Investigating The Structural Dynamics of Therapeutic Proteins Using Time Resolved Hydrogen Deuterium Exchange Mass Spectrometry

MARWA ABDALLA ABDELMAKSOUD MOHAMED ELSABAAWY

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

GRADUATE PROGRAM IN CHEMISTRY
YORK UNIVERSITY
TORONTO, ONTARIO

June 2023

Abstract

Biopharmaceutical products have drastically grown as essential treatments in clinical settings. Hence, the drug industry is moving towards protein-based therapeutics with increasing demands for parallel advancement in bioanalytical technologies.

This thesis centers around displaying the versatility of coupling front-end analytical methods to mass spectrometers. We implement time-resolved electrospray ionization hydrogen-deuterium exchange (TRESI-HDX) and ion mobility spectroscopy (IMS) MS for uncovering new and exciting insights into therapeutic protein dynamics and protein-substrate interactions on millisecond timescale. Using automated continuous flow injection (CFI) has enhanced the analytical throughput of our results.

Collectively, the implications for these results widen the fundamental understanding of the studied proteins. Understanding the crosstalk between dynamics and function may help scientists to design new and improved therapeutics.

Acknowledgments

Firstly, I would like to show my deep appreciation to my supervisor, Dr. Derek Wilson, for his support, discussions, and guidance. I would also appreciate my committee members Dr. Audette Gerald, Dr. Mark Bayfield, and Dr. Yi Sheng, for their time and advice.

I would like to thank Dr. Maxime Rossetto for his support at the proteomics lab and his diligent work on the protein digest. A big thank you to Dr. Joseph Anacleto for his help in developing a tool to help with my project. I really appreciate the advice and input for his mass spec support.

Also, Jazz Pharmaceuticals and Mitacs program for giving me the opportunity to explore the industry side of research. I would like to thank my managers at Thermofisher Scientific for their support and for being welcomed and appreciated for the work I have done during my time on site.

Last but not least, my deepest gratitude to my family for their endless support during my studies; I am forever grateful to my parents and siblings.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ALL–Acute Lymphoblastic Leukemia

ASNase– Asparaginase

CCS –Collisional Cross Section

CEM – Chain Ejection Model

CFI–Continuous Flow Injection

CFIT– Continuous Flow Injection with Trap column

CID – Collision Induced Dissociation

CRM – Charged Residue Model

CRAPome– Contaminant Repository for Affinity Purification

CQAS– Critical Quality Attributes

2-D gel–2 Dimensions gels

D_2O- Deuterium Oxide

Da – Dalton

DVT– Deep Vein Thrombosis

E.coli – *Esherichia coli*

ErA– *Erwinia chrysanthemi* L-asparaginase

ESI – Electrospray Ionization

ELISA- Enzyme Linked Immunosorbent Assay

FA – Formic Acid

FDA–Food and Drug Administration

HCP–Host Cell Protein

HDMS –High Definition Mass Spectrometer

HDX – Hydrogen Deuterium Exchange

IEM – Ion Evaporation Model

IgG–Immunoglobulin G

IM – Ion Mobility

kDa – Kilo Dalton

kV – Kilovolts

LC – Liquid Chromatography

LMWH – Low Molecular Weight Heparin

MALDI–Matrix-Associated Laser Desorption

MCP – Multichannel Plate Detector

MW –Molecular Weight

m/z – Mass To Charge Ratio

MS – Mass Spectrometry

MoA– Mechanism of Action

NMR – Nuclear Magnetic Resonance

PAI-1–Plasminogen activator Inhibitor 1

PAS-Proline Alanine Serine

PA200– Proline Alanine 200

PEG– Poly-Ethylene Glycol

PDB – Protein Data Bank

PTS– Post-Thrombotic Syndrome

PMMA–Poly Methyl Methacrylate

QIT – Quadrupole Ion Trap

QTOF – Quadrupole Time of Flight

QQQ – Triple quadrupole

RCL– Reactive Center Loop

RT – room temperature

RF–Radio Frequency

S.C– Subcutaneous

Serpin –Serine Protease Inhibitor

SMB– Somatomedian B Domain

TLC – Thin Layer Chromatography

TRESI – Time Resolved Electrospray Ionization

TOF – Time of Flight

TWIMS – Travelling Wave Ion Mobility Spectrometer

TIC– Total Ion Chromatogram

t-PA– Tissue Type Plasminogen Activator

u-PA– Urokinase Type Plasminogen Activator

UFH– Unfractionated Heparin

UniDec–Universal Deconvolution

UPLC– Ultra Performance Liquid Chromatography

WHO- World Health Organization

VN– Vitronectin

Chapter 1

Introduction

1.1 Mass Spectrometry

1.1.1 Current Perspective

Mass Spectrometry (MS) plays a vital role as a high-throughput analytical technique, from identifying toxins to the design and quality control of pharmaceuticals. The applications are endless. The history of MS was initiated in the 20th century by Wilhelm Wien's discovery that particles with positive charges are deflected using a strong magnetic field. Enlightenment in MS technology was yet revealed by J.J. Thompsons who separated different neon isotopes using electric and magnetic fields.[1, 2] Transferring the analytes into the gas phase, i.e., ionization, is needed to determine the mass. Thanks to the Nobel prize winners Koichi Tanaka and John Fenn, who developed the soft ionization techniques; Matrix-Associates Laser Desorption Ionization (MALDI) and Electrospray Ionization (ESI), expanding the analytical aspects of mass spectrometers to large biomolecules.[3,4] Three main components cooperate to run the mass spectrometer: an ion source, a mass analyzer, and a detector. The sample molecule is converted into gas-phase ions by the ion source. A mass analyzer separates ionized analytes according to their mass-to-charge ratio (m/z). A detector records the number of ions at each m/z value.

1.1.2 Electrospray Ionization

Electrospray ionization (ESI) is a soft-ionization technique in which a high voltage (2-6 kV) is applied to the solution passing through a metal capillary under atmospheric pressure. This results in positive and negative ion formation at the capillary tip.[5] Positive ion mode leads to the accumulation of positive ions at the tip, which gets reduced at the curtain plate. The cone shape, named the Taylor cone, is created just before small droplets overcome surface tension and break off because of increasing coulombic attraction. Lastly, these small droplets are evaporated into the gas phase while traveling through the air in the atmospheric pressure zone (Figure 1).[5,6]

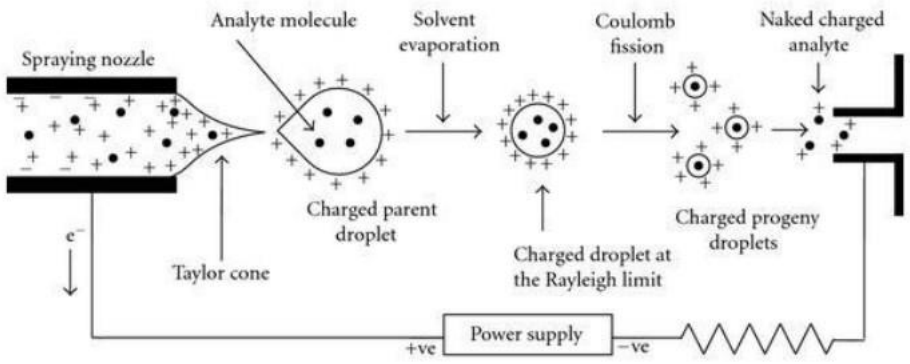


Figure 1. Schematic depiction of an ESI source operated in positive mode. An analyte solution passes through a metal spray nozzle whose tip is held at a high voltage (2-6 kV). The electric potential created at the capillary tip causes the dispersion of the sample solution creating a Taylor cone, resulting in the ejection of positively charged parent droplets due to overpowering columbic forces. During transit towards an oppositely charged curtain plate in the mass spectrometer, the interplay between solvent evaporation and sequential jet fission events results in a reduction in droplet size until naked charged analytes are transferred into the mass spectrometer. Adapted from Banerjee & Mazumdar *et al.,* 2012.[7]

There are three theories explaining how the soft ionization of proteins and other analytes occurs: the charged residue mechanism (CRM), the ion evaporation model (IEM), and the third theory of the chain ejection model (CEM) (Figure 2).[8,9] In the CRM, a charged droplet containing one ion within a liquid droplet is expelled by coulombic fission or completely evaporates from the liquid exterior in a cycle until the ion remains with little to no solvent. The charge of the ion and droplet size must not exceed Rayleigh's limiting charge and can be calculated from Rayleigh's equation.[10]

On the other side, in the IEM model, low molecular weight analytes are ejected from the droplet. Droplets shrink by evaporation until the field strength is large enough for ions to be ejected from the droplet. The energy associated with the large electric field compensates for the energy required for the solvated ion expulsion.

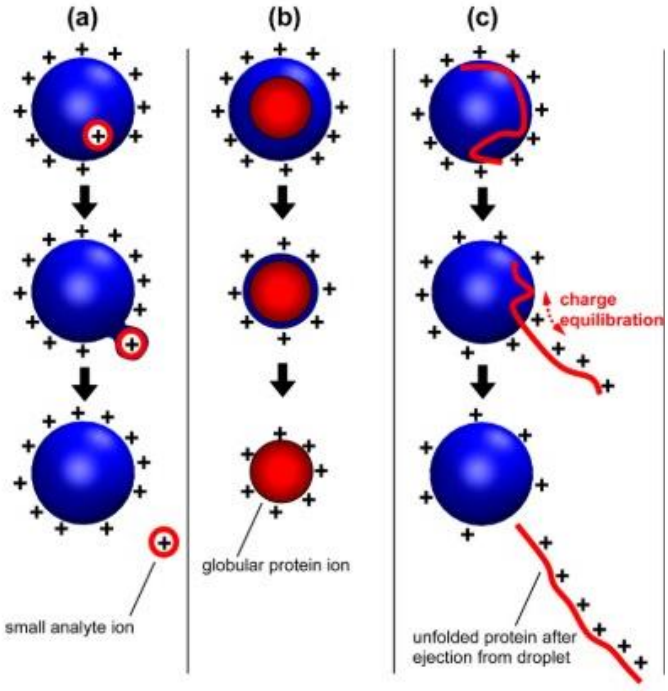


Figure 2. Summary of ESI mechanisms. (a) IEM: Small ion ejection from a charged nanodroplet. (b) CRM: Release of a globular protein into the gas phase. (c) CEM: Ejection of an unfolded protein. Charge equilibration in panel c is indicated by red arrows. Adapted from Konermann *et al.,* 2013.[11]

Finally, CEM is a combination of the original two proposals, in which the macromolecular analyte ionization behavior is dependent on its conformation; species that are hydrophobic or unfolded will show IEM-like ejection, but polymers that are hydrophilic or folded will behave as the CRM model.[12] Native globular proteins are believed to follow the CRM; charged polar residues face the solvent and are more prone to fission events. Smaller analytes follow the IEM driven by hydrophobic interactions. The CEM represents the behavior of proteins under non-native conditions, for example, protein unfolding during electrospray.

1.1.3 Nano-ESI

Conventional ESI has pressure-driven flow rates in the μl/min range under which initial droplet sizes are in the ranges of tens of μm. Thus, the electrospray process takes place in hundreds of ms. Recently, modern electrospray sources incorporate in-line heating, and collateral gas flows to improve sample evaporation in the ionization step.[13] However, potential fractions of analytes are lost because of insufficient time for desolvation during the ESI process. Alternatively, the usage of nano-liquid chromatography and narrow-bore capillary tip emitters lowers the flow rates. The lower flow rates (typically in the nl/min range) result in significantly smaller initial droplet

diameters in the hundreds of nm range and thus ionize faster in tens of μs compared to the hundreds of ms with the conventional ESI.[14] Using these nanospray ionization conditions is beneficial to lower sample consumption, display a better signal response, and have a higher tolerance for sample impurities.[15]

1.1.4 Hybrid Mass Spectrometers

The post-ionization step is based on the type of mass spectrometer thus the type of mass analyzer is the major differentiator between instruments. Mass analyzers separate ions based on their mass-to-charge (m/z), but each mass analyzer has distinctive features in terms of resolution, mass range, speed, and sensitivity. For example, orbitrap, triple quadrupole (QQQ), and quadrupole ion trap (QIT). Hybrid mass spectrometers are a combination of mass analyzers and mass filters.[16]

In this work, the setup used for the experiments is that of Quadrupole time-of-flight (QTOF) mass analyzer for structural MS applications and quadrupole-Orbitrap MS for proteomics. QToF mass spectrometers are robust instruments that are commonly used in protein characterization studies due to their high mass accuracy and sensitivity.[17] QToF instruments are composed of quadrupole 0 (Q0) and Q1 which are in the front for ions focusing and mass selection and transmission. Q2 for parent ion fragmentation at the collision cell, then finally, the ions are passed on to TOF for mass separation and a detector component.[18]

Quadrupoles separate ions based on m/z stability and oscillating electric fields while ions are passing through a radio frequency field between the parallel rods. Ions are focused in Q0 as a beam as it enters the Q1 chamber. Q1 acts as a mass filter and is controlled under certain rf and dc voltages to permit specific m/z transmission window.[18] The ions are then accelerated to the pressurized Q2 chamber where collisional induced dissociation (CID) takes place. Ions are then reaccelerated and focused by optics into the ToF mass analyzer where they are separated by kinetic energy and velocity.[18]
ToF analyzers use electric fields for ions acceleration thus calculating the time taken for the ions to reach the detector. Accelerated ions reach the detector at different times because they have

different m/z ratios. The micro-channel plate (MCP) detector propagates the signal, and the final output is displayed on the computer software as mass spectra.[18]

An additional feature of the QToF is an ion mobility cell that can be positioned directly after the first quadrupole. Selected- mass ions are separated in time, giving detailed information on their size.[17,18]

1.2 Ion Mobility

Ion Mobility (IM) is an analytical technique that separates gas phase ions based on their size and shape. Mass spectrometry cannot separate ions by size, and vice versa,, IM cannot separate ions by mass. Therefore, the combination of ion mobility and mass spectrometry provides highly complementary and additional orthogonal analytical advantages. IMS coupled to MS set-ups can separate ions by both m/z as well as their collisional cross-section area (CCS).[19,20]

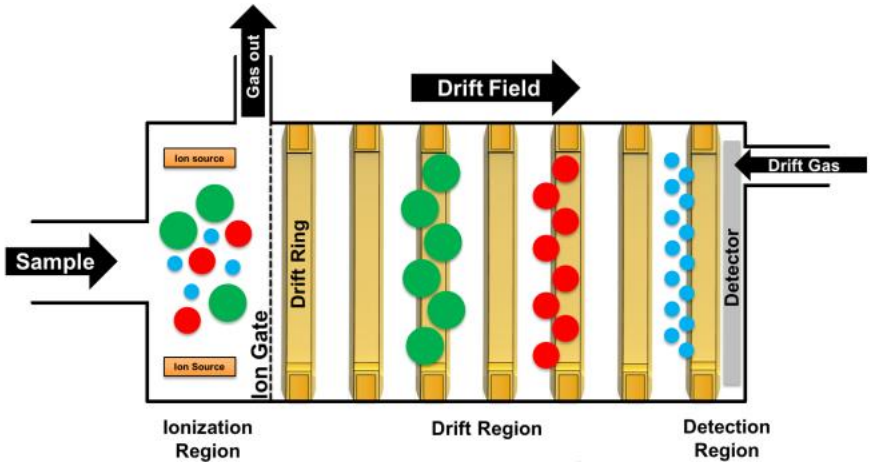


Figure 3. Representation of a drift-tube ion mobility spectrometer. Ionizes analytes (green, blue and red spheres) are trapped for some time behind an ion gate. Once the potential on the gate is dropped, the analyte enters the drift region and undergoes separation, ions migrate at different drift velocities to the detector. Adapted from Cumeras *et al.,* 2015.[21]

A gas-filled cylinder (drift tube) where ions are passed through it under the influence of a weak electric field (Figure 3). Traveling ions interact with a uniform electric field in which ions undergo continuous rounds of acceleration-deceleration events due to collisions with drift gas molecules[19]. While an ion cloud travels between the entry point into the drift tube and the detector, it acquires an average drift velocity (V_d), which is described in the following equation:

$$\mathrm{V}d = {}^{d}/_{t_D} \quad (1)$$

Where td refers to the drift time and d is the distance.

The mobility of an ion (K), which crosses the drift tube under an electric field (E), is related to its drift velocity (V_d).

$$K = {}^{V_d}/_{E} \quad (2)$$

The time of an ion to traverse through the tube is correlated with the ion's collisional cross section (CCS) where some ions will travel faster (smaller CCS), and others will be slower (bigger CCS) because of fewer interactions with the drift gas (fewer collisions). Thus, the ion mobility (K) is usually reported as reduced mobility coefficient (K_0) where the ion mobiles at standard temperature and pressure. Using the following equation:

$$K_0 = \frac{L^2}{t_D V} \times \frac{273.15}{T} \times \frac{P}{760} \quad (3)$$

Where L is the drift tube length, V is the applied voltage across the drift cell, T refers to the gas temperature in kelvin, and P is the gas pressure in torr.[22] Therefore, the mobility of the ion can be converted to its corresponding CCS in the gas phase.[23]

1.2.1 Traveling Wave Ion Mobility Spectrometry

A different type of ion mobility named traveling wave IMS (TWIMS) is comprised of a strong electric field applied sequentially one segment at a time through stacked rings ion guides. Alternating phases of rf voltages are applied to the rings. So, while the electric field passes through the ions, it creates a wave in which the ions can surf .[24] This method was developed by Waters for their Synapt HDMS system (Figure 4).[25] The drift cell in the TWIMS technology consists of a chain of planar electrodes that are perpendicular to the ion transmission path with each adjacent electrode having opposite phases of RF voltage.

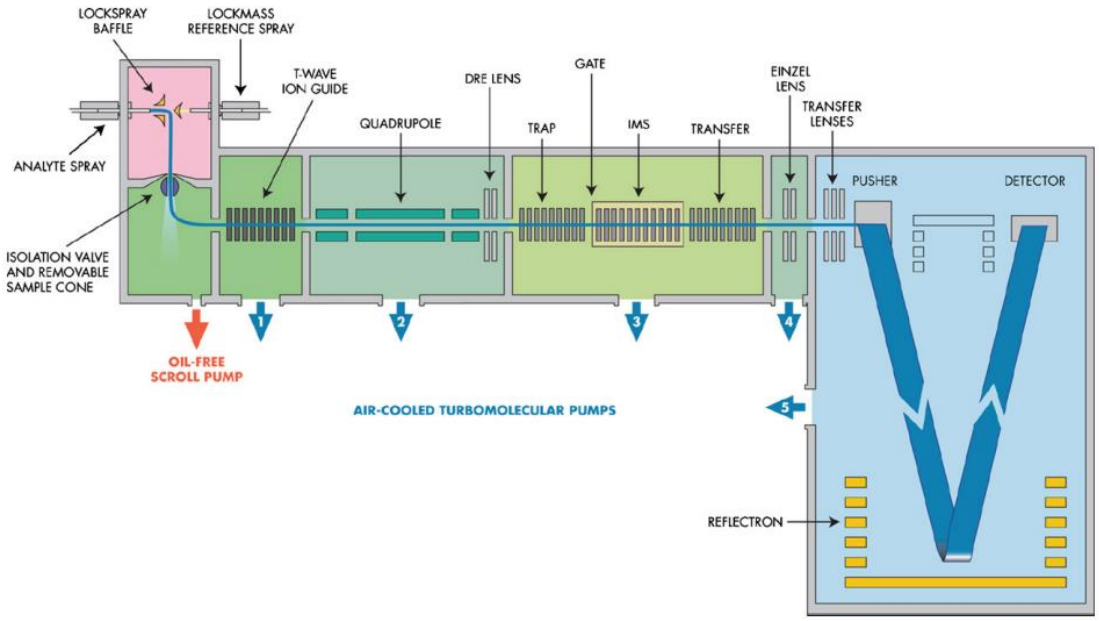


Figure 4. Diagram of the WATERS Synapt HDMS system. The ion mobility cell is between the quadrupole and the time of flight mass analyzer. Adapted from Pringle *et al*., 2007.[26]

To propel ions in waves, TWIMS uses a repeating sequence of transient dc pulses on channels of rings along the length of the device (Figure 5). At regular time intervals, this creates hill motion that is applied to the following pair of electrodes where the ions are moving from the potential hills and carried along the tube with the waves. Therefore, lower mobility ions take longer drift time and vice versa.[25]

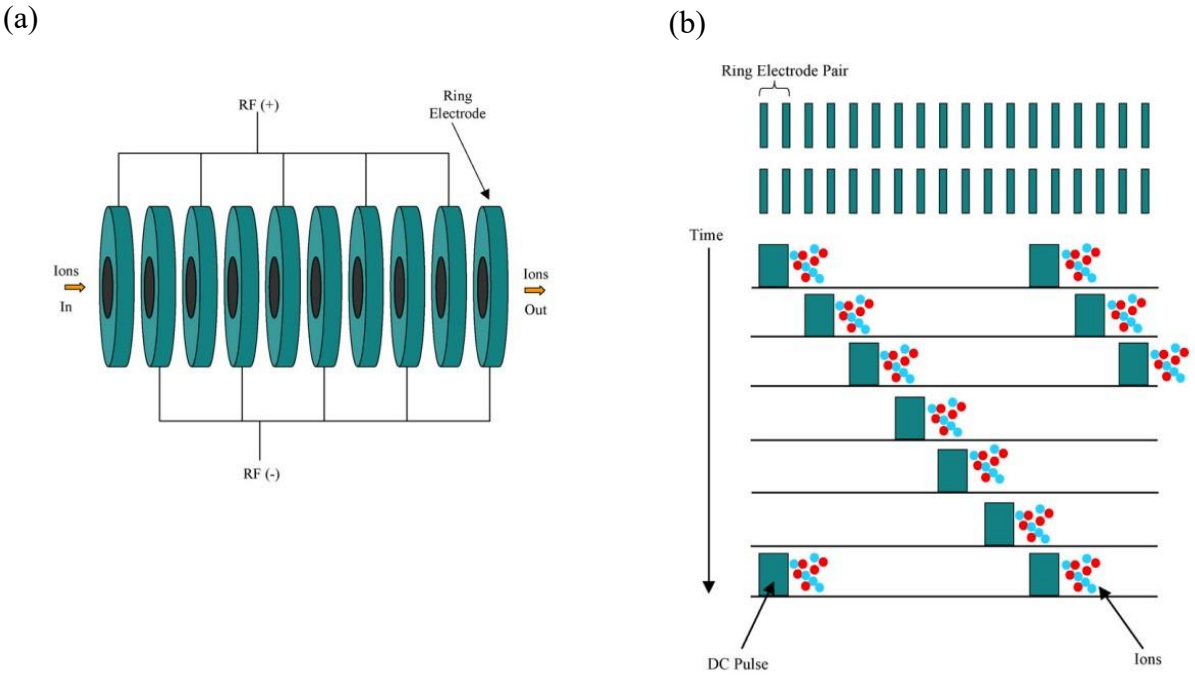


Figure 5. Illustration of traveling wave ion mobility. (a) stacked rings of ion guide and (b) represents ion propulsion in the background gas. Adapted from Pringle *et al*., 2007.[26]

Compared to the conventional IMS, the electric field in TWIMS is not constant and that makes it difficult to calculate Ω from td. Users need to use compounds with known Ω previously derived from conventional IM to calibrate the TWIMS instrument, which subsequently rules the Ω of the ions of interest.[25]

1.3 Hydrogen-Deuterium Exchange Mass Spectrometry (HDX-MS)

Proteins are fundamental biomolecules as they control vital biological processes. However, proteins are dynamic and consistently flexing, with their motions linked to their biological function. Therefore, studying protein dynamics at the molecular level is essential in biochemistry. Several biochemical techniques, such as nuclear magnetic resonance (NMR) and X-ray crystallography have been used to study protein structure in native conditions. However, these techniques have had limitations in studying protein motions. For example, NMR requires high protein concentrations, is limited by protein size and is difficult to use for protein complexes.[27]

To overcome these challenges, Hydrogen deuterium exchange coupled with mass spectrometry (HDX-MS) has become a powerful non-static technique for studying activity-linked dynamic motions in proteins. HDX is a structure-dependent chemical reaction first recorded by Kaj Ulrik Linderstrøm-Lang using ultracentrifugation and further developed by Englander *et al.* in HDX measurements to single amide resolution by NMR.[28] In 1993, HDX was first coupled to LC-MS, providing the protocols for HDX-MS workflow. The principle of Hydrogen deuterium exchange (HDX) is an isotopic exchange reaction in which deuterium (D_2O) replaces a labile hydrogen atom located in the backbone amide of the protein[29] (Figure 6). The reaction is then quenched using a low acidic pH and then digested using pepsin. The correlation between the structural dynamics of a protein and the rate of deuterium uptake is dependent on four main factors:

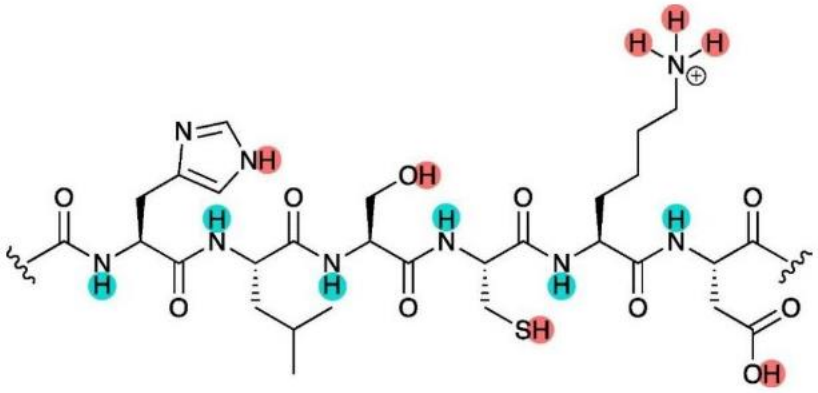


Figure 6. Exchangeable amide and side-chain hydrogens in a peptide. Exchangeable backbone amide hydrogens are depicted in teal, and the labile side-chain protons are depicted in salmon-pink color. Adapted from Oganesyan and colleagues.[30]

Solvent accessibility, intramolecular hydrogen bonding, pH, and temperature, with the latter two controlled experimentally. Rapid isotopic exchange can occur with the protein side-chain, i.e., any hydrogen bound to sulfur, oxygen, or nitrogen. However, for probing subtle changes in higher-order structure, amide hydrogens are the most monitored in HDX experiments[29,31].

The reaction is cooled to 0°C, and the pH is reduced to 2.5; both decrease the rate of amide hydrogen exchange by about 4-5 orders of magnitude compared to normal physiological conditions (pH 7 and at 37°C).

Therefore, these quenching conditions are typically used to minimize the loss of the deuterium label (i.e., back-exchange) in HDX-MS experiments (Figure 7).

1.3.1 Top-down (global) and bottom-up (local) approaches for HDX

HDX experiments can be applied using two approaches: top-down and bottom-up. In the top-down setup, intact protein is detected directly by ESI MS for global deuterium exchange analysis. The results are analyzed as mass increases versus HDX time. The global HDX setup helps give information about the overall flexibility of proteins.

In the bottom-up technique, the protein is digested into peptides by acid-resistant pepsin. This approach does not require specialized equipment, and there is no limitation on protein size. Hydrogen atoms (H) are exchanged with heavier deuterium isotopes (D), resulting in a mass increase of one unit per exchange event, which can be detected as a shift of the mass spectrum. The longer the peptide is, the more exchangeable backbone amide hydrogens. Results are presented as a percentage of deuterium uptake (% uptake) for the number of deuterated backbone amide hydrogen positions relative to the total number of exchangeable hydrogens (# exchangeable hydrogens = # residues - # proline residues -1 for the N-terminal amide)[32,33].

In both approaches, the mass spectrometer tracks deuterium uptake as a function of time, either on the intact protein (global) or individual peptide (local) level. In stable folded protein regions, HDX is affected by the solvent abstraction of the proton.

Folded protein (F_H) must unfold to expose its amide backbone (U_H) in order to undergo hydrogen exchange (U_D) and refold to the deuterated state (F_D).

$$F_H \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} U_H \xrightarrow[D_2O]{k_2} U_D \underset{k_1}{\overset{k_{-1}}{\rightleftharpoons}} F_D$$

The rate of protein unfolding is governed by the rate of protein opening k_1, the rate of refolding occurs at rate k_{-1}, and the chemical rate of hydrogen exchange is k_2. Folded proteins exhibit EX2 kinetics whereby the refolding rate is faster than the rate of hydrogen exchange ($k_{-1} >> k_2$). As the protein refolds faster than deuterium can be exchanged, every instance of protein exposure to deuterium does not necessarily lead to incorporation at the backbone. Therefore, the mass spectrum reports the average population that is capable of exchanging backbone amide hydrogens with deuterium at the time of exposure. As deuteration continues, the spectrum will shift towards higher m/z.

The observed rate of deuterium uptake is controlled by the effects of hydrogen bonding and solvent accessibility. Tightly folded regions are less dynamic, and exhibit reduced uptake than loops or other disordered protein regions.[29,34,35]

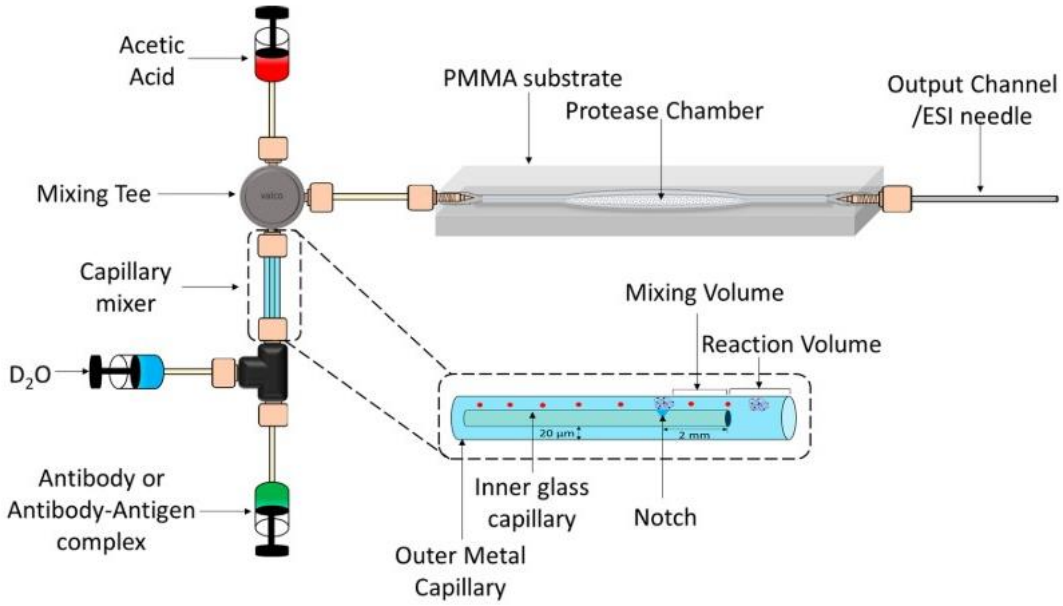


Figure 7. Schematic representation of the microfluidic device used for time-resolved electrospray ionization hydrogen−deuterium exchange mass spectrometry. Gastight syringes are used to transport solutions through the capillaries. Protein samples (antibody or antibody−antigen complexes) are transported through the inner capillary while the D_2O flows through the outer capillary of the time-resolved kinetic mixer where the HDX reaction takes place. A Valco mixing T is used to connect the kinetic mixer with the capillary that transports acetic acid to allow mixing and subsequent quenching of the reaction. The solution is then transported to the protease chamber, where digestion takes place before entry into the mass spectrometer via a metal capillary that serves as an ESI probe. Adapted from Brown and colleagues.[36]

1.4 Therapeutic Proteins

1.4.1. L-Asparaginase in leukemia treatment

Acute lymphoblastic leukemia (ALL) treatment outcome has improved dramatically due to the development of effective treatments and well-designed protocols[37,38]. Among the drugs used is L- asparaginase (ASNase), an enzyme that has been clinically acceptable as an anti-tumor agent. Also, it is a cornerstone in the treatment protocols of ALL.
L-asparagine is an essential requirement for the cells to synthesize protein. Lymphatic tumor cells need high amounts of asparagine to maintain their rapid malignant growth. Therefore L-asparagine is a vital amino acid to growing cancer cells. In contrast, normal cells are independent of their requirement as they can be synthesized in amounts sufficient for their metabolic needs with their own enzyme L-asparagine synthetase.[39]

L-asparaginase catalyzes the hydrolysis of L-asparagine into L-aspartic acid and ammonia. Thus, it deprives tumor cells of an essential growth factor, and they fail to survive. Thus, the using this enzyme as a potent antileukemic drug. Depletion of L-asparagine results in cytotoxicity for leukemic cells. Some brand names for L-asparaginase are RYLAZE, ONCASPAR, ELSPAR, and ERWINASE.[40] The FDA has approved L-asparaginase for acute lymphoblastic leukemia and lymphosarcoma treatment.

1.4.1.1 Proline-Alanine (PA) Long-Acting Technology

A significant challenge of biopharmaceutical proteins is their quick clearance from circulation through kidney filtration, which strongly hampers efficacy in animal studies and human therapy.[41,42] PEGylation is a method in which a chemical coupling with a synthetic polymer poly-ethylene glycol (PEG). This method has emerged as an accepted technology to develop biologics that exercise prolonged action. However, PEGylation shows many drawbacks that hamper its clinical use and bioprocess development, limiting its applicability for drug discovery[43] (i) Loss of bioactivity of the biological (ii) Commercially available activated PEG derivatives can be expensive and have inherent polydispersity associated with difficulties in product analysis (iii) Poor bioavailability of subcutaneously administered PEGylated proteins due to the waxy behavior of highly concentrated solutions. (iv) Growing evidence on the

immunogenicity of PEG and (v) Unnatural PEG polymer lacks biodegradability that can lead to tissue accumulation, such as renal tubular vacuolation.

A biological alternative to PEGylation is the designed technology known as "PASylation" developed by the Skerra group.[43,44,45] They designed genetically encodable three amino acid sequences (Pro, Ala, and Ser) that form long hydrophilic and highly soluble polypeptide chains, thus offering comparable biophysical properties to PEG (Figure 8). PASylation has similar biophysical features compared with PEGylation, strong and tunable PK long-acting effects, and high target-binding activity.[43]

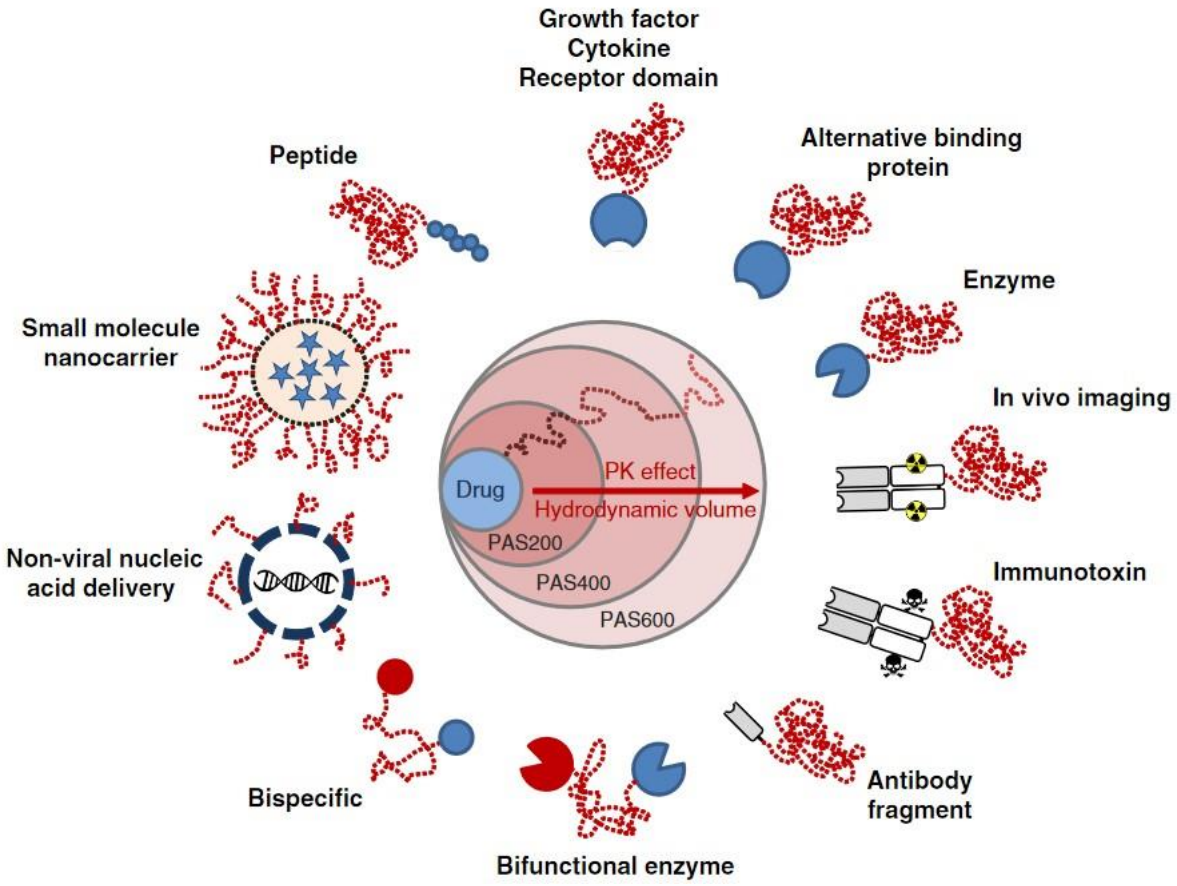


Figure 8. PASylation technology therapeutic application. PAS moiety length increases lead to expanding the hydrodynamic volume of the drug, thus slowing down glomerular filtration. This strategy extends plasma half-life and thus can be used in various classes of biologics ranging from classical recombinant proteins over peptides, alternative binding proteins, enzymes, and

soluble receptors up to antibody fragments for in vivo imaging or targeted tumor therapy. Beyond that, PASylation offers opportunities for bispecific biopharmaceuticals, coupled biocatalysts, as well as DNA and small molecule nanocarriers. Adapted from Binder *et al*., 2017.[46]

1.4.2. Plasminogen Activator Inhibitor in Blood disorders

Plasminogen activator inhibitor 1 (PAI-1) is a glycoprotein that belongs to the serine protease (serpin) superfamily. Serpins have essential regulatory physiological functions such as blood coagulation, fibrinolysis, hemostasis, and inflammation. Serpins have similar overall structural folds and highly conserved proteins. Many serpins inhibit serine proteases, and all inhibitory serpins share a common inhibitory mode of action.[47,48]

Structural studies, mainly X-ray, of serpins and serpin complexes have progressed in recent years, however, the greatest number of studies has been on the uncomplexed states. The structure of serpins is characterized by a solvent-exposed reactive center loop (RCL) or the hinge which forms the binding site for the target protease. The cleavage of P1-P1' bond in the RCL peptide initiates a rapid translocation of the protease to the opposite side of the serpin.[49] While folding, the cleaved RCL is inserted into the core protein and the loop becomes no longer solvent-accessible; the active site of the protease becomes distorted and thus unavailable for interaction with target proteases. This confirmation is called the inactive "latent" conformational state. Serpins usually adapt such latent state in rare occasions only observed by heating, introduced mutations, or when it is exposed to certain buffer conditions.[50,51]

PAI-1 is unique in that it spontaneously inserts its RCL into the body protein without prior proteolytic cleavage of the RCL, thus it adapts a latent form. Therefore, under physiological conditions, PAI-1 is by far the most labile to the confirmational shift into the latent state which occurs over a half-life of 1-2 hours. X-ray diffraction has been daunting to crystallize wild-type PAI-1 in the active conformation. Using 14-1B, an introduced stabilized tetramutant (N150H/K154T/Q319L/M354I) has been used as a stable model for wild-type PAI-1 in structural studies. 14-1B has shown a significant delay in latency transition (half-life 145 h) in comparison to the active wild-type PAI-1, thus increasing the structural stability (Figure 9).[50]

PAI-1 substrates such as the somatomedin B domain (SMB) of Vitronectin (VN) and Mab-1 monoclonal antibody slow the latency transition of PAI-1 thus prolonging its activity. For instance, it has been shown that binding of VN to active PAI-1 slows the latency transition by 1.4 fold. The SMB alone slows the latency transmission of PAI-1 wild type by 3-fold and increases the thermal stability of 14-1B by 20°C. Substrate binding is expected to protect against isotopic exchange by blocking the binding interface however, the protected regions extend beyond the interface. SMB domain binding drastically protects against exchange in peptides spanning s2A-hE, hE-s1A, and the N-terminal half of hF constituting or proximal to the binding interface. Furthermore, SMB protects against exchange in peptides of hA, hB-hC, hD, C-terminal hF-hF-s3A loop, hI, and hI-s5A.

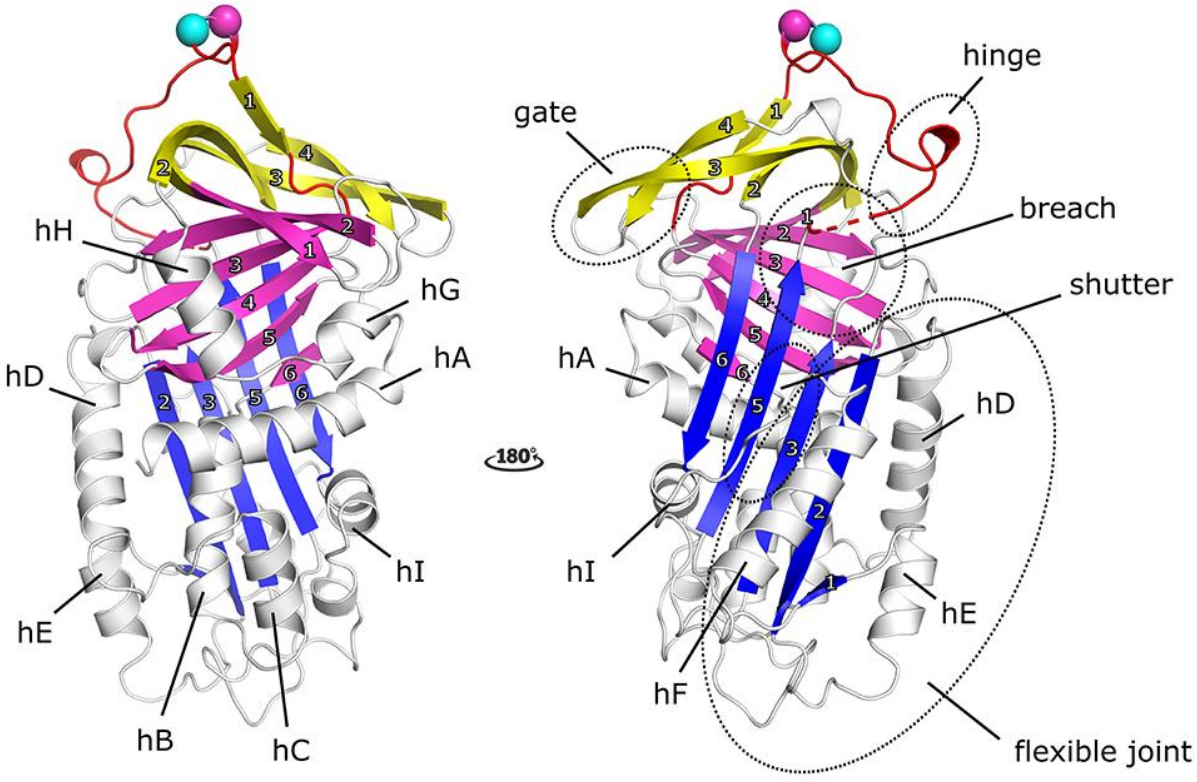


Figure 9. Crystal structure of the active form of plasminogen activator inhibitor-1 (PAI-1) [PDB ID 6ZRV].[52] PAI-1 shows the evolutionarily conserved topology of serpins. The secondary structures (α-helices and β-strands) are indicated in the colors corresponding to three β-sheets (A–C) and nine α-helices (hA-hI). β-sheet A, B, and C are shown in blue, magenta, and yellow, respectively, with numbers labeling the individual strands. The α-helices are indicated in the figure. The reactive center loop

(RCL) of PAI-1 connects strand 5 of β-sheet A (s5A) to strand 4 of β-sheet B (s4B) and comprises strand 1 of β-sheet C (s1C). The RCL is shown in red, with the reactive center Arg346 (P1) and Met347 (P1′) represented by a magenta and cyan sphere, respectively. Other important domains that control and modulate PAI-1 conformational changes (the gate, hinge, breach, shutter, and flexible joint regions) are also indicated. Adapted from Sillen *et al.,* 2020.[51,52]

Therefore, SMB binding confers stabilization on peptides that are positioned at the bottom pole of the molecule.[50,53,54]

The Mab-1 antibody also delays the PAI-1 latency transition, but it was shown that the Mab-1 epitope on PAI-1 is topographically far from the SMB binding sites. Mab-1 antibody binding encompasses E53 in hC, Q56 in hc, and D305 in the hI-B5A loop. Despite the similar effect between substrates, it is not completely understood the mechanism of substrate binding stabilizes PAI-1 and delays the latency transition.[50,55]

1.4.2.1 Heparin and Plasminogen Activator Inhibitor

A major healthcare problem is post-thrombotic syndrome (PTS) which is common in every 2 to 6 patients affected with deep venous thrombosis (DVT). An inflammatory condition that results in thickened and damaged tissue veins. The current standard treatment evolves around using low-molecular-weight heparin (LMWH).[56]

It has been proven that the localized presence of heparin appears to contribute to the non-thrombogenic properties of the vessel wall.[57,58] Thus, treatment with LMWH alters the vein wall response to DVT. In addition, it was observed that LMWH protects against vein wall fibrosis in a PAI-1 dependent manner.[57] Lack of PAI-1 was associated with accelerated venous thrombosis resolution but no protection from fibrosis. The biological mechanisms responsible for this phenomenon are not fully understood.

1.5 Research Objectives

The principal objective is to implement mass spectrometry (MS)-based workflows in emerging areas where MS can fulfill analytical needs. Firstly, Structural MS techniques were carried out as time-resolved electrospray ionization hydrogen-deuterium exchange (TRESI-HDX-MS) to characterize conformational dynamics in modified L-asparaginase drug candidates.

Analyzing charge variants of modified L-asparaginase batches provided by the manufacturer Jazz Pharmaceuticals is critical for characterizing and monitoring the quality attributes of the final product. Using sensitive structural MS techniques will help understand PASylation tag modification's impact on L-asparaginase activity.

Secondly, this work aims to understand the influence of LMWH binding to PAI-1 dynamics. We hypothesize that LMWH governs the PAI-1 stability conformations through protein dynamics. We seek to solve many unanswered questions using HDX. To investigate how LMWH binding increase or decrease the PAI-1 dynamics in regions vital for latency transitions and if this binding can stabilize or destabilize PAI-1 and delay or accelerate latency, The understanding of the PTS pathophysiologic process can potentially lead to the identification of biochemical pathways and molecular targets to allow therapeutic intervention. Lastly, these insights will contribute to elucidating the structure information and the characterization of biological systems with the end goal of biopharmaceutical development.

Chapter 2

Heparin binds to Plasminogen Activator Inhibitor-1 (PAI-1)

2.1 Introduction

2.1.1 PAI-1 Biological Function

Serpin Superfamily consists of over 1,500 inhibitory and non-inhibitory proteins which are broadly distributed in humans, animals, plants, viruses, and bacteria. Eukaryotic serpins have been divided into 16 groups, termed A-P, with clades A-I representing human serpins.[59] (Table 1) Despite Serpins structural similarity, they are functionally diverse. Non- Inhibitory Serpin function as hormone transporters, molecular chaperons or tumor suppressors, however the majority of serpins are inhibitors.

Plasminogen activator inhibitor type-1 (PAI-1) belongs to the serpin superfamily and is one of the most important inhibitors of the plasma fibrinolytic system. The inhibitory properties of PAI-1 make it crucial to maintain a strict balance between fibrinolysis and coagulation in the plasminogen activator system. Bleeding happens when plasminogen is activated to plasmin via its two activators urokinase type plasminogen activator (u-PA) and tissue-type plasminogen activator (t-PA). The activity of both is controlled by specific plasminogen activator inhibitors (PAIs). Three types of PAIs are PAI type I, or called the endothelial cell PAI (clade E serpin), PAI type 2 (PAI-2) and known as a placental-type PAI (clade B), and PAI type 3 (PAI-3) (clade I).[60,61] Among these types, the rapid acting PAI-1 inhibits the plasminogen activators and controls the active plasmin levels therefore stopping bleeding and subsequent clot breakdown. PAI-1 is an acute response protein that is secreted at low concentrations (5-20 ng/ml), and it has a fast turnover in blood.

Excessive inhibition of PAI-1 expression shifts the balance towards extreme bleeding while elevated levels of PAI-1 can result in occluded blood vessels and a thrombotic state. Imbalance of PAI-1 concentrations is associated with pathological conditions including thrombosis, cardiovascular diseases and angiogenesis and cancer progression.[51] (Figure 10)

PAI-1 is glycoprotein of 379 amino acid residues with an apparent molecular weight of 48 kDa. It is produced by a variety of cells, once synthesized it is mainly stored in platelets and when needed it gets secreted to extracellularly (to blood flow) or deposited in the subendothelial

matrix. Circulating PAI-1 is released as active form or associated with a co-factor named Vitronectin (VN) as a tight 1:1 complex. VN is a thermostable glycoprotein that stabilize PAI-1 in the active form. Studies have found that VN and PAI-1 dysregulation have been implicated in a number of overlapping disease states.

2.1.2 PAI-1 Inhibitory Mechanism

Serine protease inhibitor (Serpin) proteins are essential inhibitory proteins in multiple physiological conditions. All Serpins share a common inhibitory mode of action that uses spontaneous conformational change to control inhibitory activity. They specifically cleave peptide bonds by a thermodynamically favorable hydrolysis reaction.

PAI-1 is a well conserved molecule that has 16 total strands organized into three β-sheets (named A-C, with enumerated strands termed as s(#)A, s(#)B, and s(#)C) and nine α-helices (termed hA-hI). PAI-1 carries a surface-exposed reactive center loop or RCL protruding from the top of the molecule. This loop region carries residues that are recognized as a "pseudo" substrate by the target proteases (uPA or tPA). RCL has 26 residues long (331-356, designated as P16-P10') which act as a substrate mimicking peptide sequence (Arg 346-Met 347 designated as P1-P1').[60]

Initial binding of the proteases (enzyme, E) to RCL part of the PAI-1 (inhibitor, I) forms a non-covalent 1:1 Michaels Complex (EI). Following the formation of this docking complex, the P1-P1' bond is cleaved to generate an acyl-enzyme intermediate (E~I). Following this intermediate complex, a rapid and full insertion of the amino-terminal portion of the RCL inserts itself into the center of β sheets A to form an extra strand (fourth, s4A). This major conformational change leads to a 70^{o}A translocation of the protease to the opposite side of the serpin. In this final serpin-proteases complex, the enzyme is trapped and severely disordered resulting in the formation of an irreversible inhibitory complex (E-I).[51] (Figure 10) PAI-1 inhibits proteases rapidly in a reaction rate constant between 10^6 and 10^7 M^{-1} S^{-1}.

Table 1: Types of Human Serpin

Serpin	Alternative name(s)
SERPINA1	Antitrypsin
SERPINA2	Antitrypsin-related protein
SERPINA3	Antichymotrypsin
SERPINA4	Kallistatin (PI4)
SERPINA5	Protein C inhibitor (PAI-3)
SERPINA6	Corticosteroid-binding globulin
SERPINA7	Thyroxine-binding globulin
SERPINA8	Angiotensinogen
SERPINA9	Centerin
SERPINA10	SERPINA10 Protein Z-dependent proteinase inhibitor
SERPINA11	XP_170754.3
SERPINA12	Vaspin
SERPINA13	XM_370772
SERPINB1	Monocyte neutrophil elastase inhibitor
SERPINB2	Plasminogen activator inhibitor-2 (PAI2)
SERPINB3	Squamous cell carcinoma antigen 1
SERPINB4	Squamous cell carcinoma antigen 2
SERPINB5	Maspin
SERPINB6	Proteinase inhibitor-6 (PI6
SERPINB7	Megsin
SERPINB8	Cytoplasmic antiproteinase 8 (PI8)
SERPINB9	Cytoplasmic antiproteinase 9 (PI9)
SERPINB10	Bomapin (PI10)
SERPINB11	Epipin
SERPINB12	Yukopin
SERPINB13	Heasdpin (P13)
SERPINC1	Antithrombin
SERPIND1	Heparin cofactor II
SERPINE1	Plasminogen activator inhibitor 1
SERPINE2	Protease nexin I (PI7)
SERPINE3	Hs.512272
SERPINF1	Pigment epithelium
SERPINF2	Alpha-2-antiplasmin
SERPING1	C1 inhibitor
SERPINH1	47kDa heat-shock protein
SERPINI1	Neuroserpin (PI12)
SERPINI2	Myoepithelium-derived serine proteinase inhibitor (PI14)
SERPINB13	Headpin (PI13)

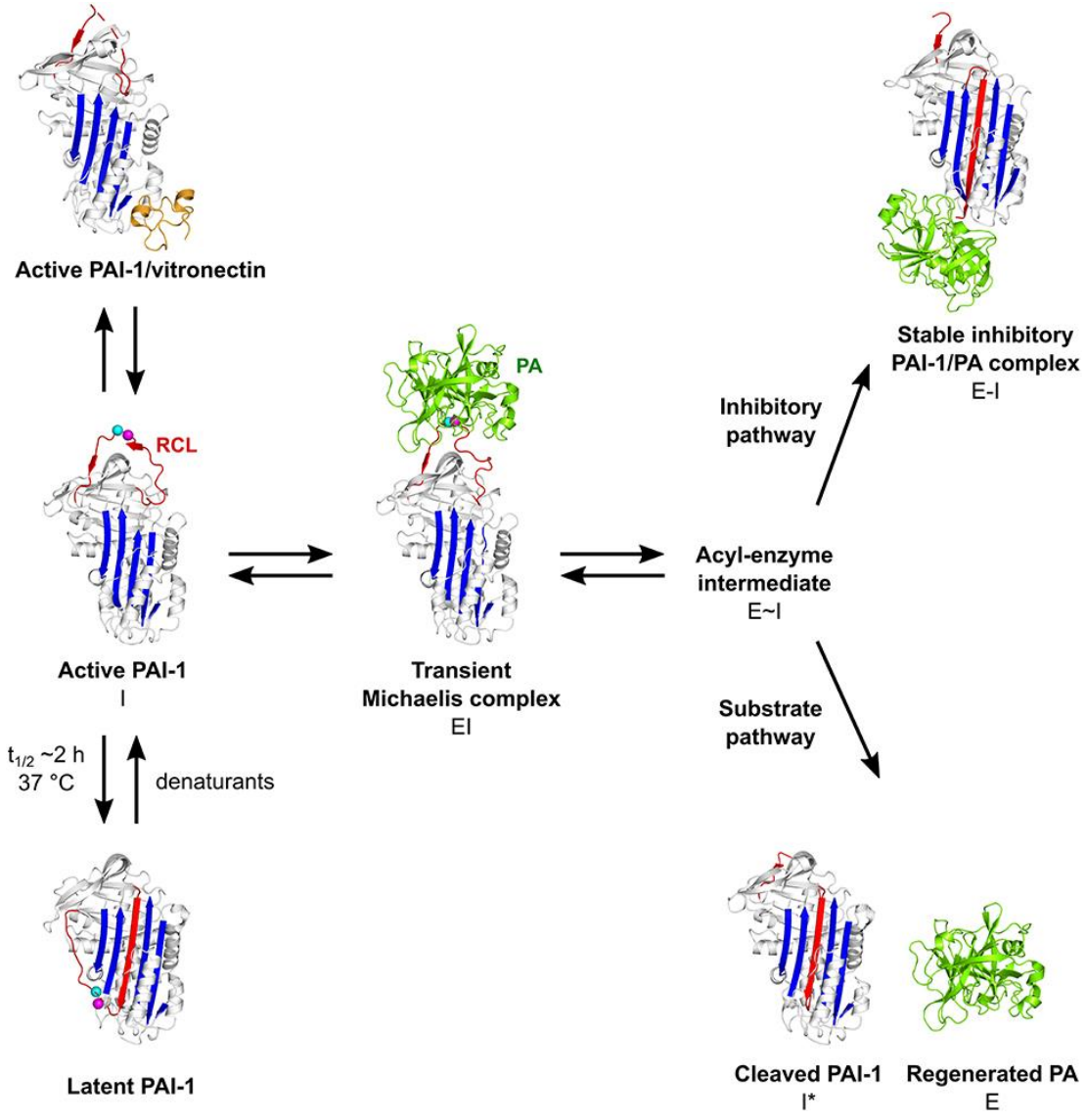


Figure 10. An overview scheme for the PAI-1 conformations. PAI-1 and its interactions with plasminogen activators (PAs, E) and cofactor vitronectin. Following the formation of a non-covalent PAI-1/PA Michaelis complex (EI), the P1-P1′ bond is cleaved to generate an acyl-enzyme intermediate (E~I). Next, two possible pathways take place: either the formation of an irreversible inhibitory complex (E-I) or the generation of cleaved PAI-1 (I*) due to the hydrolysis of the acyl-enzyme intermediate. PAI-1 is shown in white; the central b-sheet A of the PAI-1 molecule in blue; the flexible reactive center loop (RCL) in red, and Arg346 and Met347 (P1-P1′) of the reactive center are indicated by magenta and cyan spheres, respectively. The PA is shown in green. Vitronectin is shown in orange. PDB structures 1DVN (45), 1DB2 (46), 5BRR (47), 3EOX (48), 1EZX (49), 1H4W (50), and 1OC0 (51). Figure adapted from Sillen et al.[51,52]

The following regions have been defined as important spots in modulating PAI-1 conformational changes; the hinge region (P15-P9 of the RCL), the breach region (top of β-sheet A), the shutter domain (central parts of s3A, s5A, and N-terminal of hB), the gate region (s3C and s4C) and lastly the flexible joint region (hD, hE, hF, and s1A).

This PAI-1 conformational transition is known as the stressed (S) to relaxed (R) transition since the RCL cleavage of native inhibitory serpins leads to the formation of a thermodynamically favorable stable conformation.[62] That makes the serpins consider suicide or irreversible inhibitors. Therefore, native serpins are trapped in an intermediate metastable state rather than their most stable state conformation.

The rate of the RCL insertion (K_{lim}) into the protein body is a rate limiting step to determine the pathway of the acyl-enzyme intermediate (E~I). If the rate of peptide insertion is very slow, intact protease can escape the conformational trap post RCL cleavage, leaving as active protease. In later case, the PAI-1 will exhibit substrate-like pathway.[62] RCL insertion rate can be influenced by the pH, temperature, presence of co-factors, mutations and other external conditions.

2.1.3 PAI-1 latency transition

PAI-1 has the tendency to spontaneously insert its RCL into the protein body, without prior proteolytic cleavage of the RCL. The loop becomes no longer solvent accessible and becomes unavailable to target protease whereby it adapts to a latent or inactive conformation state. Under physiological conditions, this transition occurs over a half-life of 1-2 hours. Such spontaneous transition to the latent state happens rarely in the serpin family and was observed over a longer time span in only few serpins. That makes PAI-1 a unique molecule which uses spontaneous conformational change as a regulatory mechanism to shut down its inhibitory activity (automatic shutdown).

Structural studies are available for both active and latent conformations, however this transition is not completely understood. Many important regions of PAI-1 are communicating in an orchestrated manner to achieve this latency transition. [63]

Latency transition starts when a strand (s1C) (located within the gate region) becomes detached. Next, the gate region (s3C, s4C, s3B) will open up to allow the uncleaved RCL to pass through where RCL is partially inserted (up to P11) at the central of β sheet A (the breach and central part of the shutter regions). Opening of the later regions helps to facilitate loop insertion, forming a fourth beta strand (s4A). Concurrently, hA helix is transit change to accommodate this conformation change.[63,64] This is so called the pre-latent state, which is proven to co-exist with active PAI-1 in solution. The molecule can exist in the pre-latent state for an extended period of time, after which the full insertion of the RCL will continue. However, the full insertion of the remaining P6-P4 region of the RCL is blocked by the lower part of the shutter region(s3A) and helix F (hF) loop which overlies β-sheet A, posing a high energy barrier. It was found that hF temporarily unfolds to allow RCL full insertion. Lastly, hF returns to its native position overlies β-sheet A and irreversibly locks PAI-1 in its latent state. This process is irreversible under physiological conditions. Unlike experimentally, the inhibitory properties can be reversed by treating the latent PAI-1 with denaturants or negatively charged phospholipids.

Calorimetric studies show the enthalpy of the molecule is the same during its latency transition, therefore the PAI-1 spontaneous conformational change to latent state is entropically driven. Unlike the inhibitory reaction with proteases, the latency process is energetically silent.

2.1.4 Diverse approaches to control PAI-1 Levels

Hydrogen-deuterium exchange (HDX) studies have found that active PAI-1 is highly dynamic which plays a key role in the latency conversion. How long the PAI-1 molecule remains active prior converting to the latent state is controlled by many factors.

Firstly, in vivo stabilization by PAI-1 interaction with its endogenous ligands. Glycoprotein vitronectin (VN) is secreted from platelets and is abundant in plasma and extracellular matrix. It interacts with active PAI-1 to direct circulating PAI-1 to blood clot for thrombolysis. This interaction happens with high affinity association (K_D~0.1-1 nM), once bonded; the active PAI is stabilized at least 2-folds and delayed to its latency transition.[58,65] The N-terminal somatomedin B domain (SMB) of vitronectin is the primarily site for this binding to happen where it binds through the flexible joint region (hE, hF and s1A). Later, the X-ray structure of the SMB bonded

to PAI (PDB 1OC0) revealed more specific interaction interface. However, studies found that VN is still able to bind to PAI-1 when SMB domain is explicit, but the binding has lower affinity (K_D~50 nM).

VN stabilizes the lower half of the PAI-1 which causes allosteric modulation in multiple regions away from the SMB binding. Altogether, reduce the structure flexibility of PAI-1 and slow down the rate of RCL sliding movement into the molecule body. Consequently, this delays its latency and increases its active state half-life. Another protein called α_1-acid glycoprotein was found to stabilize the active PAI-1. The binding takes place in different regions of VN binding regions. This interaction is less stable, but it adds more understanding of PAI-1 during acute phase reactions.

Secondly, inducing random and site-directed mutagenesis increases PAI-1 stability. These mutants are mainly created to gain insights into the structure/function relationship in PAI-1 and to investigate its regulatory mechanism. Using these stabilized active mutants has essentially helped to determine the PAI-1 structure conformations (active, latent, and cleaved PAI-1) and its complex with biological ligands. The first mutant to be successfully crystallized in the active confirmation is the PAI-1 14-1B quadruple variant (N150H, K154T, Q319L, M154I). Another mutant is the PAI-1-stab which has an extra fifth mutation (N150H, K154T, Q301P, Q319L, M154I). Lastly, the recent crystal structure of active PAI1- W175F that is singly mutated in the N-terminal hinge of prior to the RCL. Single substitution induces moderate structure stabilization however the combination of many mutations results in significant stability with extended half-lives up to 450 h. The tertramutant version (14-1B) has extended the active PAI-1 half life to 145 h and increased its thermal stability compared to active wildtype PAI-1.[50]

Finally, some external conditions during recombinant PAI-1 protein purification can affect its rate of latency transition. Temperature, pH and ionic are thought to play a key role. For example, PAI-1 is more stable in the active form as pH drops from the physiological pH to pH 5. Decrease in the pH results in imidazole protonation which pinpoints that histidine residues might be responsible for PAI-1 pH dependent stability. Salt stabilization was observed as an anion-binding site in the tertramutant (14-1B) crystal structure. However, this phenomenon could not be observed in the recent PAI—W175F. Thus, it leads to a controversial hypothesis of the salt effect

on latency. Besides, studies have shown that type I metal ions (Ca^{+2}, Mg^{+2}, and Mn^{+2}) and type II (Co^{+2}, Cu^{+2}, and Ni^{+2}) can help to stabilize or destabilize the PAI-1 in presence or absence of VN, simultaneously.

2.1.5 PAI-1 inhibitors and heparin

PAI-1 plays a pivotal role in many pathophysiological conditions; therefore, research efforts have developed some PAI-1 inhibitors. Developing PAI-1 inhibitors have been extensively studied in efforts to translate it to human usage but the efforts continue. On the other hand, some marketed therapeutics have been shown to decrease PAI-1 plasma levels.

Pronounced effects on PAI-1 molecule have been witnessed using small molecules, RNA aptamers, monoclonal antibodies, and synthetic peptides. The developing PAI-1 inhibitors focus on three mechanisms to inactivate the molecule; directly, by blocking the initial interaction between PAI-1 and proteases, or indirectly by either blocking the final inhibitory complex (PAI-1 substrate like behavior) or by accelerating PAI-1 transition from its active to latent state.

For instance, antibodies (MAb-2,6, 33HIF7) binding epitopes facilitate latency transition by binding on the hF region and disrupt hF translocation to form a substrate like form of PAI-1. Likewise, MAb-8H9D4 blocks the binding site of the proteases by binding to the PAI-1 shutter region and render the initiator in a substrate like fashion. Additionally, some RNA aptamers were developed to interfere between PAI-1 and its binding partners. R10-4 and R10-2 interfere with PAI-1/tPA complex only but not with PAI-1/uPA.

Since PAI-1 is the main plasminogen activator inhibitor, it plays a key role in the pathophysiology of vascular endothelial cells. In healthy young adults, PAI-1 may be a marker of cardiovascular risk factors. As mentioned earlier, low PAI-1 is associated with bleeding tendency while raised levels lead to thrombosis. In case of emergency thrombolytic events, treatment with low-molecular weight heparin is initiand. Heparin works as an anti-coagulating agent and to alter veins wall response to thrombus. It exerts its effect by accelerating the inhibitory binding of antithrombin to its target proteases (thrombin and factor Xa) and by transferring tissue factor pathway inhibitor from vascular endothelium into the blood. However,

several clinical studies have reported that heparin can also interact with other proteins that are relevant to its pharmacological effect.

A significant decrease of PAI-1 plasma level was observed after LMWH was administered for patients with hemodialysis. This decrease was correlated with increasing the heparin dose and concluded an enhanced anti-coagulation activity. Moreover, in deep-vein (DVT) thrombotic induced animal model, researchers have found that treatment with LMWH protect against vein wall fibrosis in a PAI-1 dependent manner. It concluded that inhibition of PAI-1 in DVT is a potential treatment strategy to accelerate the clearance of thrombus. In addition, it was found that unfractionated heparin (UFH) added to human endothelial cells decreases the endothelial PAI-1 expression.[66,67,68] Hence, the ability of heparin to reduce PAI-1 can be an additional mechanism to prevent thrombosis.

Despite the potential inhibitory effect of heparin on PAI-1, the mechanism remains unclear. Revealing how both molecules interact is an attractive therapeutic approach for cardiovascular diseases. To understand these complexes, we need to reveal the structure plasticity of PAI-1 interacting with heparin. In this dissertation, we further probe the interaction between heparin and PAI-1 on the molecular level.

2.2 Experimental Methods

2.2.1 Reagents and 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. Recombinant human active wild-type PAI-1 and recombinant human active PAI-1 stabilized by four mutations (i.e., N150H, K154T, Q319L, and M354I) were obtained from Molecular Innovations. Fondaparinux is a synthetic heparin pentasaccharide obtained from Amsbio.

2.2.2 Samples preparation

Protein samples were diluted to a final concentration of 10 μM in 100mM ammonium acetate for practical use. Active wildtype PAI-1 and active tetramutant clone 14-1B of PAI-1 were subjected to isotopic exchange of amide hydrogen for deuterium in the presence or absence of a 2-fold molar excess of Fondaparinux. All exchange reactions were conducted at 4 °C where the conversion of active PAI-1 to latent PAI-1 is very slow (t1/2 = 12 h), and the amount of active PAI-1 converted to latent PAI-1 during the 2 h HDX experiment is therefore negligible. PAI-1 stock solutions with Fondaparinux (10 μM PAI-1 and 20 μM Fondaparinux) and without Fondaparinux (10 μM PAI-1) were prepared and allowed to incubate for 30 min on ice.

2.2.3 Materials for mass spectrometry experiments

Pepsin, deuterium oxide (D_2O), ammonium acetate (≥99.0%), and high purity acetic acid (≥99.7%) were purchased from SigmaAldrich.

2.2.4 Pepsin Preparation

Pepsin from porcine mucosa and NHS-activated agarose beads (2:5 w/w) were resuspended in coupling buffer (0.1 M sodium phosphate, 0.15 M NaCl, pH 4) and rotated for 17 hours at 4°C. The unbound pepsin was aspirated, and beads were resuspended in a blocking buffer (1 M Tris–HCl, pH 4) and then rotated at room temperature for 1 hour. The beads were washed five times with 10% acetic acid (v/v) via centrifugation at 1000 x g for 2 minutes and stored at 4°C.

2.2.5 TRESI-HDX and updated version (CFIT)

The microfluidic device used for time-resolved ESI (TRESI) measurements was built as previously described. [30] Briefly, a VersaLaser™ was used to etch a proteolytic chamber into a poly(methyl methacrylate) (PMMA) surface measuring 5.1 cm × 2.0 cm × 1.2 cm. A rapid-mixing device was made by inserting an inner glass capillary (outer diameter of 151.0 μm) inside a 28-gauge metal capillary (inner diameter of 178.8 μm), creating an intercapillary space of 27.8 μm. The end of the glass capillary was sealed using the VersaLaser, and a notch was cut 2 mm from the sealed end. Hydrogen deuterium exchange (HDX) labeling reactions were carried out in the volume between the inner and outer capillaries. For localized measurements, the TRESI mixer was connected to an acid channel through a valco t-mixer, followed by on-chip proteolytic digestion. The updated version of TRESI-HDX-MS was developed by Joseph Anacleto[69]. Automated Continuous Hydrogen Deuterium Exchange Mass Spectrometry and an added trap column (CFIT), details in (Appendix A, Figure A1).

2.2.6 Data Analysis

Peptide sequences were confirmed via MS/MS with analysis in mMass. Deuterium uptake per peptide was calculated using Mass Spec Studio[70] by matching experimental deuterium uptake to the theoretical isotopic distribution for the native peptide sequences. Changes in deuterium uptake were summed across 3 time points (labeling times of 1, 15, and 45 seconds) and averaged across technical triplicates from two different preparations (total n=3 replicates). Changes in deuterium uptake were considered significant if they exceeded one time the standard deviation (1σ). Protein structures were shown using PyMol v1.5.0.4 (Schrödinger, LL).

2.3 Results

2.3.1 Investigating Active Mutated PAI-1 Local Flexibility

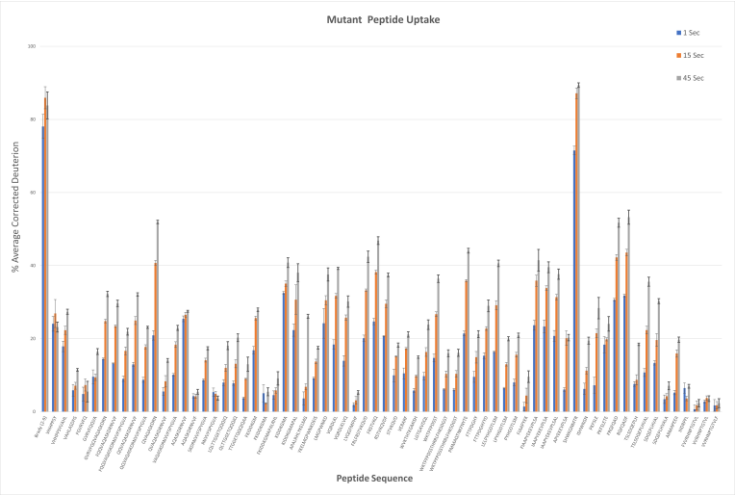
The backbone amide hydrogens exchange with deuterons was done by dilution of mutated type PAI-1 into deuterated buffer at 4 °C. After 1,15, and 45 sec, the exchange reaction was quenched, and the protein digested with pepsin. Subsequently, the peptic peptides were analyzed with MS, and deuterium contents were calculated and plotted against the exchange time (Figure 11).

We find that some PAI-1 peptides exhibit high relative deuteration levels after exchange for just 15 sec. These peptides represent hB, hC, hD, and hE (30-40, 46-59, 79-92, and 114-123 respectively). This is illustrated by peptide 46-59 covering the shutter region. Thus, these regions have relatively weak structural protection against isotopic exchange in the active conformation of PAI-1 (Figure 11).

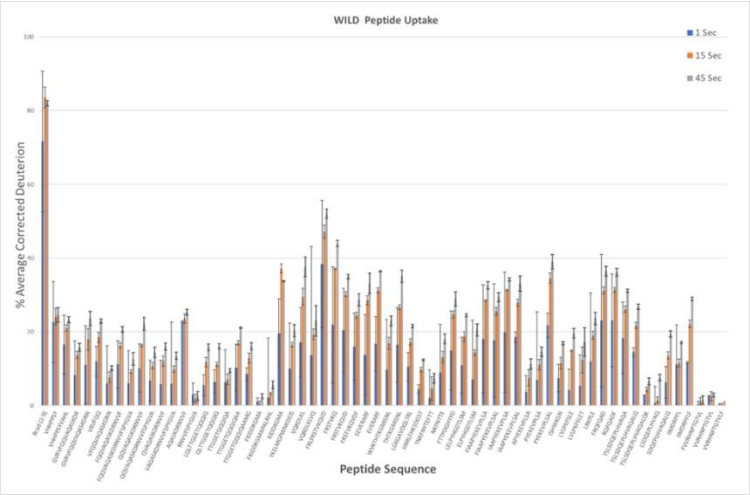
Fast exchange kinetics were also observed at the proximal region of β3A and β5A. These peptic peptides represent the proximal parts of the shutter region as illustrated by peptide 160-170 and 300-320. Likewise, peptide 200-212 and 239-265, which comprises the Gate region in PAI-1 structure (β3C and of β4C) show a fast exchange rate after exchange for 5 sec. These observations highlight substantial structural flexibility in the shutter, Flexible joint, and gate regions of the PAI-1.

The local HDX properties of wild PAI-1 were also explored (Figure 11). Most peptides were more exposed to isotopic exchange in the wild form than in the corresponding mutant form. Exception is peptide (299-306) and (310-322) were more protected against isotopic exchange in the wild than in the mutant form of PAI-1.

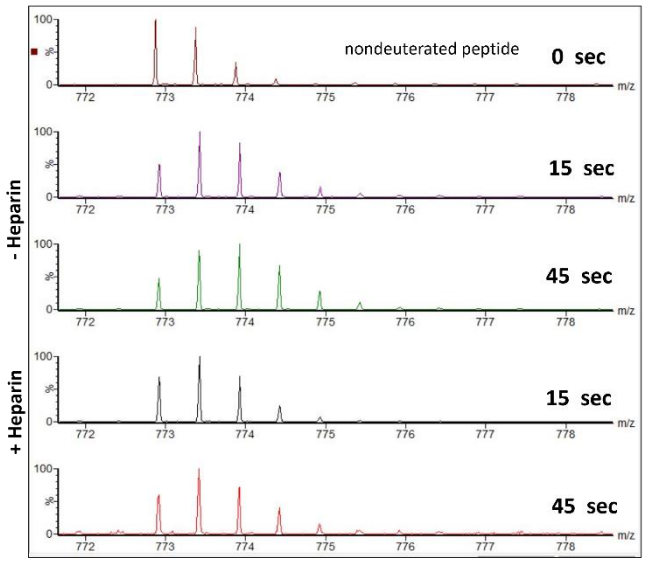
(a)



(b)



(c)



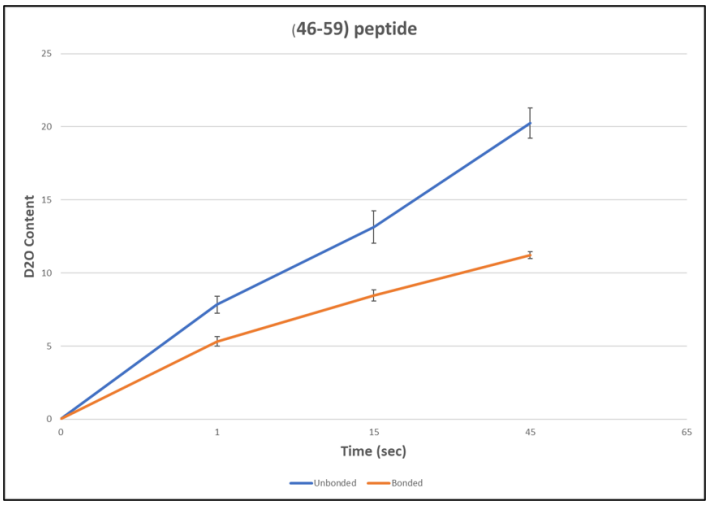


Figure 11. Apo states of PAI-1 protein Deuterium Exchange.
(a) Peptide uptakes in mutant PAI-1 and (b) Wild PAI-1 after1, 15 and 45 sec exchanges.
(c) Example of a deuterium –labeled PAI-1 peptide. D_2O Uptake Plot: MS spectra of the peptide 46-59 shutter region (m/z 772.9) ion signal after HDX for 15 and 45 sec on PAI-1 in the presence and absence of LMWH. The spectrum of 0 sec represents nondeuterated peptide.

2.3.2 Effect of Fondaparinux Binding on PAI-1 Plasticity

Next, we investigated the impact of LMWH (Fondaparinux) binding on the local flexibility of active PAI-1. Deuterium uptake for representative peptic peptides is shown in Figure 13 and mapped on the three-dimensional structure in Figure 12.

LMWH binding to PAI-1 induces a widespread attenuation of deuterium uptake in PAI-1. Peptides experiencing the strongest protection against isotopic exchange upon LMWH occupancy are generally confined to the lower part of PAI-1.

Peptides covering the Flexible joint binding interface (Figure 13) are shielded from the solvent and act at slower exchange rates, including peptides 99−113 (β2A−hE), 114−126 (hE−β1A), 127−134, and 139−151 (bottom and top part of hF loop). Peptides remote from the LMWH binding site were also impacted. These peptides are 46−59 (hB and hC), 64−92 (hD), 200-212, 239-265, 273−278, 300−306, and 310−320 (hI and the loop connecting hI with β5A).

`

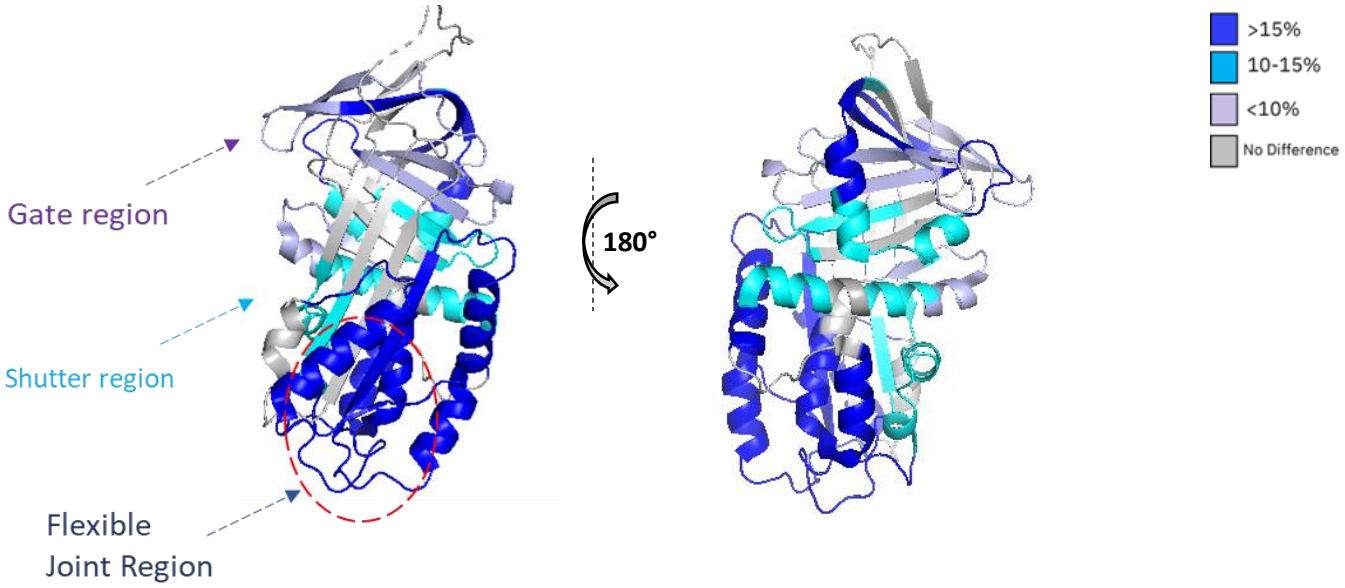
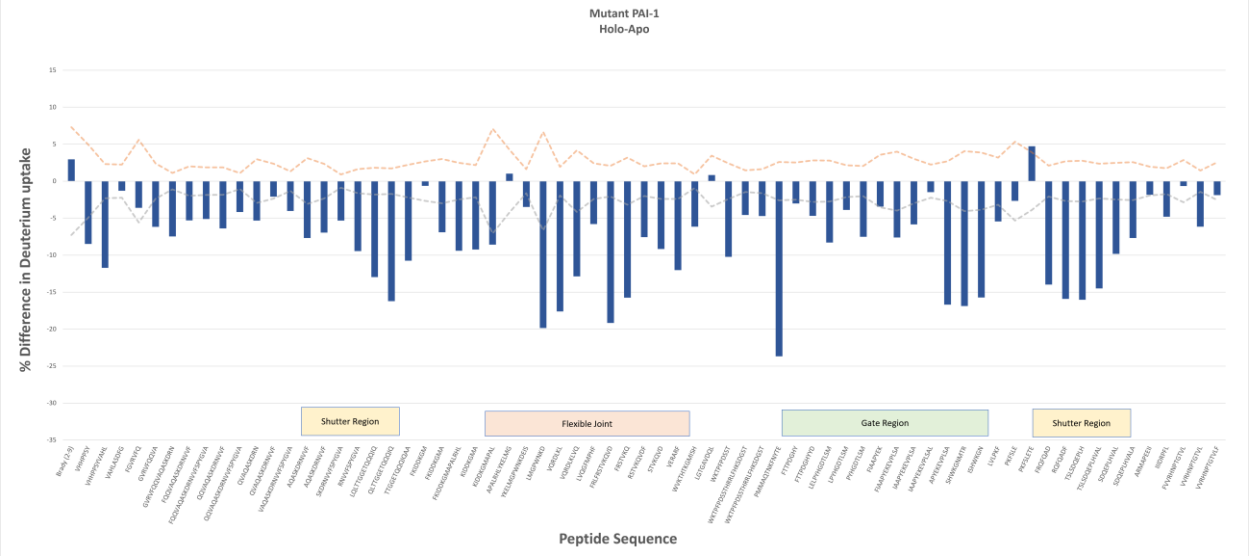


Figure 12. 3-D structure mapping of the maximal difference in relative deuterium content of mutant PAI-1 with and without LMWH. The back and front view of relative deuterium content of mutated PAI-1 mapped on PAI-1 structure, blue regions indicate the LMWH binding sites where Deuterium content is lowest. PDB 3Q02

(a)



(b)

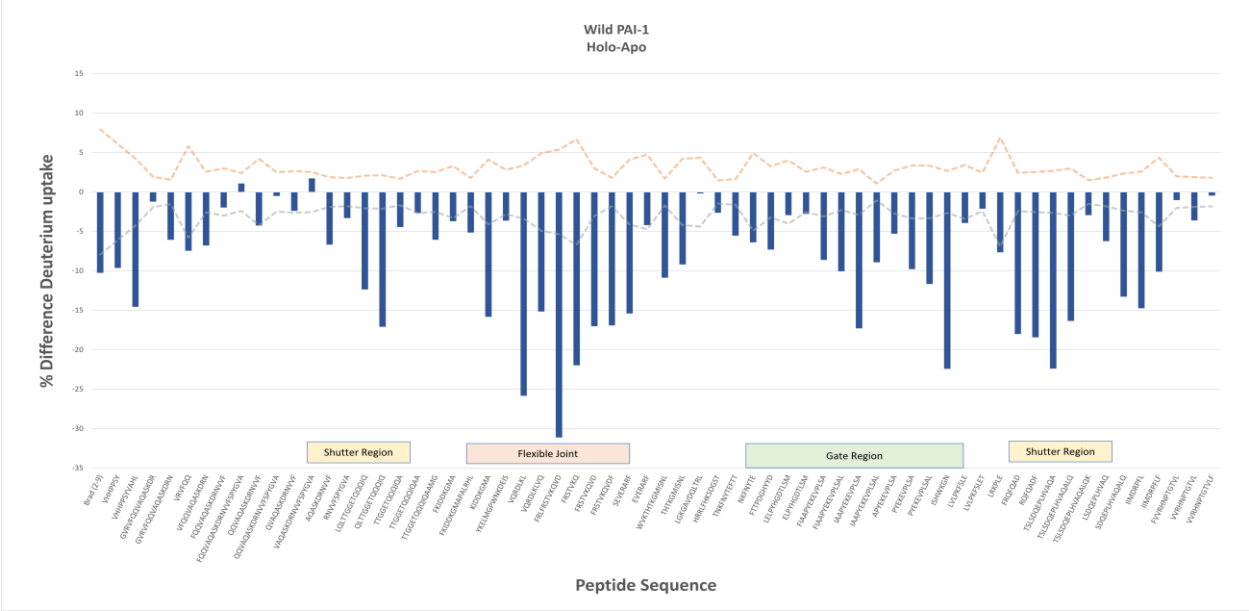


Figure 13. Maximal difference in relative deuterium content of PAI-1 with and without LMWH

(a) and (b) LMWH binding to PAI-1, mutant and wild respectively, induces a significant attenuation of deuterium uptake. Peptides covering the Flexible joint region has the strongest protection against isotopic exchange [peptides (99-105), (114-125)]. Some changes are also noticed in regions away from LMWH binding site, where modest decrease in D_2O uptake take place [peptides (206-213), (239-249); gate area]. Lastly, Protection was also observed in Shutter area [peptide (307-319)] Difference in D_2O uptake is mapped on 14-1B mutated PAI-1 structure. Only Moderate protection was observed in the mutated PAI-1 when compared to wild. Sum of three timepoints and Error bar is 1 sigma significance threshold.

Likewise, protection observed in peptides covering hD (79−92), hB and hC, and hI and the following loop (286−298 with 299−306) are in all cases isotopic exchange decreased significantly. Thus, conformational change takes place in these structural regions upon LMWH binding. To summarize, our data show that binding of LMWH to active PAI-1 results in dramatic stabilization in the entire lower half of the PAI-1 molecule and in some select parts in the upper half of the molecule.

2.3.3 Structural Plasticity of wild type and mutant PAI-1

In the unbounded states, we observed moderate protection from deuterium exchange in the tetramutant 14-1B PAI-1 compared to the wild type. Both forms of protein act equally when binds to heparin but some plasticity differences were found in the apos states.

Mutated PAI-1 has less deuterium uptake at parts close to the mutation sites; peptide 139-151 which is the bottom of hF loop, near the flexible joint region. The wild PAI-1 structure is more dynamic in that part (more deuterium uptake). That is where N150H, K154T, and Q319L mutations residues are close. Likewise, peptide 273-282 (β3C and β6A strands), which is located close to the M354I mutation, show more protection in the mutated form. Some peptides located further away from the mutated sites were also protected in 14-1B compared to wild-type PAI-1 (Figure 14). The hB and hC loops (49-61 peptide) and the top loop of hE and β1A strand (114−125 peptide). On the other side, the β5A peptide (299-306) spanning in the shutter region had more dynamic pattern in the mutated PAI-1 while less dynamic in the wild type.

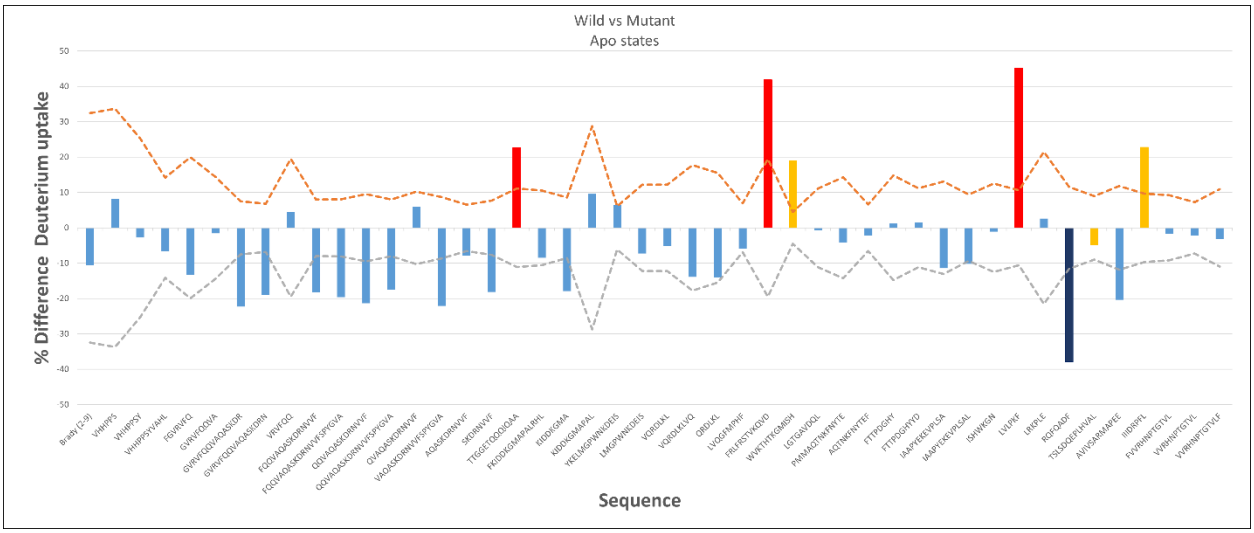


Figure 14. Apo states maximal difference in relative deuterium content between wild type and 14-1B PAI-1. Unbounded mutant and wild are protected similarly from deuterium exchange except for some regions in the structure. Peptides, near the mutation residues, covering the hF loop and β6A (139-151, 273-282) strand are more protected in the 14-1B. Other parts at hB, hC, hE and β1A (49-61,114-125) are also protected despite being remote from the mutation sites. Decrease in protection was only observed in the β5A peptide (299-306) spanning in the shutter region. The mutated residues are colored yellow, read means increase and back means decrease.

2.4 Discussion

2.4.1 Allostery effect: LMWH Decrease dynamics in the SMB Domain

We have observed that the same regions that are more dynamic in the absence of LMWH, become stabilized in its presence. The flexible joint region, gate region, and shutter region are known to be involved in steps in the mechanism for the structural activation of PAI-1. The observed changes in dynamics were in predominantly overlapping regions affected by the binding of the SMB domain including the flexible joint region, and the N-terminal part of hF. Additionally, we observed changes due to LMWH addition in the area underlying the shutter strands in hB-hC and the proximally located hI. That indicates the allosteric effect of heparin on PAI-1 molecule.

Importantly, the altered dynamics were not a global effect but were localized structurally. That means some dynamic, as well as protected regions of PAI-1 were unaffected as a result of

LMWH binding in our data. Apart from the conformational changes of the lower half of PAI-1, LMWH also stabilizes a few areas in the upper half of the molecule.

2.4.2 LMWH binding: the effect on latency transition.

Prior research was conducted on PAI-1 in the presence of RNA aptamers and Mab-1 antibodies to evaluate their stabilizing effects.[50,71] These molecules bind to PAI-1 in the flexible joint region as well as to hC in the case of the antibody. Mab-1 binding stabilizes PAI-1 by 1.4-fold , while the aptamers do so by 3-4 fold . Taken together, these findings add to the evidence that the dynamics of hB, hC, and hI have an important role in the PAI-1 activation process..
Observation of these peptides covering this region hB, hC and S6A (residues 46−61 and 273-282), we find out these peptides are protected in all states, with and without the LMWH binding. These regions are known to have a factor in the latency transition mechanism. However, small changes in the next peptide 30-40 which also directly connected to the hB- and hC parts.To study this effect in details, a stability study is needed to determine the impact of the LMWH binding on active structure. Conclusively, a possible crosstalk between these regions plasticity and impacting the latency conversion of PAI-1.

2.5 Conclusions and future studies

In summary, the HDX based investigations show the PAI-1 structural plasticity with and without binding heparin. These details provide novel insight into the flexible and inflexible regions in the molecule.

The data reported here clearly show that binding of the Fondaparinux to PAI-1 causes dramatic and widespread stabilization in most of the lower half of the serpin. This information provides an insight for further mechanics of Fondaparinux to enhance its anti-coagulant activity.

Moreover, an insight to develop new targets for PAI-1 inhibition. This HDX analysis provides detailed picture of PAI-1 dynamic that cannot be obtained by X-ray crystal structure analysis. Future work will continue to further explore the effect of this binding upon latency transition, and longer time points from seconds to minutes.

Chapter 3

Proline-Alanine tagged L-asparaginase

3.1 Introduction

3.1.1 L-Asparaginase: Essential medication for leukemia

L-Asparaginase is an integral chemotherapeutic enzyme used in the treatment of acute lymphoblastic leukemia (ALL), Hodgkin's disease, acute myelocytic leukemia and melanosacroma.[38] Besides its clinical benefits, L-asparaginase also has a potential role in food industry as a food processor.[72] Given its importance, the World Health Organization (WHO) implement it in its list of Essential Medicines. Tumor-inhibitory properties of ASNase were first discovered by Kidd in 1953, when he observed that lymphoma-bearing mice, treated with Guinea-pig serum, underwent complete tumor regression. A decade later, Broome's experiments completely elucidated the asparaginase role in antitumor action (Figure 15).[73]

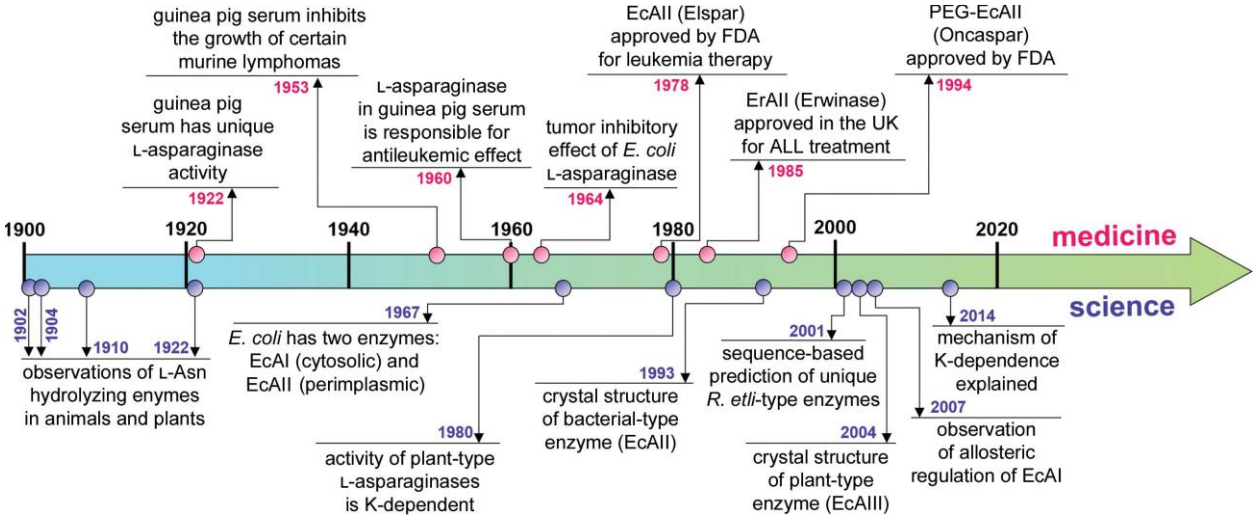


Figure 15. Discovery of l-asparaginases throughout history. The turning point of L-asparaginase usage was in 1960s, when the mechanism of action was first described by Broome. Figure adapted from Joanna I. Loch, 2021.[73]

L-Asparaginase can be found in mammals, but also can exist in birds, bacteria, plants and yeasts. But not all asparaginases have anti-cancer activity. Classification of L-asparaginases based on its source of origin.[73] The first L -asparaginase enzyme discovered was named after the source

organism. Later, the nomenclature divides L-asparaginases into three classes. Class 1 (bacterial source) has type 1 and type 2, type 1 enzymes which are expressed in the cytosol. Type 2 enzymes are secreted to the periplasm and are produced under anaerobic conditions. Class 2 (plant based) are either K- dependent and independent and both alongside the fungal based class 3 (type 4) enzymes are thermostable. Lastly, type 5 enzymes are thermolabile and their expression is induced by the presence of l-Asn. Bacterial sources from *Escherichia coli* (E.coli) and *Erwinia chrysanthemi* L-asparaginase (ErA) and their derivates have been used as the only preparations for medical purposes.[74] The enzyme derived from E.coli is used as a first line therapy, while the Erwinia

derived L-asparaginase is kept for patients who develop hypersensitivity reactions to the previous form.

Both normal and malignant cells require the amino acid L-asparagine for its metabolic needs. However normal cells can synthesis L-asparagine independently. Healthy cells utilize transaminase enzyme which converts oxaloacetate into an intermediate aspartate, that is later transfers an amino group from glutamate to oxaloacetate producing alpha-ketoglutarate and aspartate (Figure 16).[74] Finally, aspartate is converted to asparagine-by-asparagine synthetase enzyme. Cancer cells cannot synthesize the asparagine because of the absence of L-asparaginase synthetase enzyme, therefore are totally dependent on the exogenous supply of asparagine for their viability and reproduction. Consequently, L-asparaginase administration drains all circulating asparagine, which leads to starvation of neoplastic cells and ultimately their die off.

3.1.2 Structural insights of L-Asparaginase

Two isoenzymes of L-asparaginase have been identified, namely type I and type II. Both types are characterized by enzymatic activity for both L-asparagine and L glutamine.[75] However, type II asparaginase has a higher specific action against L-asparagine. Type II asparaginase particularly shows antitumor activity and is used in chemotherapy guidelines.

Many research efforts have been going to elucidate the structure of L-asparaginase on the molecular level. It has been found that l-asparaginase exists as a tetramer however monomeric, dimeric and hexameric forms are also found when isolated from different sources.[72] Molecular

structure of ErA asparaginase enzyme is a homo-tetramer (ABCD) of 35.6 kDa, made up of four identical monomers (A to D). Each monomer has 327 amino acids with 14 alpha-strands, 8 B-helices and two domains, N and C-terminal domains. The active site is located in between two adjacent monomers (A and C, B and D).[76] The tetramer has four identical subunits, and the whole molecule is regarded as dimer of dimers. The four active sites are located between the N- and C-Domain, comprising of Thr 12, Tyr 25, Asp 90 and Lys 162 residues (Figure 17) are responsible for the enzyme catalytic activity.

Drugs of L-asparaginase are available in the market with growing expected rate of 16 % from 2022 to 2028. L-asparaginase contributes to 40% of total enzyme world demand and one third of global healthcare protocols in antileukemia and anti-lymphoma agents.[77]

Many factors influence the enzyme therapeutic properties that includes sources of L-asparaginase, upstream and downstream bioprocessing, rate of hydrolysis, resistance in tumor cells, serum clearance and pharmacological factors. In the manufacture, three different types of as asparaginases have been introduced to treat ALL: Firstly, native E. coli asparaginase or Elspar, Secondly a pegylated version of asparaginase derived from E.coli Oncospar and Lastly, E. Chrysanthemum based native asparaginase.[78] Since ASNase therapeutics are derived from bacteria, they are highly immunogenic. Therefore, alternate formulation, the PEG ASNase, has been developed in order to reduce the immunogenicity and also number of infusions.

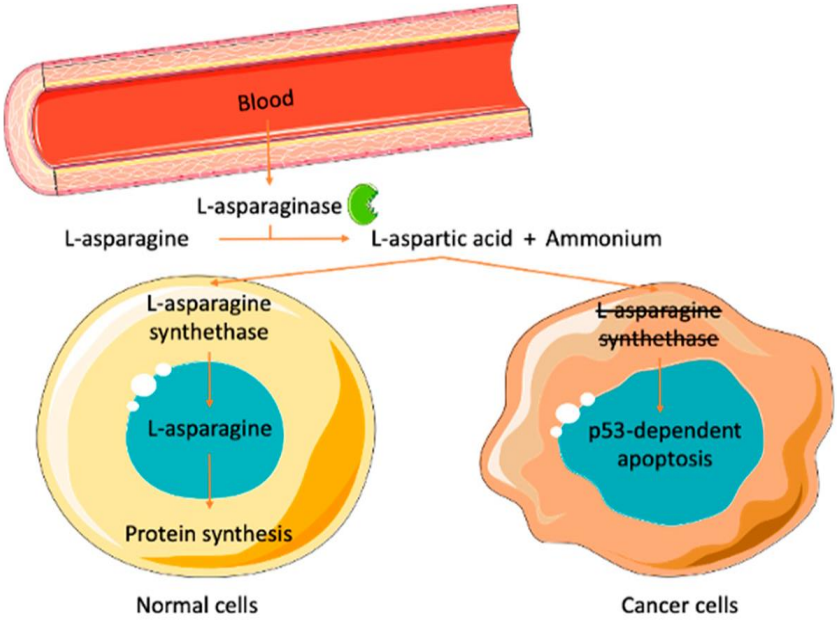


Figure 16. Mode of action of asparaginase. Adatapted from João C. F. Nunes, 2020.[6]

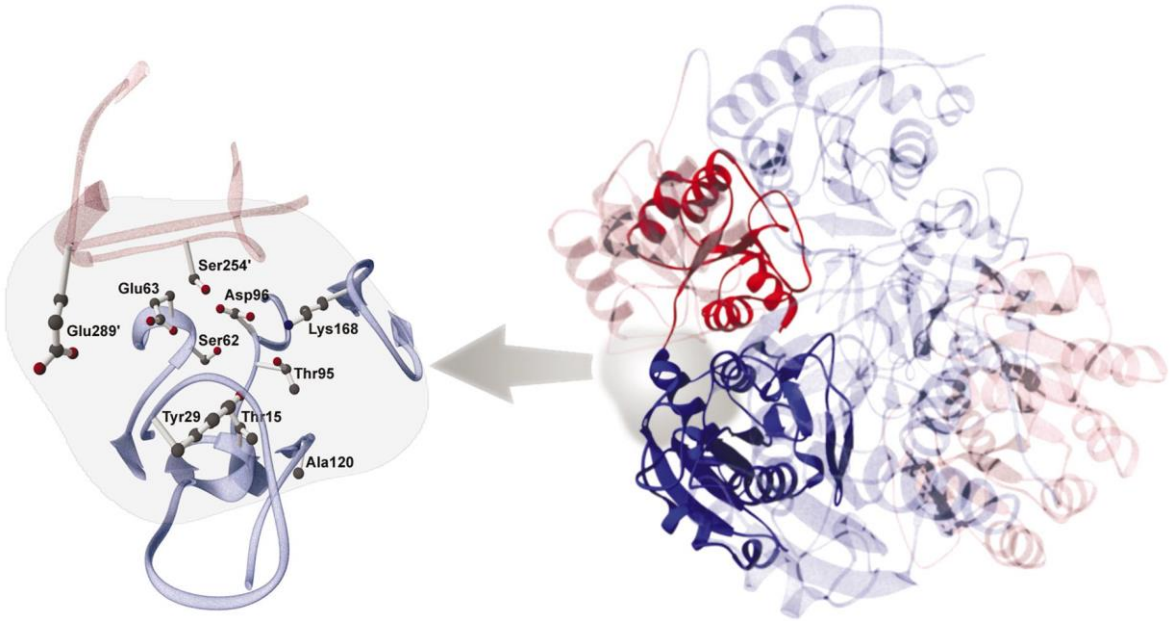


Figure 17. Schematic representation of the tetrameric enzyme ErA molecule. Solid colors represent the monomers. In all four monomers the N-terminal domains are drawn in blue and the C-terminal domains in red. The location of one of the four equivalent active sites is indicated by a semi-transparent gray blob. The active-site pocket is formed from the fragments of two domains contributed by different monomers. PDB 1O7J. Adapted from Jacek Lubkowski, 2002.[79]

3.1.3 PASylation technology: Biological alternative to PEGylation

PEGylation is the chemical coupling of the protein with synthetic polymer polyethylene glycol (PEG). The PEG-ASNase has an improved pharmacokinetic profile with a longer half-life of about 1 week, while native E.coli and Er.A asparaginases have a half-life of 1.3 and 0.65 days.[80,41] Therefore, due to the short half-life of Erwinia asparaginase, higher doses and frequency are required to ensure adequate serum enzyme activity.

While PEGylation has emerged as recognized technology for the development of long-acting biologics, it shows more disadvantages that limit its application in bioprocess and drug development research.[43,45] For example: chemical coupling in vitro and the need to make corresponding modifications to the protein of interest which corresponds to the high cost of synthesis. Moreover, commercially available activated PEG derivatives inherent polydispersity

thus the associated difficulties in product analysis. In addition, the biological material frequently losses its biological activity, poor bioavailability of PEGylated protein administered subcutaneously (S.C) and the lack of biodegradability of non-natural PEG polymers that lead to tissue accumulation and further immunogenicity. Seeking a biological alternative to PEGylation, PASylation technology was recently introduced.

PASylation is a genetically encodable amino acid sequence conjugated with the protein of interest. This amino acid sequence forms a naturally occurring unstructured polypeptide chain with high solubility and above all no charge, offering biophysical properties comparable to PEG.

Three amino acids Pro, Ala and Ser (PAS) have an efficient biosynthesis of long hydrophilic and highly soluble polypeptide chains exhibiting properties similar to PEG.The length of the chain can range from 100-600 units which increases the hydrodynamic volume of the conjugated drug. Thus, prolong plasma half-life of the molecule and decrease its clearance.[43]

This natural polypeptide polymer shares similar biophysical properties of the synthetic PEG. In other words, PAS tail is highly soluble and adapts expanded random coil structure in aqueous solution while with lower viscosity and more hydrophilic. That explains the high bioavailability of administrated PASylated therapeutics via injection in animal and clinical model.[81,46] The most significant benefit of PAS sequences is that they are non-immunogenic, stable in blood plasma, and swiftly destroyed by intracellular proteases following cellular absorption, preventing vacuole accumulation. Indeed, PASylation delivers the advantages of PEGylation while avoiding major PEG drawbacks such as bioactivity loss, polydispersity, and cellular disposition.

3.1.4 Host Cell Proteins (HCP) In Biopharmaceutical Industry

Protein based therapeutics are produced in host cell lines, and thus, the host cell endogenous components are also produced along with the biotherapeutic protein sequence. Host cell proteins (HCP) are a complex group of proteins characteristic of the host cell, which differ significantly in molecular mass, isoelectric point and structure.[82] The presence of HCP in biologics can potentially impact the product safety and can elicit an immune response in patients, leading to adverse reactions and decreased efficacy of the product. In addition, some HCP have known to degrade the therapeutic protein formulation content, leading to a short shelf-life impacting product quality.

HCP are process-related impurities which may copurify with the drug product. According to the FDA guidance, HCP must be removed to trace levels during drug purification. Regulation authorities have identified HCP as critical quality attributes (CQAS) which affect patient safety. During risk assessment many factors are taken into account such as: route of administration, dose and frequency.[82,83] Therefore, control strategies are needed during biopharmaceutical process development to identify HCP individually, quantify their acceptable range, and to determine any process changes that need to be made to remove HCP present at elevated levels.

3.1.5 Analytical Methods for HCP Detection and Quantification

Identifying HCP in biotherapeutic products provides valuable information to develop safe products during process optimization. However, qualitative determination of HCP is not sufficient solely to control these impurities. Quantitative or semi-quantitative measure is a must to recognize all HCPs species. Currently, the techniques used in the industry are either immunospecific methods such as biosensors, western blotting and ELISA, or non-immunospecific methods as in electrophoresis, 2D chromatography, and MS.[84] Ideally, the technique used should have a short analysis time, and a high-throughput format.

Immunospecific methods detect HCPs by using polyclonal anti-HCP antibodies which detect a broad range of different proteins from the host cell proteome. The enzyme-linked immunosorbent assay (ELISA) is commonly used and considered the golden standard in the detection of HCPs. However, the HCP ELISA result is a total measurement of HCP; it does not provide information on specific proteins found in the product and, as a result, provides an incomplete view of the impurity profile.[83] Moreover, the results can be incomparable across labs due to differences in HCP populations from different host cell lines and differences in the immunoreactivity against HCP for different ELISA kits.

In non-immunospecific methods, each technique has a different principle in detecting HCP. For example, in gel electrophoresis, 2-Dimensions gels (2D) separate proteins according to the isoelectric point followed by separation based on size. The detection of HCPs using gel electrophoresis is limited to proteins with molecular weights ranging from 5 to 1000 kDa. Using this method, one should select a sample matrix compatible with the gel to recover the protein

sample and to remove interfering agents. In addition, various stains and buffers are used in the electrophoresis protocols. Thus, the need for a range of repetitive runs to ensure the reproducibility of the results. All these challenges in the gel application make it be used as more of a complementary method to identify HCP.[84]

Mass spectrometry (MS) has emerged as a valuable tool in HCP analysis. Liquid chromatography with tandem mass spectrometry (LC-MS/MS) is used for identifying and quantifying individual HCP.[85] Proteomic techniques are used as an orthogonal tool to HCP ELISA, thus providing detailed information about specific HCP in the biotherapeutic samples. With full analysis of product HCP, one can evaluate its potential biological activity on the product quality. Information about a particular HCP is helpful in risk assessments to determine the risk to manufacturing, product stability, and patient safety. Moreover, figuring out any process alterations needed to get rid of HCP existing at high levels can be done with the help of studying the distinctive qualities of each HCP.

3.1.6 Rational of the study

Jazz Pharmaceuticals is a developer of oncology drugs with global reach in the United States of America, Ireland and substantial development operations in Canada. One of the candidates currently under development at Jazz is the L-asparaginase enzyme to treat ALL. Substantial progress has been made at characterizing the biological activity of this molecule, however, its size and biophysical complexity has made detailed molecular characterization challenging. The proposed Jazz product is called JZP-341 and it is a recombinant Er.A expressed with pseudomonas with the Genetic Fusion of an N-terminal 201 amino acid polypeptide sequence comprised of proline and alanine residues (PA200) to extend the half-life of the asparaginase protein in vivo (Figure 18).

JZP-341 is a non-disulfide bonded, tetrameric protein consisting of four identical polypeptide chain subunits with a combined molecular weight of 204 kDa. It converts asparagine to aspartic acid to starve the tumor cells. JAZZ is developing a recombinant crisantaspase product by microbial expression in *Pseudomonas fluorescens* as a long acting Er.A for the acute lymphoblastic leukemia (ALL) indication. It is in early phase program and plan to file IND by

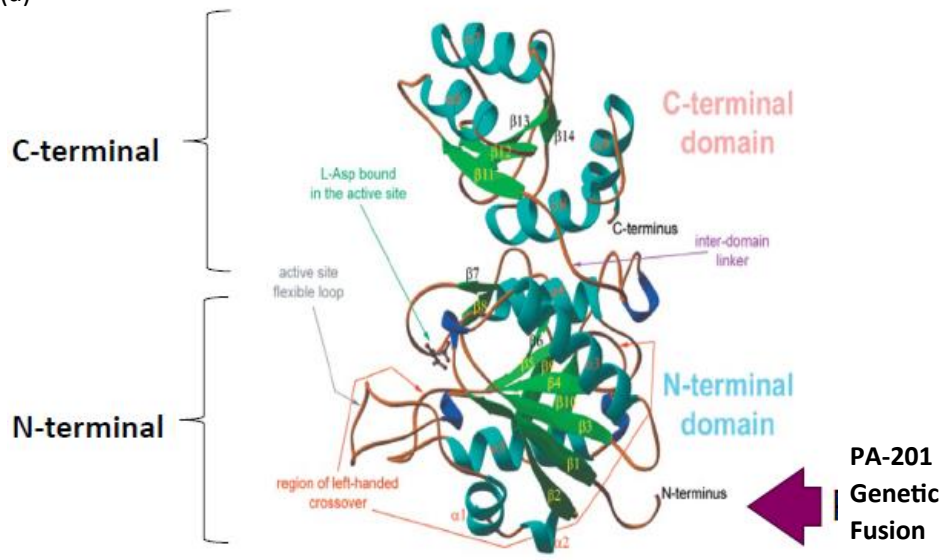
the company. While JZP-341 has been characterized in a broad biochemical sense by Jazz, its structural and functional characteristics at the molecular level, including its molecular mechanism of action (MoA), are not well understood.

Molecular characterization is an increasingly critical part of the drug development pipeline, especially for complex biotherapeutics like JZP-341, where the ‘active ingredient’ corresponds not to a single entity but a distribution of configurations resulting from diverse post-translational modifications, disulfide states, truncations, and conformational states. In these cases, a thorough characterization of the ensemble of species is required by regulators.

During manufacturing of the new PA-tagged protein (JZP-3410), the company noticed different charges of the in-process samples taken during testing. Therefore, a concern was raised in terms of the PA tag effect of structural dynamics that might happen during in-process manufacturing. The company provided us with different samples collected at different stages of in-process manufacturing to test the structural dynamics of the molecule.

The proposed work is designed to provide a detailed characterization of the JZP-341 protein dynamics in various formulations. Structural MS techniques were carried out as time-resolved electrospray ionization hydrogen-deuterium exchange (TRESI-HDX-MS) to characterize conformational dynamics in modified L-asparaginase drug candidates. Analyzing charge variants of modified L-asparaginase batches provided by the manufacturer Jazz Pharmaceuticals is critical for characterizing and monitoring the quality attributes of the final product. Using sensitive structural MS techniques will help understand PASylation tag modification's impact on L-asparaginase activity.

(a)



(b)

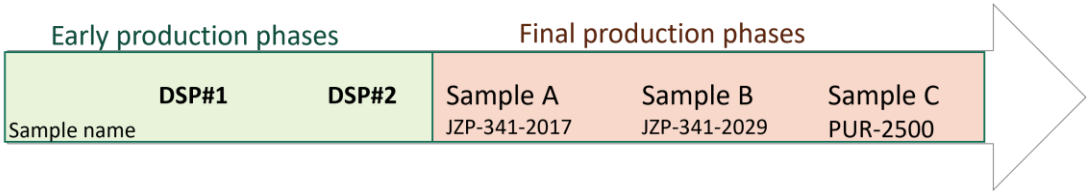


Figure 18. (a) Scheme of the PA-infused *Erwinia chrysanthemi* L-asparaginase (Er. A).Structure adapted from Aghaiypour, 2001.[75]

(b) In-process sampling during the bioprocess of PA-L-asp manfacturing

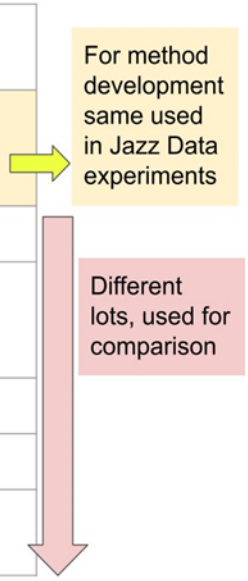
3.2 Experimental Methods

3.2.1 Sample and supplies

Jazz Pharmaceuticals provided five tagged protein samples of modified L-asparaginase as follow in Table 2: DSP1, DSP2, sample A (JZP-341-201701), sample B (JZP-341-201701) and sample C (RS PUR2500-1)

Table 2: Material as provided by Jazz Pharmaceuticals.

Samples arrive ID	Protein Conc	Formulation	Amount available
JZP-341-201701	**1 g/L**	**20 mM Histidine, 150 mM Arginine, 0.02% PS 80, pH 6.0**	**50 ml**
JZP-341-202902	9.85 g/L	same	2 ml
RS PUR2500-1	1.08 g/L	10 mM Na2HPO4, 137 mM NaCl, 2.7 mM KCl, 1.8 mM KH2P04, pH 7.4	10 ml
DSP1	0.22 g/L	20 mM TRIS, 5mM Nacl, 2mM EDTA	23 ml
DSP2	0.53 g/L	same	9.4 ml
Formulation buffer	N.A	20 mM Histidine, 150 mM Arginine, 0.02% PS 80, pH 6.0	10 ml



The five tagged protein samples have the same protein composition but in different buffer compositions. The major difference is each sample was collected at a different in-process step as follows: Early production phases (DSP1 and DSP2), final manufacturing process (Samples A, B and C). Sample A (JZP-341-201701) is mainly used in our experiments and for method development upon the client request. The remaining samples are used for comparison. In addition, a control sample was provided named “untagged core” which is made of L-asparaginase protein (manufactured by Jazz) without the tail (unmodified). Buffers used for sample preparation were made in-house using ultrapure water using a Millipore Milli-Q Advantage A10 system. Chemicals were purchased from Sigma-Aldrich unless otherwise specified.

3.2.2 Sample preparation

Using Amicon Ultra 0.5 mL Centrifugal Filters with a 30 kDa cutoff. Protein samples were buffered exchanges into 100 mM ammonium acetate buffer, pH=7. Samples were diluted to a final concentration of 5 µM in 100mM ammonium acetate for practical use.

3.2.3 Shotgun Proteomics

Jazz pharmaceuticals provided us with sample D (JZP341-221103) for HCP investigation. Proteomics analysis was done by York University Proteomics lab. Briefly, the sample was digested with pepsin using the method previously described. Sample was buffered exchanged into 100 mM ammonium acetate, pH=7, using a 30 kDa MW cut-off filter. The concentration of the sample was measured after buffer exchange using a Nanodrop. The digest was collected for analysis. Each digested sample was analyzed using a Waters ACQUITY UPLC system coupled to a Thermo Scientific Q Exactive Plus mass spectrometer. (Figure 20)

3.2.4 Data Processing for Shotgun Protein Identification

The resulting LC-MS/MS data was searched against the *Pseudomonas fluorescens* and the Crapome database reference proteome database using Proteome Discoverer version 2.1 (Thermo Scientific) for data processing. Results were filtered based on the number of unique peptides and peptide length. Proteins that had two unique peptides and unique amino acid sequences with high quality MS/MS spectra and peptide length >6 amino acids were accepted as true positive identifications.

3.2.5 Materials for mass spectrometry experiments*

As mentioned previously at chapter 2. Briefly, Pepsin, deuterium oxide (D_2O), ammonium acetate (≥99.0%), and high purity acetic acid (≥99.7%) were purchased from SigmaAldrich.

3.2.6 Pepsin Preparation*

Pepsin from porcine mucosa and NHS-activated agarose beads (2:5 w/w) were resuspended in coupling buffer (0.1 M sodium phosphate, 0.15 M NaCl, pH 4) and rotated for 17 hours at 4°C. The unbound pepsin was aspirated, and beads were resuspended in blocking buffer (1 M Tris–

HCl, pH 4) and then rotated at room temperature for 1 hour. The beads were washed five times with 10% acetic acid (v/v) via centrifugation at 1000 x g for 2 minutes and stored at 4°C.

3.2.7 Microfluidic Device Fabrication*

The microfluidic device used for time-resolved ESI (TRESI) measurements was built as previously described[86,87]. Briefly, a VersaLaser™ was used to etch a proteolytic chamber into a poly(methyl methacrylate) (PMMA) surface measuring 5.1 cm × 2.0 cm × 1.2 cm. A rapid-mixing device was made by inserting an inner glass capillary (outer diameter of 151.0 μm) inside a 28-gauge metal capillary (inner diameter of 178.8 μm), creating an intercapillary space of 27.8 μm. The end of the glass capillary was sealed using the VersaLaser, and a notch was cut 2 mm from the sealed end. Hydrogen deuterium exchange (HDX) labeling reactions were carried out in the volume between the inner and outer capillaries. For localized measurements, the TRESI mixer was connected to an acid channel through a valco t-mixer, followed by on-chip proteolytic digestion. Continuous Flow Injection (CFI) was used as a comparison to the chip system[69] (Figure 19)

3.2.8 Data Analysis*

Peptides sequences were confirmed via MS/MS with analysis in mMass[88]. Deuterium uptake per peptide was calculated using Mass Spec Studio[70] by matching experimental deuterium uptake to the theoretical isotopic distribution for the native peptide sequences. Changes in deuterium uptake for ‘unmodified l-asparaginase’ (Core protein without the PA tag) versus ‘Tagged protein’ states were summed across 3 time points (labeling times of 5, 10, and 50 seconds) and averaged across technical triplicates from two different preparations (total n=6 replicates). Changes in deuterium uptake were considered significant if they exceeded two times the standard deviation (2σ). Protein structures were shown using PyMol v1.5.0.4 (Schrödinger, LL).

** (3.2.5-3.2.8 same method as mentioned previously at Chapter 2)*

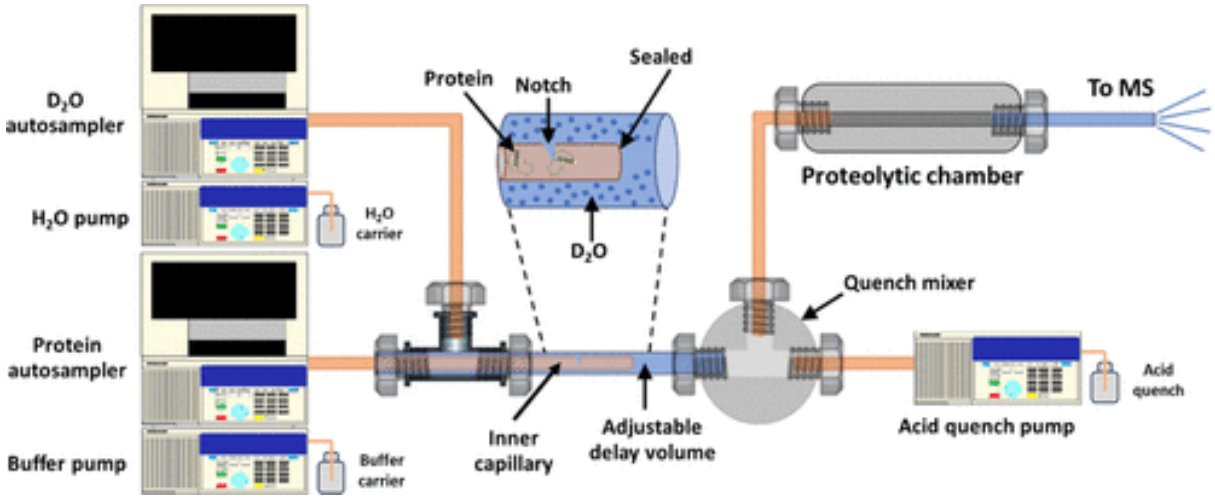


Figure 19. Automated CFI-TRESI-HDX setup. LC pumps deliver continuous solvent delivery, the protein and D_2O plugs delivered via autosamplers. Inside the mixing capillary, an inner capillary is sealed at the end and notched to allow D_2O and protein mixing. Low millisecond to second HDX labeling times can be achieved based on the position of the inner capillary within the outer capillary, size of capillaries, and flow rates. Adapted from Anacleto 2023.[69]

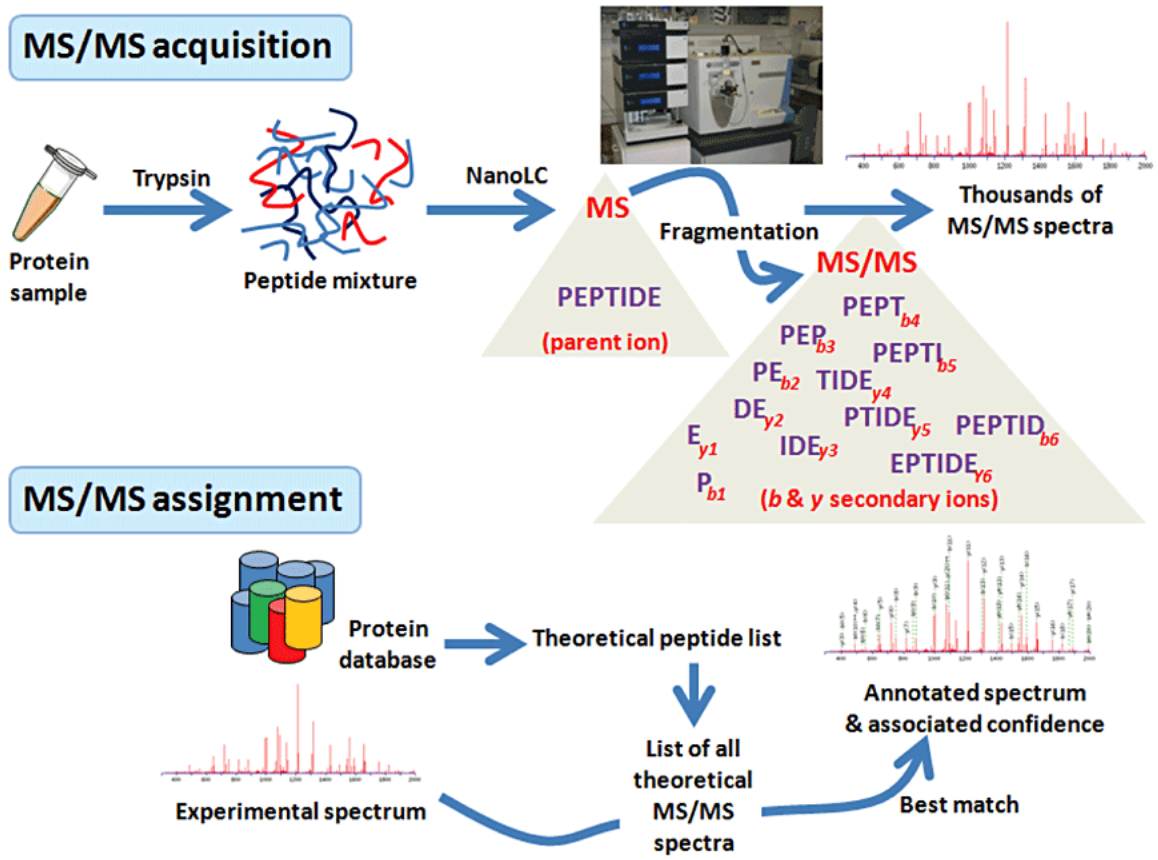


Figure 20. Workflow of shotgun proteomics. Pepsin digestion of the protein sample is followed by separation on a reverse phase chromatographic column connected to a tandem mass spectrometer. Each peptide's precursor mass is determined by the MS, which then chooses a predetermined number of the strongest parent ions for collisional activation-induced fragmentation. The result is an MS/MS spectrum with ions that are fragments of peptide sequences. In order to assign a peptide sequence from which peptide and protein identity may be determined with a certain degree of confidence, one must compare experimental and theoretical MS/MS spectra. Adapted from Armengaud, 2012.[32]

3.3 Results and Discussion

3.3.1 Native structure by mass spectrometry

To better understand the structure of PA-tagged protein, we ran electrospray ionization (ESI) mass spectrometry for both the core untagged protein sample and the PASylated Asparaginase (JZP-341). The molecular weight of the untagged protein was confirmed through mass spectrometry.

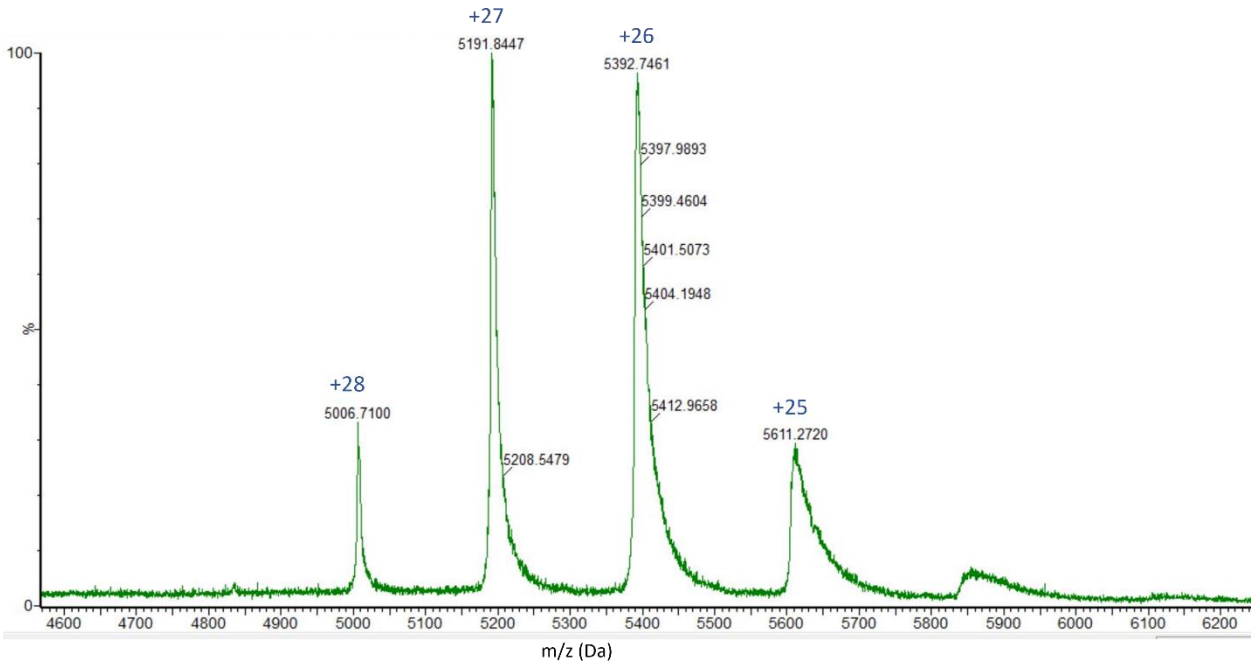
Figure 21 indicates that four main charge states were visible (+28, +27, +26, and +25). Molecular weight can be obtained directly from any of these peaks. The untagged protein exhibits a typical 'native' mass spectrum for a large multimeric protein, corresponding to a relatively low charge and a relatively narrow distribution associated with charge residue ionization.

The population is more than 90% tetrameric, based on the relative peak intensities of the tetramer vs. the monomer. The peaks associated with the monomer also appear to ionize via a charged residue mechanism, which is consistent with a folded protein.

Interestingly, the PA tail added as a PASylation tag to improve in vivo stability of this protein radically changes the native ESI mass spectrum. In particular, PA-tagged L-asparaginase (Sample A; JZP-341-201701) exhibits a mixture of intense, highly charged monomeric, dimeric, trimeric, and tetrameric species (Figure 21, a).

This suggests that the PASylated protein has undergone ionization primarily by chain ejection rather than charge residue, which would imply that the ionization mechanism is dominated by the unstructured PA 'tail', in line with previous observations made with isolated PAS sequences[45].

(a)



(b)

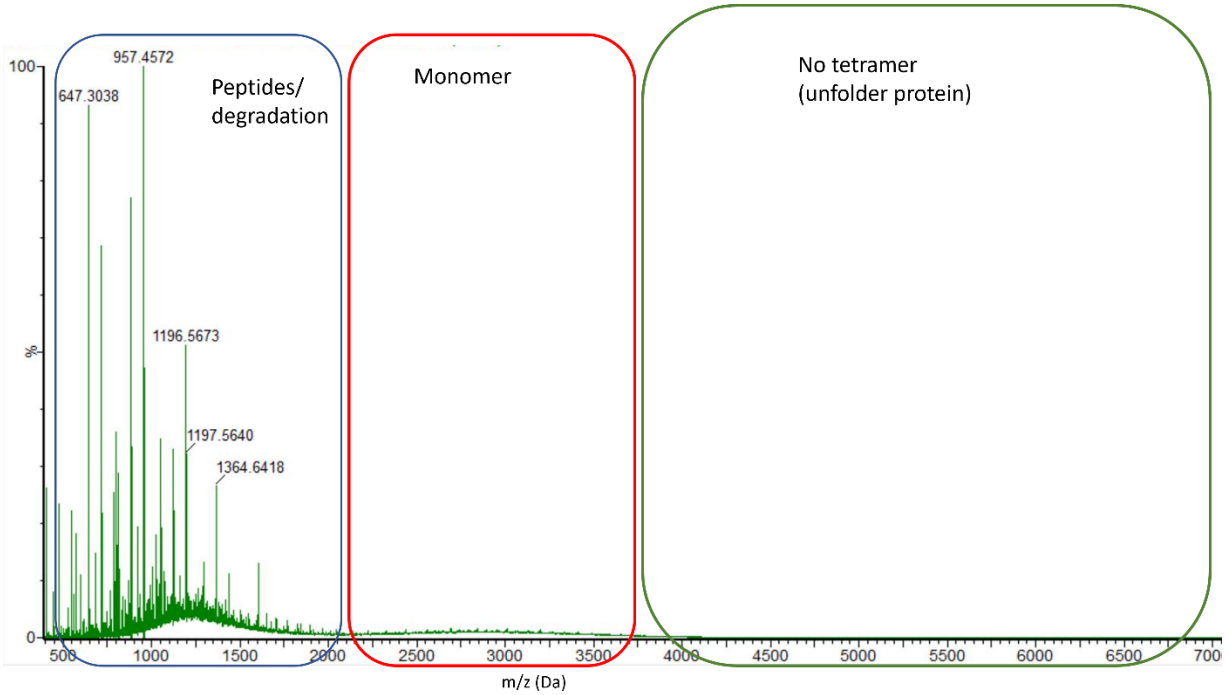


Figure 21. Mass spectrum for L-asparaginase protein samples. a) Untagged L-asparaginase, four dominant charge states are observed, with +27 being the most intense. b) PA-tagged L-asparaginase (Sample A; JZP-341-201701) ionized in different mechanisms, and its structure unfolds in the ESI conditions.

3.3.2 TRESI HDX MS revealed structural dynamics of JZP-341 PA-tagged L-asparaginase.

To compare the effect of the PA-tag on the dynamic of the protein, we investigated the structural dynamics of untagged core and modified l-asparaginase samples. we used bottom-up TRESI-HDX approach. Two setups of TRESI-HDX were used: the traditional microfluidics chip system and the automated CFI.

The experiment was carried out on a microfluidic device that facilitates variable HDX labeling times for three timepoints (2, 4 and 20 s), with on-chip digestion and electrospray ionization. Testing 'untagged core' and PA- tagged sample A (JZP-341-201701). Figure 22 shows insignificant difference in deuterium uptake between the untagged and PA tagged l-asparaginase. Interestingly, PA-tag attachment did not induce any significant decreases or increases in deuterium uptake in the core region. Therefore, the new molecule of (JZP-341-201701) is equivalently dynamic as of the core structure, and conclusively has no impact on the active sites (Figure 23).

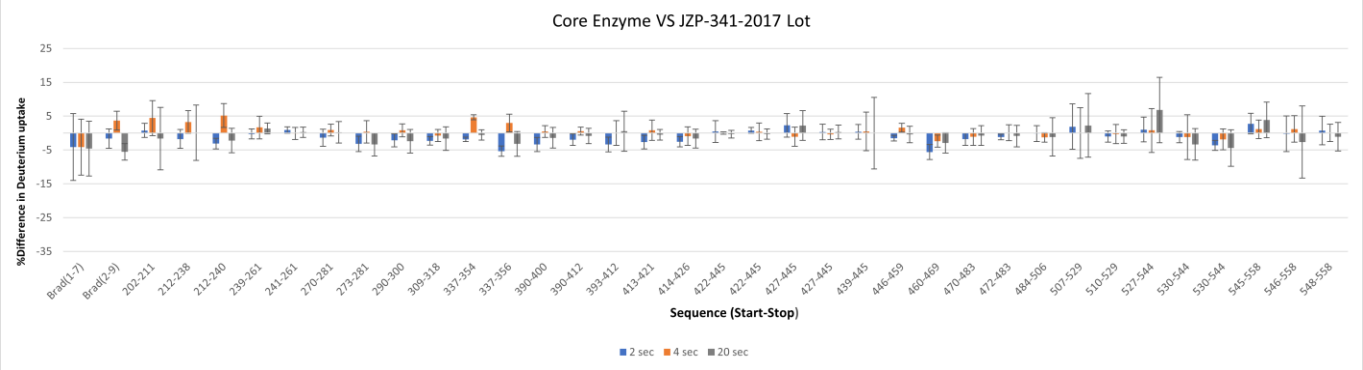


Figure 22. Difference in deuterium uptake between "untagged core" and PA tagged protein (sample A (JZP-341-201701)). Insignificant differences between two samples were observed. The same deuterium uptake in the two samples. Error bars are 2σ significance threshold.

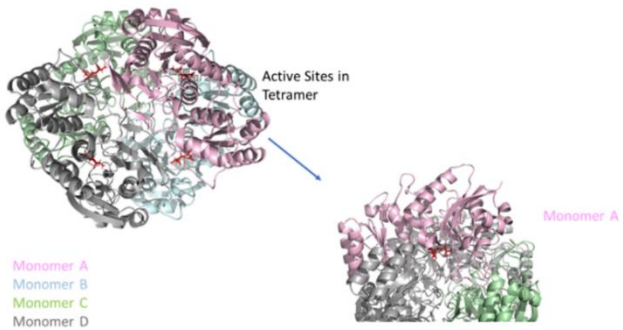


Figure 23. Scheme of the active structure of *Erwinia chrysanthemi* L-asparaginase (Er.A) by PyMole. Each monomer has a unique color. Four active sites (Thr 12, Tyr 25, Asp 90 and Lys 162 residues) are shown in red sticks. PDB 1O7J

Using the same setup, we tested the differences between two modified l-asparaginase samples, DSP#1 and DSP#2. Bothe samples were buffered exchanged with 100 mM ammonium acetate (pH=7) and final concentration was measured by Nanodrop. The protein samples were quenched and digested before MS analysis so that the deuterium exchanged information was locked to a specific peptide. To digest, pepsin was used as described previously and to quench the H-D reaction, 10% acetic acid (pH=2.5) was used. Comparison of these samples was carried out in the CFI setup below.

CFI setup makes continuous data collection and automated operation measurements. The microfluidic chip has been replaced by conventional LC mixing tees and a short PEEK microcolumn. The column is filled with immobilized acid protease-mediated digestion, and the volume used is ~ 65 µL (length: 80 mm/ width:0.04 inch/ vol: 64.9 ul) to achieve adequate digestion and reduce carryover. The protein and D_2O are supplied via separate LC pumps with independent autosamplers.

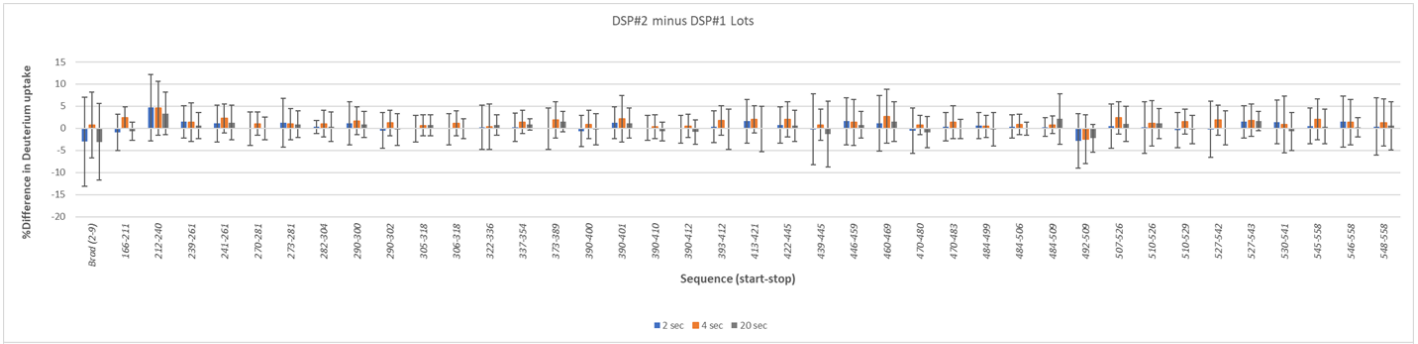
The protein mixture passes through a fused silica “inner capillary that is blocked at the distal end. Fluid escapes from this capillary through a small notch cut 2 mm from the distal end, forcing the protein solution to mix with D_2O in the narrow intercapillary space. The volume between the end of the inner capillary and the subsequent acid mixing tee can be adjusted by moving the inner capillary within the outer capillary, which allows for the acquisition of labeling times. Same timepoints were taken from 2, 4 and 20 s. This adjustment can be made in a discrete increments by withdrawing the inner capillary from the quench mixer, After labeling, the HDX reaction is acid-quenched using a conventional microLC mixing tee, and the quenched solution is passed through an acid protease column to generate the labeled peptides whose deuterium uptake is analyzed. Then, peptides are directly injected into the mass spectrometer.

For optimization, in this setup we used 100 mM ammonium formate as the protein buffer (pH 7.0), and 3% formic/40 mM ammonium (pH=2.5) to the quench the D-H exchange reaction. H-D Exchange timepoints are acquired by manually withdrawing the inner capillary within the outer capillary. The data is collected in triplicate for three timepoints (2, 4 and 20 s).

We observed, increase in the peptide sequence coverage by using CFI and optimizing the condition using formic acid as the quench buffer. Coverage increases from 59% to 65% compared to the chip set up.

The difference in deuterium uptake of the two samples, DSP#1 and DSP#2, shows insignificant differences in the protein dynamics (Figure 24). Despite different charges in both samples, the structural dynamics of the protein in DSP#1 and DSP#1 sample are the same. In conclusion, the charges variances noticed by Jazz pharmaceuticals during the in-process testing is unrelated to the PA tag attached to the protein. Our data comes in agreement with other testing done by (Appendix B Figure B3).

(a)



(b)

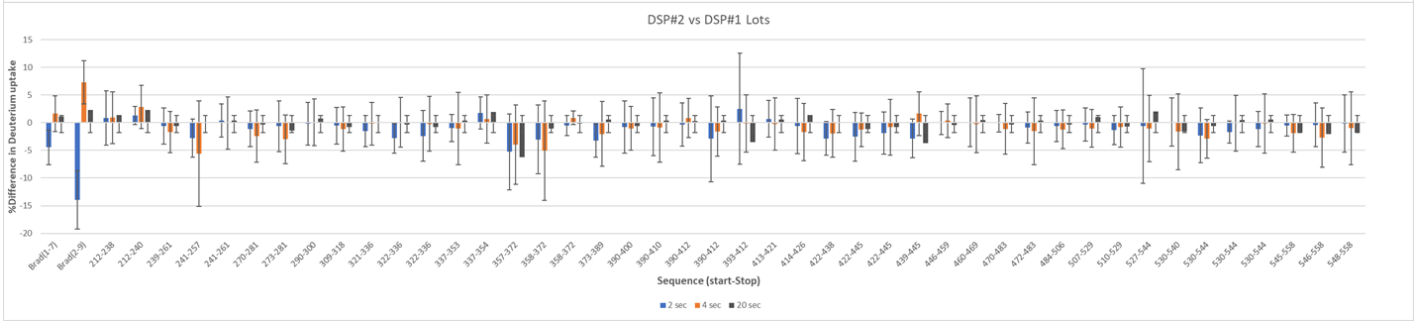


Figure 24. CFI-TRESI-HDX vs Chip system data analysis.

Show that the difference in deuterium uptake in two tagged protein samples are the same (DSP#1 and DSP#2). a) Chip and pump injections set up and (b) Continuous Flow Injection (CFI) set up. In both set ups, the PA tag shows insignificant differences between the two samples.

3.3.3 Proteomics Analysis

During the JZP-341 protein manufacturing, Jazz identified impurity that was not expected in the drug substance. Jazz experimental evidence of a new impurity but conflicting evidence whether that it is an HCP or a product variant. Unknown band at ~70 kDa was seen in the referred sample. The possible impurity is visible on western blot runs. Three western blots were done: in the first gel a different HCP kit antibody was used and showed a band at ~ 70 kDa. The second gel an anti-core enzyme antibody kit was used and showed several asparaginase bands including one at a band at ~ 70 kDa. The third gel an anti PAS tail antibody kit was used and showed many bands including the same 70 kDa band (Appendix B FigureB1).

To clear this myth, we run MS based analysis to identify these extra bands. The company sent sample D (JZP341-221103) to investigate at the York University proteomics lab. Sample D (JZP341-221103) was prepared as mentioned then digested by pepsin in our lab and sent to the Proteomics Center for complete profiling. Briefly, the digested sample is separated on a reverse phase chromatographic column coupled to a tandem MS instrument (top panel).

The MS measures the precursor mass of each peptide, selecting a pre-defined number of the most intense parent ions for fragmentation by collisional activation. The output is a MS/MS spectrum containing fragments ions corresponding to a peptide sequence. A protein database is theoretically digested with pepsin, resulting in a list of all possible peptides and their associated MS/MS spectra. A comparison between experimental and theoretical MS/MS spectrum enables assignment of peptide sequence from which peptide and protein identity can be inferred within a level of confidence.

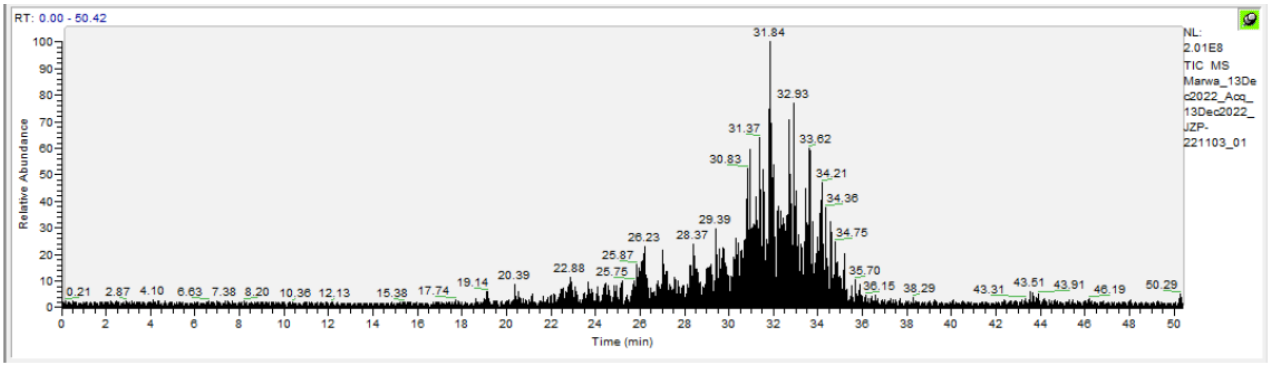
For analysis, we used the host organism *Pseudomonas fluorescens* database alongside the Contaminant Repository for Affinity Purification (CRAPome). This is a large database of standardized negative controls, aggregated from several leading labs specializing in affinity purification mass spectrometry. This database provides a qualitative and semiquantitative description of the propensity of a protein to be ‘nonspecific’ in a protocol.

`

Visual inspection of total ion chromatogram (TIC) from sample D (JZP341-221103) digests show no unusual peptide intensity has appeared (Figure 25). A peptide mapping of the identified

protein sequence was performed from mass spectrometry data and the resulting TIC comprising annotated peptide species is Table 3.

(a)



(b)

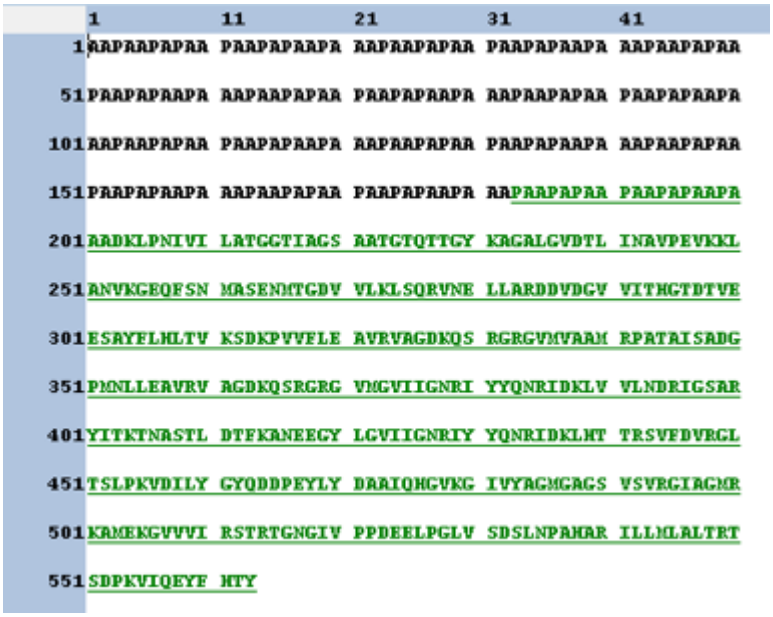


Figure 25. (a) Total ion chromatograms and (b) amino acid sequence coverage of L-asparaginase from *Pseudomonas Fluorescens*, 66% of the sequence were covered by pepsin peptides.

The results from protein queries employing CRAPome database showed L-asparaginase as the major protein in sample D (JZP341-221103) with 66% coverage. PEPA pig was used as pepsin and added during digestion. Keratin was also detected but is most probably introduced during processing. Besides pepsin and keratin, contaminants from sample processing, no further proteins were identified within the acceptance parameters.

Table 3: CRAPome database protein results for sample D (JZP341-221103).

Protein FDR Confidence: Sequest HT	Sequence	Description	# Unique Peptides	MW [kDa]	Coverage [%]
High	AAPAAPAPAAPAAPAPAAPAAAPAAPAPAAPAAI	Protein of interest	164	55.1	66
High	MKWLLLLSLVVLSECLVKVPLVRKKSLRQNLIKNGKI	PEPA_PIG (leucylaminopeptidase E.coli)	10	41.3	24
High	SRQFSSRSGYRSGGGFSSGSAGIINYQRRTTSSSTR	K2C1_HUMAN (Kertain)	4	65.8	8
Low	MSCRQFSSSYLSRSGGGGGGGLGSGGSIRSSYSRF	K1C9_HUMAN (Kertain)	2	62.1	4

Modified L-asparaginase was identified at ~55 kDa with 164 peptides of 66% sequence coverage. The sequence coverage of modified L-asparaginase is significantly higher than for the two other identified proteins, pepsin and keratin. No major protein ~ 70 kDa was identified as HCP. Our data is in alignment with work done by FyoniBio for Jazz, working on the same sample with different method using trypsin (Appendix B Table B1)

To further elucidate the unknown 70 kDa band, we used Universal Deconvolution of mass spectra (UniDec)[89]. This is software which provides a quick, robust, and flexible deconvolution of mass spectra and ion mobility-mass spectra. UniDec analysis shows four major states in the sample D (JZP341-221103), which corresponds to monomer, dimer, trimer, and the native tetrameric structure of the modified l-asparaginase (Figure 26).

The observed molecular weight of ~50kDa matches the calculated mass of the protein construct comprising an N-terminal polypeptide tail. Interestingly, dimer peak intensity is the highest among the other species. Since the dimer peak is more intense, a possible dimer without a tail can exist in the sample in very untraceable amount (70 kDa without the PA tag).

Therefore, the unknown 70 kDa band is possibly corresponding to a degradant of the product and not as HCP. That comes as a logical conclusion, as was seen in the western blot data from the company (Figure 25, b).

Taken all together, the identified protein L-asparaginase is the major compound in sample D (JZP341-221103) and no other contaminating protein was detected. The expected size of the L-asparaginase construct, comprising a poly-PA tail is 51 kDa and matches our proteomics analysis. The unknown ~70kDa band is not an HCP and more likely a degraded species of the product.

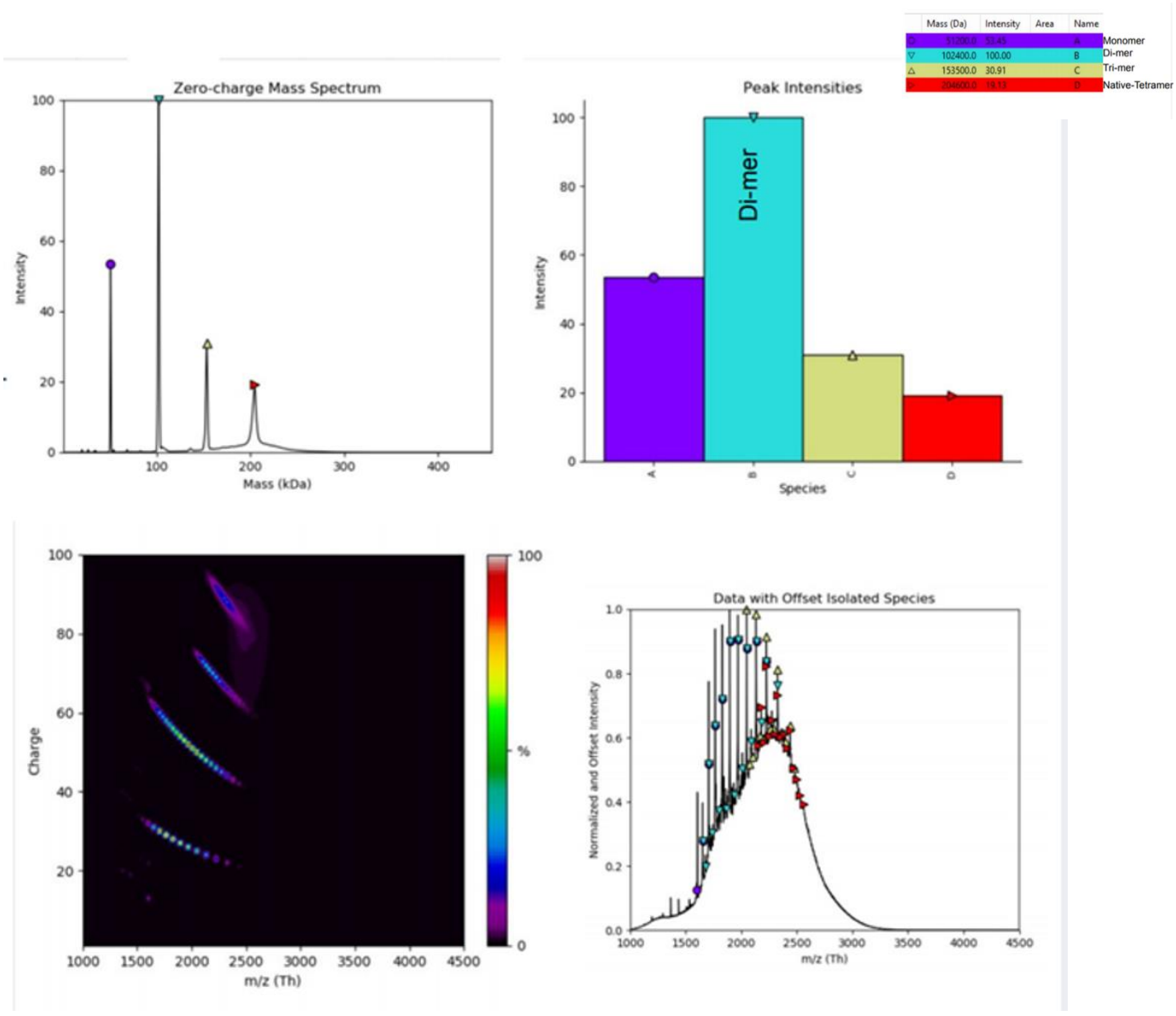


Figure 26. UniDec software deconvolution. Dimer is the most intese peak and corresponding to ~ 100 kDa of the PA-tagged l-asparaginase molecule.

3.4 Conclusions

This study aimed to study the impact of PA-tag on l-asparaginase protein dynamics. Our current analysis demonstrates that the PA-tail fragment did not change the core protein dynamics using TRESI-HDX to provide identical stability with untagged l-asparaginase.

With TRESI-HDX showing that both L-asparaginase and PA-L-asparaginase show identical results in protein dynamics, it would be interesting to test the PA tag relation to charge variants among different batches. Global HDX is a helpful method to see the overall change in protein samples.

The tag is composed mainly of proline and alanine amino acids. Proline has been known to have cis and trans isomerization structures.[90,91,92] Tag composition of proline may hypothetically have impacted the charge variance among batches and needs to be tested. Also, the way of ionization of each molecule is of interest. This raises two questions that we will explore: What is the minimum length of unstructured 'tail' required for a 'native' protein to favor chain ejection over charge residue, and (2) can PA tails be used as 'ionization enhancement tags' to facilitate detection of large or otherwise 'difficult to ionize' protein complexes?

Chapter 4

Conclusion and Future Work

The work presented here represents successful implementation of the HDX coupled to mass spectrometer for investigating therapeutic protein dynamics. In Chapter 1 we applied the HDX principles and methods couples to this tool. Ion mobility, and hybrid mass spectrometers are introduced and categorized into three significant areas: pre-ionization, in-ionization, and post ionization, representing areas on a mass spectrometer where research efforts has accumulated throughout the history. Our focus is dedicated to time-resolved methods for characterizing our therapeutic proteins.

In Chapter 2, The optimized CFI-HDX-MS method is ready for deployment in monitoring of automated workflow. We revealed the structural interaction between LMWH and PAI-1, our data provide a detailed image where it cannot be captured by x-crystallography. Fondaparinux is a potential inhibitor for controlling PAI-1 overexpression in thrombus cases. Future work to complement the results for the longer timepoints and test the effect on latency transition. Also, different types of commercial heparin can be tested to better understand conformational changes with different sources of heparin.

In Chapter 3, the conformational dynamic of modified l-asparaginase was investigated by bottom-up approach. Refinement of sample digestion strategies for proteomics analysis also facilitated more comprehensive analysis of the samples. Proteomics was performed to elucidate host cell protein presence or absence. After specifying that PA tag has no impact on the core protein dynamics, thus the PA tag is an inert polymer with no significant impact on protein conformation. The PA tag causes the protein to ionize by chain ejection model hence can work as an enhancement tag for ionization enhancement studies. Future studies in this part can implement the required length for this tail for ionization enhancement and to potentially discover new fundamentals of ESI for macromolecules.

References

1. Donnan, F. G. Rays of positive electricity and their application to chemical analyses. By Sir J. J. Thomson, O. M. F.R.S. Second edition. Pp. x + 237. London: Longmans, Green and Co., 1921. Price 16s. *Journal of the Society of Chemical Industry* **42**, 861–861 (1923).
2. Griffiths, J. A Brief History of Mass Spectrometry. *Anal. Chem.* **80**, 5678–5683 (2008).
3. Tanaka, K. *et al.* Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **2**, 151–153 (1988).
4. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64–71 (1989).
5. Kebarle, P. & Verkerk, U. H. Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrometry Reviews* **28**, 898–917 (2009).
6. Kebarle, P. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J Mass Spectrom* **35**, 804–817 (2000).
7. Banerjee, S. & Mazumdar, S. Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *Int J Anal Chem* **2012**, 282574 (2012).
8. Felitsyn, N., Peschke, M. & Kebarle, P. Origin and number of charges observed on multiply-protonated native proteins produced by ESI. *International Journal of Mass Spectrometry* **219**, 39–62 (2002).
9. Ahadi, E. & Konermann, L. Surface Charge of Electrosprayed Water Nanodroplets: A Molecular Dynamics Study. *J. Am. Chem. Soc.* **132**, 11270–11277 (2010).
10. Banerjee, S. & Mazumdar, S. Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *Int J Anal Chem*, 282574 (2012).
11. Konermann, L., Ahad, E., Rodriguez, A. D. & Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **85**, 2–9 (2013).
12. Susa, A. C., Xia, Z., Tang, H. Y. H., Tainer, J. A. & Williams, E. R. Charging of Proteins in Native Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **28**, 332–340 (2017).
13. Allen, M. H. & Vestal, M. L. Design and performance of a novel electrospray interface. *J Am Soc Mass Spectrom* **3**, 18–26 (1992).
14. Kebarle, P. & Verkerk, U. H. Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom Rev* **28**, 898–917 (2009).

15. Körner, R., Wilm, M., Morand, K., Schubert, M. & Mann, M. Nano electrospray combined with a quadrupole ion trap for the analysis of peptides and protein digests. *J Am Soc Mass Spectrom* **7**, 150–156 (1996).

16. Glish, G. L. & Burinsky, D. J. Hybrid mass spectrometers for tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* **19**, 161–172 (2008).

17. Tang, X. J. *et al.* Investigation of protein-protein noncovalent interactions in soybean agglutinin by electrospray ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* **8**, 750–754 (1994).

18. Chernushevich, I. V., Loboda, A. V. & Thomson, B. A. An introduction to quadrupole-time-of-flight mass spectrometry. *J Mass Spectrom* **36**, 849–865 (2001).

19. McDaniel, E. W., Martin, D. W. & Barnes, W. S. Drift Tube-Mass Spectrometer for Studies of Low-Energy Ion-Molecule Reactions. *Review of Scientific Instruments* **33**, 2–7 (1962).

20. Mason, E. A.; McDaniel, E. W. Transport Properties of Ions in Gases; Wiley, 1988.

21. Cumeras, R., Figueras, E., Davis, C. E., Baumbach, J. I. & Gràcia, I. Review on Ion Mobility Spectrometry. Part 1: Current Instrumentation. *Analyst* **140**, 1376–1390 (2015).

22. Mason, E. A. & Schamp, H. W. Mobility of gaseous lons in weak electric fields. *Annals of Physics* **4**, 233–270 (1958).

23. Verbeck, G., Ruotolo, B., Sawyer, H., Gillig, K. & Russell, D. A fundamental introduction to ion mobility mass spectrometry applied to the analysis of biomolecules. *J Biomol Tech* **13**, 56–61 (2002).

24. Giles, K. *et al.* Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Communications in Mass Spectrometry* **18**, 2401–2414 (2004).

25. Pringle, S. D. *et al.* An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *International Journal of Mass Spectrometry* **261**, 1–12 (2007).

26. Pringle, S. D. *et al.* An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *International Journal of Mass Spectrometry* **261**, 1–12 (2007).

27. Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. The Energy Landscapes and Motions of Proteins. *Science* **254**, 1598–1603 (1991).

28. Lindorff-Larsen, K., Best, R. B., DePristo, M. A., Dobson, C. M. & Vendruscolo, M. Simultaneous determination of protein structure and dynamics. *Nature* **433**, 128–132 (2005).

29. Englander, S. W., Mayne, L., Bai, Y. & Sosnick, T. R. Hydrogen exchange: the modern legacy of Linderstrøm-Lang. *Protein Sci* **6**, 1101–1109 (1997).

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

31. Smith, D. L., Deng, Y. & Zhang, Z. Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J Mass Spectrom* **32**, 135–146 (1997).

32. Giladi, M. & Khananshvili, D. Hydrogen-Deuterium Exchange Mass-Spectrometry of Secondary Active Transporters: From Structural Dynamics to Molecular Mechanisms. *Frontiers in Pharmacology* **11**, (2020).

33. Harrison, R. A. & Engen, J. R. Conformational insight into multi-protein signaling assemblies by hydrogen-deuterium exchange mass spectrometry. *Curr Opin Struct Biol* **41**, 187–193 (2016).

34. Englander, S. W., Sosnick, T. R., Englander, J. J. & Mayne, L. Mechanisms and uses of hydrogen exchange. *Current Opinion in Structural Biology* **6**, 18–23 (1996).

35. Konermann, L., Pan, J. & Liu, Y.-H. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* **40**, 1224–1234 (2011).

36. Brown, K. A., Lento, C., Rajendran, S., Dowd, J. & Wilson, D. J. Epitope Mapping for a Preclinical Bevacizumab (Avastin) Biosimilar on an Extended Construct of Vascular Endothelial Growth Factor A Using Millisecond Hydrogen–Deuterium Exchange Mass Spectrometry. *Biochemistry* **59**, 2776–2781 (2020).

37. Egler, R. A., Ahuja, S. P. & Matloub, Y. L-asparaginase in the treatment of patients with acute lymphoblastic leukemia. *J Pharmacol Pharmacother* **7**, 62–71 (2016).

38. Pieters, R. *et al.* L-asparaginase treatment in acute lymphoblastic leukemia: A focus on Erwinia asparaginase. *Cancer* **117**, 238–249 (2011).

39. Chand, S. *et al.* A comprehensive review on microbial L -asparaginase: Bioprocessing, characterization, and industrial applications. *Biotechnology and Applied Biochemistry* **67**, 619–647 (2020).

40. Jazz Pharmaceuticals. An Open-Label, Multicenter Study of Recombinant Crisantaspase Produced in Pseudomonas Fluorescens (RC-P) in Patients With Acute Lymphoblastic Leukemia (ALL)/Lymphoblastic Lymphoma (LBL) Following Hypersensitivity to E. Coli-derived Asparaginases. Clinicaltrials.gov registration number: NCT04145531 (2022).

41. Melik-Nubarov, N. S. *et al.* PEGylated recombinant L-asparaginase from Erwinia carotovora: Production, properties, and potential applications. *Appl Biochem Microbiol* **53**, 165–172 (2017).

42. Meneguetti, G. P. *et al.* Novel site-specific PEGylated L-asparaginase. *PLoS ONE* **14**, 0211951 (2019).

43. Schlapschy, M. *et al.* PASylation: a biological alternative to PEGylation for extending the plasma half-life of pharmaceutically active proteins. *Protein Eng Des Sel* **26**, 489–501 (2013).

44. Breibeck, J. & Skerra, A. The polypeptide biophysics of proline/alanine-rich sequences (PAS): Recombinant biopolymers with PEG-like properties. *Biopolymers* **109**, e23069 (2018).

45. Breibeck, J. *et al.* PAS-cal: a Generic Recombinant Peptide Calibration Standard for Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **25**, 1489–1497 (2014).

46. Binder, U. & Skerra, A. PASylation®: A versatile technology to extend drug delivery. *Current Opinion in Colloid & Interface Science* **31**, 10–17 (2017).

47. Binder, B. R. *et al.* Plasminogen activator inhibitor 1: physiological and pathophysiological roles. *News Physiol Sci* **17**, 56–61 (2002).

48. Lin, H. *et al.* Therapeutics targeting the fibrinolytic system. *Exp Mol Med* **52**, 367–379 (2020).

49. Wind, T., Hansen, M., Jensen, J. K. & Andreasen, P. A. The molecular basis for anti-proteolytic and non-proteolytic functions of plasminogen activator inhibitor type-1: roles of the reactive centre loop, the shutter region, the flexible joint region and the small serpin fragment. *Biol Chem* **383**, 21–36 (2002).

50. Trelle, M. B. *et al.* Hydrogen/Deuterium Exchange Mass Spectrometry Reveals Specific Changes in the Local Flexibility of Plasminogen Activator Inhibitor 1 upon Binding to the Somatomedin B Domain of Vitronectin. *Biochemistry* **51**, 8256–8266 (2012).

51. Sillen, M. & Declerck, P. J. Targeting PAI-1 in Cardiovascular Disease: Structural Insights Into PAI-1 Functionality and Inhibition. *Frontiers in Cardiovascular Medicine* **7**, (2020).

52. Sillen, M., Weeks, S. D., Strelkov, S. V. & Declerck, P. J. Structural Insights into the Mechanism of a Nanobody That Stabilizes PAI-1 and Modulates Its Activity. *Int J Mol Sci* **21**, 5859 (2020).

53. Egelund, R. *et al.* A regulatory hydrophobic area in the flexible joint region of plasminogen activator inhibitor-1, defined with fluorescent activity-neutralizing ligands. Ligand-induced serpin polymerization. *J Biol Chem* **276**, 13077–13086 (2001).

54. Declerck, P. J. *et al.* Purification and characterization of a plasminogen activator inhibitor 1 binding protein from human plasma. Identification as a multimeric form of S protein (vitronectin). *J Biol Chem* **263**, 15454–15461 (1988).

55. Egelund, R., Schousboe, S. L., Sottrup-Jensen, L., Rodenburg, K. W. & Andreasen, P. A. Type-1 plasminogen-activator inhibitor -- conformational differences between latent, active, reactive-centre-cleaved and plasminogen-activator-complexed forms, as probed by proteolytic susceptibility. *Eur J Biochem* **248**, 775–785 (1997).

56. Kearon, C. *et al.* Antithrombotic therapy for VTE disease: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* **141**, e419S-e496S (2012).

57. Obi, A. T. *et al.* Low-molecular-weight heparin modulates vein wall fibrotic response in a plasminogen activator inhibitor 1-dependent manner. *Journal of Vascular Surgery: Venous and Lymphatic Disorders* **2**, 441-450.e1 (2014).

58. Ehrlich, H. J., Keijer, J., Preissner, K. T., Gebbink, R. K. & Pannekoek, H. Functional interaction of plasminogen activator inhibitor type 1 (PAI-1) and heparin. *Biochemistry* **30**, 1021–1028 (1991).

59. Yasar Yildiz, S., Kuru, P., Toksoy Oner, E. & Agirbasli, M. Functional Stability of Plasminogen Activator Inhibitor-1. *ScientificWorldJournal* **2014**, 858293 (2014).

60. Law, R. H. P. *et al.* An overview of the serpin superfamily. *Genome Biol* **7**, 216 (2006).

61. Leader, B., Baca, Q. J. & Golan, D. E. Protein therapeutics: a summary and pharmacological classification. *Nat Rev Drug Discov* **7**, 21–39 (2008).

62. Gong, L. *et al.* Crystal Structure of the Michaelis Complex between Tissue-type Plasminogen Activator and Plasminogen Activators Inhibitor-1 *. *Journal of Biological Chemistry* **290**, 25795–25804 (2015).

63. Cazzolli, G. *et al.* Serpin latency transition at atomic resolution. *Proc Natl Acad Sci U S A* **111**, 15414–15419 (2014).

64. Petersen, M., Madsen, J. B., Jørgensen, T. J. D. & Trelle, M. B. Conformational preludes to the latency transition in PAI-1 as determined by atomistic computer simulations and hydrogen/deuterium-exchange mass spectrometry. *Sci Rep* **7**, 6636 (2017).

65. Kost, C., Stüber, W., Ehrlich, H. J., Pannekoek, H. & Preissner, K. T. Mapping of binding sites for heparin, plasminogen activator inhibitor-1, and plasminogen to vitronectin's heparin-binding region reveals a novel vitronectin-dependent feedback mechanism for the control of plasmin formation. *Journal of Biological Chemistry* **267**, 12098–12105 (1992).

66. Kellici, T. F., Pilka, E. S. & Bodkin, M. J. Therapeutic Potential of Targeting Plasminogen Activator Inhibitor-1 in COVID-19. *Trends in Pharmacological Sciences* **42**, 431–433 (2021).

67. Naumnik, B., Pawlak, K. & Mys'liwiec, M. Unfractionated heparin but not enoxaparin causes delayed plasma PAI-1 depletion in hemodialysis patients: a prospective study. *Clin Appl Thromb Hemost* **15**, 84–91 (2009).

68. Baluta, M. M. & Vintila, M. M. PAI-1 Inhibition – Another Therapeutic Option for Cardiovascular Protection. *Maedica (Bucur)* **10**, 147–152 (2015).

69. Anacleto, J. *et al.* Apparatus for Automated Continuous Hydrogen Deuterium Exchange Mass Spectrometry Measurements from Milliseconds to Hours. *Anal. Chem.* **95**, 4421–4428 (2023).

70. Rey, M. *et al.* Mass spec studio for integrative structural biology. *Structure* **22**, 1538–1548 (2014).

71. Madsen, J. B. *et al.* RNA aptamers as conformational probes and regulatory agents for plasminogen activator inhibitor-1. *Biochemistry* **49**, 4103–4115 (2010).

72. Vimal, A. & Kumar, A. Biotechnological production and practical application of L-asparaginase enzyme. *Biotechnology and Genetic Engineering Reviews* **33**, 40–61 (2017).

73. Loch, J. I. & Jaskolski, M. Structural and biophysical aspects of L -asparaginases: a growing family with amazing diversity. *IUCrJ* **8**, 514–531 (2021).

74. Batool, T., Makky, E. A., Jalal, M. & Yusoff, M. M. A Comprehensive Review on L-Asparaginase and Its Applications. *Appl Biochem Biotechnol* **178**, 900–923 (2016).

75. Aghaiypour, K., Wlodawer, A. & Lubkowski, J. Structural Basis for the Activity and Substrate Specificity of *Erwinia chrysanthemi* L -Asparaginase ᵎ. *Biochemistry* **40**, 5655–5664 (2001).

76. Radadiya, A., Zhu, W., Coricello, A., Alcaro, S. & Richards, N. G. J. Improving the Treatment of Acute Lymphoblastic Leukemia. *Biochemistry* **59**, 3193–3200 (2020).

77. Robertson, J., Barr, R., Shulman, L. N., Forte, G. B. & Magrini, N. Essential medicines for cancer: WHO recommendations and national priorities. *Bull World Health Organ* **94**, 735–742 (2016).

78. Chiu, M., Taurino, G., Bianchi, M. G., Kilberg, M. S. & Bussolati, O. Asparagine Synthetase in Cancer: Beyond Acute Lymphoblastic Leukemia. *Front. Oncol.* **9**, 1480 (2020).

79. Lubkowski, J., Dauter, M., Aghaiypour, K., Wlodawer, A. & Dauter, Z. Atomic resolution structure of Erwinia chrysanthemi L-asparaginase. *Acta Crystallogr D Biol Crystallogr* **59**, 84–92 (2003).

80. Yari, M., Eslami, M., Ghoshoon, M. B., Nezafat, N. & Ghasemi, Y. Decreasing the immunogenicity of Erwinia chrysanthemi asparaginase via protein engineering: computational approach. *Mol Biol Rep* **46**, 4751–4761 (2019).

81. Schiefner, A., Walser, R., Gebauer, M. & Skerra, A. Proline/alanine-rich sequence (PAS) polypeptides as an alternative to PEG precipitants for protein crystallization. *Acta Crystallogr F Struct Biol Commun* **76**, 320–325 (2020).

82. Zhu-Shimoni, J. *et al.* Host cell protein testing by ELISAs and the use of orthogonal methods. *Biotechnology and Bioengineering* **111**, 2367–2379 (2014).

83. Molden, R. *et al.* Host cell protein profiling of commercial therapeutic protein drugs as a benchmark for monoclonal antibody-based therapeutic protein development. *mAbs* **13**, 1955811 (2021).

84. Hogwood, C. E., Bracewell, D. G. & Smales, C. M. Measurement and control of host cell proteins (HCPs) in CHO cell bioprocesses. *Current Opinion in Biotechnology* **30**, 153–160 (2014).

85. Reisinger, V., Toll, H., Mayer, R. E., Visser, J. & Wolschin, F. A mass spectrometry-based approach to host cell protein identification and its application in a comparability exercise. *Anal Biochem* **463**, 1–6 (2014).

86. Brown, K. A., Rajendran, S., Dowd, J. & Wilson, D. J. Rapid characterization of structural and functional similarity for a candidate bevacizumab (Avastin) biosimilar using a multipronged mass-spectrometry-based approach. *Drug Test Anal* **11**, 1207–1217 (2019).

87. Nouchikian, L. Comparing Solution-Phase and Gas-Phase Protein Stability Using Ion Mobility and Differential Mobility Mass Spectrometry (Publication No. **36327**) [Master Dissertation, York University]. ProQuest Dissertations Publishing (2019).

88. Strohalm, M., Hassman, M., Kosata, B. & Kodícek, M. mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Commun Mass Spectrom* **22**, 905–908 (2008).

89. Marty, M. T. *et al.* Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).

90. Kern, D., Schutkowski, M. & Drakenberg, T. Rotational Barriers of cis/trans Isomerization of Proline Analogues and Their Catalysis by Cyclophilin. *J. Am. Chem. Soc.* **119**, 8403–8408 (1997).

91. Kawagoe, S., Nakagawa, H., Kumeta, H., Ishimori, K. & Saio, T. Structural insight into proline cis/trans isomerization of unfolded proteins catalyzed by the trigger factor chaperone. *Journal of Biological Chemistry* **293**, 15095–15106 (2018).

92. Melis, C., Bussi, G., Lummis, S. C. R. & Molteni, C. Trans−cis Switching Mechanisms in Proline Analogues and Their Relevance for the Gating of the 5-HT3 Receptor. *J. Phys. Chem. B* **113**, 12148–12153 (2009).

Appendix A

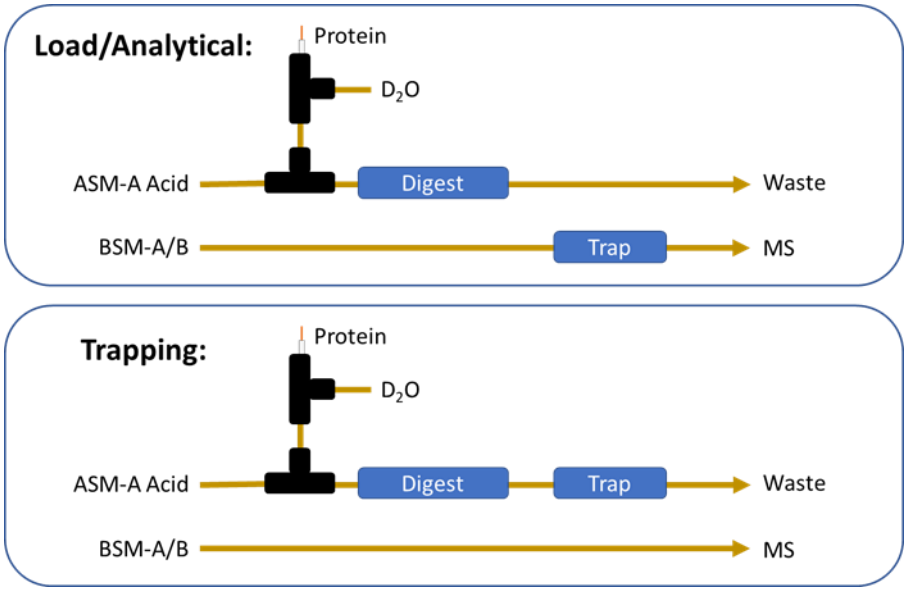
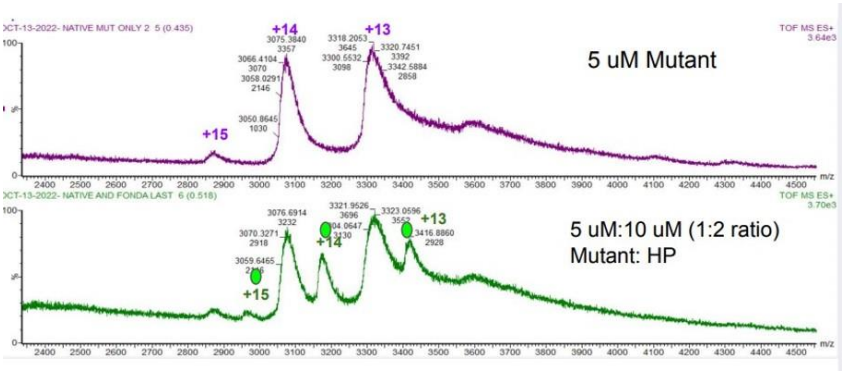


Figure A1: CFIT-TRESI connected directly to digest column, injection valve not used. Digested protein is trapped and "washed" on the Trap Column. Peptides then rapidly elute off the Trap directly to the MS. New CFIT-TRESI Peak Shape Simulator

Figure A2 (a) Native MS of bonded and unbonded PAI-1



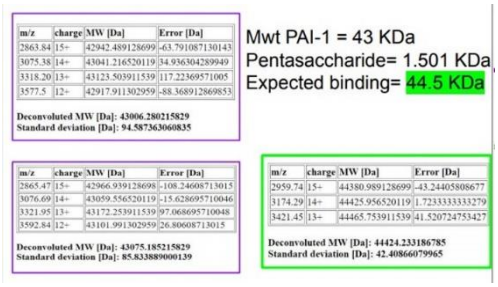
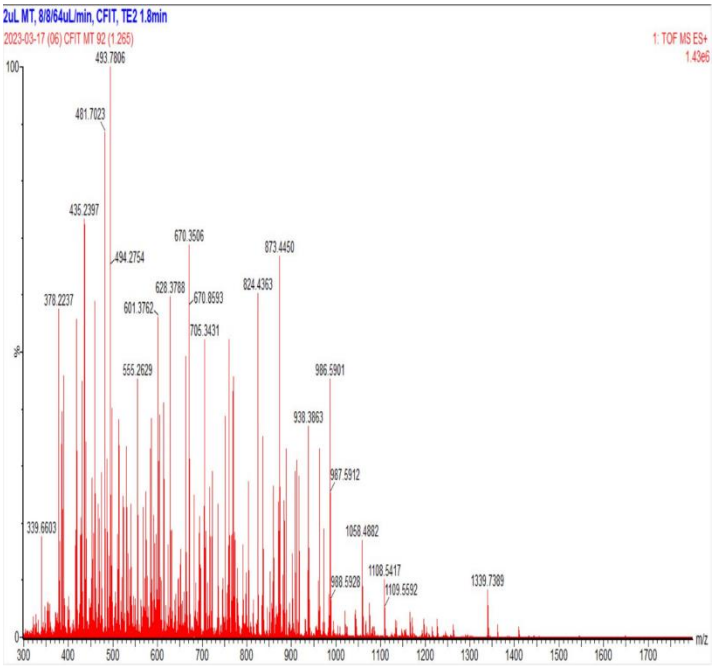


Figure A2 (b) Digest spectrum of PAI-1 Mutant and Wild respectively.



Appendix B

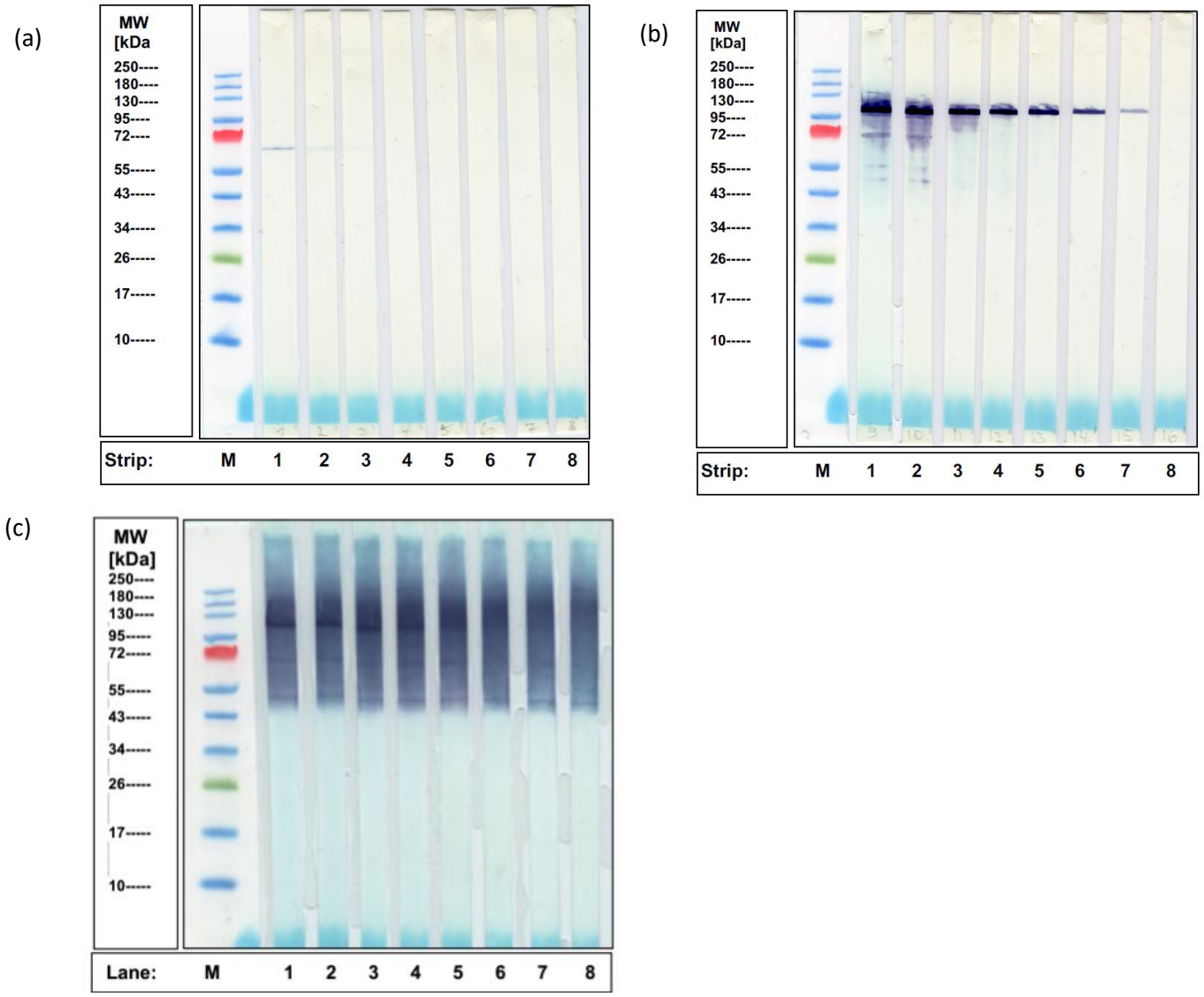


Figure B1. Jazz Pharmaceuticals Western Blot.

(a) Western blot of a different HCP kit antibody (QC 12/14 Final Bulk and immunostaining with anti-JP-HCP-total-IgG, lot: CE-131020-01).
(b) Western blot of anti-core enzyme (QC 12/14 Final Bulk and immunostaining with p-rab-anti-cris, lot: P 24039).
(c) Western blot of JP_final bulk and immunostaining with anti PAS tail antibody

The molecular weight (MW) is indicated on the left (pres-tained protein standard, M). Protein amount per strip: ~0.5 µg. Strips 1-8: Antibody concentrations: 15, 10, 5, 2, 1, 0.5, 0.2, 0.1 µg/mL

Table B1: SwissProt protein results for the sample D using trypsin digestion, run by FyoniBio GmbH.

Accession	Protein	MW [kDa]	Score	#Peptides	SC [%]
P06608|ASPG_DICCH	L-asparaginase OS=Dickeya chrysanthemi OX=556 GN=ansB PE=1 SV=1	37.6	2065.5	36	82.8
P00761|TRYP_PIG	Trypsin OS=Sus scrofa OX=9823 PE=1 SV=1	24.4	619.0	19	25.1
A5A6M6|K2C1_PANTR	Keratin, type II cytoskeletal 1 OS=Pan troglodytes OX=9598 GN=KRT1 PE=2 SV=1	65.4	569.0	13	21.4
P35908|K22E_HUMAN	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens OX=9606 GN=KRT2 PE=1 SV=2	65.4	515.5	9	15.2
P13645|K1C10_HUMAN	Keratin, type I cytoskeletal 10 OS=Homo sapiens OX=9606 GN=KRT10 PE=1 SV=6	58.8	359.0	10	19.3
P35527|K1C9_HUMAN	Keratin, type I cytoskeletal 9 OS=Homo sapiens OX=9606 GN=KRT9 PE=1 SV=3	62.0	260.0	6	12.7

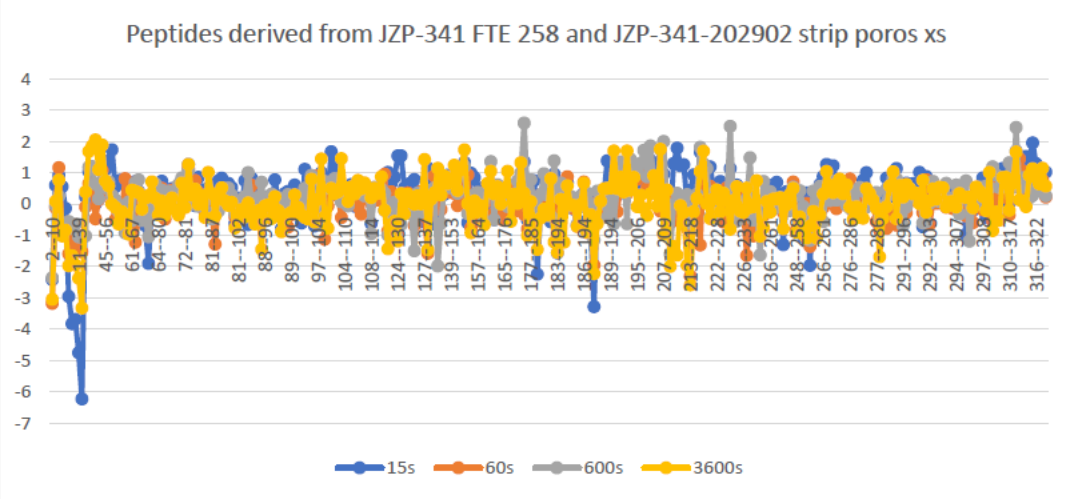


Figure B2. Done by NovaBioassys. Difference plot showing the maximum deuterium uptake percentage difference (ΔD%, y axis) for all the peptides derived from JZP-341 (JZP-341 FTE 258 vs JZP-341-202902 strip poros xs). NovaBioassay used Pepsin/protease XIII and prolyl endoprotease digestion and LC-MS. They recovered 100% sequence coverage was achieved for JZP-458 RM-M-0009 and three JZP-341 samples core enzyme part. They found the maximum difference was less than 3% for most of the peptides, indicating no significant conformational differences are observed among those samples.