J., Wilson, D.J.: Protein Folding Mechanisms Studied by Time-Resolved Electrospray Mass Spectrometry. *Biotechniques*. **40**, 135–141 (2006).

43. Zhu, S., Shala, A., Bezginov, A., Sljoka, A., Audette, G., Wilson, D.J.: Hyperphosphorylation of Intrinsically Disordered Tau Protein Induces an Amyloidogenic Shift in Its Conformational Ensemble. *PLoS One*. **10**, e0120416 (2015).

44. Rob, T., Gill, P.K., Golemi-Kotra, D., Wilson, D.J.: An electrospray ms-coupled microfluidic device for sub-second hydrogen/deuterium exchange pulse-labelling reveals allosteric effects in enzyme inhibition. *Lab Chip*. **13**, 2528 (2013).

45. Donovan, K.A., Zhu, S., Liuni, P., Peng, F., Kessans, S.A., Wilson, D.J., Dobson, R.C.J.: Conformational Dynamics and Allostery in Pyruvate Kinase. *J. Biol. Chem.* **291**, 9244–56 (2016).

46. Deng, B., Zhu, S., Macklin, A.M., Xu, J., Lento, C., Sljoka, A., Wilson, D.J.: Suppressing allostery in epitope mapping experiments using millisecond hydrogen / deuterium exchange mass spectrometry. *MAbs*. **9**, 1327–1336 (2017).

47. Mistarz, U.H., Brown, J.M., Haselmann, K.F., Rand, K.D.: Simple setup for gas-phase H/D exchange mass spectrometry coupled to electron transfer dissociation and ion mobility for analysis of polypeptide structure on a liquid chromatographic time scale. *Anal. Chem.* **86**, 11868–11876 (2014).

48. Rand, K.D., Pringle, S.D., Morris, M., Brown, J.M.: Site-specific analysis of gas-phase hydrogen/deuterium exchange of peptides and proteins by electron transfer dissociation. *Anal. Chem.* **84**, 1931–1940 (2012).

49. Rand, K.D., Pringle, S.D., Iii, J.P.M., Fadgen, K.E., Engen, J.R.: Guide for the Examination of Protein Conformations. **81**, 10019–10028 (2009).

50. Israelsen, W.J., Vander Heiden, M.G.: Pyruvate kinase: Function, regulation and role in cancer. *Semin. Cell Dev. Biol.* **43**, 43–51 (2015).

51. Eschweiler, J.D., Rabuck-Gibbons, J.N., Tian, Y., Ruotolo, B.T.: CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions. *Anal. Chem.* **87**, 11516–11522 (2015).

52. Polasky, D.A., Dixit, S.M., Fantin, S.M., Ruotolo, B.T.: CIUSuite 2: Next-Generation Software for the Analysis of Gas-Phase Protein Unfolding Data. *Anal. Chem.* **91**, 3147–3155 (2019).

53. Haynes, S.E., Polasky, D.A., Dixit, S.M., Majmudar, J.D., Neeson, K., Ruotolo, B.T., Martin, B.R.: Variable-Velocity Traveling-Wave Ion Mobility Separation Enhancing Peak Capacity for Data-Independent Acquisition Proteomics. *Anal. Chem.* **89**, 5669–5672 (2017).

54. Hernández, H., Robinson, C. V.: Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat. Protoc.* **2**, 715–726 (2007).

55. Timasheff, S.N.: Solvent effects on protein stability. *Curr. Opin. Struct. Biol.* **2**, 35–39 (1992).

56. Tehei, M., Madern, D., Pfister, C., Zaccai, G.: Fast dynamics of halophilic malate dehydrogenase and BSA measured by neutron scattering under various solvent conditions influencing protein stability. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14356–61 (2001).

57. Srinivas, V.R., Reddy, G.B., Ahmad, N., Swaminathan, C.P., Mitra, N., Surolia, A.: Legume lectin family, th “natural mutants of the quaternary state”, provide insights into the relationship between protein stability and oligomerization. (2001).

58. Bhanuprakash Reddy, G., Srinivas, V.R., Ahmad, N., Surolia, A.: Molten Globule-like State of Peanut Lectin Monomer Retains Its Carbohydrate Specificity IMPLICATIONS IN PROTEIN FOLDING AND LEGUME LECTIN OLIGOMERIZATION*. (1999).

59. Shelimov, K.B., Clemmer, D.E., Hudgins, R.R., Jarrold, M.F.: Protein Structure in Vacuo: Gas-Phase Conformations of BPTI and Cytochrome c. (1997).

60. Scarff, C.A., Thalassinos, K., Hilton, G.R., Scrivens, J.H., Scrivens, J.H.: Travelling wave ion mobility mass spectrometry studies of protein structure: biological significance and comparison with X-ray crystallography and nuclear magnetic resonance spectroscopy measurements. *Commun. MASS Spectrom. Rapid Commun. Mass Spectrom.* **22**, 3297–3304 (2008).

61. Bornschein, R.E., Niu, S., Eschweiler, J., Ruotolo, B.T.: Ion Mobility-Mass Spectrometry Reveals Highly-Compact Intermediates in the Collision Induced Dissociation of Charge-Reduced Protein Complexes. *J. Am. Soc. Mass Spectrom.* **27**, 41–49 (2016).

62. Niu, S., Ruotolo, B.T.: Collisional unfolding of multiprotein complexes reveals cooperative stabilization upon ligand binding. *Protein Sci.* **24**, 1272–81 (2015).

63. Hyung, S.J., Robinson, C. V., Ruotolo, B.T.: Gas-Phase Unfolding and Disassembly Reveals Stability Differences in Ligand-Bound Multiprotein Complexes. *Chem. Biol.* **16**, 382–390 (2009).

64. Nettleton, E.J., Sunde, M., Lai, Z., Kelly, J.W., Dobson, C.M., Robinson, C. V: Protein Subunit

Interactions and Structural Integrity of Amyloidogenic Transthyretins: Evidence from Electrospray Mass Spectrometry.

65. Byrne, D.P., Vonderach, M., Ferries, S., Brownridge, P.J., Eyers, C.E., Eyers, P.A.: cAMP-dependent protein kinase (PKA) complexes probed by complementary differential scanning fluorimetry and ion mobility-mass spectrometry. *Biochem. J.* **473**, 3159–3175 (2016).

66. Tokuriki, N., Tawfik, D.S.: Stability effects of mutations and protein evolvability. *Curr. Opin. Struct. Biol.* **19**, 596–604 (2009).

67. Tang, K.E.S., Dill, K.A.: Native protein fluctuations: The conformational-motion temperature and the inverse correlation of protein flexibility with protein stability. *J. Biomol. Struct. Dyn.* **16**, 397–411 (1998).

68. Teilum, K., Olsen, J.G., Kragelund, B.B.: Protein stability, flexibility and function. *BBA - Proteins Proteomics*. **1814**, 969–976 (2011).