Mass Spectrometry Methods for Enhanced Characterization of Biological Systems

LISA RACHEL SZYMKOWICZ

A THESIS SUBMITTED TO

THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULLFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN CHEMISTRY

YORK UNIVERSITY

TORONTO, ONTARIO

April 2019

ABSTRACT

As the pharmaceutical industry moves towards larger protein-based therapeutics over chemically synthesized small molecules, there are increasing demands for parallel advancement in the technologies used for quality control of these biologically-derived products (biopharmaceuticals). This work describes the implementation of mass spectrometry (MS)-based methods in three areas where MS is well-positioned to fulfill unmet analytical needs. First, time-resolved electrospray ionization hydrogen-deuterium exchange (TRESI-HDX) and ion mobility spectroscopy (IMS) MS were used to study conformational dynamics in transient Cytochrome *c*-lipid interactions. MS was also implemented for proteomic characterization of in-process and purified samples from acellular Pertussis vaccines. Further development of a targeted MS method is described for quantitation of two residual protein toxins, adenylase cyclase toxin (ACT) and dermoneucrotic toxin (DNT), and a glycopeptide, tracheal cytotoxin (TCT) from *Bordetella pertussis*. Collectively, the results highlight advancement in the application of structural and quantitative proteomics for more effective characterization of next generation biopharmaceuticals.

ACKNOWDLEDGMENTS

I would like to thank my supervisors, Dr. Derek Wilson and Dr. Andrew James, for providing the unique opportunity to explore both academic and industry research projects as an MSc student. Not only are you both great mentors, you provided an ideal balance of freedom and direction which has been instrumental to my development. Derek, thank you creating a space for me to learn and grow as researcher. You've always pushed me to "just go for it" and gave me liberty to explore my own scientific curiosities even when they fell outside your area of primary interest (i.e. proteomics). Andrew, thank you for inspiring me to pursue mass spectrometry. Not many people can see something that catches their eye in a seminar as an undergraduate co-op student and get an opportunity to pursue an MSc while learning to do what they saw. You took me under your wing and continue to challenge me to grow in what I know. I would also like to acknowledge my committee members Dr. Yi Sheng and Dr. Mark Bayfield for valuable feedback and support.

At Sanofi Pasteur, I would like to recognize my first manager, Michael Leach, and scientist Elena Newman, for teaching the foundation of working in industry and encouraging me to keep my eyes open to discover my interests. I would also like to thank Neil Blackburn, Marina Kirkitadze, and Bruce Carpick for sharing their knowledge and wisdom along the way. To my colleagues in the MS TAG group, Maxime Rossato and Manon Fradin, thank you for the ongoing support and inspiring me to continue to work on my French.

To Peter Liuni and Shaolong Zhu, thank you the informative discussions and teaching me the importance of balance in the life of a grad student/researcher. Having you both around as friendly faces having "been there and done that" has meant more than I can express. To Cristina Lento, thank you for all the edits, technical and moral support since day #1 in the Wilson lab. Special thanks to the other group members who have made my time in and out of the lab so memorable – Kerene Brown, Xiaojing Huang, Ruth Knox, Banafsheh Mehrazma, Luci Nouchikian, Irina Oganeysan, and John van Nostrand.

Finally, I would like to thank my parents for their unwavering love, care and support, without which none of this would be possible.

TABLE OF CONTENTS

CHAPTERS

ABBREVIATIONS

ABC – Ammonium bicarbonate
ACN - Acetonitrile
ACT - Adenylate Cyclase Toxin
AGC – Automatic Gain Control
aP – acellular Pertussis
AQUA – Absolutely quantified
atm – attamole (10^{-18})
C18 – Carbon 18
CID – Collision Induced Dissociation
CL - Cardiolipin
Cyt *c* – Cytochrome c
Da – Dalton
DC – Direct currect
DDA – Data dependent acquisition
DHPC - Dihexanoylphosphatidylcholine
DIA – Data independent acquisition
DMPC – Dimyristoylphosphatidylcholine
DNT – Dermoneucrotic Toxin
ELISA – Enzyme-linked immunosorbent assay
ESI – Electrospray Ionization
EU - European
FA – Formic acid
Fmol – femtomole (10^{-15})
FHA – filamentous hemagglutinin
FIM - fimbriae 2,3
HCD – Higher energy collision induced disassociation
HCP – Host cell protein
HDX – Hydrogen Deuterium Exchange
HF – High field
hr - hour
HPLC- High Performance Liquid Chromatography
HRAM – High resolution, accurate mass
IAA – Iodoacetamide
i.d. – inner diameter
IMS – Ion Mobility Spectrometry
kDa – Kilo Dalton
kV – Kilovolts
LC – Liquid Chromatography
LC-MS – Liquid Chromatography with Mass Spectrometry
LC-MS/MS – Liquid Chromatography with Tandem Mass Spectrometry
LOD – limit of detection
LLOQ – lower limit of quantitation
maxIT – Maximum Injection Time
MeOH - Methanol
min - minute
MRM – Multiple Reaction Monitoring
MS – Mass Spectrometry, mass spectrometer
MS/MS – Tandem mass spectrum/ spectrometry
MW – Molecular weight
m/z – mass to charge ratio
nanoLC – nano Liquid Chromatography
NMR – Nuclear Magnetic Resonance
o.d. – outer diameter

PC - Phosphocholine
Ph. Eur – European Pharmacopoeia
PMMA – Polymethyl Methacrylate
ppm – part per million
PRM – Parallel Reaction Monitoring
PRN - Pertactin
PT – Pertussis toxin
QE – Q Exactive
QTOF – Quadrupole Time of Flight
QqQ – Triple quadrupole
RF - Radio frequency
RP-HPLC – Reverse phase HPLC
rpm – revolutions per minute
RT – Retention time, room temperature
PTM - Post-translational modification
SDS - Sodium Dodecyl Sulphate
sec - second
S-trap – Suspension trap
TCEP - Tris(2-carboxyethyl) phosphine hydrochloride
TCT – Trachael cytotoxin
TEAB – Triethylammonium bicarbonate
TFA – Trifluoroacetic acid
TOF – Time of Flight
TRESI – Time Resolved Electrospray Ionization
UPLC – Ultra-Performance Liquid Chromatography
UV – Ultra Violet
v/v – volume per volume
XIC – Extracted Ion Chromatogram
w/w – weight per weight

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

Chapter 1: Introduction

1.1 – Mass Spectrometry

1.1.1 Current Perspectives

Mass spectrometry (MS) as technique for measuring the mass of molecules has existed for over 100 years, starting with the separation of Ne isotopes using a "parabola spectrograph" by Nobel Prize winner J.J Thomson in 1913.[1,2] Mass spectrometers have developed into remarkably versatile analytical instruments with a unique ability to transfer molecules into the gas phase and accurately measured molecular masses according to their specific mass-to-charge ratio (m/z). Until the late 1980s, MS was primarily used to study low molecular weight chemical compounds, as no technique existed to transfer intact biomolecules into the gas phase. This changed when K. Tanaka and J. B. Fenn developed Nobel prize winning soft ionization techniques, matrix-assisted desorption ionization (MALDI)[3] and electrospray ionization (ESI).[4] These soft ionization techniques together with fundamental improvements to ion sampling and detection have increased the analytical power of MS in an unprecedented range of applications. Today, mass spectrometers have evolved into a backbone technique in the analytical sciences, with widespread use in the fields of biology,[5] medicine,[6] and chemistry[7].

1.1.2 Electrospray Ionization

Electrospray ionization (ESI) is the soft-ionization technique of choice for online analysis of solution-phase biomolecules. Amendable to coupling with high-performance liquid chromatography (HPLC)-based separation techniques, ESI also facilitates hyphenated HPLC-MS workflows which are particularly advantageous for complex biological samples.[8] In a typical ESI source, ions are generated by passing a solution containing the analyte(s) of interest through a thin metal capillary. A high electric potential (2-6 kV) at the tip of the capillary causes formation

of a Taylor cone. In positive ion mode, charging occurs via protonation resulting in accumulation of positively charged particles at the tip of the Taylor Cone.[9] As the ejected parent droplets fly towards an oppositely charged plate inside the mass spectrometer, solvent evaporation in the gas-phase causes the droplets to shrink. The charge-to-volume ratio in the droplet increases until columbic repulsion destabilizes the structure resulting in droplet fission.[10] Parent and offspring droplets undergo numerous evaporation/jet fission cycles as shown in Figure 1.1. Described by the Rayleigh equation, the charge limit prior to a fission event (QRy) is based on the electrical permittivity (ε0), surface tension (γ), and radius of the droplet (R).[11–13]

$$QRy = 8\pi(\varepsilon 0 \gamma R^3)^{1/2}$$

Electrospray ionization is thought to occur via two mechanisms, namely the 'Charged Residue model' (CRM)[14], and 'Ion Evaporation model' (IEM)[15]. With considerable experimental evidence to support the validity of both, there is a consensus that small analytes form droplets with very small radii (< 10 nm diameter). Favouring IEM, these small droplets gain charges as they enter the gas phase and solvent is evaporated.[16] In contrast, larger molecules such as folded proteins are thought to follow CRM and undergo multiple fission/evaporation cycles until the remainder of the solvent evaporates leaving a 'residue' of charge. Of relevance to 'bottom-up' MS workflows for protein analysis, peptides from proteolytic digests may fall in the grey area between the IEM and CRM mechanisms of ionization, directly impacting absolute MS signal response.[17] Generally, IEM predicts more efficient ionization of peptides with less hydrophobicity-driven surface; a trend that holds up for small hydrophobic peptides in-solution.[18] Larger peptides are thought to mimic the ionization process for small, globular proteins with post-ionization declustering and ion transmission having the greatest influence on observed signal intensity.[17,18]

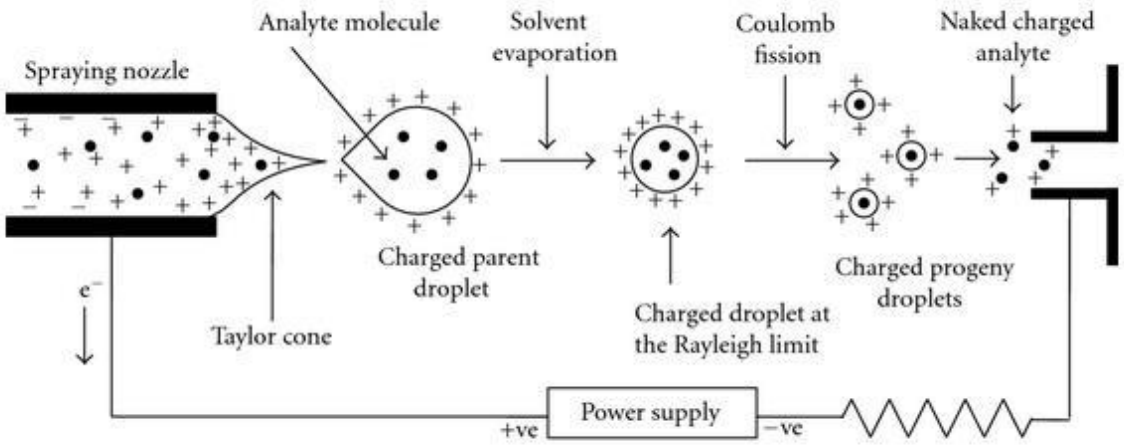


Figure 1.1. Schematic depiction of an ESI source operated in positive mode. Analyte solution passes through a metal spraying nozzle whose tip is held at a high voltage (2-6 kV). Electric potential created at the capillary tip causes the dispersion of the sample solution creating a Taylor cone, resulting in ejection of positively charged parent droplets due to overpowering of columbic forces. During transit towards the an oppositely charged curtain plate in the mass spectrometer, interplay between solvent evaporation and sequential jet fission events results in reduction in droplet size until naked charged analytes are transferred into the mass spectrometer. Adapted from Banerjee & Mazumdar, 2012.[10]

1.1.3 Nano Electrospray Ionization

Conventional ESI involves pressure-driven flow rates in the 1-100 µL/min range for capillary LC and from 0.1 µL/min to 1 mL/min for conventional flow LC. Modern electrospray sources incorporate in-line heating and collateral gas flows to hasten evaporation during the ionization process.[19,20] Despite these additions, a substantial fraction of analytes are lost due to lack of desolvation during the ESI process. An alternative approach involves the use of lower flow rates (typically 100-500 nL/min), achieved via nano-liquid chromatography (nanoLC) and narrow-bore pulled tip emitters. Droplets formed under these 'nanospray' conditions are significantly smaller (hundreds of nm), and thus ionize in tens of µs instead of the hundreds of ms with conventional ESI flowrates.[21] The decrease in solvent flow also enables reduced sample

consumption and softer ionization conditions. These advantages together with boosted in baseline sensitivity lead to the first major application of nanoESI with hyphenated nanoLC coupled ESI-MS methods (nanoLC-MS) proteomics.[22,23]

1.1.4 Hybrid Mass Spectrometers

Mass spectrometers usually consist of an ion source, mass analyzer, and data processing electronics. Instrument platforms are increasingly being built with dual MALDI/ESI source compatibility and electron multiplier or time to digital converter (TDC) detectors. Therefore, it is often the type of mass analyzer that is the major differentiator between instrument platforms, and their suitability for downstream applications.

All mass analyzers separate ions based on their mass-to-charge (m/z). However, each mass analyzer has unique properties pertaining to mass range, analysis speed, resolution, sensitivity, ion transmission, and dynamic range. Ion trap (IT), Orbitrap, and ion cyclotron resonance (ICR) mass analyzers separate ions based on their m/z resonance frequency. Excelling at high resolution tandem mass spectrometry (MS/MS), these trap-based instruments are often limited in dynamic range of detection.[24] Quadrupoles (Q) use m/z stability and oscillating electrical fields to selectively stabilize the paths of ions passing through a radio frequency (RF) field between 4 parallel rods. Configurations such as the triple quadrupole (QqQ) contain two mass selecting quads in parallel. QqQ instruments provide particularly robust quantitation and specificity when used with low-energy collision-induced dissociation (CID), but with lower resolution than trap-based analyzers. Lastly, time-of-flight (TOF) analyzers use an electric field to accelerate the ions through a common potential and measures the time taken to reach the detector. With a theoretically infinite mass range, TOFs provide the fastest scan times and are

particularly useful for analysis of large biomolecules; however, are limited in precursor selectivity and quality of MS/MS fragment ion analysis.[25]

Hybrid mass spectrometers are denoted by the combination of mass analyzers and mass filters in order of appearance along the ion path. Configurations such as the triple quadrupole (QqQ), quadrupole linear ion trap (QLIT or LTQ), Quadruple Time of Flight (QTOF), TOF-TOF, and quadrupole-Orbitrap, are increasingly common; each having their own distinct analytical characteristics.[26] For the coming work, the focus will be on the quadrupole-Orbitrap MS for proteomics (Figure 1.2), and the QTOF for structural MS applications (Figure 1.3).

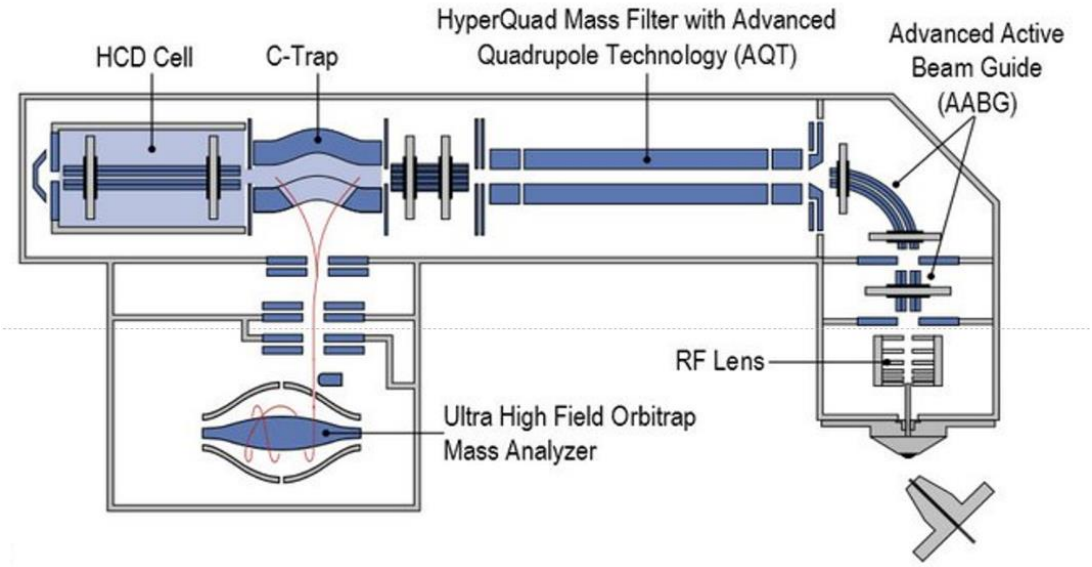


Figure 1.2. Schematic representation of the Q Exactive High Field (HF) Quadrupole-Orbitrap Mass Spectrometer. Ions entering from electrospray ionization (ESI) source are focused by radio frequency (RF) and a bent flatapole, removing neutrals from the ion beam. A mass selective quadrupole enables selection of ions at a specified m/z range towards storage in Curved Linear Trap (C-trap) or fragmentation cell. Fragments ions are generated in the high-energy collision induced dissociation (HCD) cell and detected in the High Resolution Accurate Mass (HRAM) Orbitrap mass analyzer. Adapted from Scheltema and colleagues, 2014.[27]

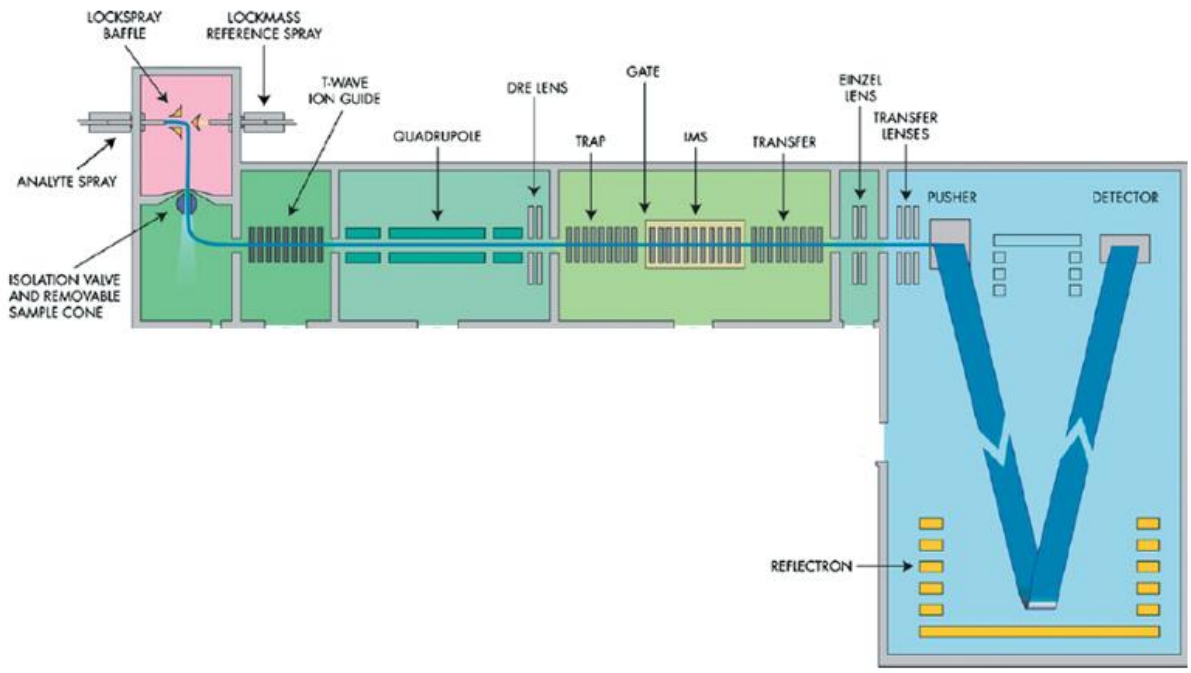


Figure 1.3. Schematic representation of the Synapt G1 High Definition Mass Spectrometer. A traveling-wave (T-wave) ion guide focuses incoming ions with trapping and release of parent ions for mass selection in the quadrupole. Mass selected ions are then trapped prior to the ion gate of the ion mobility (IMS) cell which separates analytes based on their cross-section as they collide with inert gas. Ions exiting the IMS are focused again by a T-wave ion guide as they are transferred into time-of-flight (TOF) analyzer. Fragmentation by collision-induced-dissociation (CID) can be performed before mobility in the trap, or after mobility in the transfer region of the instrument. Ion mobility separation adds an additional dimensionality, enabling detection of ions by mass and size. Adapted from Pringle and colleagues, 2007.[28]

1.2 Mass Spectrometry in Structural Biology

1.2.1 Existing Biochemical Techniques for Studying Protein Structure and Dynamics

The ability to identify proteins and to determine structure-function relationships is central to the life sciences. The amino acid sequences of proteins serve as a link between their coding genes and their physiological roles as components in complex cellular regulatory networks. Before to the emergence of genomics, chemical or enzymatic assays were most commonly used to characterize highly purified proteins. Direct or indirect evidence of protein structure was generally deduced through low-resolution spectroscopic techniques including circular dichroism, intrinsic protein fluorescence, UV-Vis, and infrared (IR) spectroscopy.[29]

With increased performance and versatility of the instrumentation, MS emerged as the technique of choice for determining amino acid sequence of peptides and proteins by the mid-1990s.[30] Despite the close relationship between protein sequence and structure, adoption of MS was significantly slower for higher order structure characterization. Though often hindered by the upfront challenges associated with getting a protein crystal, X-ray crystallography remained the classical approach for protein structure determination.[31] However, several advancements in instrumentation in the 1980s and 1990s made nuclear magnetic resonance (NMR) amenable for structural characterization of proteins in addition to small molecules.[31] The main advantage of NMR over X-ray crystallography is the ability to analyze proteins under more physiological conditions, thereby retaining the structural flexibility observed in-solution. Of note, the hydrogen deuterium exchange (HDX) methodology, described for MS-based applications in the subsequent section, can also be analyzed by NMR. Nonetheless, HDX NMR is restricted by the size limitations of NMR (typically less than 25 kDa), and is generally insensitive to transient, yet biologically-relevant protein conformations occurring on sub-second timescales.[31]

1.2.2 Hydrogen Deuterium Exchange

HDX is a structure-dependant chemical labelling reaction that was first observed by Kaj Ulrik Linderstrøm-Lang using ultracentrifugation[32], and further explored by Englander *et al.* in HDX measurements to single amide resolution by NMR.[33] In 1993, HDX was first coupled to LC-MS, providing the foundation the "bottom-up" HDX-MS workflow that is most commonly used today.[34] Briefly, the protein was reacted in excess deuterium resulting in the exchange of covalently attached, backbone amide hydrogens with deuterium atoms in solution. The reaction was quenched by lowering the pH with acid then digested using pepsin. Cool acidic conditions (pH 2.5, 0°C) were used to minimize back exchange in amide backbone positions. The labelled peptides were further separated by LC and deuterium incorporation calculated from the mass shift detected by MS.[35]

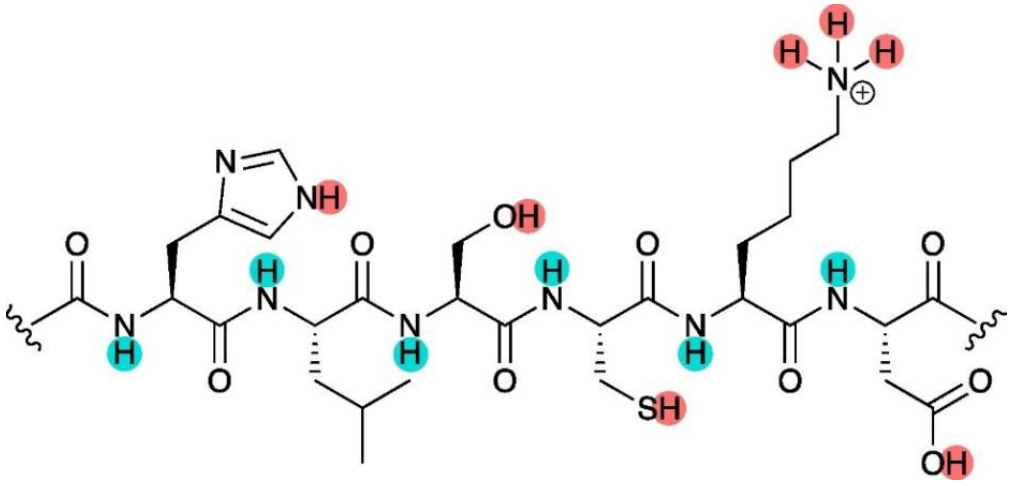


Figure 1.4. Exchangeable amide and side-chain hydrogens in a peptide. Exchangeable backbone amide hydrogens are depicted in teal and the labile side-chain protons are depicted in salmon-pink color. Adapted from Oganesyan and colleagues, 2018.[35]

In a protein, rapid isotopic exchange can occur with any hydrogen bound to sulfur, oxygen or nitrogen. Amide hydrogens, found uniformly across the peptide backbone, are the most commonly monitored in HDX experiments for probing subtle changes in higher order structure (Figure 1.4). At physiological pH, exchange is base-catalyzed where the backbone amide is deprotonated by nucleophilic attack of a deuterium hydroxide forming an amidate ion. The amidate ion is subsequently reprotonated with a deuterium atom from the deuterium oxide (D_2O) solution (Figure 1.5). Cooling the reaction to 0°C and reducing the pH to 2.5 decreases the rate of amide hydrogen exchange by approximately 4-5 orders of magnitude compared to pH 7 and 23°C. Consequently, these "quenching conditions" are typically used to minimize loss of the deuterium label (i.e. back-exchange) in HDX-LC-MS experiments.[34,36]

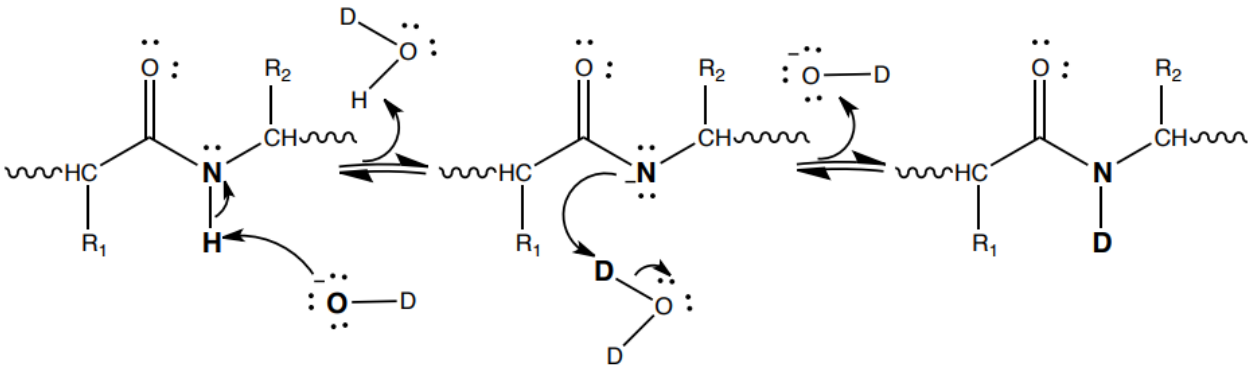


Figure 1.5. **Mechanism of base-catalyzed backbone amide hydrogen-deuterium exchange (HDX).** At neutral pH, HDX is primary driven by base catalysis where nucleophilic attach by deuterium hydroxide deprotonates a covalently attached hydrogen at the peptide amide position. Subsequent re-protonation of the amidate ion by D_2O leaves an isotopically heavy deuterium atom that can be differentiated by MS.

The rate of exchange is dependent on four main factors: intramolecular hydrogen bonding, solvent accessibility, pH and temperature, with the later two controlled experimentally. The method can be applied 'globally' to give information on the overall flexibility of proteins or 'locally' at the peptide level. Spatial localization arises from differential deuterium uptake due to varying degrees of structural protection. Replacement of a hydrogen atom (H) with a heavier deuterium isotope (D) results in a mass increase of one unit per exchange event that can be detected as a shift of the mass spectrum. Since longer peptides have more exchangeable backbone amide hydrogens, results are typically expressed 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).[32,37]

In both global and local experiments, the mass spectrometer monitors deuterium uptake as a function of time, either on the intact protein (global) or individual peptide (local) level. In stably folded protein regions, the rate limiting step for HDX is the solvent abstraction of the proton. Under physiological conditions where F and U are the folded and unfolded states, respectively, folded proteins exhibit EX2 kinetics whereby they rate of refolding is much greater than the rate of unfolding (k-1 >> k2). Consequently, the observed rate of deuterium uptake is dominated by the effects of hydrogen bonding and solvent accessibility. Tightly folded regions with α-helices and β-sheets are generally less dynamic, and therefore exhibit reduced uptake compared to loops or other disordered protein regions.[38,39]

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

1.2.3 Time-Resolved Hydrogen Deuterium Exchange

Conventional HDX as described in 1.2.2 is a powerful tool for studying a variety of systems including monoclonal antibodies, antibody-drug conjugates, protein-protein interactions, and protein-ligand interactions.[40] However, the timescale of conventional HDX is generally limited to steady-state equilibrium conditions with labelling times from 10 seconds to several hours. In cases were the protein structure is very dynamic, such as intrinsically disordered proteins and transient interactions, functionally-relevant intermediate structures are often missed. To circumvent this limitation for studying pre-steady state conditions, millisecond HDX-MS techniques were developed using stopped-flow rapid mixing devices coupled to ESI-MS.[41,42] The early version of the capillary-based mixing device was a fixed length mixer which required physically changing the reaction capillary to obtain different timepoints. For improved consistency, flow rates could be also be adjusted to alter the reaction time; however, this caused considerable pressure fluctuations and made it challenging to obtain stable, time-resolved results (Figure 1.6).[43]

In 2003, the 'fixed mixer' apparatus evolved to include two concentric capillaries forming an adjustable reaction chamber that allowed for continuous-flow reactions on a millisecond timescale.[41] Reaction times were adjusted by withdrawing the inner capillary to increase the dead-volume between the mixing point and the onset of ESI for introduction to the MS (Figure 1.7). This approach, known as Time-Resolved Electrospray Ionization Mass Spectrometry (TRESI-MS) was later incorporated onto a microfluidic chip in the Wilson lab at York University[44] with on-chip quenching and proteolytic digestion for local HDX-MS analysis.[45]

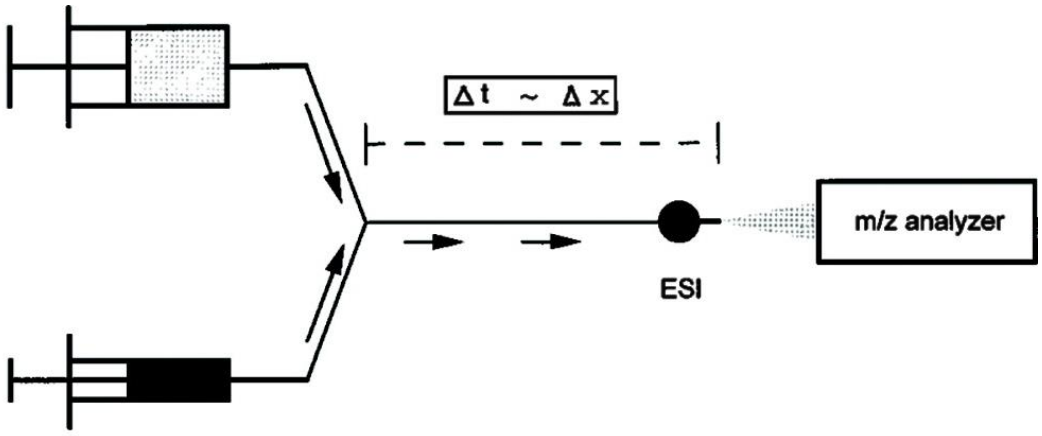
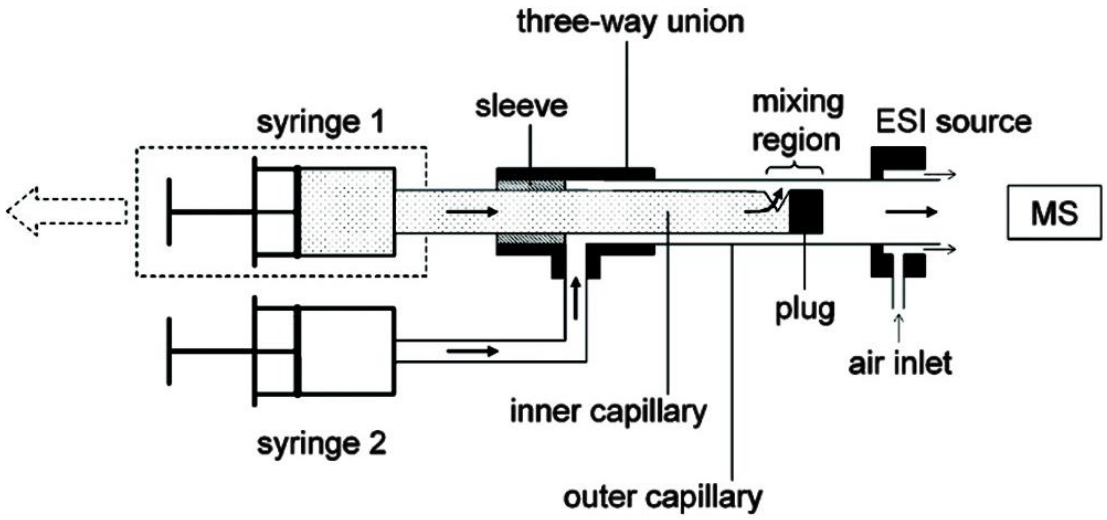


Figure 1.6. Schematic of the first time-resolved electrospray ionization MS device using a capillary-based mixer. Syringes carrying the reactants of interest meet and react within a static, or fixed, mixing tee. Reaction is controlled by the length of the reaction capillary between the mixing tee and electrospray source where reaction time (t) is proportional to the length of capillary (x). Adapted from Konermann et al. 1997.[46]

Figure 1.7. Continuous-flow time-resolved electrospray ionization device featuring two



concentric capillaries and adjustable reaction volumes. Solution flowing through the inner capillary exits through a notch produced 2 mm from the inner capillary tip and mixes with the solution flowing through the outer capillary. Multiple time points achieved by manual or automated pullback of the inner capillary relative to the outer capillary. Adapted from Wilson and Konermann, 2003.[41]

1.2.4 Ion Mobility Spectroscopy

Ion Mobility Spectrometry (IMS) is an analytical technique that separates gas phase ions based on their size and shape. The power of the approach is rooted in its speed and compatibility with LC-MS for isolating isobaric species with different ion conformations. While IMS-MS does not provide localized structural information, it has the advantage of being sensitive to co-existing populations that would otherwise overlap on a mass-to-charge (m/z) spectrum.

IMS is based on the time it takes for ions to travel through a gas-field chamber, called a drift tube, upon which a weak electric field has been applied. Collisions with neutral buffer gas delay ion movement, giving rise to different mobilities. Ions with a larger surface area will experience more collisions and thus take longer to transverse the drift relative to smaller, more compact ions of the same mass and charge.[47] IMS-MS results contain an additional degree of dimensionality with arrival times corresponding to each species detected on the mass spectrum (Figure 1.8). This arrival time distribution reflects the conformational range of the analyzed species, and can be converted into collision cross-sections for ions in the gas-phase.[48] Sharp peaks typically reflect a single protein conformation whereas broader peaks can represent multiple conformational states.[49]

The Waters Synapt G1 High Definition MS was the first commercial MS instrument to incorporate hardware for IMS-MS.[47] The travelling-wave IMS cell contains a series of stacked ring ions guides that builds on the principles of drift tube mobility separation described above. Briefly, application of alternating phases of RF voltage and superimposed direct current (DC) voltage creates a "wave" which differentiates ions of different mobilities. Ions with higher mobility are able to "surf" the wave more effectively thereby reducing their transit time through the IMS cell.[28,49]

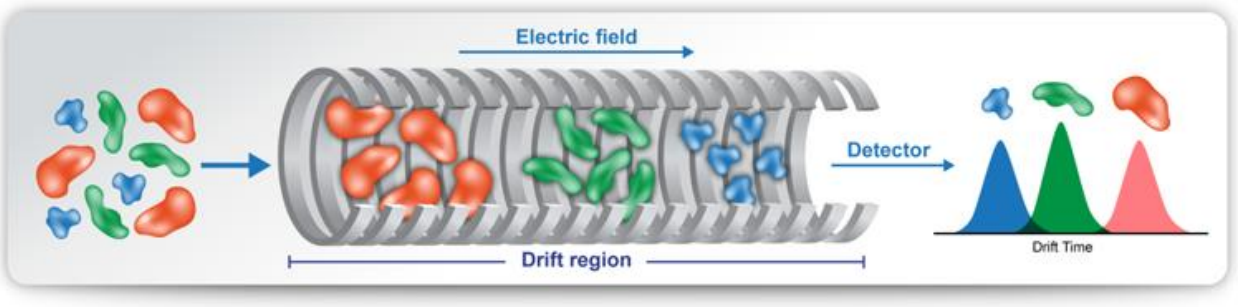


Figure 1.8. Schematic representation of IMS-MS. Ions enter a drift tube with an applied electric field and migrate at different drift velocities to the detector. Measurements of arrival time distributions reflect the conformational range of the examined species. Adapted from Gili Ben-Nissan and Michal Sharon, 2018.[50]

1.2.5 Applications at Protein-Membrane Interfaces

Many cellular signalling events are controlled by the selective recruitment of protein complexes to membranes surfaces. For membrane proteins found at this critical interface, biological activity often depends on transient changes in protein conformation due to binding interactions with lipids in membrane bilayers.[51,52] Structural studies of membrane proteins by conventional techniques such as NMR and X-ray crystallography have advanced thanks to the development of small, versatile model membrane systems including micelles, bicelles and liposomes.[53–58] Though detailed, the pictures captured by these methods typically reflect the populated, ground-state structure, and not transient protein conformations such as those induced by peripheral interactions with phospholipids in membrane bilayers.[59,60] Structure-function studies of peripheral membrane proteins also contend with the natural heterogeneity of biological membranes, and the influence of different lipid head groups on the polarity of membrane surfaces.[61–63]

1.3 – Proteomics

The availability of full genome sequences for thousands of organisms has greatly contributed to the growth of MS-based proteomics. Rather than focusing on particular protein family, proteomics aims to provide a top-down view of all proteins in an organism, including post-translational modifications (PTMs) and protein-protein interactions. This level of global protein information about the proteome can be combined with other "big data" techniques such as genomics[64], lipidomics[65], transcriptomics[66], and metabolomics[67] for a systems biology approach to understanding an organisms' life cycle.[52]

1.3.1 "Top-down" versus "Bottom-up" Approaches to Protein Analysis

MS has emerged as a core tool for large-scale protein analysis, owing to rapid advancements in LC separation, data acquisition, and bioinformatics analysis. MS-based characterization of proteins is divided into two types of analysis: "top-down" referring to intact proteins, or "bottom-up" approaches in which peptides are released via enzymatic digestion (proteolysis) prior to LC-MS analysis. Bottom-up analysis of a protein mixture is commonly referred to as "Shotgun proteomics". Coined by the Yates lab[68,69] due to analogy to shotgun genomic sequencing, shotgun proteomics provides an indirect measure of a protein, making use of tandem mass spectrometry (MS/MS) sequencing of peptides derived from proteolytic digestion.[70] Bottom-up workflows are typically more sensitive than top-down approaches, offering more efficient ionization and less division of signal among different charge states. Whereas shotgun proteomics workflows employ enzymatic digestions that can be applied universally across sample types, top-down methods are typically restricted to relatively pure samples, due to limits in fractionation and fragmentation capabilities.[71]

1.3.2 Shotgun Proteomics

1.3.2.1 Nano-liquid chromatography (nanoLC)

While conventional LC-MS techniques typically make use of 2.1 mm internal diameter (i.d) columns and flow rates of 100-1000 µL/min, nanoLC benefits from downsizing to sub-100 µm columns and nanoflow rates (< 500 nL/min) with a theoretical gain in absolutely sensitivity of approximately 1000-fold.[72,73] This boost in baseline sensitivity is rooted in improved ionization efficiency and fewer co-eluting peptide precursors. The benefits of nanoLC come at the expense of less robust system performance and longer acquisition times. For shotgun proteomics where the goal is to maximize proteome coverage, these trade-offs are outweighed by the ability to identify more proteins in complex samples. Therefore, nanoLC remains a fixture proteomics where upmost sensitivity is required.[23]

1.3.2.2 Proteolytic Digestion

In a typical bottom-up proteomics experiment, protein mixtures are enzymatically digested with a protease producing a peptide mixture. This digestion step is most commonly performed using trypsin due to its efficient and specific cleavage at arginine and lysine residues. The resulting tryptic peptides are generally short for good sequencing via collision induced fragmentation. Having a basic arginine (Arg) or lysine (Lys) residues at the C-terminus also makes tryptic peptides ideal for chromatographic separation by reverse phase high pressure liquid chromatography (RP-HPLC). Nevertheless, it has become apparent that trypsin, like all other proteases, can skip a cleavable residue resulting in so-called missed cleavages.[74] Using only one protease also limits the number of identifiable post-translational modifications (PTMs) with different extents of PTM coverage based on the type of modification, amino acid residue, and position in the protein sequence.[75]

In response to these limitations, there has been emergence of complementary proteases for LC-MS preparation (Table 1.1).[76] For example, working in a Trypsin/Lysozyme C (Lys-C) mix has been shown to overcome the proteolytic resistance of tightly folded proteins, reduce the number of missed cleavages, and prevent tryptic autolytic activity.[77] To maximize coverage, protocols up to six proteases have been implemented including the following alternatives to trypsin: chymotrypsin, LysN, AspN, GluC, ArgC and pepsin (albeit with non-specific cleavage properties).[76]

Table 1.1. Complementary proteases for sample digestion in bottom-up proteomic analyses.

Protease	Specificity	Optimal pH	Approximate MW
Trypsin	Carboxyl side of Arg and Lys	pH 8.0	23.3 kDa
Asp-N	Amine side of Asp and Cys	pH 6.0 - 8.5	24.5 kDa
Glu-C	Carboxyl side of Glu and Asp	pH 4.0 - 7.8	29.0 kDa
Lys-C	Carboxyl side of Lys	pH 8.5	28.0 kDa
Arg-C	Carboxyl side of Arg	pH 7.5 - 8.5	26.5 kDa

1.3.2.3 Data-dependent acquisition (DDA)

After enzymatic digestion, peptides are separated by RP-HPLC followed by shotgun MS analysis (Figure 1.9). This is mostly commonly performed via a data dependant acquisition (DDA) method in which the "topN" most abundant precursors are selected for collision induced fragmentation. The generated tandem mass spectra (MS/MS) spectra contain fragment ions corresponding to the sequence of the tryptic peptide (Figure 1.10). Peptide identification is achieved using bioinformatics search tools such as Mascot[62], Sequest[78] and MS Amanda[79] which compare experimental MS/MS spectra against theoretical peptide MS/MS spectra generated from in silico digestion of a protein database .The identity of 100s-1000s of proteins can be inferred by matching identified peptides to protein sequences (based on sequence homology).

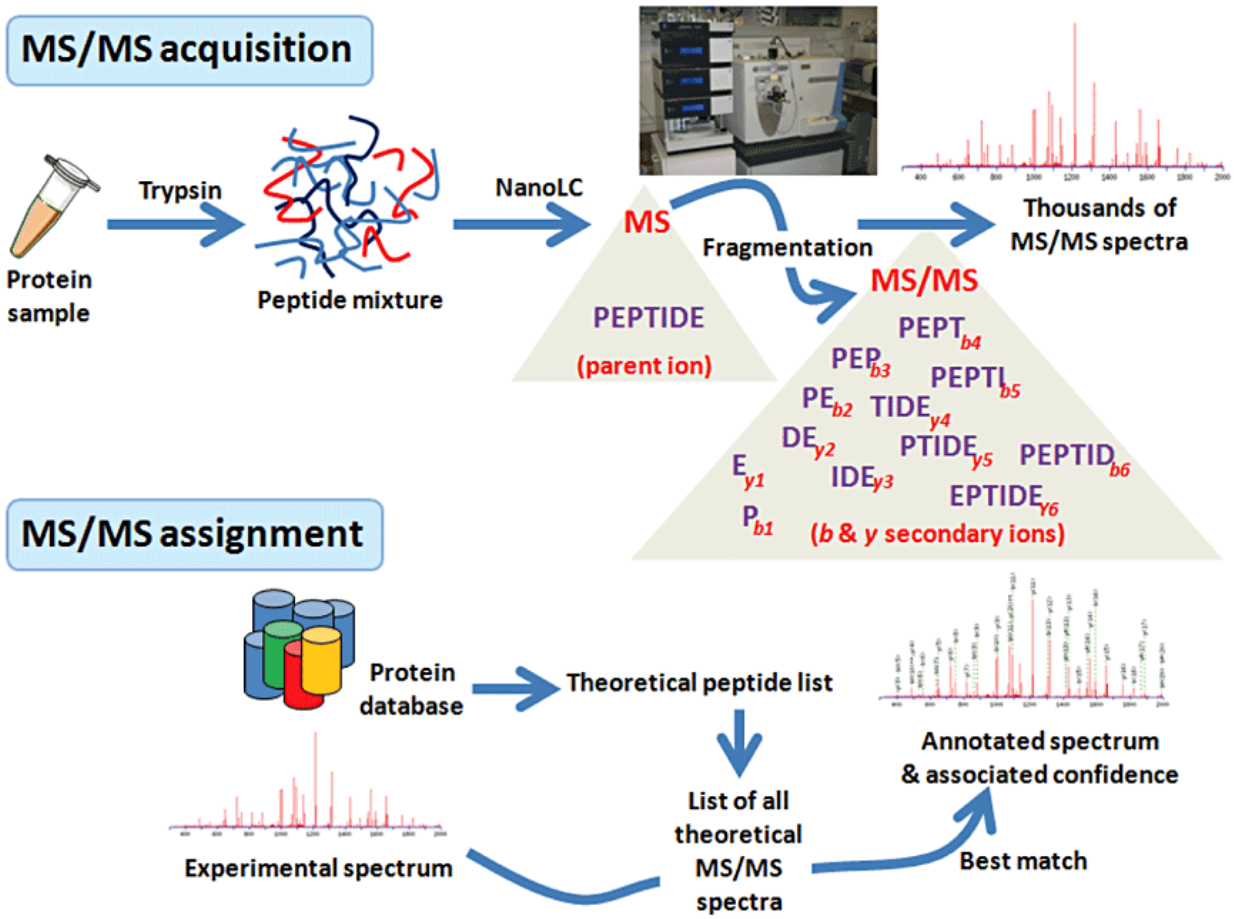


Figure 1.9. Shotgun proteomics workflow for protein identification by nano-liquid chromatography tandem mass spectrometry (nanoLC-MS/MS). Protein sample is digested with trypsin and is separated on a reverse phase chromatographic column coupled to a tandem MS instrument (top panel). The MS measures the 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 trypsin, 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. Adapted from Armengaud, 2012.[32]

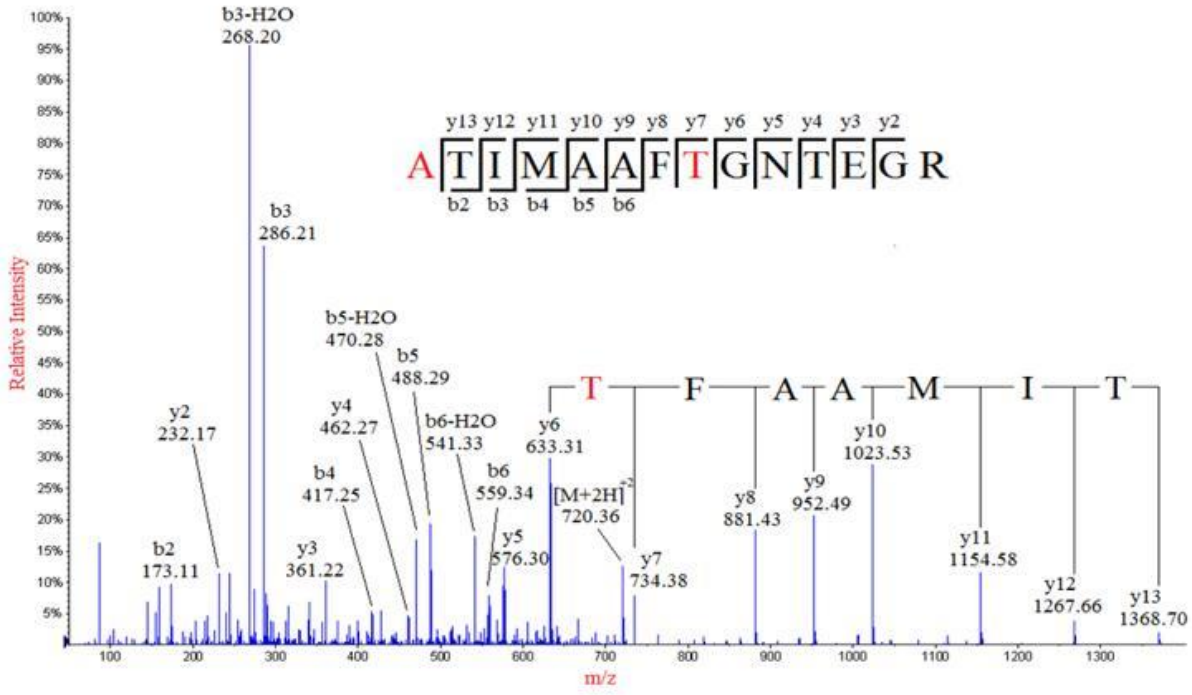


Figure 1.10. Example of a CID/HCD fragment ion (MS/MS) spectra for a tryptic peptide. **‘**y-ions’ correspond to fragments in the C to N-terminal direction whereas ‘b-ions’ sequence residues in from the N to C-terminus of the peptide. Adapted from Ma and Johnson, 2012.[81]

1.3.2.4 Limitations in Shotgun Proteomics Analyses

Though providing a good global snapshot of proteins in a sample, shotgun proteomics methods are inherently limited in the breadth of coverage, reproducibility, and dynamic range of detection. Most biological samples contain a mixture of aqueous and membrane-associated proteins, spanning over 5-10 orders of magnitude in abundance.[82] Despite ongoing improvements, MS instruments are only capable of up to 5-6 orders of magnitude of dynamic range.[7] Furthermore, the most commonly used solution-based digestion workflows bias proteome coverage towards proteins which are water soluble, and hence enzymatically digestible. Therefore, many present-day proteome maps are missing protein sequences despite ongoing development efforts.[83,84]

Another limitation in shotgun proteomics is the data-dependant nature of precursor selection. Conventional DDA workflows favour detection of the most abundant proteins as they generate the most intense precursor peptide signals. For Orbitrap instruments which are the most commonly used platform in proteomics, dynamic range in the survey scan is further limited by the charge holding capacity of the curved linear ion trap (C-trap) prior to detection.[26] Consequently, peptides which co-elute with more intense features may not be selected for fragmentation, and thus will not be detected by a database search. While the dynamic range and reproducibility of fragmentation can be enhanced by using data-independent acquisition (DIA) strategies on quadrupole-Orbitrap hybrid instruments[85,86], fundamentally a lack of MS detection using shotgun methods cannot be definitely linked to absence below a consistent LOD due to semi-stochastic precursor selection. This ultimately leaves shotgun MS with poor run-to-run reproducibility and limited capabilities beyond relative quantitative analyses, despite efforts to develop universal approaches for precursor-based quantitation.[87]

1.3.3 Targeted Proteomics

Limitations in DDA for absolute quantitation have propelled the development of targeted proteomic methods—namely, selected reaction monitoring (SRM), also known as multiple reaction monitoring (MRM) on triple quad MS instruments (QqQ), and parallel reaction monitoring (PRM) on quadrupole-Orbitrap MS platforms. As in bottom-up DDA workflows, proteins are enzymatically digested to peptides and analyzed by RP-HPLC. A key difference is the use of an inclusion list such that the sensitivity of the instrument is specifically harnessed for analytes of interest rather than the most abundant precursors. Targeted MS acquisitions methods provide the reproducibility that shotgun workflows lack, and facilitate absolute quantitation when coupled with appropriate internal standards.[88]

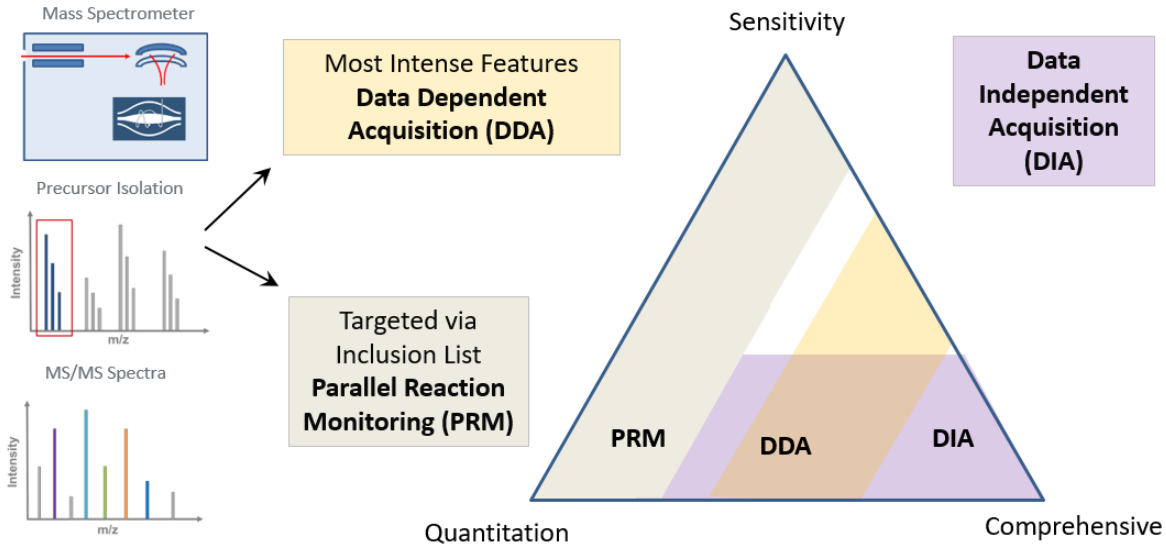


Figure 1.11. Comparison of shotgun and targeted acquisition methods for proteomic analysis. Data-Dependent Acquisition (DDA) and Parallel Reaction Monitoring (PRM) workflows differ in how precursors are isolated for fragmentation. Using the most intense features as in a DDA workflow, results in a more comprehensive analysis whereas targeting precursors by PRM enhances sensitivity and enables absolute quantitation.

In a targeted MS workflows, precursor ions are isolated using a quadrupole (Q1) and subsequently fragmented by collision induced dissociation (CID) or high-energy collision induced dissociation (HCD) creating sequence-specific b and y-ion fragment ions (Figure 1.10).[89] In a conventional targeted analysis performed on a QqQ MS instrument, the first and third quadrupole are used to selectively isolate pre-defined precursor-fragment ion pairs called transitions. Following fragmentation in the collision cell (2nd quadrupole), 3-5 abundant fragment ions per precursor are sequentially isolated by the 3rd quadrupole. The presence of the fragment ions transitions is monitored over chromatographic retention time via extracted ion chromatogram (XIC) traces which are subsequently used for quantition.[90–92]

The speed and specificity from analyte-specific fragmentation patterns makes SRM/MRM workflows the gold standard for MS-based quantitation.[93] Nevertheless, these targeted MS method on QqQ instruments require significant upfront development and a priori knowledge to define the precursor-fragment ion pairs for analysis. Additionally, SRM/MRM acquisitions are specific to fragment ions from targeted analytes. Consequently, acquisitions are effectively blind all other components in the sample, making it difficult to identify and account for potential matrix effects.

1.3.3.1 Parallel Reaction Monitoring

More recently, rising popularity of quadrupole-Orbitrap instrument platforms has led to emergence of parallel reaction monitoring (PRM) as an alternative quantitative MS method.[94] In PRM, the third quadrupole of the QqQ is substituted with a high resolution accurate mass (HRAM) Orbitrap mass analyzer, enabling parallel detection of all fragment ions via full MS/MS scans. PRM methods can also be combined with MS survey scans to enable evaluation of sample complexity and matrix effects using the total ion chromatogram (TIC) for a given sample. Analysis of precursor-fragment ion transitions is performed post-acquisition such that transitions can easily be refined in the event of background interference (Figure 1.12).

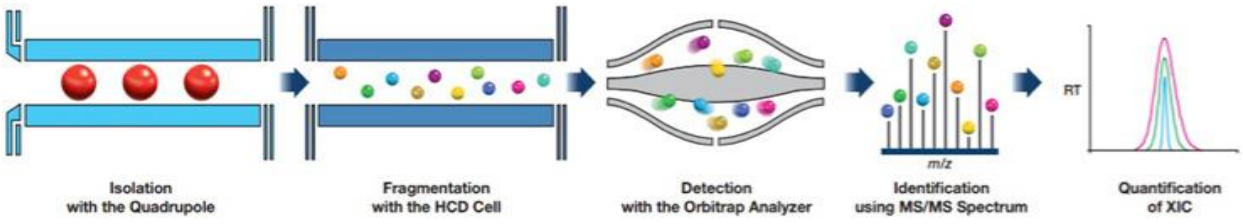


Figure 1.12. Parallel Reaction Monitoring on an Orbitrap MS. Precursors isolated in the quadrupole are subject to fragmentation by HCD and are detected in an MS/MS spectrum by the Orbitrap. Post-acquisition analysis of extracted ion chromatograms (XIC) enables quantitation relative to a spiked internal heavy peptide standard.

1.3.3.2 Stabile Isotope Labelled Peptide Standards for Absolute Quantitation

MS-based quantitation makes use of stable isotope dilution (SIL) for absolution quantitation. The principle is based on absolutely quantified heavy-labelled peptide standards, called AQUA peptides, which are spiked into the sample as internal standards.[95] AQUA peptides are identical to their native peptide counterparts down to the base atoms. As a result, they have the same ionization efficiency and co-elute chromatographically. Presence of a ^{13}C, ^{15}N heavy labelled Arg or Lys at the C-terminus creates a mass shift to the endogenous peptides. Native peptide concentration is calculated from the ratio of unlabelled (native)-to-heavy peptide, and the known amount of spiked heavy peptide standard.[96]

Selection of the surrogate peptide sequence for use with a matching AQUA peptide is arguably one of the most challenging, yet critical steps in developing of sensitive targeted MS assay. Though there are several strategies for guiding surrogate peptide selection, most employ a priori knowledge of protein sequence combined with previously acquired data (usually from a shotgun proteomic analysis). Using this so-called data-driven strategy biases selection towards the most abundant peptides at the survey scan or MS1 level. However, MS1 intensity does not always translate into specificity and sensitivity at the fragment ion (MS2) level used in MRM/PRM quantitation. Furthermore, shotgun analysis of an individual sample provides no information on digestion kinetics, peptide stability or potential matrix effects; all critical attributes in a quantitative MS assay. As a result, there remains room for alternative approaches for surrogate peptide selection.

1.4 Research Objectives

The main objective of this work is the implementation of mass spectrometry (MS)-based workflows in emerging areas where MS is well-positioned to fulfill unmet analytical needs. Chapter 2 will focus on the application of structural MS techniques including travelling-wave ion mobility (IMS-MS) and time-resolved electrospray ionization hydrogen-deuterium exchange (TRESI-HDX-MS) for characterization of conformational dynamics in protein-lipid interactions. The use of these sensitive structural MS techniques will facilitate mechanistic understanding of electrostatic interactions involved in the recruitment of a peripheral membrane protein to a phospholipid bilayer. Shifting away from biochemical interactions, Chapter 3 will focus on strategies for contending with complex, lipid-containing samples for more complete coverage of membrane-associated proteins in shotgun proteomics workflows. Lastly, Chapter 4 will highlight a synthetic peptide screening approach in the development of a targeted nanoLC-MS/MS method for absolute quantitation of residual toxins in vaccines samples from *Bordetella pertussis (B.pertussis)*. Collectively, the highlighted workflows will contribute to enhanced characterization of biological systems with the end goal of accelerated biopharmaceutical development.

Chapter 2: Deciphering electrostatically-driven lipid binding modes in the cytochrome c-membrane interaction using structural mass spectrometry

2.1 Introduction

The interaction between 12 kDa Cytochrome c (Cyt *c*) and the inner-mitochondrial membrane represents a paradigm for the challenges associated with mechanistic studies of peripheral membrane proteins. Primarily known for its redox activity in the electron transport chain,[97] release of Cyt *c* from the mitochondria to the cytosol is a critical first step in the initiation of mitochondrial-induced apoptosis.[98–102] Cyt *c*'s pro-apoptotic function is rooted in its interaction with cardiolipin (CL), a negatively charged (anionic) phospholipid making up approximately 25% of lipids found in the mammalian mitochondrial membrane.[103,104] When bound to CL, Cyt *c* makes use of H_2O_2 in the mitochondria to catalyse the oxidation of CL, which in turn decreases the strength of the CL-mediated Cyt *c*-membrane interaction.[105] This peroxidase-induced detachment from the mitochondrial membrane enables Cyt *c* to release into the cytoplasm where it participates in apoptosome complex formation and downstream activation of caspases required for programmed cell death.[101,106,107]

Since it was shown that CL-bound Cyt *c* acts a peroxidase in the early stages of apoptosis,[108–110] it has been thought that interactions with membrane phospholipids induce conformational changes in Cyt *c* enabling it's functional shift from electron carrier to pro-apoptotic catalyst.[110–112] Higher order structural characterization of Cyt *c*-membrane interactions (CD, FTIR and Raman) have shown loosening of the tertiary structure and disruption of the Met80-heme ligation upon binding to anionic phospholipids.[113–116] In these structure-function studies, CL has generally been mixed with other phospholipids in liposomes in an attempt to mimic the composition of the inner mitochondrial membrane. Of the mitochondrial phospholipid classes, phosphatidylcholine (PC) is the most commonly used phospholipid in combination with CL.[63, 64] Containing a primary amine bound to a

phosphate head group (Figure 2.1), PC lipids bear a positive charge and neutral pH, and thus are not expected to have significant electrostatic affinity for cationic Cyt *c*. Therefore, it has long been presumed that CL/PC liposososomes are suitable mimetics for studying role of CL in the the Cyt *c*-membrane interaction.[117] However, recent studies have shown Cyt *c* binding to zwitterionic phospholipids, including PC, with liposomes of varying acyl chain length and CL composition.[119,120] Consequently, the question of electrostatic involvement in the Cyt *c*-membrane remains a fundamental gap in our mechanistic understanding of Cyt *c* as peripheral membrane protein and its role in mitochondrial-induced apoptosis.

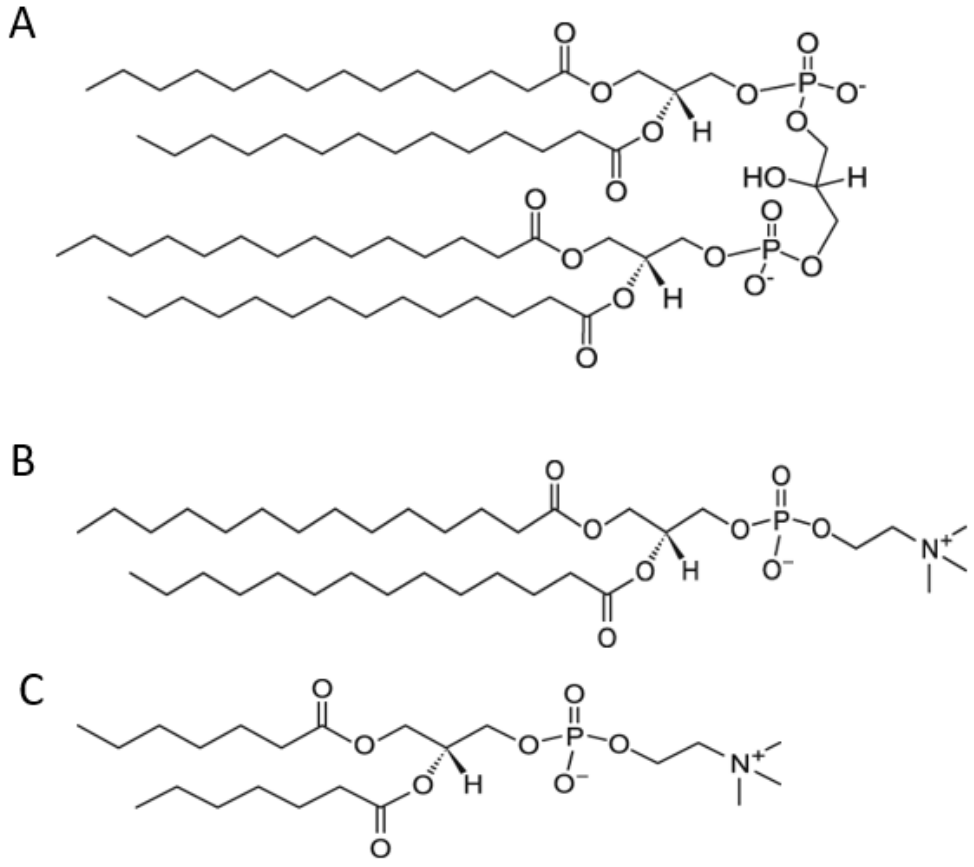


Figure 2.1. Chemical structures of CL and PC phospholipids analyzed with Cyt *c*.
A) 1', 3'-bis [1,2-dimyristoyl-sn-glycero-3-phospho]-glycerol (14:0 Cardiolipin), B) 1,2-Dimyristoyl-sn-Glycero-3-Phosphocholine (DMPC), and C) 1, 2-Dihexanoyl-sn-Glycero-3-Phosphocholine (DHPC).

Early work characterizing the phospholipid-mediated Cyt *c*-membrane interaction established two binding sites, namely the A-site and C-site, with differing affinities for negatively-charged phospholipids.[121] The interaction at the A-site was proposed to be electrostatic in nature, involving positively charged residues (possibly at Lys72/Lys73), whereas the C-site was suggested to participate in hydrophobic interactions with one of the acyl chains of the phospholipid interacting with the highly conserved Asn52 residue.[122–124] More recent experimental evidence has provided grounds for a more peripheral binding mechanism including two independently acquired time-resolved florescence energy transfer (TR-FRET) studies showing conformational interchange between extended and compact CL-bound conformers on the bilayer surface**.**[125,126] Following a two-stage mechanism, the Cyt *c*-membrane interaction has been proposed to involve electrostatic forces for initial establishment, followed by activation of pro-apoptotic peroxidase activity.[124] A summary of previously proposed modes of Cyt *c* – membrane association is shown in Figure 2.2.

Functional studies of Cyt *c* have shown increased peroxidase activity at high CL/Cyt *c* ratios.[110,118,127] Interestingly, examination of absorbance at 695 nm as a function of peroxidase activity showed no change in biological activity despite Met80-heme bond disruption at higher CL/Cyt *c* ratios.[110] The lack of correlation between Met80-heme bond disruption and peroxidase activity suggests that CL-induced protein unfolding is not a prerequisite for peroxidase activity. This observation is further supported by NMR and FTIR measurements showing that Cyt *c* retains its secondary structure upon interaction with phospholipid membrane surfaces.[111,116,128] Collectively, these results demonstrate that despite extensive structural characterization, understanding of the dynamics involved in the establishment of the Cyt *c*-membrane interaction remains limited.

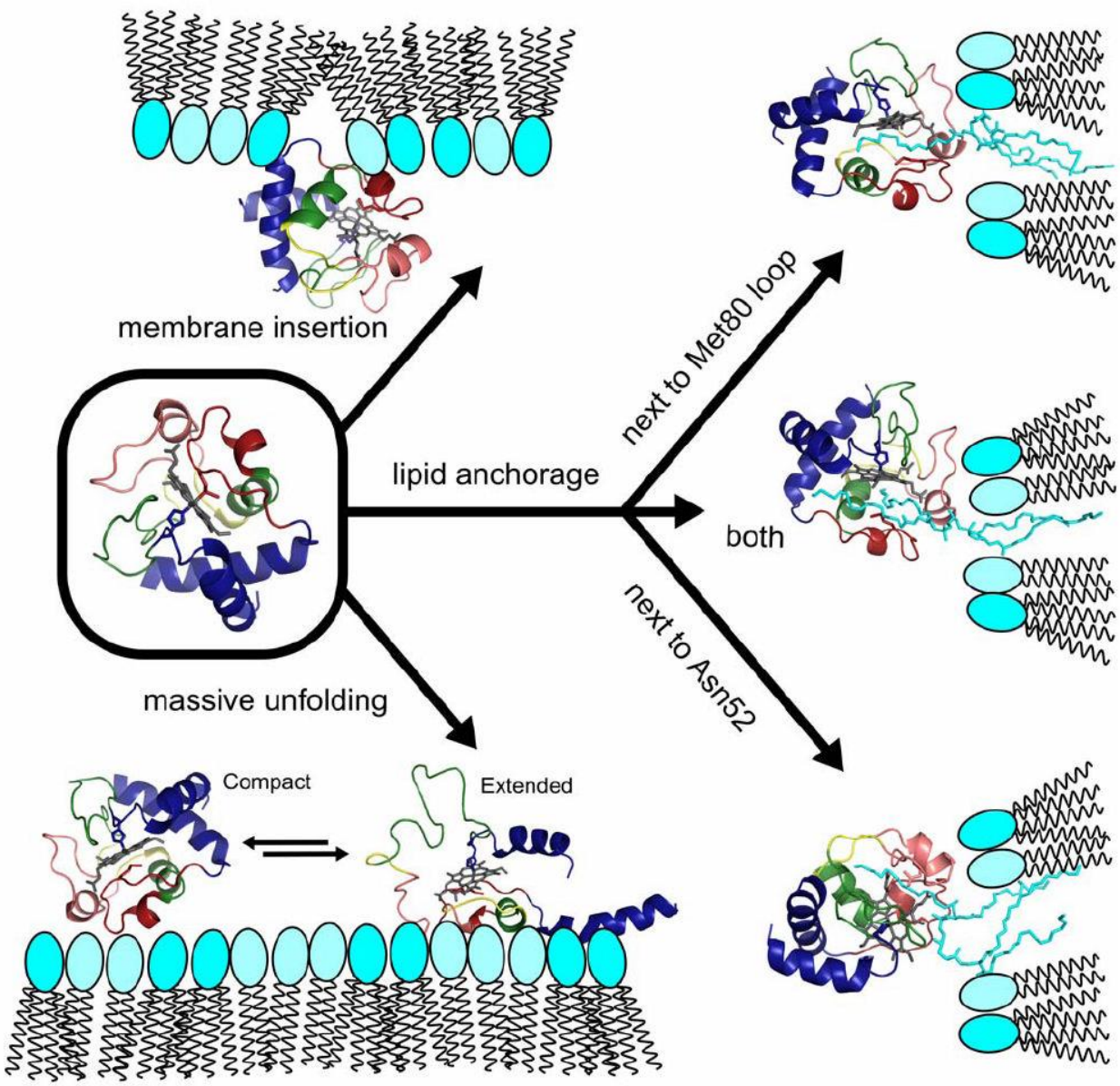


Figure 2.2. Proposed modes of CL-mediated association of Cyt *c* with the membrane bilayer. Previously suggested models of the Cyt *c*-membrane interaction include insertion of the protein into the membrane[129], acyl chain insertion into the protein[121,122,124], and a peripheral insertion on the membrane surface with unfolding of the protein[125,126]. Adapted from Muenzner and Pletneva, 2014.[114]

To examine the basis for CL-specific functionality in the Cyt *c*-membrane interaction, the conformational dynamics of Cyt *c* were studied in the presence of anionic cardiolipin (CL) and cationic phosphatidylcholine (PC) phospholipids. The role of electrostatic interactions in the initial formation of the Cyt *c*-membrane complex was investigated by florescence spectroscopy, ion mobility mass spectrometry (IMS-MS) and hydrogen-deuterium exchange mass spectrometry (HDX-MS) to gain insight into the unique specificity of anionic CL in this pro-apoptotic interaction.

2.2 Experimental Methods

2.2.1 Reagents and Supplies

Cytochrome *c* (Cyt *c*) from equine heart (≥99%), deuterium oxide (D_2O − 99.99%), acetic acid (99.7%), and pepsin from porcine mucosa were purchased from Sigma Aldrich (St. Louis, MO). NHS-activated agarose beads were purchased from Fisher Scientific (Mississauga, ON). 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), 1,2-dihexanoyl-sn-glycero-3-phosphocholine (DHPC) and Cardiolipin (ammonium salt) lipids were purchased from Avanti Polar Lipids Inc. (Alabaster, AL). Solutions were prepared with HPLC water (Fisher Scientific), excluding lipid stocks which were prepared in ethanol (≥ 99%) and frozen at ≤ -20°C.

2.2.2 Fluorescence Spectroscopy

Intrinsic fluorescence of Cyt *c* in the presence of phospholipids was probed using a BioTek Synergy H4 Hybrid Reader. Florescence measurements were obtained under conditions analogous to those used for global and local TRESI-HDX (1:10 protein:lipid ratio, 10% EtOH (v/v) in H_2O, pH 7). Emission spectra were acquired with an excitation wavelength of 280 nm over 300-400 nm (bandwidth = 9.0 nm, read height = 8 nm). For titration experiments, protein-lipid preparations were analyzed in triplicate at protein-to-lipid ratios of 1:1, 1:2, 1:4 and 1:10.

Spectroscopy measurements were acquired at 22°C ± 0.5°C with buffer only, lipid only and protein only controls at each concentration tested.

2.2.3 Protease 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 protease was aspirated, and beads resuspended in blocking buffer (1 M Tris–HCl, pH 4) 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.

2.2.4 Microfluidic Device Fabrication

The microfluidic device used for time-resolved ESI (TRESI) measurements was built as previously described.[130] 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 cut 2 mm from the sealed end. Hydrogen deuterium exchange (HDX) labelling reactions were carried out in the volume between the inner and outer capillaries. For global measurements, the outer capillary was connected directly to the heated electrospray source; whereas, for localized measurements, the TRESI mixer was connected to an acid channel through a valco t-mixer followed by on-chip proteolytic digestion (Figure 2.3).

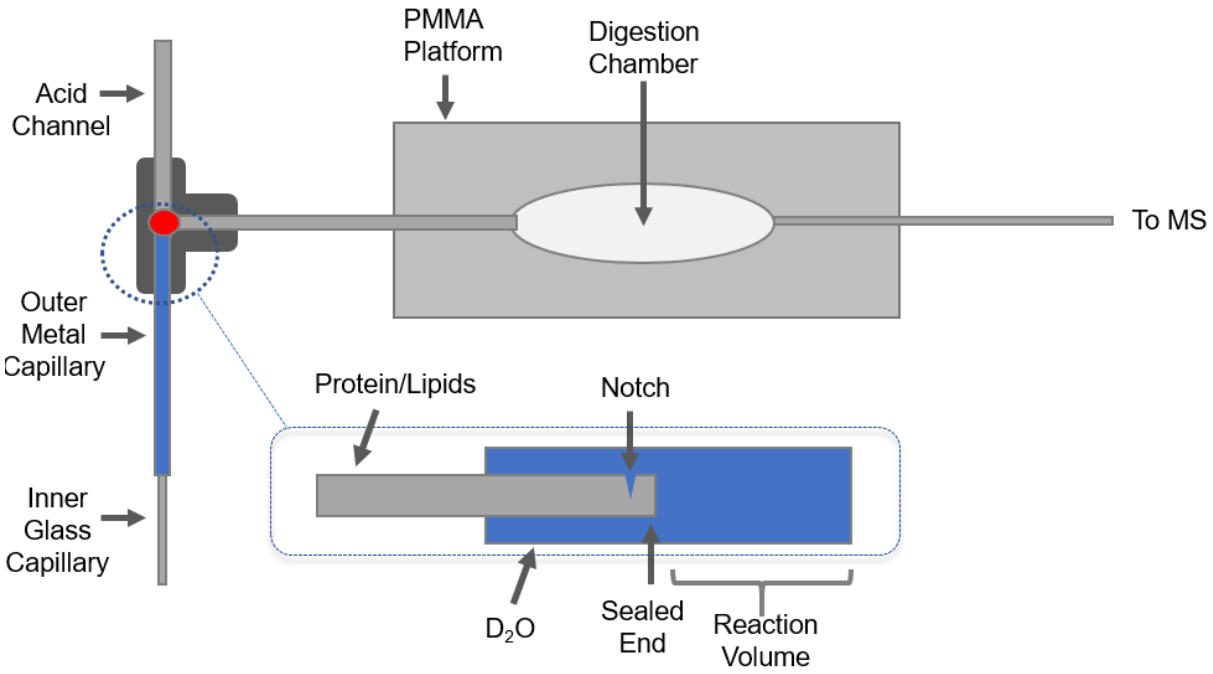


Figure 2.3. Schematic representation of the capillary-based mixer used for TRESI-HDX-MS. Protein-lipid solution in an inner glass capillary flow through a notch cut 2 mm from the sealed end and mixes with deuterium flowing in the outer metal capillary. The deuterium labelling reaction is quenched by lowering the pH with acid followed by online digestion by immobilized pepsin beads on a poly (methyl methacrylate) (PMMA) platform. Peptides are introduced into the mass spectrometer by electrospray ionization. For global measurements, labelled protein was infused directly into the mass spectrometer without acidification or protease digestion. Deuterium incorporation is measured by a delta (Δ) mass shift on the mass spectrum, enabling characterization of structural changes from transient protein-lipid interactions.

2.2.5 Global HDX

50 μM Cyt *c* only or Cyt *c*-phospholipid solution in a 1:10 protein:lipid ratio were infused into an inner glass capillary at 2 μL/min, and mixed with D_2O in the outer metal capillary at a flow rate of 8 μL/min. The reaction volume in the mixing chamber was increased by adjusting the position of the inner capillary relative to the outer capillary, thereby increasing the labelling time before direct introduction to a Synapt G1 Q-TOF MS. The instrument was operated at a source voltage of 3500 V

in positive ion mode, and samples scanned over a range of 400–4000 m/z. Time-dependent deuterium incorporation was monitored at reaction times from 0.28 to 2 seconds by withdrawing the inner capillary from the end of the outer capillary in 1 mm increments (total flow rate of 10 μL/min). The amount of deuterium incorporated into the protein (backbone amide and sidechains) was calculated from the difference between the theoretical and observed (deuterium labelled) molecular masses for Cyt *c*, measured as a delta (Δ) *m/z*, for the protein only and in the presence of DMPC or CL phospholipids. A background polysiloxane at 445.12003 m/z was used as an internal lock mass in the mass shift calculations (Δ *m/z* x *z* = Δ Da). HDX uptake results were averaged for the 6+, 7+ and 8+ charge states across triplicate acquisitions at each timepoint.

2.2.6. Ion Mobility Spectrometry

Traveling-wave ion mobility (IMS-MS) experiments were performed on a Waters Synapt G1 Q-TOF instrument using electrospray ionization (ESI). Nitrogen gas was used for both the source and ion mobility T-wave cell, whereas argon gas was used in the trap region of the mass spectrometer. The instrument was operated in positive ion mode with a capillary voltage of 3000 V, sampling cone of 30 V, source temperature of 80 °C and cone gas flow of 25 L h^{-1}. IMS-MS measurements were carried out in triplicate on Cyt *c* only and in a 1:10 ratio with CL or DMPC phospholipids. Protein-lipid solutions were infused at rate of 5 μL/min and mass spectra were acquired for 5 minutes over a mass range of 400–4000 m/z. Optimal IMS separation was obtained using 350 ms^{-1} IMS wave velocity and 11 V IMS wave height, with trap and transfer energies set to 10 V. Transfer T-wave velocity of was held at 180 ms−1 with a wave height of 8 V. IMS-MS results were analysed using MassLynx version 4.1 (Waters).

2.2.7. Local TRESI-HDX Measurements

Local hydrogen-deuterium exchange measurements (HDX) were performed using a microfluidic TRESI device as depicted in Figure 2.2. Sample solutions were prepared with 50 μM Cyt *c* only or Cyt *c* pre-incubated with CL, DMPC or DHPC in a 1:10 protein-to-lipid mixture. Cyt *c*-lipid solutions were infused into the inner capillary at 2 μL/min and mixed with 2 μL/min D_2O in the outer metal capillary. The HDX reaction was quenched by the introduction of 10% acetic acid at 12 μL/min of (pH 2.5) followed by online digestion with pepsin. Labelled peptides were sprayed directly into a Waters Synapt G1 Q-TOF instrument operated in positive ion mode with a source voltage of 3000 V. and mass range of 400-1500 m/z.

2.2.8. Data Analysis

Peptides were identified using the FindPept tool on the ExPASy Proteomics server.[131] and sequences confirmed via MS/MS with analysis in mMass.[132] Deuterium uptake per peptide was calculated using Mass Spec Studio[133] by matching experimental deuterium uptake to the theoretical isotopic distribution for the native peptide sequences. Changes in deuterium uptake for 'protein & lipid' versus 'protein only' states were summed across 3 timepoints (labelling times of 0.5, 1 and 2 seconds) and averaged across technical triplicates from two different preparations (n=6 replicates in total). Changes in deuterium uptake were considered significant if they exceeded two times the standard deviation (2σ), with propagation of error from the 'protein & lipid' versus 'protein only' % D_2O uptake measurements. Protein structures were rendered using PyMol v1.5.0.4 (Schrödinger, LL) and mapped onto the protein crystal structure for equine Cyt *c* (pdb: 1hrc).

2.3 Results & Discussion

The binding of cationic phosphatidylcholine (PC) and anionic cardiolipin (CL) to Cyt *c* was probed by monitoring instrinsic fluorescence in the presence of phospholipids. Due to the close proximity of Cyt *c*'s heme moeity to the lone tryptophan residue (Trp59), changes in the emission spectrum could be attributed solely to the alterations in the enviroment of tyrosine residues upon phospholipid binding. Following excitation at 280 nm, the emission spectra for Cyt *c* showed an apex peak at 303 nm, consistent with the the instrinsic fluorescence characteristics of tyrosine.[134] In the presence of both PC and CL phospholipids, the intensity of instrinsic fluorescence decreased upon establishment of protein-lipid interactions, resulting in quenching of fluorescence signal (Figure 2.4).

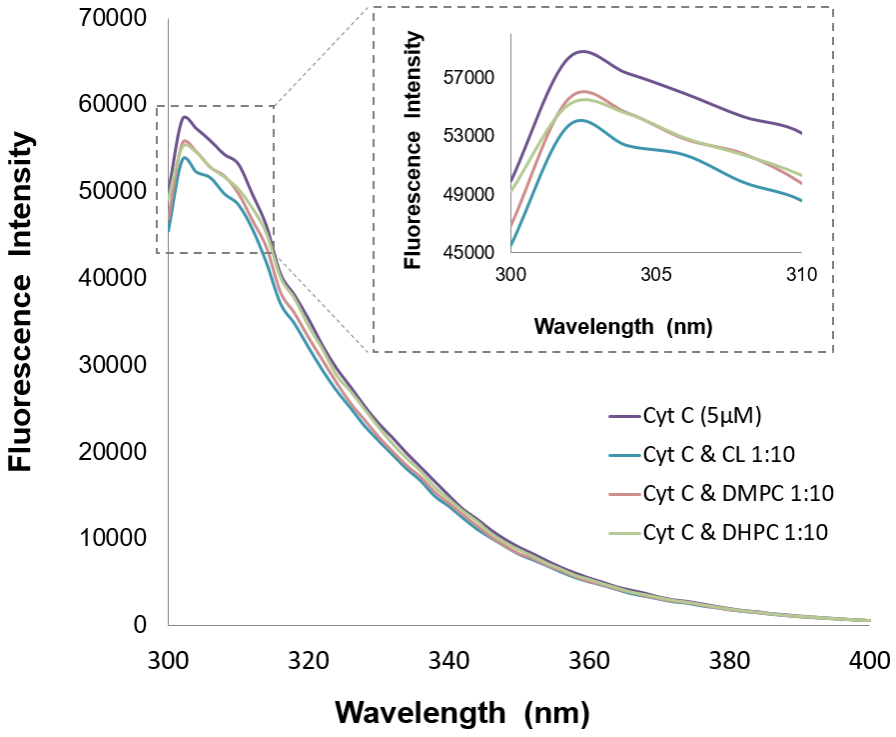


Figure 2.4. Fluorescence spectra for Cyt *c* with CL, DMPC and DHPC phospholipids. Apex signal from tyrosine residues (303 nm) in Cyt *c* quenched upon binding to phospholipids.

Though fluorescence spectroscopy lacks the resolution to distinguish between tyrosine residues in the Cyt *c* sequence (Tyr48, Tyr67 and Tyr97), the magnitude of the quenching effect was observed to increase as a function of the protein to-lipid ratio (Figures 2.5-2.7). Comparing decreases in fluorescence intensity at 1:1, 1:2, 1:4 and 1:10 protein-lipid ratios, negatively charged CL showed the greatest magnitude of fluorescence quenching relative to Cyt *c* alone (Figure 2.5). Based on the relative magnitudes of the tyrosine fluorescence response, it can be infered that CL has a stronger affinity for Cyt *c* than positively charged DMPC and DHPC. This is consistent with what would be expected for an electrostatically-driven interaction.

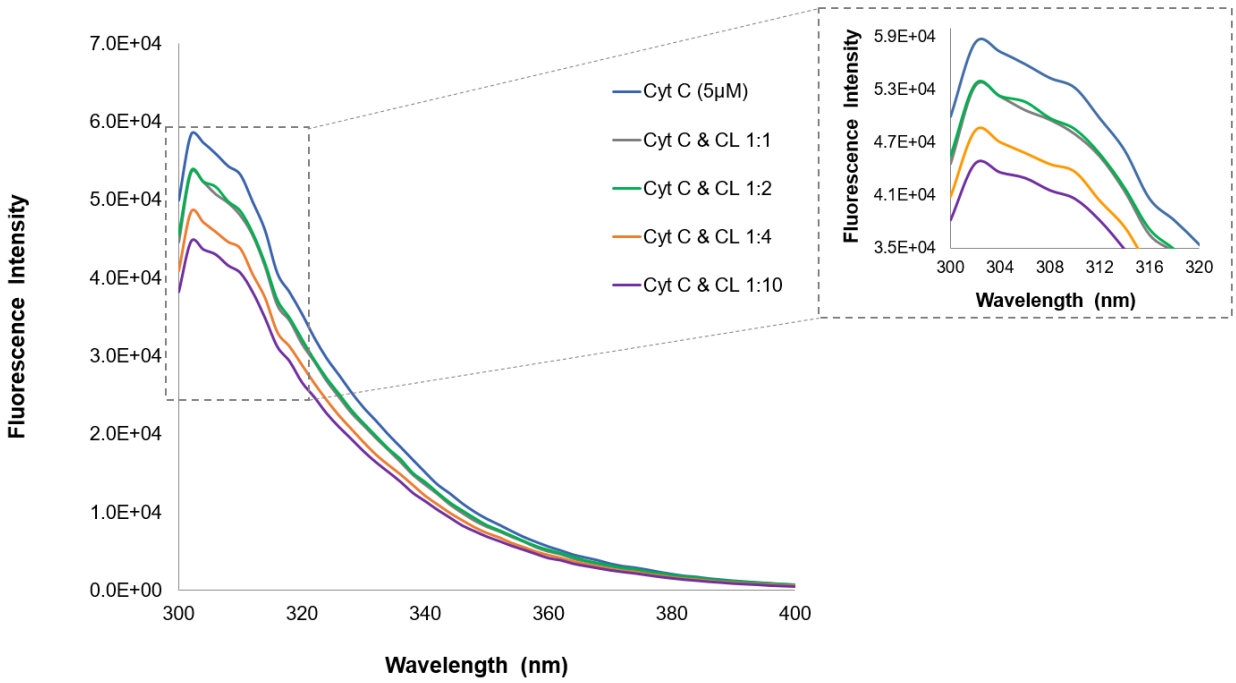


Figure 2.5. Fluorescence spectra for Cyt *c* at increasing concentrations of CL lipids. Apex signal from tyrosine residues (303 nm) decreased as a function of increasing CL concentration, demonstrate proof of Cyt *c*-CL interaction.

Differing only in acyl chain length, positively charged DMPC and DHPC showed similar extents of florescence quenching at 1:10 protein-to-lipid ratios (Figures 2.4). At a 1:1 protein-to-lipid ratio, Cyt *c* with DMPC did not show any significant quenching relative to Cyt *c* only. Additionally, no significant difference was noted between the 1:2 and 1:4 Cyt *c*-to-DMPC ratios (Figure 2.6). In contrast, the signal for Cyt *c* with DHPC decreased at just a 1:1 ratio. The fluorescence quenching effect increased sequentially as a function of protein-to-lipid ratio up to the 1:10 ratio used for IMS-MS and TRESI-HDX (Figure 2.7).

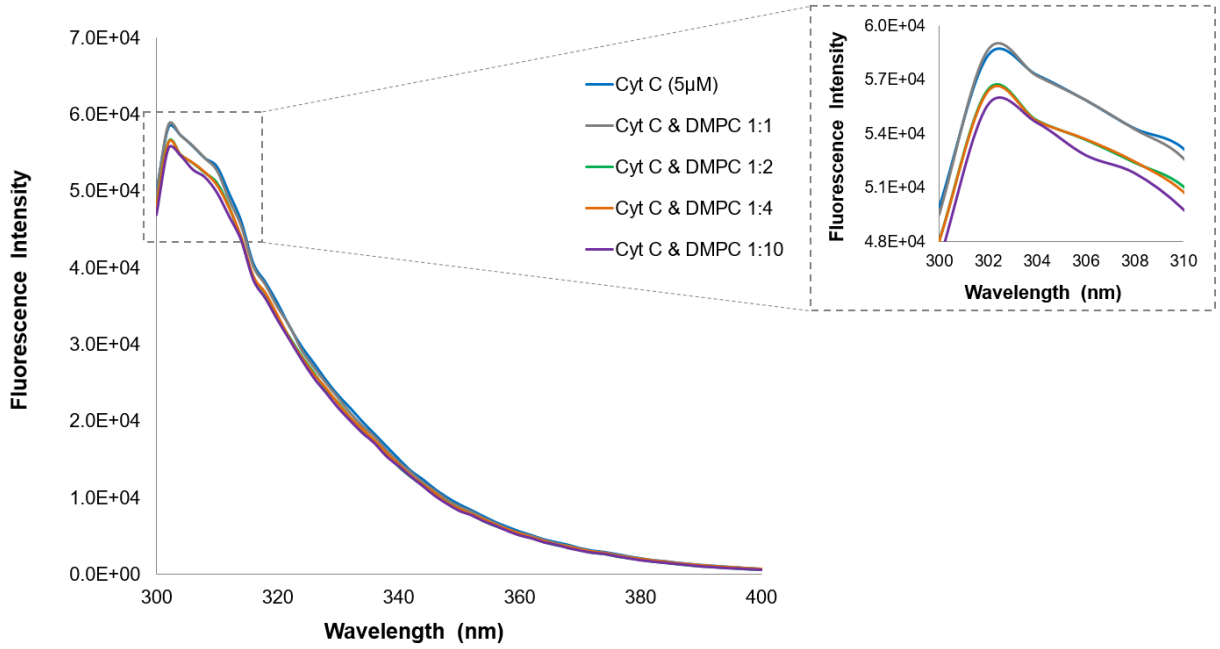


Figure 2.6. Fluorescence spectra for Cyt *c* at increasing concentrations of DMPC phospholipids. Apex signal from tyrosine residues (303 nm) decreased sequentially using 1:2, 1:4 and 1:10 protein-to-DMPC ratios.

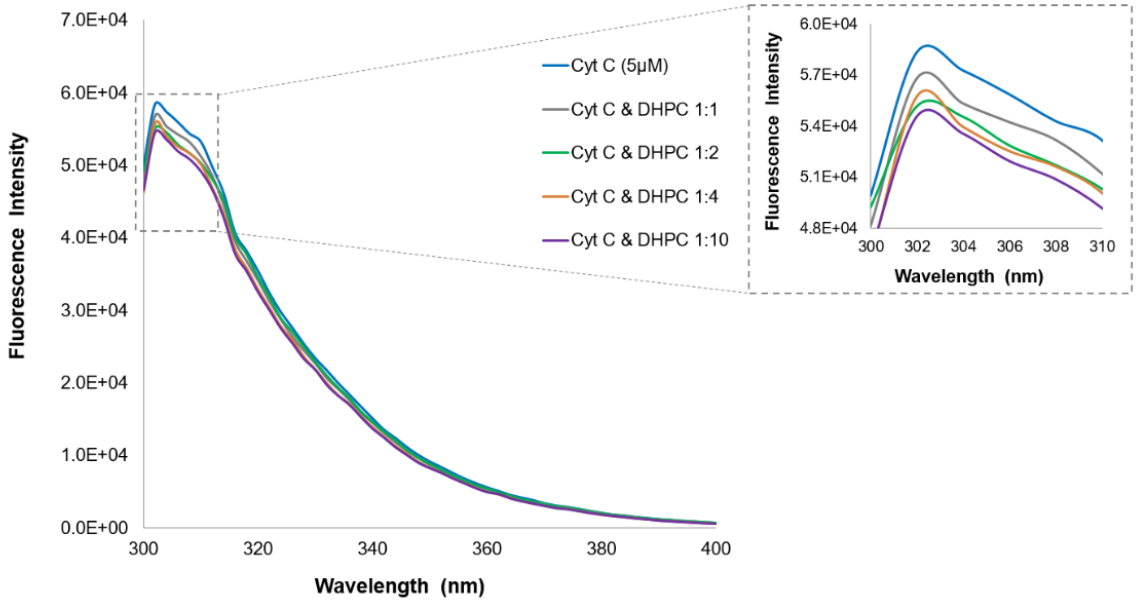


Figure 2.7. Fluorescence spectra for Cyt *c* at increasing concentrations DHPC phospholipids. Apex signal from tyrosine residues (303 nm) decreased as a function of quenching from binding to DHPC phospholipids.

The lipid only controls (Figure 2.8) showed no significant difference in intrinsic fluorescence at concentrations used in the titration analysis. Peak fluorescence signal in the lipid only controls was noted to overlap the tyrosine emission spectra observed for Cyt *c* in the presence of phospholipids (Figure 2.4-2.7). The intensity of fluorescence response in the absence of protein can be attributed to light scattering as previously described for phospholipid-based model membrane systems.[135,136] Common strategies for mitigating light scattering in studies of membrane proteins include measuring typothophan emission at 350 nm instead of tyrosine fluorescence at 303 nm, and the use of additives to increase the refractive index.[137] For the Cyt *c*-lipid system, the presence of the heme moeity quenched fluorescence at 350 nm precluding analysis of trypothophan fluorescence upon lipid binding. The presence of heme protein also increased the refractive index of the system enabling a relative assessment of Cyt *c*-lipid interactions by tyrosine fluorescence quenching (Figure 2.4).

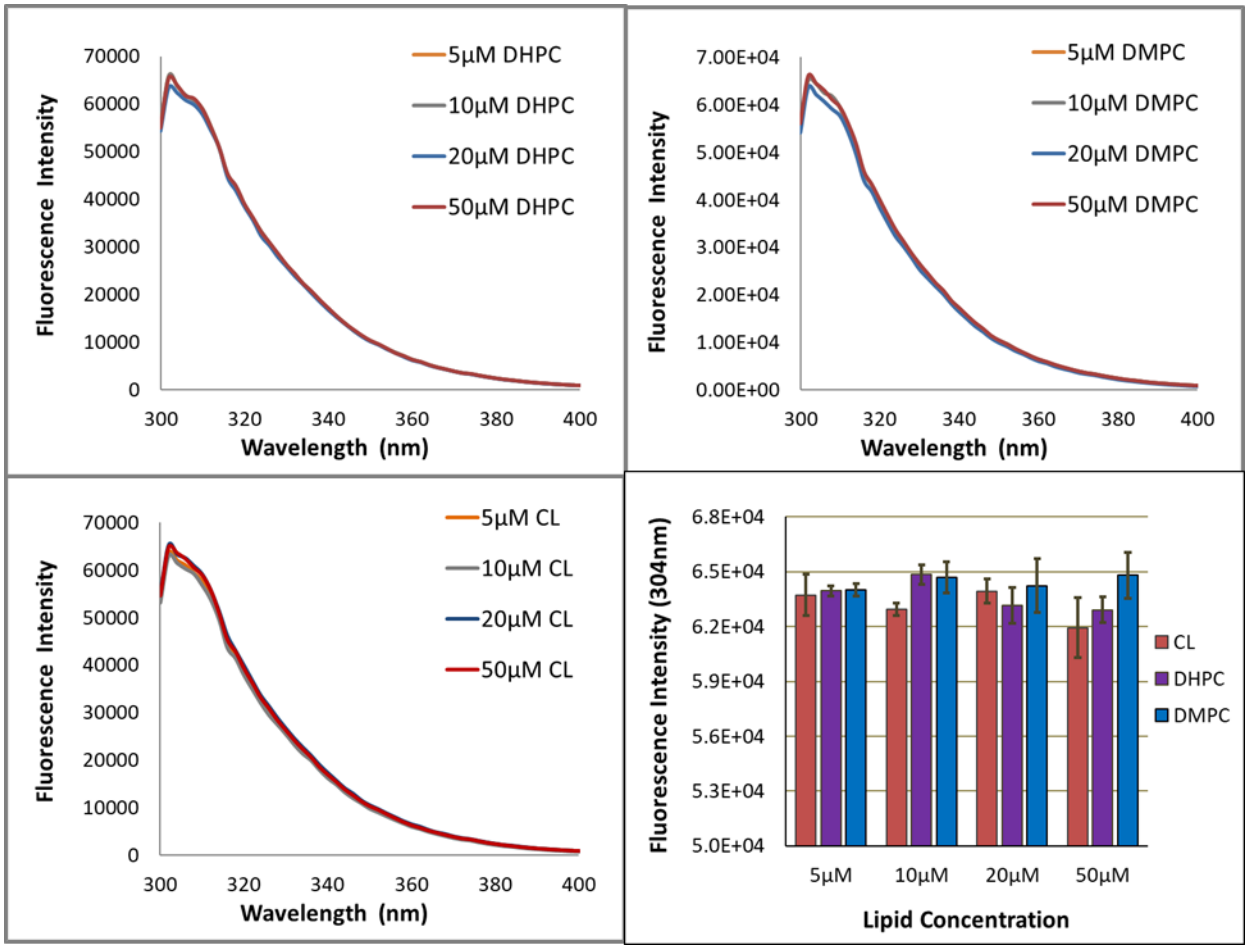


Figure 2.8. Fluorescence spectra for lipid only controls at concentrations used in titration analysis. Error bars reflect standard deviation in measured fluorescence from triplicate preparations. No significant difference in intrinsic fluorescence at increasing concentrations of CL, DHPC or DMPC lipids alone.

To investigate class-dependent differences in Cyt *c*-lipid binding highlighted by fluorescence spectroscopy, global time-resolved hydrogen-deuterium exchange (TRESI-HDX) measurements were performed on Cyt *c* pre-incubated with anionic cardiolipin (CL) or cationic phosphatidylcholine (PC) at a 1:10 protein:lipid ratio. The "time-resolved" aspect of the approach facilitates characterization of transient structural changes using millisecond-to-second D_2O exposure times in contrast to conventional HDX workflows carried out on a minutes to hours timescale.[38,138,139]

Using a capillary-based TRESI mixer, time-dependent deuterium incorporation was monitored from 0.28 to 2 seconds, with the hydrogen exchange rate dependent primarily on hydrogen bonding (i.e. secondary structure) and the degree of solvent exposure, as influenced by electrostatic interactions with phospholipids.[35,37] Globally, Cyt *c* was observed to take up more deuterium in the presence of DMPC, indicative of increased dynamics relative to the protein alone. In contrast, Cyt *c* showed a persistent decrease in deuterium incorporation in the presence of CL suggesting a more compact protein conformation upon CL binding (Figure 2.9).

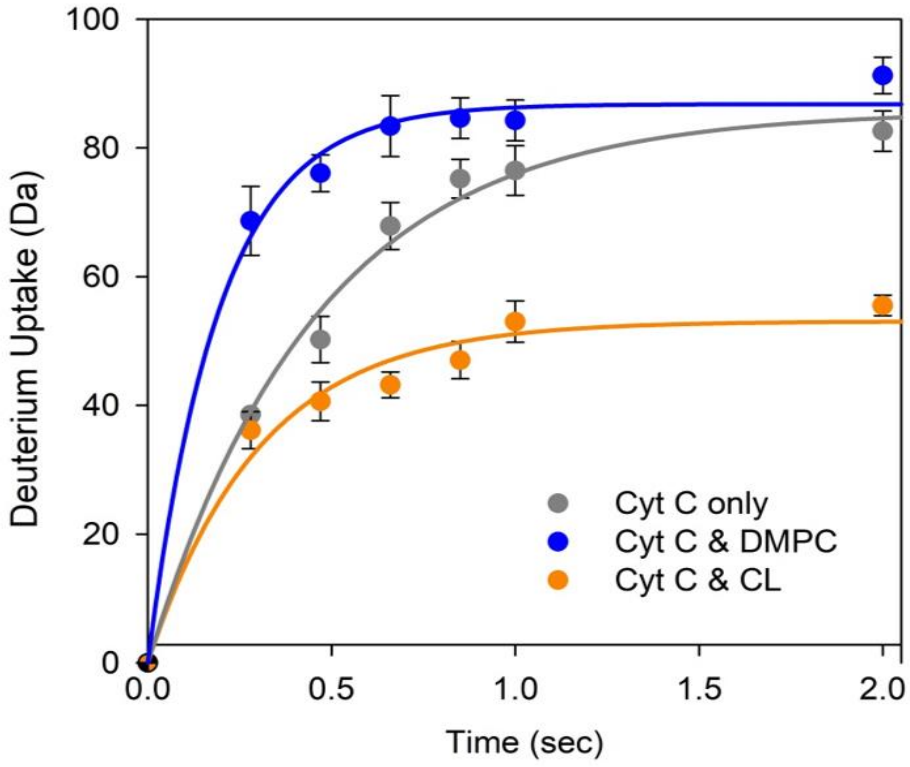


Figure 2.9. **Global HDX results for Cyt *c* in the presence of CL and DMPC phospholipids.** Results represent average deuterium incorporation from 6+, 7+ and 8+ charge states over labelling times from 0.28 to 2 seconds fitted as an exponential rise to maximum, f=a*(1-exp(-b*x)). Error bars represent ± standard deviation from n=3 replicates. See Appendix 1 for curve fitting analysis from Global HDX results.

Ion mobility spectrometry-MS (IMS-MS) was used to provide further insight into the conformational heterogeneity in the lipid-bound Cyt *c* ensemble as previously suggested by time-resolved florescence (FRET) experiments.[125,126] As a gas-phase analogue of gel electrophoresis, ion mobility separates ions based on their size, shape and charge as they traverse the ion mobility cell. Shorter drift times are indicative of a faster mobility and more compact gas-phase protein structure, whereas longer relative mobility times corresponded to more extended protein conformations.[47,140] For Cyt *c* only (pH 7), the observed drift time chromatogram showed a predominant 7+ charge state with a Gaussian distribution centered at 6.55 ms ± 0.06 ms (Figure 4). This is consistent with previous IMS and native MS studies of Cyt *c*.[141,46,142,143]

The addition of phospholipids resulted in the formation of two newly dominant sub-populations, one with increased mobility and an apex drift time of 7.05 ms ± 0.07 ms, and a second peak with decreased mobility at 6.04 ms ± 0.11 ms (Figure 4). Though both sub-populations are seen with the addition of CL or DMPC, the more compact conformation with decreased mobility was observed to be more heavily populated in the presence of CL. This agrees with a negative drift time differential of - 0.51 ms for 'Cyt *c* & CL' – 'Cyt *c* only' considering the apex of the corresponding peaks in the drift time chromatogram (Figure 2.10). The more extended conformation with increased mobility was dominant in the presence of DMPC, corresponding to a positive drift time differential of 1.01 ms for 'Cyt *c* & DMPC' – 'Cyt *c* only' (Figure 2.10) These IMS results agree with the global HDX results showing a more extended protein conformation with positively charged DMPC, and more compact structure with negatively charged CL (Figure 2.9).

Evidence for differential conformer populations between lipid classes is further supported by the charge state distribution observed through native mass spectrometry (MS). Consistent with previous MS characterization of native Cyt *c*,[143,144] the protein only spectra was predominantly comprised of the 7+ charge state. Incubation with negatively charged CL promoted the formation of lower charge states (6+ and 7+) during the electrospray ionization process. In contrast, Cyt *c* in presence of DMPC showed an increased proportion of the 8+ charge state (Figure 2.10), in line with of a more extended protein structure during the solution to gas-phase transition.

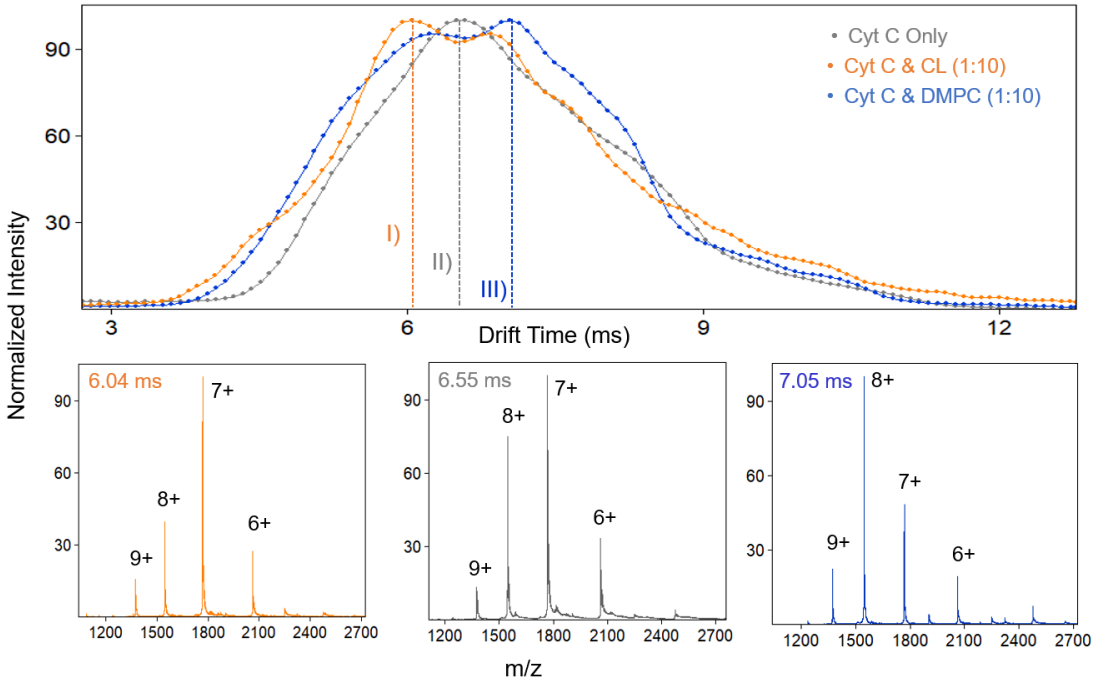


Figure 2.10. Ion mobility drift time chromatograms for Cyt *c* only and with CL or DMPC phospholipids. Cyt *c* with CL exhibited a shift in population towards a more compact protein conformation (decreased drift time) relative to Cyt *c* only. Cyt *c* with DMPC showed a more ended conformation with a longer drift time. Native MS spectra of Cyt *c* show differential charge state distributions consistent with the conformational heterogeneity observed by Global HDX and IMS-MS.

Following protease digestion by pepsin and MS/MS analysis, 99% sequence coverage was obtained for Cyt *c* including heme contact sites (Cys14, Cys17, His18, Met80), and residues implicated in the extended lipid anchorage model (Asn52, Lys72, and Lys73). Overall, the local measurements probed by TRESI-HDX agree with the differences observed in Global HDX analysis of the intact protein. CL-bound Cyt *c* showed decreased levels of deuterium uptake in peptides spanning covalent thioester linkages (Cys14 & Cys17) and axial ligands (His18 & Met80), reflecting a more compact Cyt *c* conformation relative to non-lipid bound Cyt *c*. This is consistent with previous solid-state NMR and FTIR measurements of Cyt *c* bound to membrane conformers with preservation of secondary structure upon peripheral interaction with the membrane interfaces.[111] Interestingly, CL did not induce any significant decreases in deuterium uptake in the region spanning the C-site (including Asn52) previously implicated in the formation of hydrophobic interactions and/or hydrogen bonding.[122,124] In regards to the electrostatic interactions proposed by the lipid anchorage model,[122,123] the TRESI-HDX results showed decreased uptake consistent with prior identification of the A-site (Lys72/73) as a lipid-binding region. However, decreased deuterium uptake around Lys72/73 was observed with both positively-charged (PC) and negatively-charged (CL) phospholipids. This suggests that the interactions at this proposed (A-site) for lipid-binding may not be solely driven by electrostatic forces.

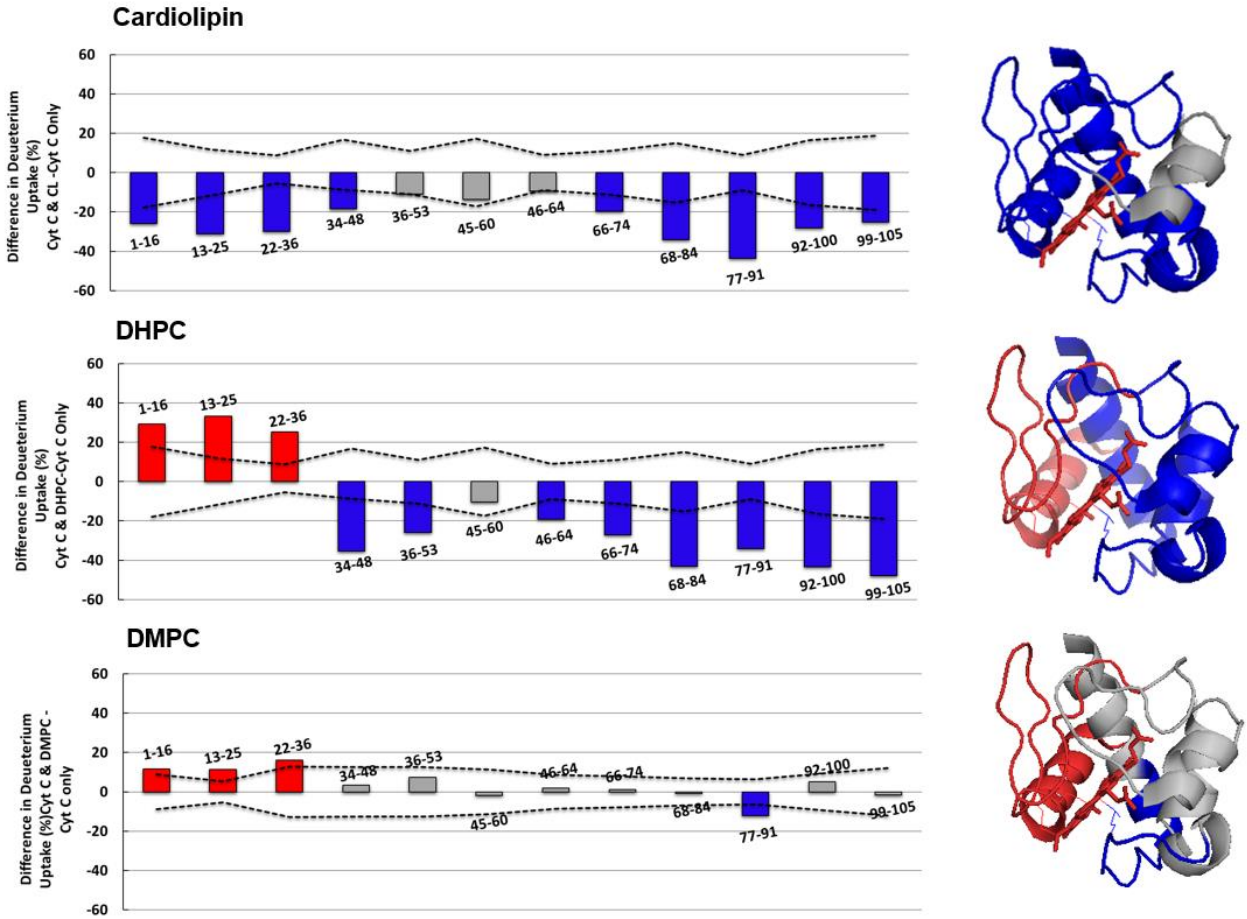


Figure 2.11. TRESI-HDX measurements comparing Cyt *c* with CL, DMPC or DHPC phospholipids. Differences between 'protein & lipid' versus protein only summed across labelling times of 0.5, 1 and 2 seconds. For graphs, positive (red) bars indicate increased uptake in the presence of phospholipids whereas negative (blue) bars indicate decreased deuterium uptake relative to protein only. 95% confidence intervals (2σ) depicted in grey (dashed lines) for differences observed in each peptide across n=6 measurements. Peptides exhibiting differences below the 2σ significance threshold are denoted in grey. Using the same colour scheme, results were mapped onto the protein crystal structure for equine Cyt *c* (pdb: 1hrc).

With DMPC and DHPC, Cyt *c* showed increased uptake in peptides spanning the N-terminus and heme contact residues (Cys14/Cys17/His18) located in the proximal face of the heme pocket. Comparing differences in uptake values between the positively charged PC phospholipids, it was noted that at the same Cyt *c*:lipid molar ratio (1:10), DMPC induced less significant differences in deuterium uptake relative the shorter PC lipid, DHPC. Given that acyl chain length is the only structural difference between DMPC and DHPC, this is likely indicative of different affinities, as previously shown with fluorescence spectroscopy under the same experimental conditions (Figure 2.6-2.7).

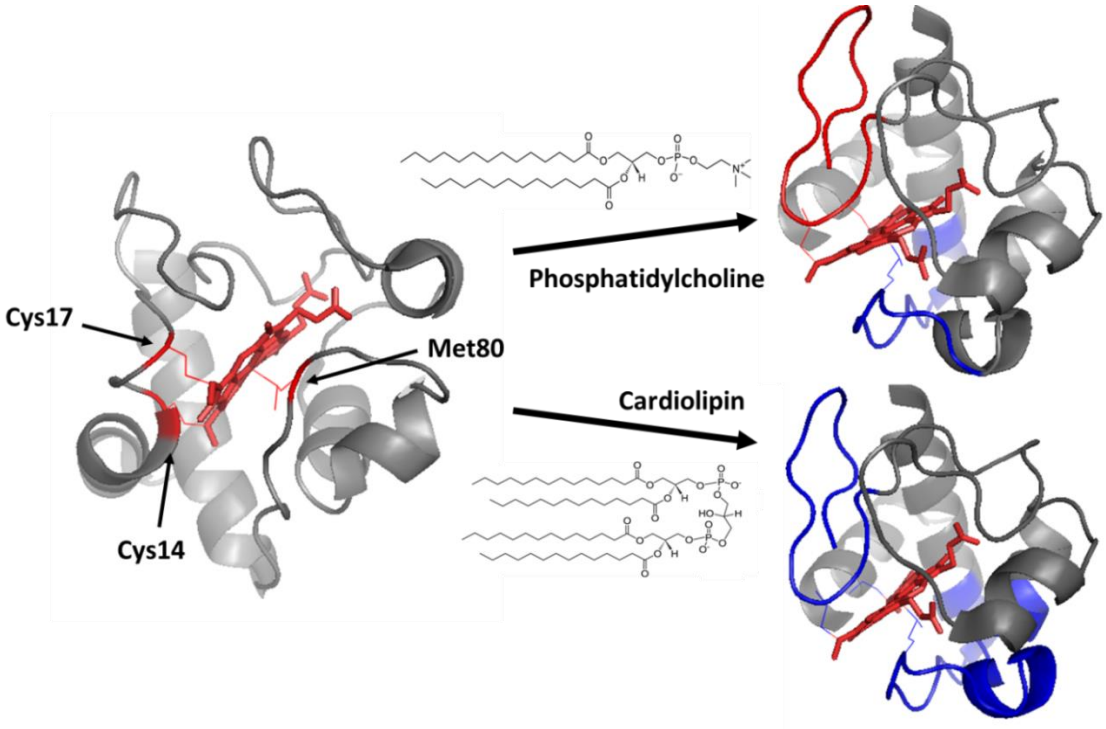


Figure 2.12: Proposed electrostatic priming effect for Cyt *c* with PC and CL phospholipids. TRESI-HDX results for DHPC and CL centered on heme contact residues. Peptides coloured in blue represent a significant decrease in uptake relative to Cyt *c* only, whereas those in red represent a significant increase in uptake. Based on HDX results, electrostatic interactions with negative charged CL promote a tightening in Cyt *c* whereas PC lipids increase dynamics on the proximal face of the heme binding pocket containing Cys14 and Cys17.

From a mechanistic perspective, peroxidases typically contain a penta-coordinated heme group with a vacant (distal) site where H_2O_2 can bind and initiate the catalytic cycle.[145] Native Cyt *c* is paradoxical in that it is hexa-coordinated with His18 and Met80 axial ligands and four pyrole nitrogens, yet still exhibits peroxidase activity. It has been proposed that interaction with phospholipid bilayers induce a change in Cyt *c* conformation to a penta-coordinated heme state that is capable of CL peroxidation.[146–148] Cyt *c* first associates with the membrane then gains peroxidase active as a result of H_2O_2-induced covalent modifications. Recent experimental evidence has shown oxidation at Tyr67 or Met80, [149,150] and side chain carbonylation at Lys72/73[151] to be pre-requisite modifications for Cyt *c* peroxidase activity. This two-stage activation model is also further corroborated by the well-established initial lag phase in Cyt *c*-CL peroxidase kinetics.[149]

Interestingly, the peptide regions implicated in modification-driven activation of peroxidase activity (Tyr67, Lys72/Lys73, Met80) showed decreased uptake upon binding to phospholipids. Though observed with both PC and CL lipids, the peptide spanning the Met80-heme bond (77-91) showed the largest decrease in the presence of CL. This decrease in dynamics in the region spanning Met80-heme ligation may facilitate activation by making H_2O_2-induced modification of Cyt *c* more energetically favourable as intracellular H_2O_2 levels rise from submicromolar to submillimolar in the early stages of apoptosis.[56,57] In contrast, cationic DMPC and DHPC destabilize heme contacts that are structurally needed to form a penta-coordinated conformation of Cyt *c*. Thus, conformational priming from interactions with PC lipids likely makes peroxidase catalysis less energetically favourable by preventing binding of H_2O_2 and/or stabilization of the oxidase heme (Fe=O) intermediate (Figure 2.13).

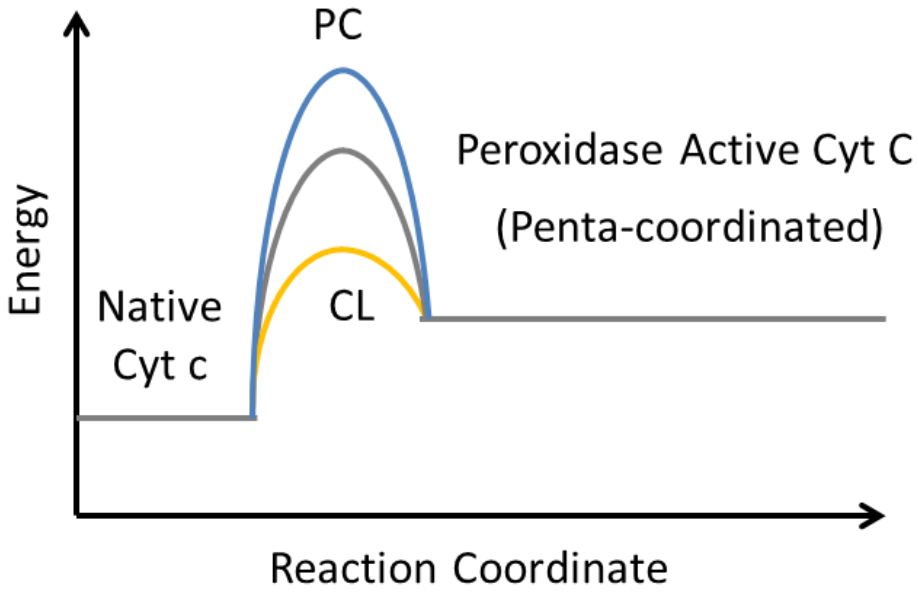


Figure 2.13. Proposed reaction coordinate diagram based on conformational dynamics from electrostatic interactions with PC and CL phospholipids. CL primes native, hexa-coordinated, Cyt *c*, facilitating interactions required to modify Cyt C into peroxidase-active state. PC lipids induce partial unfolding of heme binding pocket preventing stabilization of the oxidase heme (Fe=O) intermediate for peroxidase catalysis.

2.4 Conclusion

Taken together, TRESI-HDX analysis of Cyt *c* with PC and CL lipids demonstrates that electrostatic contributions of different lipid classes should be considered when designing experiments using model membrane surfaces. In probing exclusively for electrostatic interactions, the results support a two-stage model in which lipids prime Cyt *c* peroxidase activation by influencing conformational dynamics in the establishment of the Cyt *c*-membrane assembly. Global HDX, IMS-MS and TRESI-HDX analyses are in agreement, with PC lipids causing unfolding of one face of the heme pocket, impacting Cyt *c*'s ability adopt the penta-coordination heme conformation that is a prerequisite for peroxidase activity. Cyt *c* was also observed to adopt a more compact protein conformation when bound to CL, with no evidence of global unfolding by HDX or IMS-MS. This compact CL-bound Cyt *c* structure may facilitate further activation of peroxidase activity in Cyt *c* as required for mitochondrial-induced apoptosis. Collectively, this work demonstrates TRESI-HDX as a powerful tool for investigating conformational dynamics in lipid binding with adequate sensitivity for probing transient changes at protein-membrane interfaces.

Chapter 3 – Optimization of proteomic methods for studying complex biological samples from *Bordetella Pertussis*

3.1 – Introduction

While studies of biochemical processes such as the pro-apoptotic Cyt *c* pathway described in the previous chapter provide value at a mechanistic level, it is important to understand that these types of interactions do not occur in isolation. Greater understanding of complex cellular networks requires larger scale analyses of proteins and their functions. Fortunately, 20 years of development has made holistic profiling routinely accessible via LC-MS-enabled shotgun proteomics.

3.1.1 Proteomics in Vaccine Development

Applications of shotgun proteomics techniques have led to tremendous advances in our understanding of human diseases. However, substantial gaps still remain in medically-relevant proteomes including many microbes responsible for infectious diseases. While the success of genome projects has provided us with vast information on the genes of many pathogenic species, development of new vaccines depends on identification and characterization of potential antigens for eliciting a protective immune response.[152] Historically, the role of MS has been limited to the confirming the identity protein antigens in recombinant and viral vaccines. In addition to intact mass analysis typically used for identity, LC-MS can also provide information on primary amino acid sequence, disulfide bonds, and post-translational modifications (PTMs) such as glycosylation.[153]

Another area of biopharmaceuticals where MS is increasingly being applied is the profiling of host cell protein (HCP) impurities which are expressed during large-scale cultivation of biologics.[154] Monitoring the composition and abundance of HCPs across different stages of a manufacturing is a fundamental aspect of process development and control. Historically, HCPs have been detected using Enzyme-linked immunosorbent assays (ELISA).[155] While commonly accepted by industry and health authorities for determining relative HCP amounts (usually expressed in ppm levels relative to the main protein therapeutic), polyclonal multi-analyte ELISAs offer minimal information on the identity of individual proteins within a population of HCPs.[155] Moreover, they are often subject to matrix effects when trying to compare across different manufacturing stages. Consequently, there is increased interest in MS-based proteomics as an orthogonal approach for detecting proteins in biopharmaceuticals.[156]

The coming sections in this chapter will describe applications of shotgun proteomics for the characterization of acellular component Pertussis vaccines manufactured by Sanofi Pasteur Canada. Chapter 4 will highlight development of a targeted proteomics method for absolute quantitation of specific residual toxins from *Bordetella pertussis (B. pertussis)*.

3.1.2 *Bordetella Pertussis* Pathogenesis

Pertussis, also known as whooping cough, is a highly contagious respiratory disease caused by infection of an aerobic, gram-negative bacterium called *Bordetella pertussis (B. pertussis).* Molecular pathogenesis of *B. pertussis* is driven by virulence factors which can be categorized as adhesions or toxins based on their mechanism of action. Adhesions, such as filamentous hemagglutinin (FHA), pertactin (PRN) and fimbriae of serotype 2 or 3 (Fim 2/3) are cell-surface proteins which serve as attachment points for *B. pertussis* to epithelial tissue in the respiratory tract (Figure 3.1). Once anchored, *B. pertussis* makes use of pertussis toxin (PT), tracheal cytotoxin (TCT), adenylate cyclase toxin (ACT), and dermonecrotic toxin (DNT) to downregulate cellular and antibody-mediated immune responses.

Pertussis infections were widespread until the late 1940s, with annual case rates as high as 270 000 in the United States and fatality rates at about 10%.[157] Whole cell vaccines (wP), comprised of inactivated *B. pertussis* bacteria were introduced in the late 1940s and were hugely successful in reducing incidence rates throughout the 1950s and 1960s in areas with high vaccine coverage.[157] Adverse publicity in the 1970s led to decreased compliance and prompted the development of acellular Pertussis (aP) vaccines with proteins components purified from *B.pertussis* bacteria.[158] To date, up to five proteins have been harnessed for their immunogenic properties as components in acellular vaccines: detoxified PT, cell surface adhesion proteins (FHA & PRN), and pili-forming agglutinogens (FIM 2 & 3). Since the transition to aP vaccines in the 1990s, case rates and regional pertussis epidemics have increased steadily.[159] This has sparked renewed interest in understanding molecular mechanisms of *B. pertussis* pathogenesis with the goal of developing a new generation of safe and effective Pertussis vaccines.[160]

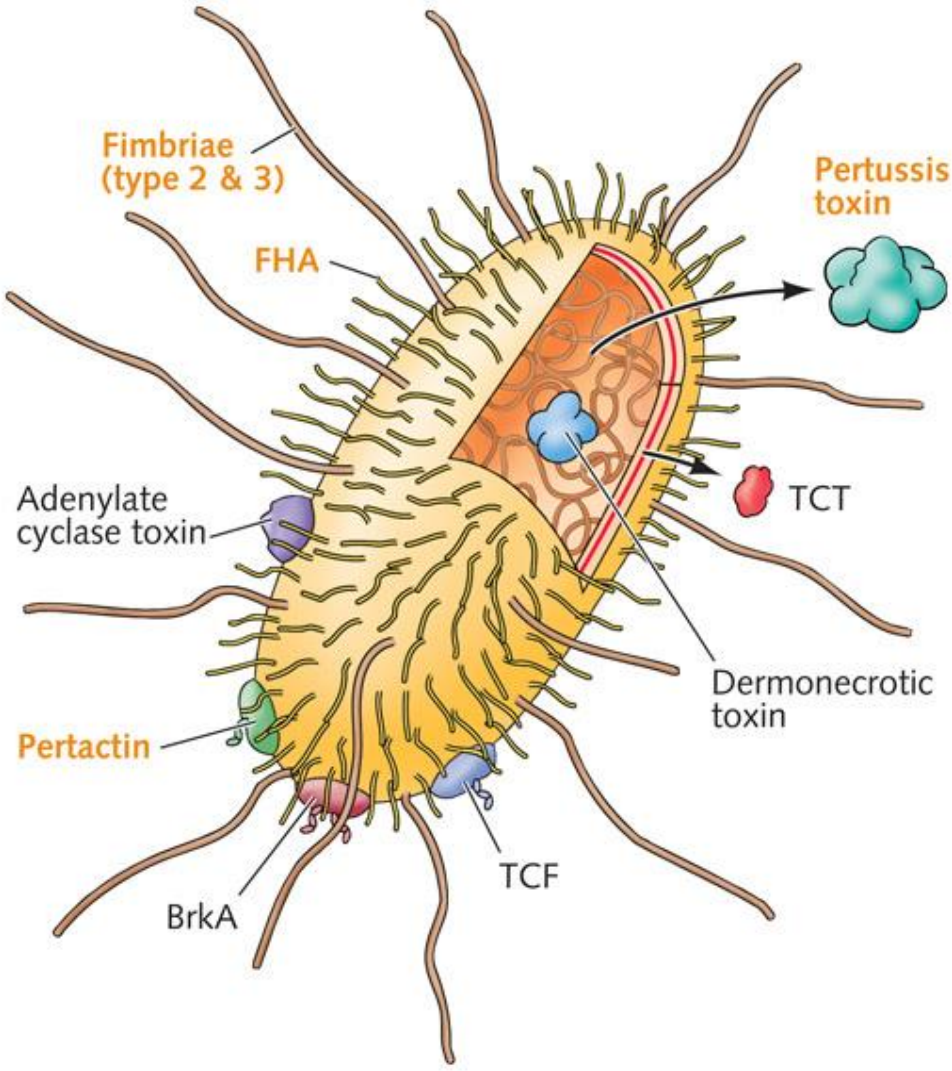


Figure 3.1. Bordetella pertussis organism with virulence factors for pathogenesis. Pertactin, Pertussis toxin, Filamentous Hemagglutinin, Fimbriae type 2 & 3 harnessed as antigen components in acellular whooping cough vaccines. Adenylate cycle toxin (ACT), dermoneucrotic toxin (DNT), and tracheal cytotoxin (TCT) are monitored as residual host cell impurities during downstream purification from fermenter harvest. Image source: Sanofi Pasteur Canada. Accessed on March 6, 2019 through public webpage: http://www.museumofhealthcare.ca/explore/exhibits/vaccinations/pertussis.html.

Advancements in shotgun proteomics have coincided with increased interest in applying LC-MS for protein identification in biopharmaceutical development.[84] Nevertheless, there remains a need for sample processing methods that are unbiased to protein solubility and can provide a greater depth of coverage from proteomic analyses. One such method is filter-aided sample preparation (FASP) first introduced by the Mattias Mann group in 2009[161] and adapted into a "suspension trap" (S-trap) format in 2014.[162] In contrast to a traditional in-solution digestion performed via sequential aqueous reactions in a test tube, S-trap enzymatically digests proteins in a spin column format.[163] The top layer contains porous silica filter media to trap particulate protein suspensions and the bottom layer holds C18 resin for online clean-up of digested peptides.[163] Advantages of spin column format include the ability to pre-concentrate samples and increases enzymatic digestion efficiency due to the physical space constraints within the pores of the column. Additionally, the ability to efficiently clean-up samples during sample loading enables the use of detergents such Sodium Dodecyl Sulfate (SDS) which was previously been limited due the incompatibility of detergents with MS analysis.

This section describes shotgun proteomics analyses performed on unpurified culture supernatant with the aim of increasing the depth of protein detection from the secretome of *B. pertussis*.[164] An evaluation of digestion protocols was then performed to assess whether the S-trap protocol could enhance HCP detection, and improve sequence coverage for hydrophobic antigens such as pili-forming FIM 2/3.

3.2 Experimental Methods

The sample digestion protocols are described based on total protein from Bradford analysis. In-solution and S-trap digestion strategies were evaluated to optimize protein detection. For both digest protocols, purified FIM antigen was reduced at 95°C for 15 min then sonicated in a 40°C water bath for 10 min to further promote solubility.

3.2.1 In-Solution Sample Digestion

150 µg of culture supernatant or 25 µg of purified FIM antigen containing both serotype 2 or 3 (Fim 2/3) were digested in triplicate using an in-solution digestion protocol. Briefly, 1 M ammonium bicarbonate (ABC) and 1% of acid-labile surfactant (Rapigest)[165] were added to samples to a final concentration of 0.1 M ABC and 0.1% Rapigest considering a 150 µL digest volume. Samples were incubated at 60°C for 30 min to facilitate disulfide bond reduction with 10 mM of Tris(2-carboxyethyl) phosphine hydrochloride(TCEP). Iodoacetamide (IAA) was then added (final concentration 20mM) for alkylation of the reduced cysteines. Following incubation for 60 min at room temperature, the reduced and alkylated samples were enzymatically digested using a Trypsin/Lys-C proteolytic enzyme mixture at an enzyme-to-substrate ratio of 1:20 (w/w) for 4 hr at 37°C. Digests were quenched by the addition of 4 µL of TFA and confirmed to have a pH of ~2 using a pH strip. The acidified digest samples were then incubated at 37°C for 30 min and centrifuged at 8,100 x g for 10 minutes to facilitate Rapigest hydrolysis.

3.2.2 S-Trap Sample Digestion

A suspension-trap (S-trap) digestion was performed as per the vendor's recommended protocol [162,163] with 25 µg of culture supernatant or purified FIM antigen samples. Briefly, the protein samples were mixed 1:1 (v/v) with 10 % SDS, 100 mM triethylammonium bicarbonate (TEAB, pH 7.55) then reduced and alkylated as described above for the in-solution digestion.

12.5 % aqueous phosphoric was added to a final concentration of 1.2 %. 6 volumes of S-Trap protein binding buffer (90% MeOH, 100 mM TEAB, pH 7.1) was then added to form colloidal protein particulate. The acidified SDS lysate/protein mixtures were loaded onto S-trap micro columns via centrifugation at 4000 x g for 2 minutes, effectively trapping the protein within the silica matrix of the spin column. Captured protein was washed by spinning through 3 volumes (150 µL) of S-Trap binding buffer. 25 µL of digestion buffer containing Trypsin/Lys C at 1:10 or 1:20 protease-to-substrate ratio (w/w) was added to the top of the column and incubated for 1 hr at 47°C (Figure 3.2). Digestions were performed in duplicate for each sample type tested. Peptides were eluted from the spin column by centrifugation at 4000 x g for 2 minutes with 40 µL of TEAB; 40 µL of 0.1% FA; and 35 µL of 50% ACN, 0.1% FA. Fractions were pooled and frozen at -80°C. Prior to LC-MS analysis, samples were diluted 3-fold to lower ACN content from ~15 to 5% for loading onto the C18 RP-HPLC column.

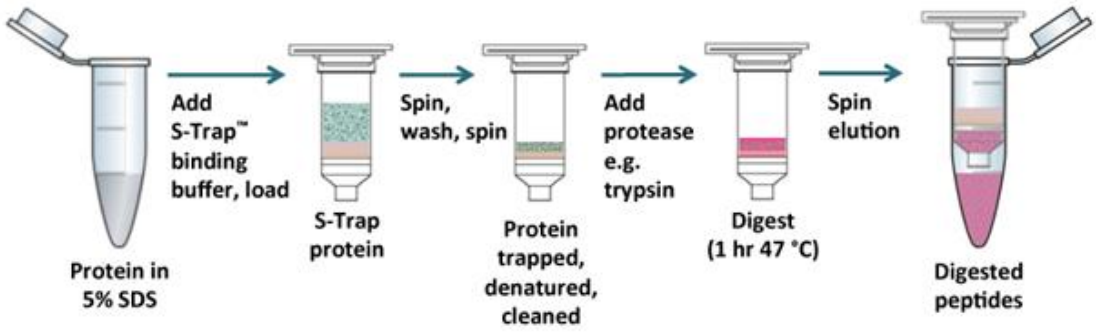


Figure 3.2. Schematic of a suspension trap (S-trap) sample digestion for LC-MS/MS analysis. A protein sample is solubilized with SDS and acidified using phosphoric acid. The acidified protein–SDS mixture is added to a methanol-based S-trap binding buffer at pH 7.1 creating a fine protein suspension that is capture by a silica (SiO_2) depth filter. The sample is washed with additional binding buffer whereby proteins are retained, and contaminants are removed in the flow through. Trapped proteins are enzymatically digested within the interstitial space of the S-trap, "falling" into a lower bed of C18. Peptides are eluted using acetonitrile followed by an optional concentration step prior to LC-MS/MS analysis.

3.2.3 Data Acquisition

Shotgun protomics analyses were performed on 425 ng culture supernant and 0.5-1 pmol of purified FIM digest samples using a UltiMate 3000 RSLCnano UHPLC system (Thermo Scientific) plumbbed in a "trap-elute" configuration (Figure 3.3). Pre-concentration and desalting were performed online using a C18 trap column (5mm x 300 µm i.d.) followed by seperation on an analytical column (PepMap100 C18 75 µm x 15cm 3µm). Loading pump mobile phases consisted of water (mobile phase A) and acetonitrile (ACN; mobile phase B) with 0.1% TFA. Nano pump mobile phases were mixed with FA as an ion-pairing agent during elution and MS analysis. nanoLC experiments were run at a flow rate of 300 nL/min with the trap column set at 40°C and the analytical column held at 30°C. 95 min and 45 min analytical gradients (1% to 35% B) were used for the cultutre supernant and FIM digest samples, respectively.

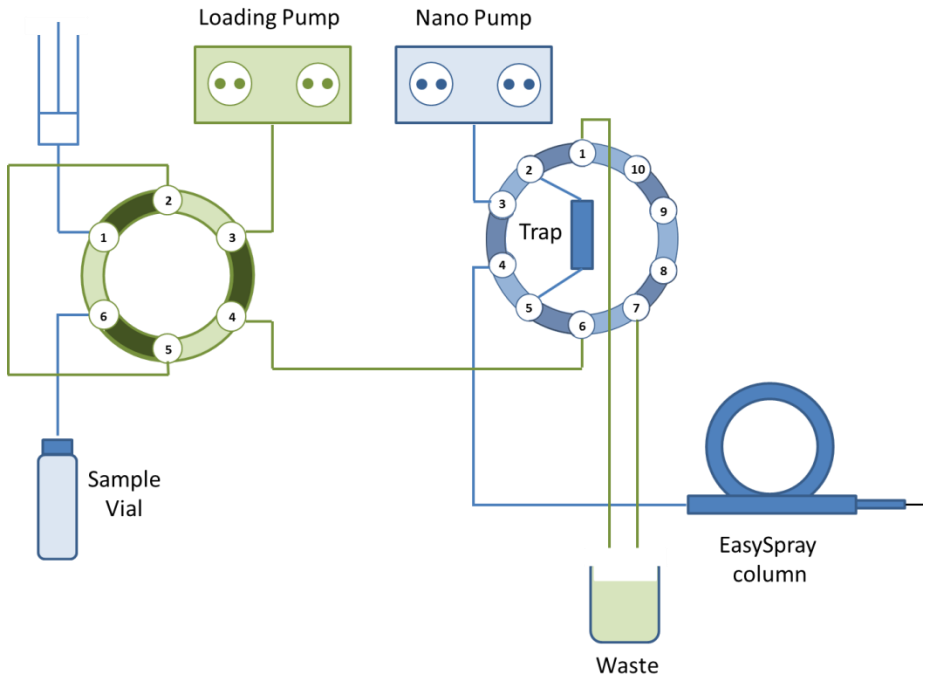


Figure 3.3. Trap-elute nanoLC setup for LC-MS/MS experiments. Online pre-concentration and desalting was performed using the loading pumps at 5 µL/min. Peptides were eluted by the nano pumps at 300 nL/min onto a EasySpray analytical column with integrated ESI emitter.

Shotgun MS acqusitions were performed on a Q Exactive HF mass spectrometer (Thermo Scientific) with a survey scan from 400-2000 m/z and up to 10 dependent MS/MS scans. The survey scan was acquired at 120, 000 resolution with an Automatic Gain Control (AGC) target of $3e^6$ and maximum injection time (maxIT) of 200 ms. Data dependent MS/MS scans were acquired at 60, 000 resolution with an AGC target of $3e^6$ and maxIT of 200 ms. Precursors were selected using a quadrupole isolation window of 1.6 m/z and fragmented by HCD using normalized collision energy of 27. Dynamic exlusion was applied for 12 seconds. Instrument tune parameters of note: capillary voltage of 1.9 kV, capillary temperature of 275°C and S lens RF voltage set to 50.0.

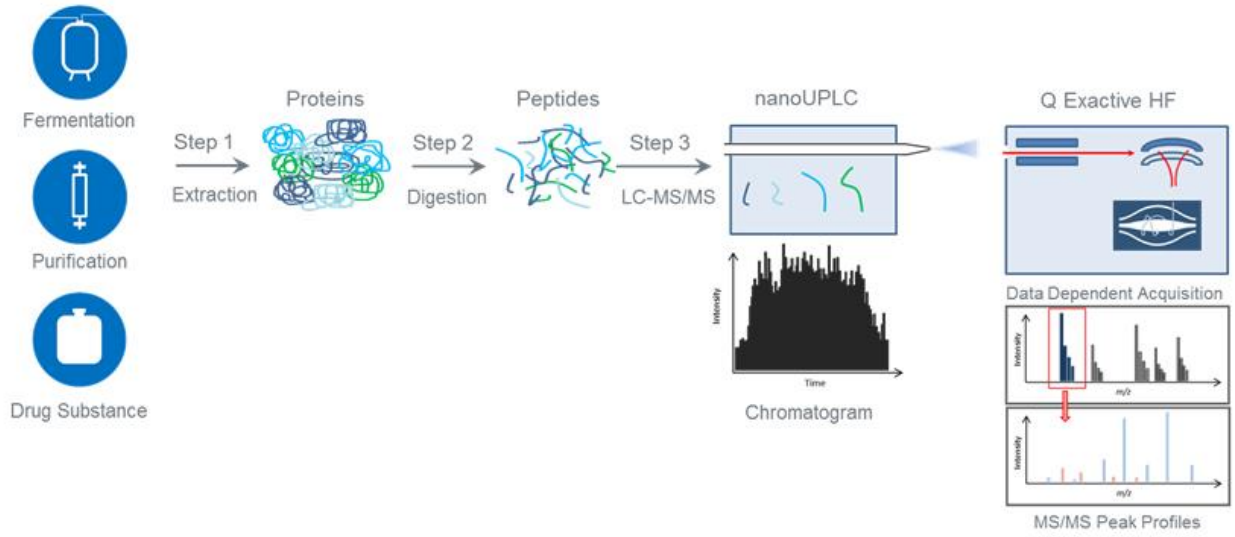


Figure 3.4. Schematic of a shotgun proteomics workflow for identification of proteins from *Bordetella pertussis*. nanoLC-MS/MS workflow for bottom-up proteomic analysis consist of a sample preparation in which proteins are extracted, solubilized and denatured with surfactact to facilitation proteolytic digestion. Peptides are then seperated by nano Ultra-Performance Liquid Chromatography (UPLC) and detected by data-dependentacqusition (DDA). Adapted from James *et al.* 2018.[166]

3.2.4 Data Processing for Shotgun Protein Identification

NanoLC-MS/MS results files were imported into Proteome Discoverer version 2.1 (Thermo Scientific) for data processing. Databases containing the following protein sequences were used for all analyses: SP antigens (67 sequences), *B.pertussis* proteome Tohama I strain (3529 sequences), and common contaminants (118 sequences).[167] Experiments were searched using the Sequest HT algorithm with trypsin as the enzyme. A maximum of 4 missed cleavages were considered with the following dynamic modifications: carbamidomethyl (C), deamidation (NQ), oxidation (M) and carbamylation (K). Precursor mass tolerance was set to 10 ppm and the fragment mass tolerance to 0.02 Da. Results were filtered based on a target decoy search with a false discovery rate of 1% at the peptide-spectral match (PSM) level.

3.3 Results & Discussion

Visual inspection of total ion chromatograms (TICs) from the in-solution and S-trap digests revealed significant "clean up" effects from the S-trap digestion protocol. Comparing the TIC profiles (Figure 3.5), the in-solution digests showed evidence of polymer contamination (80-85 min) and a late-eluting peak of hydrophobic contaminants (100-105 min) which were not present in the S-trap digests from the same starting sample. This represents efficient removal matrix components during the wash steps prior to S-trap loading; an option that is not available with traditional in-solution digestions. Both samples had approximately the same intensity of peptide-related features between 15 and 90 minutes (~$2e^9$) from a comparable 425 ng load on-column (Figure 3.5). Nevertheless, the S-trap digests (Figure 3.5b) appeared to have more defined peaks in the chromatogram and a lower signal-to-noise than the in-solution digest samples (Figure 3.5a).

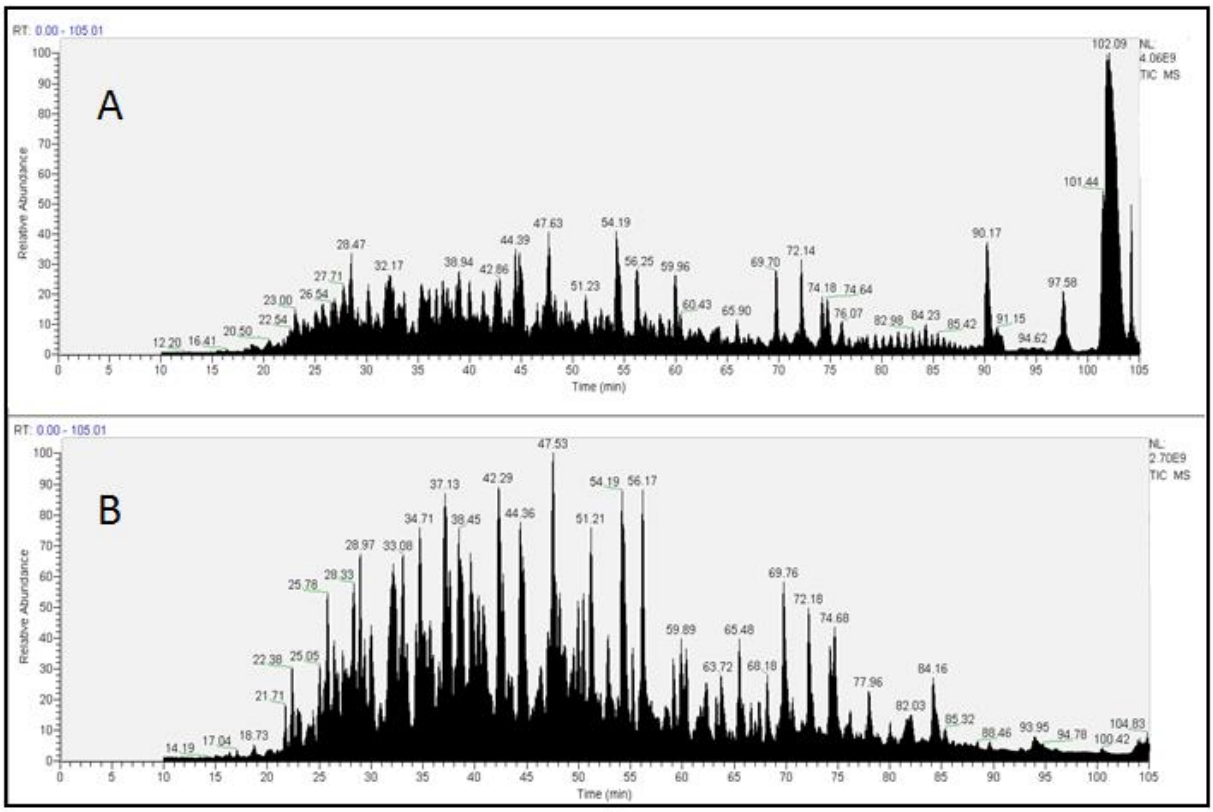


Figure 3.5. Comparison of total ion chromatograms for culture supernatant using in-solution versus S-trap digestion protocols. DDA analysis from 425 ng unpurified culture supernatant digested A) in-solution versus B) S-trap with a 1:20 (w/w) enzyme-to-substrate ratio. S-trap digestion was noted to remove late eluting salt and polymer contaminants observed in samples prepared via in-solution digestion.

DDA search results for unpurified culture supernatant identified 254 proteins belonging to 245 protein groups. Of these proteins, 24 were uniquely identified with S-trap digestion versus 3 proteins unique to the in-solution protocol (Figure 3.6). Comparison of peptide identifications showed a similar trend with increased identifications in the S-trap versus in-solution digest samples (Figure 3.6). Further comparison of the enzyme-to-protein ratios revealed no significant difference in identifications between S-trap digests performed at 1:10 and 1:20 (w/w). The results also indicate that a 1:20 ratio was appropriate given the protein load (Figure 3.6).

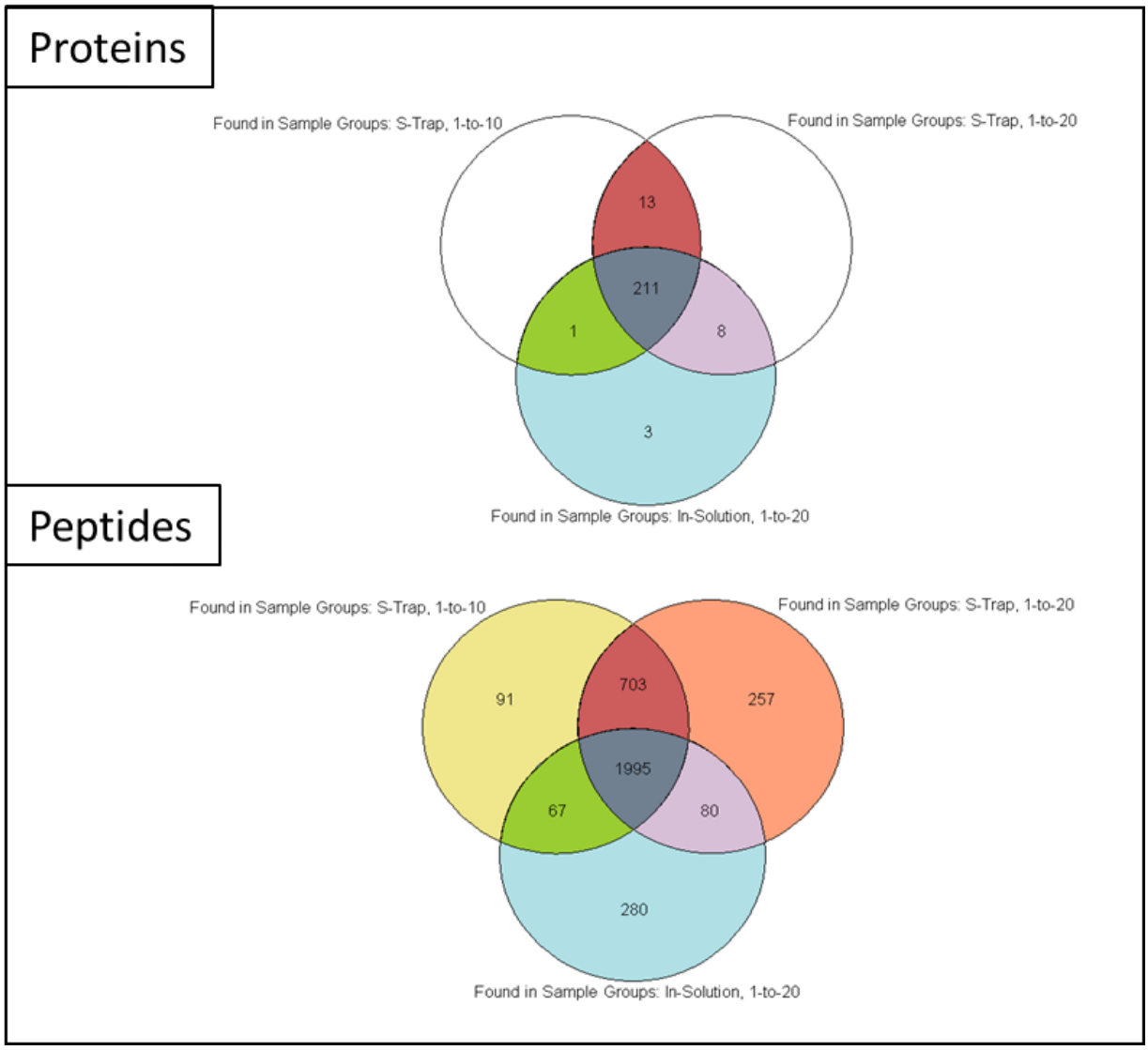


Figure 3.6. Venn diagram comparison of protein and peptide identifications in culture supernatant using in-solution versus S-trap digestion protocols. Results represent duplicate digestions under the following conditions: S-trap 1:10 (enzyme-to-substrate), S-trap 1:20 and in-solution 1:20. S-trap digestion performed with a 1:20 enzyme-to substrate ratio was observed to maximize the number of peptide and protein identifications.

Given that most antigen components are purified from the culture supernatant prior to formulation in the acellular vaccine, one would expect that all the antigen proteins would be detectable via DDA of the starting material and would be of relatively high abundance. While FHA, PT and PRN were confidently detected amongst the most abundant proteins in the sample, coverage of FIM was extremely limited with only one peptide from FIM 2 and no peptides from FIM 3 despite the use of acid-labile surfactant to promote solubility.[165] Though poor coverage of FIM can likely be attributed to its hydrophobicity and propensity for aggregation, its limited sequence coverage highlighted a potential gap in the current sample preparation protocol.

Moving to the S-trap digestion for unpurified culture supernatant improved FIM coverage to 2 peptides for FIM 2 and 5 peptides for FIM 3, corresponding to 22 % and 45% sequence coverage, respectively. Applying a detection limit of two peptides per protein, the S-trap digest protocol enabled detection of a protein that was previously below the detection threshold. Thus, implementation of the S-trap technology improved coverage from the in-solution digest profile.

As an extension of the improved detection of FIM in the unpurified culture supernatant, the S-trap digestion protocol was evaluated for purified FIM antigen as a representative hydrophobic, aggregation-prone protein sample. Leveraging biophysical characterization on the purified antigen, the temperature during the reduction step was increased from 60 to 95°C to promote the disassembly of aggregate fibril structures. To further improve solubility, FIM digestion samples were also sonicated for 10 minutes prior to digestion either in ABC buffer with 0.1% Rapigest (in-solution digest), or with 5% SDS/100 mM TEAB as per the S-trap protocol.

A side-by-side comparison of 500 fmol loaded on-column (Table 3.1) showed a 16-23% increase in sequence coverage using S-trap versus in-solution digestion protocols. The TIC traces from S-trap FIM digests were noted to have more intense peaks and fewer late-eluting contaminants than the in-solution digests, consistent with observations from the unpurified culture supernatants. Increasing the on-column load to 1 pmol of in-solution digest was necessary to obtain comparable sequence coverage to the S-trap digests for FIM 2 and FIM 3 (Table 3.1). While there were no HCPs detected with 500 fmol of in-solution digest, 3 HCPs were detected using 500 fmol of S-trap and 1 pmol of in-solution digest. The enhanced detection of HCPs is in line with increased depth of coverage using the S-trap digestion protocol.

Table 3.1. Sequence coverage results from DDA analysis of purified FIM antigen using S-trap and in-solution digestion protocols

Protein	Amount of protein loaded on-column		
	500 fmol		1 pmol
	S-trap	In-Solution	In-Solution
Fim 2	72 %	56 %	71 %
Fim 3	74 %	51 %	84 %

3.4 Conclusion

Shotgun proteomic analyses were successfully performed for the characterization of culture supernatant and purified FIM antigen from *B. pertussis*. Implementation of the S-trap digestion protocol improved proteome coverage and enabled confident detection of hydrophobic FIM antigen in the unpurified process intermediate. Effective sample cleanup and improved sequence coverage were also observed in purified FIM samples. Collectively, these results highlight the power of the S-trap technology for hydrophobic proteins and establish proof-of-concept for further use in vaccine characterization.

Chapter 4 – Development of a targeted nanoLC-MS/MS method for absolute quantitation of residual toxins from *Bordetella pertussis*

4.1 – Introduction

In contrast to recombinant vaccines, acellular Pertussis (aP) vaccines are purified from crude fermenter harvest (*B. pertussis* culture supernatant). Of the virulence factors described in Chapter 3, PT is the only toxin that has been selected along with adhesion proteins FHA, PRN and FIM as purified components in acellular vaccines.[159] Consequently, there is a need to monitor for co-purification of other toxins to ensure safety and material consistency.[168] Though traditionally monitored using enzymatic assays, animal tests, or cell-based approaches such as clustering in Chinese hamster ovary (CHO) cells[158], the residual toxins from *B. pertussis* represent a unique opportunity to apply the superior specificity and multiplexing capabilities of targeted MS to the development of next generation vaccines. In contrast to conventional shotgun analysis described in the Chapter 3, targeted MS workflows, such as parallel reaction monitoring (PRM), harness the analytical power of MS to target a list of precursors for quantitation, with limits of detection in the attamole (atm, 10^{-18}) range on-column.[88,91]

4.1.1 –Residual Toxin Background

Adenylate cyclase toxin (ACT) is 177 kDa immunogenic protein belonging to the Repeats-in-Toxin (RTX) family of bacterial toxins. Like other toxins in RTX family, ACT contains a C-terminal glycine rich domain which binds calcium enabling translocation of the N-terminal catalytic domain into target cells.[169] ACT expression is regulated by the same bvgAS operon as other *B.pertussis* virulence factors, PT, PRN, FHA and FIM.[170] Secreted from the bacterial cytosol into the extracellular medium, ACT binds to complement receptor 3 (CR3) and integrin complex (CD11b/CD18) on the surface of host cells.[170,171] Once bound, ACT translocates its

catalytic domain to the cytoplasm where it binds calmodulin resulting in conversion of ATP to cAMP. Accumulation of intracellular cAMP induces cellular apoptosis and notable toxicity.[172,173] Due to potentially adverse effects, the regulatory limit for ACT is defined as 500 ng per dose of aP vaccine as per the European Pharmacopeia (Ph. Eur) [168] and WHO guidance[174](Table 4.1).

Table 4.1 European regulatory limits for residual toxins ACT, DNT and TCT per dose of acellular Pertussis vaccine.[168]

Toxin	Molecule Type	Size	Limit in grams per dose	Limit in moles per dose
ACT	Protein	177 kDa	500 ng	2.8 pmol
DNT	Protein	160 kDa	Absence	N/Ap
TCT	Glycopeptide (small molecule)	921 Da	1.84 ng	2.0 pmol

Dermonecrotic toxin (DNT) is a 160 kDa protein toxin produced by all bacteria in the genus *Bordetella.* As the name suggests, it's notable trait is characteristic necrosis of the skin when injected into mice intradermally.[175] DNT's role during *B. pertussis* infection in humans is unknown; however, it is noted to lack an export sequence. Therefore, the presence of DNT in early stages of manufacturing is likely due to bacterial cell lysis during fermentation.[176,177]

In contrast to ACT and DNT, tracheal cytotoxin (TCT) is a 921 Da glycopeptide released from the peptidoglycan layer of the *B. pertussis* cell wall during logarithmic-phase of bacterial growth. Though most gram-negative bacteria recycle this molecule, *B.pertussis* does so poorly instead releasing TCT to the extracellular environment where it binds to ciliated epithelial cells in the respiratory tract.[178] TCT bears a unique structure which was first reported as GlcNAc1,6-anhydro-MurNAc-L-Ala-y-D-Glu-meso-diaminopimelic acid -D-Ala (Figure 4.1) by fast-atom bombardment (FAB)-MS. Its MurNAc (N-acetylmuramic acid) moiety places it in a family of

compounds of bacterial origin known as muramyl peptides.[179,180] As per Ph. Eur regulatory guidance[168], the TCT limit is 2 pmol per dose of aP vaccine (Table 4.1).

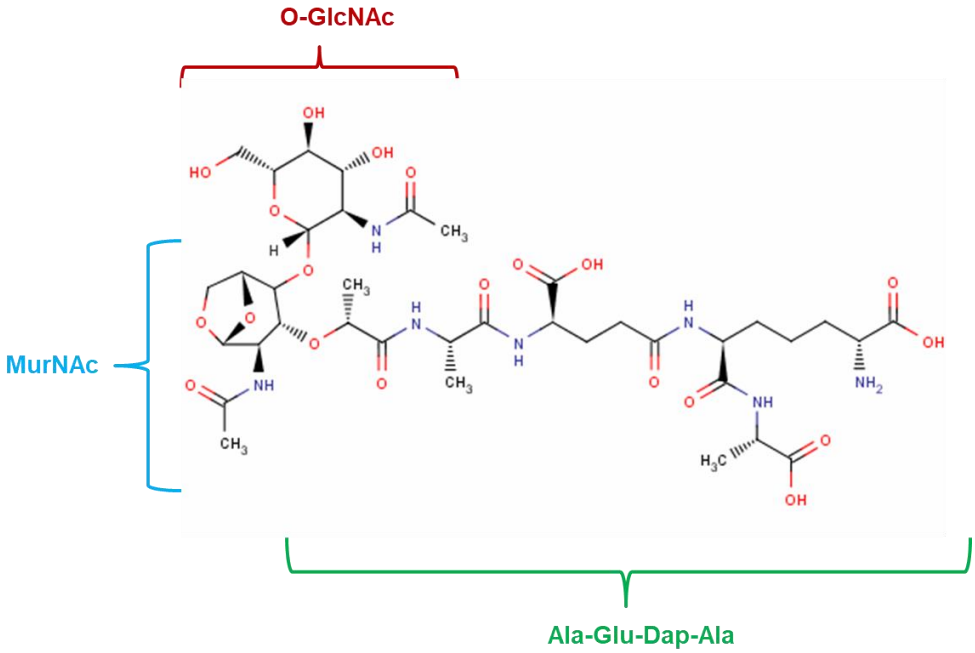


Figure 4.1. Structure of TCT, a polymeric peptidoglycan fragment released by *Bordetella pertussis*. Full name: GlcNAc1,6-anhydro-MurNAc-L-Ala-y-D-Glu-meso- diaminopimelic acid -D-Ala. Composed of disaccharide glycan moiety N-acetyl glucosaminyl (GlcNAc)-N-acetylmuramic acid (MurNAc) and a stem of amino acids. Toxicity linked to side-chain functional groups and unnatural diaminopimelic acid (Dap) residue.

4.2 – Experimental Methods

4.2.1 Design of a Targeted MS Method

Peptide PRM experiments were designed and refined in Skyline version 4.1.[181] An *in-silico* tryptic digestion was performed using the fasta sequences for ACT (P0DKX7) and DNT (Q7VTS2). Sequences were filtered to be between 8 to 16 residues in length, fully tryptic, and lack homology with any other proteins in the *B.pertussis* proteome. Easily modifiable residues (Cys, His, Met) and ragged ends (Arg/Lys repeats) were also avoided.

4.2.2 Peptide Standards:

ACT and DNT sequences identified from the in silico tryptic digestion in Skyline were ordered as crude native peptides (SpikeTides™) from JPT Peptide Technologies (GmbH, Germany). Selected "top 10" sequences from screening experiments were synthesized using JPT's spot synthesis technology[182] as crude heavy peptides with a ^{15}N, ^{13}C labelled terminal lysine or arginine (SpikeTides_L™). "Double heavy" absolutely quantified heavy labelled proteotypic peptides were custom ordered from JPT Peptides with a ^{15}N, ^{13}C labelled terminal Arg or Lys and one other heavy labelled residue per sequence (Spiketides_TQL™, ≥99% isotopic purity). Referred to as 'Qtag peptides', the Spiktides_TQL™ standards contain a UV-visible C-terminal tag (Qtag) that is cleaved by trypsin during sample digestion, thereby releasing the stable isotope labelled peptide standard for absolute quantitation. Following final standard refinement, two peptides per protein were purchased as single heavy (^{15}N, ^{13}C) AQUA™ Ultimate peptides (Thermo Fisher Rockford, IL, USA). Starting concentrations were determined by amino acid analysis (AAA) for AQUA™ Ultimate peptides (5 pmol/µL, 5% ACN in H_20) and by UV for the Qtag peptides standards.

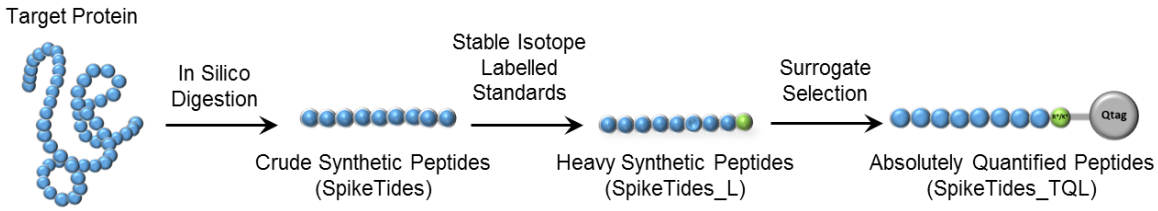


Figure 4.2. Standard-driven strategy for targeted nanoLC-MS/MS development. Outline of iterations performed from the silico tryptic digests of target proteins (ACT/DNT) through selection of surrogate peptide sequences for absolute quantitation.

4.2.3 Sample Preparation:

Heavy labelled prototypic (Qtag) peptides were resuspended in 1 M ABC 20% ACN and pooled into stock solutions containing 375 fmol/µL ACT and 75.0 fmol/µL DNT peptide standards, with 3.75 pmol/µL of a carrier peptide for stability. 15 µL of 1 M ABC (pH 7.4-7.5), 15 µL of 1% Rapigest (v/v), and 10 µL of peptide stock solution were added to 150 µg of culture supernatant determined by Bradford assay or 25 µg of purified antigen drug substance determined by UV280. Each sample was digested in triplicate. For reduction of cysteines, 3 µL of 500 mM TCEP solution was transferred to samples and incubated at 60°C for 30 minutes. For alkylation, 3 µL of 1M IAA was then added and vortexed to mix. The samples were incubated in a dark drawer for 60 minutes at room temperature. The digest solutions were brought to 150 µL total volume with LC-MS water such that the concentration of heavy Qtag standards was 25 fmol/µL and 5 fmol/µL for ACT and DNT, respectively.

Sample digestion was carried out at an enzyme-to-substrate ratio of 1:20 (w/w) for 4 hours at 37°C using a thermomixer incubator set to 300 rpm for mixing. To optimize digestion kinetics, a timecourse was also performed at 37 °C with incubation times varied from 1 to 6 hours. Digests were quenched by the addition of 4 µL of TFA and confirmed to have a pH of ~2 using a pH strip. The acidified digest samples incubated at 37°C for 30 minutes followed by 10 minutes of centrifugation at 8,100 x g to facilitate Rapigest hydrolysis. The supernatants were transferred to new tubes and stored at -80 °C until LC-MS analysis. Prior to nanoLC-MS/MS analysis, digest samples were thawed at room temperature and spiked with AQUA heavy peptides to a final concentration of 25 fmol/µL ACT and 5 fmol/µL DNT standards.

4.2.3 nanoLC-PRM Data Acquisition

Targeted parallel reaction monitoring (PRM) experiments were acquired on a Q Exactive HF mass spectrometer with an UltiMate 3000 RSLCnano UHPLC system (Thermo Scientific). Chromatographic separation was performed using a trap-elute nanoLC configuration with pre-concentration using a 5 mm x 300 µm (i.d.) C18 trap column. 1 pmol of purified antigen or 1 µg of supernatant digest was loaded onto the trap column via loading pumps run at 5 µL/min After 2 minutes of on-line desalting, the peptides were eluted onto a PepMap C18 75 µm x 15 cm Easy-Spray analytical column with integrated ESI emitter. A flow rate of 300 nL/min was used with aqueous mobile phase A and organic (ACN) mobile phase B. 0.1% TFA and 0.1% FA were used as ion-pairing reagents during the trap and elute phases, respectively. A 95-minute gradient from 1 to 35% B was used with a total run time of 117 minutes with the column oven (trap) temperature was set to 40°C and analytical column set to 30°C.

The PRM acquisition method cycled between a Full MS survey scan and up to 15 MS/MS scans with a 1.4 m/z isolation windows. Data acquisition was scheduled for ±5 minutes the observed retention time for each target peptide in a scouting run performed at the beginning of the MS acquisition sequence. Survey scans were acquired from 380-2000 m/z with a resolution of 120,000, automatic gain control (AGC) of 3×10^6, and maximum injection time (maxIT) of 50 ms. MS/MS fragmentation was performed using a normalized collision energy (NCE) of 27, maxIT of 300 ms, AGC of 2×10^5, and resolution of 120,000. Instrument tune parameters of note: capillary voltage of 1.9 kV, capillary temperature of 275°C and S lens RF voltage set to 50.0.

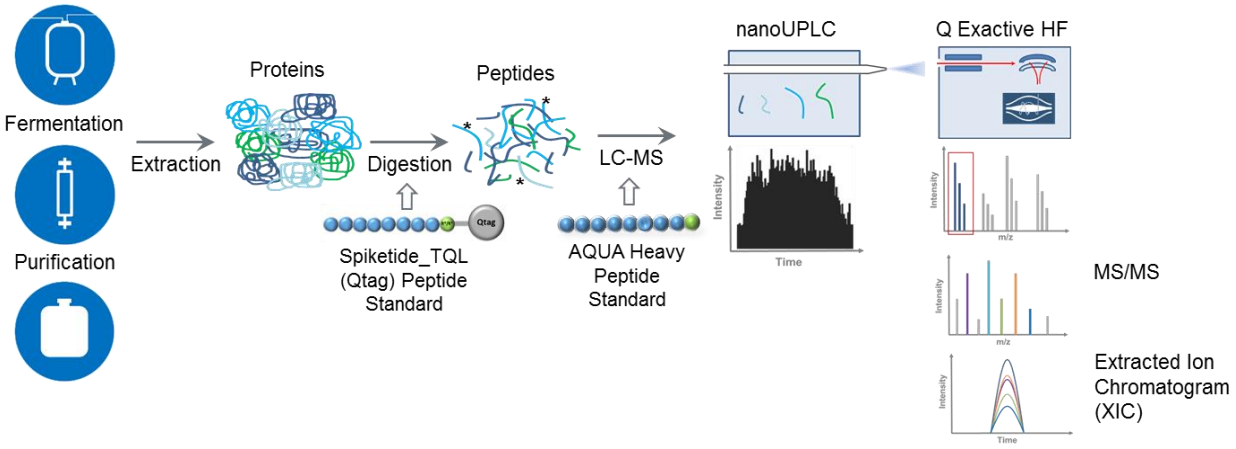


Figure 4.3. Schematic of a targeted nanoLC-PRM for absolute quantitation of residual toxins from *Bordetella pertussis*. LC-MS/MS workflow includes protein extraction, digestion, separation by nanoUPLC and scheduled parallel reaction monitoring (PRM) on a Q Exactive HF. Method can be applied across manufacturing stages for acellular Pertussis vaccines.

PRM results were imported into Skyline (version 4.1)[181] for target extracted ion chromatogram (XIC) analysis. 3-5 shared transitions were selected per peptide based on fragment ion intensity, visual determination of signal-to-noise, and co-elution of native and heavy peptide signals. The ratio of native-to-heavy peptide signal was calculated for each peptide by summing the integrated peak areas for the selected transitions. Concentrations were determined from the native-to-heavy ratio given the known concentration of heavy peptide standards spiked into the samples. Mean and standard deviations were calculated from digest triplicates acquired individually. In the optimized PRM method, one peptide sequence per protein was used for absolute quantitation and a second for confirmation of protein identity (i.e. specificity).

4.2.5. Evaluation of linearity, limit of detection, and limit of quantitation

Applying the optimized PRM method, Spiketides_TQL (Qtag) and AQUA heavy peptide standards were used via a spike-in experiment to evaluate the linearity, limit of detection (LOD) and lower limit of quantification (LLOQ). Qtag peptides containing two labelled residues (double heavy) were spiked into unpurified culture supernatant samples prior to digestion and used to normalize for run-to-run variability in absolute MS response. Ultimate AQUA (single heavy) peptides were spiked post-digestion with peptide response evaluated from 12.5 amol to 250 fmol on-column (target loads of 0.0125, 0.025, 0.0375, 0.05, 0.075, 0.100, 0.250, 2.5 and 250 fmol on-column, respectively). Samples were analyzed in triplicate in unpurified culture supernatant. Linear regression analysis was performed in Excel considering the natural logarithm of Qtag-normalized areas as a function of AQUA peptide concentration.

4.3 Results

4.3.1 Peptide Screening Results

In the absence of purified ACT and DNT reference material, a peptide standard-driven strategy was employed to identify candidate sequences to serve as surrogate peptides for targeted nanoLC-MS/MS analysis. An in silico tryptic digest in Skyline yielded 33 ACT and 32 DNT candidate sequences with physiochemical properties conducive to quantitation. These sequences were iteratively refined in 3 rounds of optimization experiments (Figure 4.4). First, crude native peptide sequences were purchased from JPT peptides and spiked into a purified antigen digest. A Full MS PRM method was used to gauge the relative signal intensities of peptide candidates at both the survey (MS1) and fragment ion (MS2) levels. Variation in signal intensity between peptides was noted, reflecting both differences ionization efficiency and relative ease of synthesis (Figure 4.4a).

MS/MS spectra acquired via PRM of the spiked native peptides were searched against the *B. pertussis* proteome using Proteome Discover version 2.1. In all, 31/33 ACT and 25/32 native DNT candidates were positively identified, confirming that the MS/MS spectra were of sufficient quality for unambiguous assignment by a database search. Sequences which were not positively identified were removed from consideration owing to poor specificity. An additional 2 sequences from ACT and 3 from DNT were removed from consideration as they showed a persistent carryover between runs on the nanoLC-MS system. From these initial screening results, the 10 best sequences per protein were selected for further evaluation.

Next, selected "top 10" peptides from ACT and DNT were ordered as crude, heavy labelled peptides and spiked into clarified harvest from fermented *B. pertussis* that was thought to contain endogenous ACT, DNT and TCT (Figure 4.4b). Based on co-elution of native and crude heavy transitions by PRM, 9/10 ACT sequences were positively detected in the fermenter harvest. In contrast, only 5/10 DNT sequences were detected in the unpurified sample, owing to substantially lower DNT levels in samples from *B. pertussis*. Detected ACT and DNT sequences were ranked based on native signal intensity with the top 5 sequences per protein chosen for further consideration as quantitative standards (Figure 4.4c).

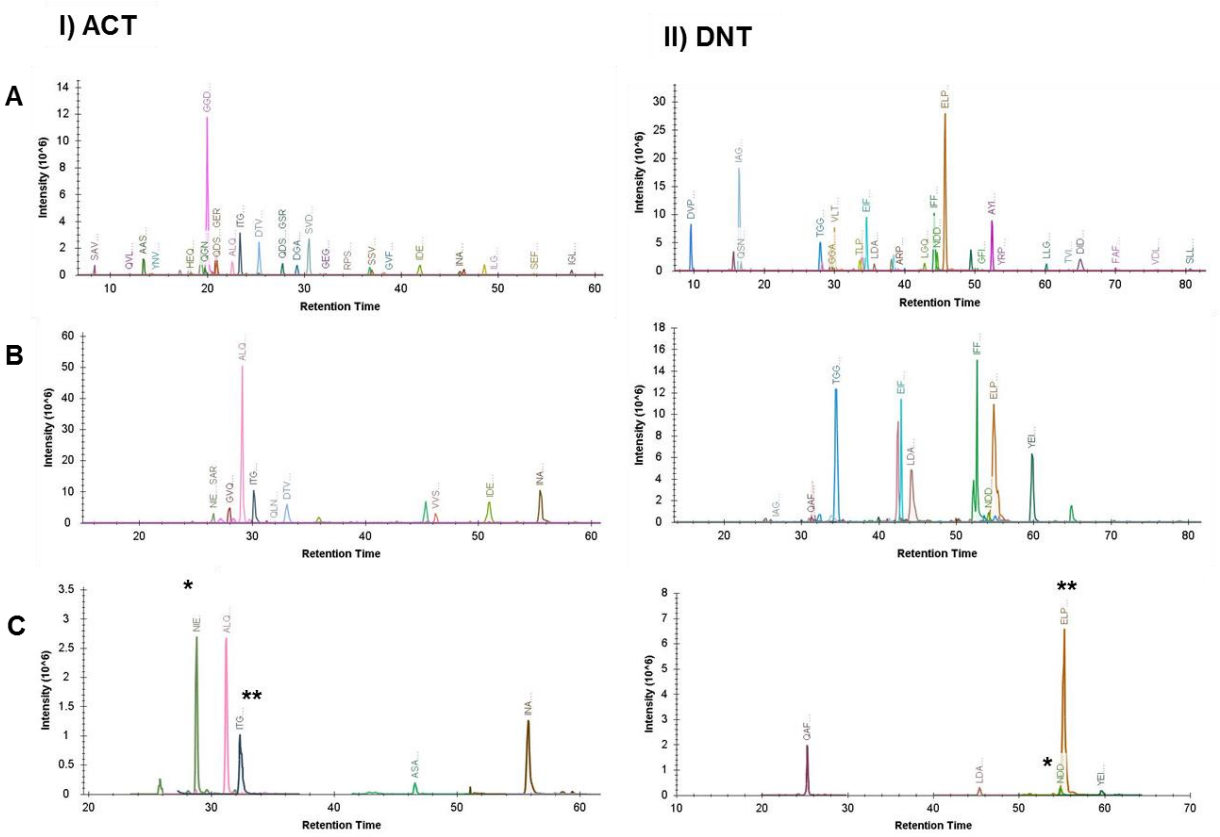


Figure 4.4. PRM traces for ACT and DNT synthetic peptides across method optimization experiments. A) Chromatogram traces for 33 ACT peptides and 32 DNT peptides screened as crude native peptides. B) Traces from second round of optimization with 10 crude heavy labelled peptides for ACT and DNT spiked into unpurified fermenter harvest. C) Final optimized traces for ACT and DNT with one peptide for one quantitation (denoted by *) and a 2nd sequence for confirmation/specificity (denoted by **).

4.3.2 TCT Detection

Detection of TCT in unpurified fermenter harvest from *B. pertussis* was evaluated by adding the TCT 1+ and TCT 2+ precursor masses (922.3839 & 461.6985 m/z) to the PRM inclusion list used for targeted analysis of ACT and DNT. TCT was positively identified by co-elution of the 1+ and 2+ precursor masses (Figure 4.5). MS/MS analysis by HCD (Figure 4.5b) generated fragment ions matching those previously observed by fast atom bombardment (FAB) MS of a TCT homologue in *Neisseria gonorrhoeae*.[183] A signature fragment series from the O-GlcNAc moiety was also detected with ions at 204, 186, 168, 144, 138 and 126 m/z.[184]

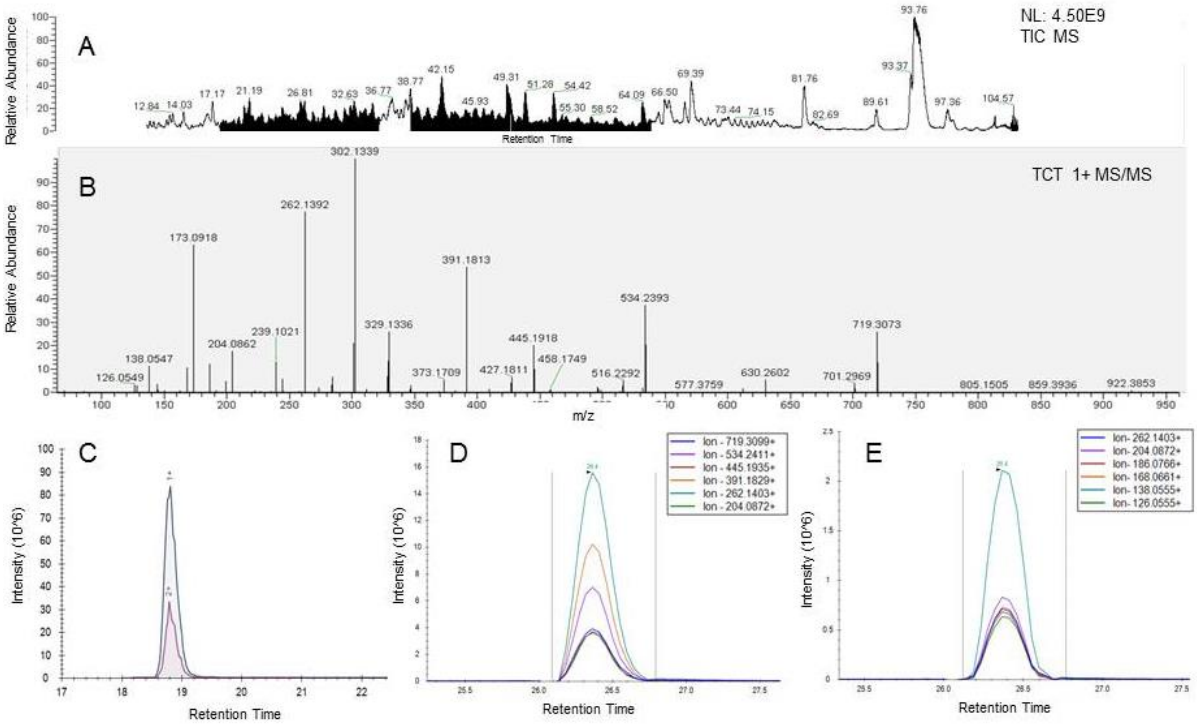


Figure 4.5. TCT detection in unpurified fermenter harvest from *Bordetella pertussis.*
A) Total ion chromatogram from unpurified *B. pertussis* fermenter harvest. B) MS/MS fragment ion spectrum from singly charged TCT precursor. C) Co-elution of intact TCT 1+ (922.3893 m/z) and TCT 2+ precursor ions (461.6985 m/z). D) Extracted ion chromatograms for TCT 1+ transitions favouring fragmentation of the peptide backbone. E) Extracted ion chromatograms for TCT 2+ transitions favouring O-GlcNAc (glycan) fragmentation.

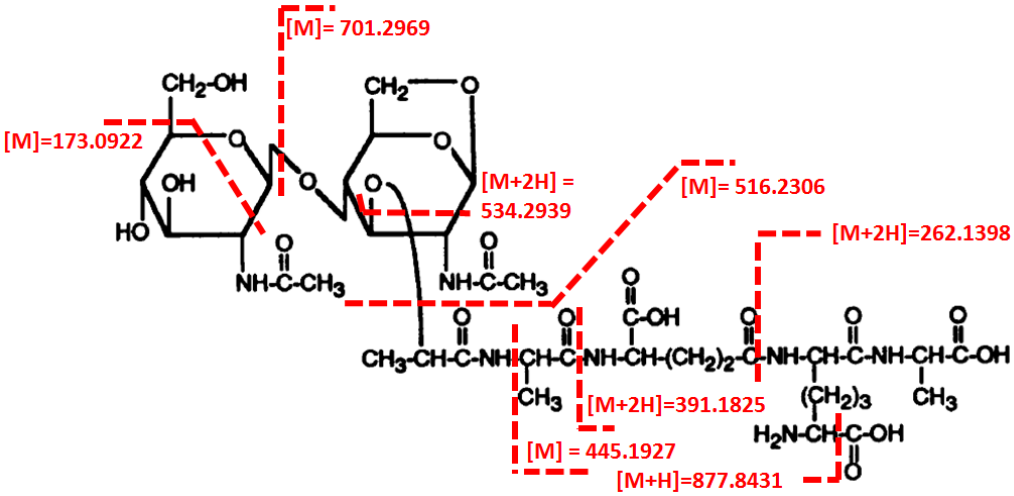


Figure 4.6. Non-GlcNAc fragments mapped onto TCT structure. Peptide fragments correspond to y-ion equivalents (262.1398 & 391.1825 m/z), loss of GlcNAc and oxygen (701.2969 m/z), and product of glycan cleavage (516.2306 m/z).

4.3.3 Targeted Method Refinement

An underlying assumption in targeted proteomics is that the level of measured peptides is a stoichiometric representation of protein levels. After narrowing down to 5 proteotypic sequences each for ACT and DNT, this assumption was tested by assessing the quantotypic properties of the surrogate sequences. Given the absence of readily available standard, and a desire to avoid handling of purified toxins, a peptide-driven strategy with absolutely quantified heavy peptide standards was employed in a third round of optimization. Using lower cost Spiketides_TQL (Qtag) standards from JPT Peptides, we were able to retain more sequences in quantitative evaluations than what would be feasible give then cost of traditional HPLC-purified AQUA peptides. Double heavy Qtag standards were spiked prior to digestion, enabling differentiation from single heavy AQUA peptide standards spiked after digestion. By requiring enzymatic cleavage to release the double heavy surrogate sequence, the Qtag peptides acted as both digestion controls and internal standards for absolute quantitation.

Additional adjustments were made to the Full MS PRM acquisition method to facilitate improved detection of targeted ACT, DNT and TCT precursors. First, the number of target peptides from ACT and DNT was refined to 5 per protein, and acquisition scheduled with 10 minute windows. This reduced the number of concurrent precursors to a maximum of 10-12 at any given point in the acquisition method. From the savings in duty cycle, it was possible to increase the maximum injection time (maxIT) from 100 ms to 300 ms per precursor per target, while maintaining collection of 8-10 points across an eluting peak for quantitation. In addition to boosting low-level detection, using longer accumulation times also justified increasing the PRM resolution from 30, 000 to 120,000, despite the slower Orbitrap scan rate. The higher resolution setting makes use of the extended ion accumulation time and provides higher selectively for extracted fragment ion chromatogram (XIC) workup.

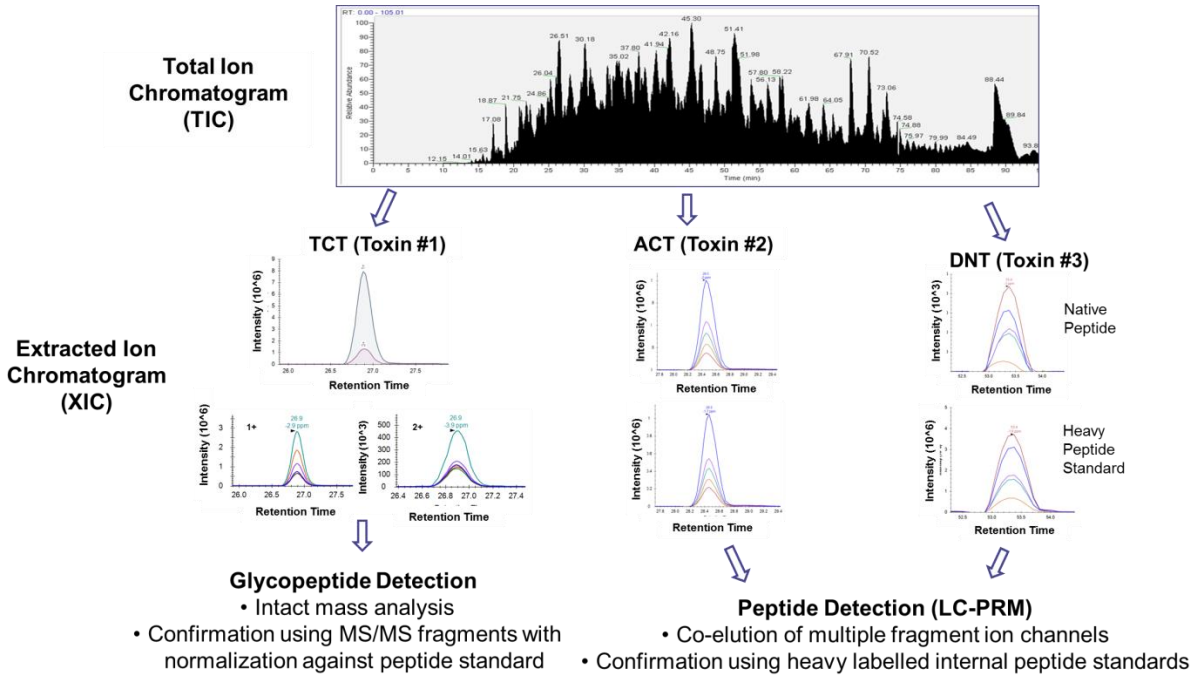


Figure 4.7. Extracted ion chromatogram traces for residual toxins ACT, DNT and TCT detected in-process culture supernatant by targeted nanoLC-PRM analysis. Absolute quantitation enabled by spiking matching heavy surrogate peptide standards for ACT and DNT.

4.3.4 Digestion Timecourse Analysis

A digestion timecourse study was used to identify the best performing tryptic peptides for absolute quantitation. Doubly heavy Qtag peptide standards were spiked into fermenter harvest and analyzed using the refined Full MS PRM method. Given variability in absolute response owing to system delivery, the ratios of native-to-heavy (Qtag) peptide standards were monitored to assess completeness of proteolytic digestions and obtain a quantitative readout from each surrogate sequence. For ACT, two of the sequences, ASALGVDYYDNVR and ALQGAQAVAAAQR, showed fast digestion kinetics with stable native-to-heavy ratios. However, their quantitative readouts appeared to be outliers relative to the other ACT sequences (Appendix 2). One sequence, NIENAVGSAR, showed a gradual increase in signal, stabilizing after 4 hours digestion time, whereas the signal for ITGDAQANVLR and INAGADQLWFAR dropped in intensity over time. From these results, NIENAVGSAR displayed the best kinetics and stability profile for absolute quantitation. ITGDAQANVLR could be used for confirmation and/or specificity (Figure 4.8). Absolute native-to-heavy response for the "top 5" ACT sequences and results from a reproducibility assessment are provided in Appendix 2.

For DNT, 2/5 proteolytic sequences detected contained N-terminal Glutamine and Glutamic Acid. Synthesis of Qtag peptide standards was deemed to be unfeasible for ELPALIGASGLR and QAFSAASSSAG due to the harsh conditions used to cleave peptides off the cellulose membrane in the SPOT synthesis method.[182] However, shotgun (DDA) analysis of the native peptides spiked post-digestion showed a low relative abundance of pyroglutamic acid modifications (< 5%). This meant the sequences were still viable for use as quantotypic standards so long as the pyroglutamic acid modification was adequately controlled.

Due to the low endogenous level of DNT, there were very limited options for sequences with detectable native DNT signal. Therefore, ELPALIGASGLR and QAFSAASSSAG were still monitored in the digestion timecourse, normalizing against a consistent concentration of crude heavy peptides.

Spiking the Qtag heavy standards at 5 fmol/µL, only 3/5 DNT sequences were confidently detected across the digestion time course based on co-elution of 2+ native peptide transitions. One DNT sequence, QAFSAASSSAGR, was observed to drop in signal intensity with longer digestion times. Digestion of ELPALIGASGLR and NDDLVSIAATYDR were noted to proceed with slower kinetics compare to the more abundant ACT sequences, with a gradual rise in signal intensity through 4 hr and 6 hr, respectively. The native-to-heavy ratio for ELPALIGASGLR was noted to be more variable across digest triplicates. Therefore, NDDLVSIAATYDR was chosen for absolute quantitation, owing to more abundant native signal intensity from endogenous DNT and less propensity for modification. ELPALIGASGLR was kept as a secondary sequence for confirmation (Figure 4.8).

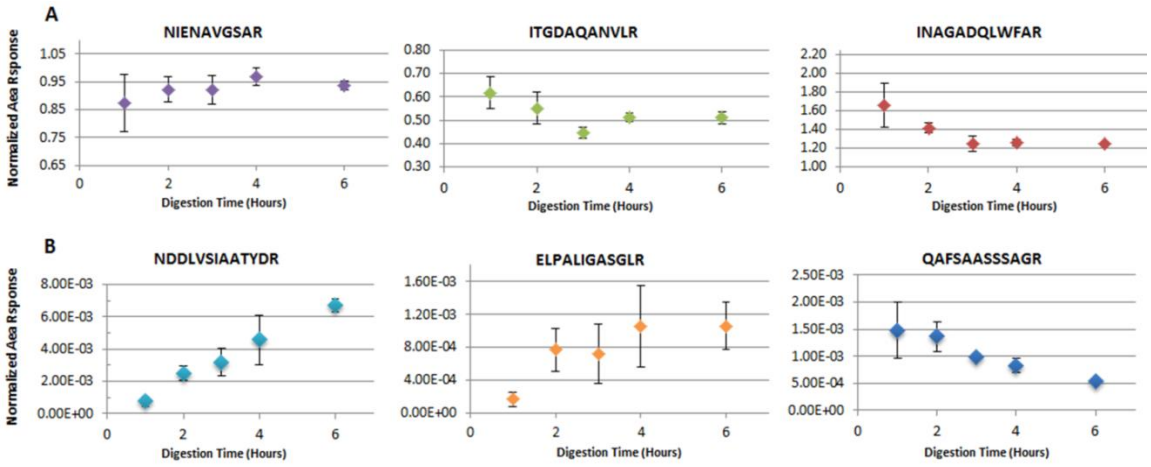


Figure 4.8. Digestion timecourse results for ACT and DNT surrogate peptides normalized against spiked heavy peptide standards. Error bars reflect +/-standard deviation from n=3 digest replicates per timepoint.

TCT signal response was monitored over the digestion timecourse to evaluate stability and confirm recovery of the small molecule glycopeptide for multiplexed analysis with protein toxins, ACT and DNT. TCT response was noted to be within variability from triplicate digests, indicating that TCT is indeed stable during sample preparation (Figure 4.9). The ratio of singly to double charged precursor was also observed to be consistent, with the singly charged precursors (TCT 1+) being approximately 3-4 fold more intense than the doubly charge TCT precursor (TCT 2+).

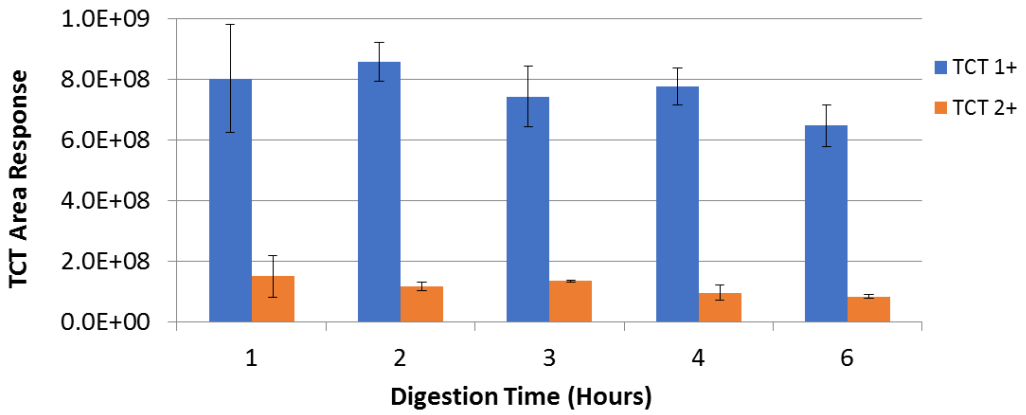


Figure 4.9. Digestion timecourse results for TCT in unpurified culture supernatant. TCT signal was response was confirmed to be within assay variability throughout digestion. Error bars reflect ± standard deviation from n=3 replicate measurements per timepoint.

4.3.5 Method Linearity & Limit of Quantitation

Following refinement of surrogate peptides to two sequences per protein, HPLC-purified Ultimate AQUA peptides were ordered for the quantitation. Contrary to the Qtag heavy peptide standards which were synthesized with two heavy labelled residues, AQUA peptides were designed to contain just one heavy, C-terminal residue per sequence. This two-tiered standard approach enabled assessment of several key performance metrics including linearity, dynamic range, limit of detection (LOD), lower limit of quantitation (LLOQ).

Linearity of response for ACT and DNT peptides was evaluated by spiking heavy labelled AQUA peptides into unpurified culture supernatant digests. Matching "double heavy" Qtag standards were used to normalized absolute response observed from the "single heavy" AQUA peptides over a range from 12.5 amol to 250 fmol on-column. A natural logarithm (ln) – ln plot of Qtag-normalized AQUA peptide peak area to AQUA peptide concentration was used to correct for heteroscedasticity across the dynamic range (i.e. unequal variability across a range of values). NIENAVGSAR from ACT was observed to be linear down to 25 amol on-column with a correlation coefficient (R^2) > 0.9987. The limit of detection (LOD) was evaluated based on co-elution of 2 or more transitions across triplicate digests considering a baseline intensity of at least 500 ion counts (as recommended by the vendor, Thermo Scientific). Using NIENAVGSAR as a surrogate peptide, the LOD for ACT was measured at 12.5 amol on-column, corresponding to 2 ng/mL (Figure 4.10).

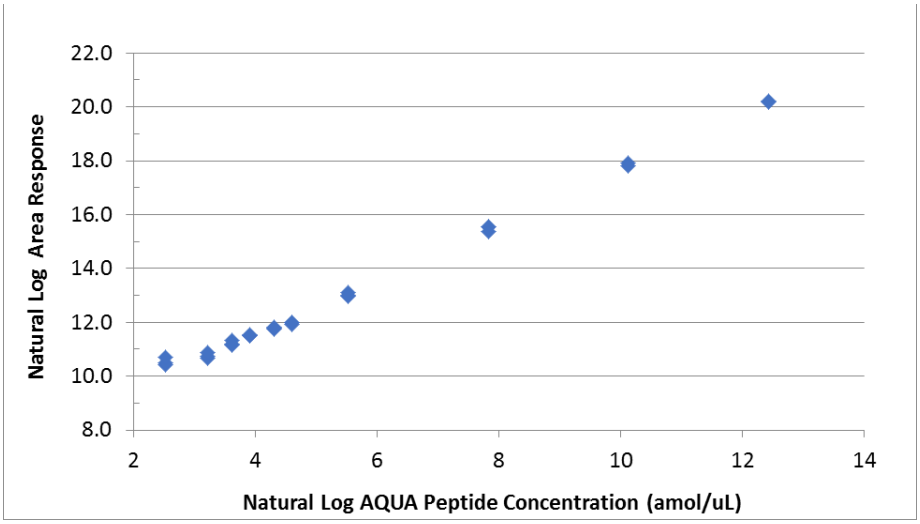


Figure 4.10. PRM response for 12.5 amol to 250 fmol of ACT AQUA peptide standard (NIENAVGSAR). Triplicates analyses plotted for each on-column concentration tested. Peptide signal response observed plateau below 25 amol on-column, indicative of loss in linearity.

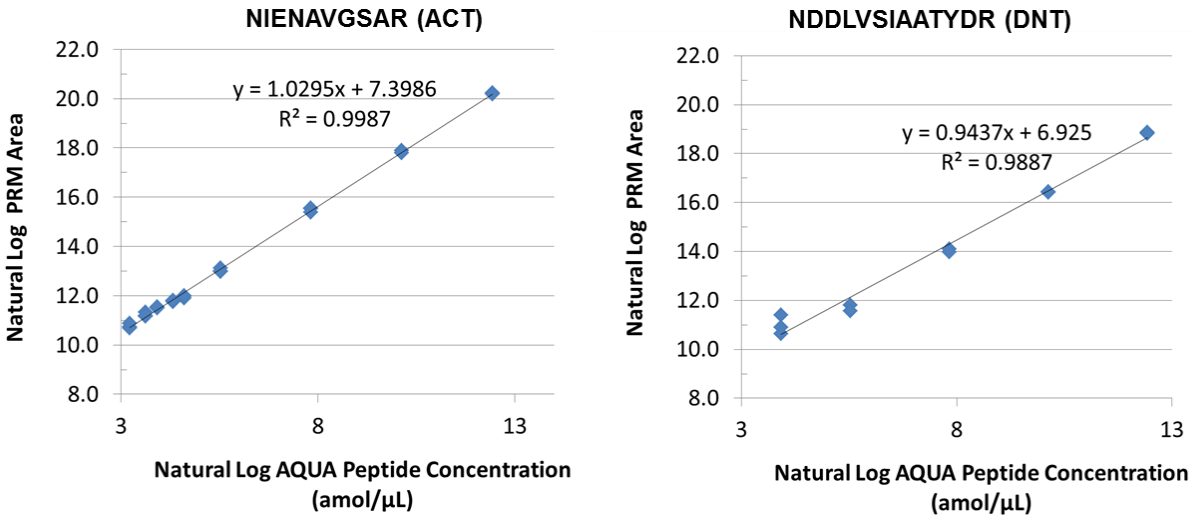


Figure 4.11. Linearity analysis for optimized ACT and DNT AQUA peptides standards. Natural logarithm of PRM peak areas from AQUA peptides fit to linear regression equation for assessment of lower limit of quantitation (LLOQ). Linearity confirmed from 25 amol to 250 fmol on-column for NIENAVGSAR (ACT) and 50amol to 250 fmol on-column for NDDLVSIAATYDR (DNT). Peptides demonstrate linearity over 5 orders of magnitude dynamic range. Triplicates analyses plotted for each on-column concentration tested.

While % CV values for NIENAVGSAR (ACT) were typically less than 5% at the highest dilutions, the % CV values for the most dilute concentrations were between 12 and 21 %. This is within acceptable limits of variability for a targeted MS assay.[185] Therefore the lower limit of quantitation (LLOQ) was set at 25 amol on-column (Figure 4.11).

Comparing the absolute response at equal amounts AQUA peptides, the optimized ACT sequences were noted to have approximately 10-fold higher intensity response than the selected DNT sequences. Lower absolute peptide response translated into higher limits for DNT in the targeted nanoLC-PRM method. The sequence proposed for quantitation, NDDLVSIAATYDR, showed linearity down to 50 amol on-column (R^2 = 0.9887) with lower detection limit at 25 amol on-column (Figure 4.11). Variability was notably higher than for ACT with % CV values at the exceeding 65% at 25 amol on-column. Considering the range in which the signal was

linearity and variability observed at the lowest concentrations tested, the LLOQ for DNT was set at 50 amol on-column, equivalent to 4 ng/mL (Figure 4.11). Note - in-solution limits for ACT and DNT differ amongst sample types based on starting concentrations and dilutions performed prior to LC-MS analysis.

4.3.6 Implementation of the optimized PRM method for simultaneous monitoring of ACT, DNT and TCT in samples from *Bordetella pertussis*

The optimized nanoLC-PRM method was implemented as an in vitro method for measuring toxin levels across culture supernatant (in-process) and final bulk product samples. Using NIENAVGSAR for absolute quantitation and ITGDAQANVLR for confirmation, ACT was measured in lots of unpurified culture supernatant at 4.0 and 11.5 µg/mL with % CV of less than 5%. Following purification, ACT was measured in bulk antigen sample at 3.4 ng/mL and 1.4 ng/mL (% CV 34-37%). DNT was measured at 7.9-10.4 ng/mL in culture supernatant (% CV 20-34 %) but was not detected in samples post-purification (Tables 4.2-4.3). TCT levels were confirmed to decrease during antigen purification; however, absolute quantitation was not possible to the lack of an internal reference standard. Reproducibility was confirmed via assessment digests replicates (% CV < 10%) and and day-to-day variability (% CV < 25%) for a given peptide sequence.

Table 4.2. Quantitative results for *B. pertussis* toxins in culture supernatant (unpurified fermenter harvest) using optimized ACT and DNT Qtag peptide standards.

Toxin	Culture Supernatant			
	Lot 1		Lot 2	
	Concentration	% CV	Concentration	% CV
ACT	4.0 µg/mL	3.9	11.5 µg/mL	4.7
DNT	7.9 ng/mL	34.2	10.4 ng/mL	20.4
TCT (Normalized Response)	8.74+08	8.6	7.94E+08	4.4

Table 4.3. Quantitative results for *B. pertussis* toxins in purified antigen using optimized ACT and DNT Qtag peptide standards.

Toxin	Purified Antigen			
	Lot 1		Lot 2	
	Concentration	% CV	Concentration	% CV
ACT	3.4 ng/mL	34.0	1.4 ng/mL	37.0
DNT	Not Detected	N/A	Not Detected	N/A
TCT (Normalized Response)	5.57E+02	8.8	4.67E+02	20.8

4.4 Discussion

The developed targeted MS method can monitor residual toxins levels during downstream purification of acellular Pertussis vaccines. Starting in the low µg/mL range in purified fermenter harvest, ACT concentrations were observed to decrease ~1000-fold between the culture supernatant and purified antigen samples. Considering a 0.5mL vaccine dose, ACT is well below the Ph. Eur. regulatory limit of 500 ng per dose.[168] For DNT, correlating absence of necrotic lesions to concentration value is beyond the scope of this thesis. Nevertheless, DNT removal during purification was confirmed with levels dropping from 8-10 ng/mL in culture supernatant to below the LOD in purified material (LOD estimated to be 4 ng/mL based on linearity analysis).

Through the targeted proteomics experiments, it was demonstrated that a synthetic peptide screening approach can be applied to accelerate development of a sensitive quantitative MS method across samples matrixes of differing complexities. Though driven initially by a desire to avoid working with purified toxins, the principles behind this standard-driven strategy to method development are applicable on other circumstances where data driven surrogate selection is not always feasible. This includes targeted analysis of novel biomarkers, protein isoforms, and other residual impurities from biologics for which there is no readily available protein standard.

The ability to start assay development without purified reference material represents a key advantage over immunochemical-based quantitation methods such as ELISAs. Though widely considered the workhorse for HCP testing for recombinant monoclonal antibodies, ELISAs are inherently limited by the specificity of the antibody.[154,155] Unlike immunochemical methods, targeted MS does not require the laborious task of developing and characterizing antibodies as critical reagents. Further, effective peptide screening can account for matrix effects, resulting in universally applicable assays that can be multiplexed to measure 100s of proteins using one method.

Starting on a quadrupole-Orbitrap provided several advantages over traditional SRM/MRM development on a triple quad (QqQ) instrument. First, the HRAM capabilities of the Orbitrap provided higher resolution MS/MS spectra with simultaneous monitoring of all fragment ions, rather than 3-4 transitions per peptide typically monitored by MRM. The ability to toggle between a Full MS survey and PRM fragment ion scans in one acquisition method was also particularly helpful when working with several sample types. Rather than being blind to salts and detergents commonly found in vaccines, the Full MS PRM method provided a high-level snapshot of everything eluting off the column. Only a quick visual check of the TIC was needed to evaluate sample complexity and rapidly identify deleterious matrix effects.

In transitioning towards absolute quantitation with the Spiketides_TQL (Qtag) peptides from JPT peptides, we benefited from the lower cost of synthesis, and thus were able to consider more sequences than would be practical using HPLC-purified AQUA peptide standards. Additionally, the option to re-quantitate the Qtag standards by UV spectroscopy immediately prior to analysis is also valuable given that heavy peptides would be considered critical standards from a regulatory perspective, and thus would be subject to monitoring requirements if implemented in routine use.

4.5 Conclusion

A synthetic peptide screening approach was successfully applied in the development of a quantitative targeted MS method for simultaneous quantitation of two residual protein toxins (ACT, DNT) and a glycopeptide (TCT) in vaccine intermediates. Digestion timecourse and replicate experiments were effective in guiding refinement of surrogate peptide sequences for absolute protein quantitation. The versatility of the assay across samples of starkly difference complexity demonstrates its value as an in vitro assay that can be employed universally for monitoring residual ACT, DNT and TCT in samples from *B. pertussis*.

Chapter 5 – Conclusion & Future Work

5.1 Conclusion

The work described above represents successful implementation of MS-based structural and quantitative proteomics techniques for analysis of complex biological systems. Electrostatic conformational priming effects were characterized for Cyt *c*-lipid interactions and demonstrated the sensitivity of TRESI-HDX for probing transient changes at protein-membrane interfaces. Refinement of sample digestion strategies for LC-MS also facilitated more comprehensive analysis of vaccine intermediates from *Bordetella Pertussis* with absolute quantitation of residual toxins. A synthetic peptide screening approach also enabled development of targeted MS method without access to purified reference material, providing significant advantages over traditional immunochemical methods including the ability to multiplex analysis of three residuals in one assay. The optimized nanoLC-MS method is ready for deployment in monitoring of residual toxins for next generation Pertussis vaccine development.

5.2 Future Work

To complement the results for the Cyt *c*-lipid interactions, conformational analysis of Cyt *c* in the presence of a membrane mimetic such as nanodiscs could be performed to better understand conformational changes occurring due to the surface area of the phospholipid bilayer. Additionally, the effect of H_2O_2, -induced covalent modifications on Cyt *c* structure could be examined by spiking H_2O_2 into the protein-membrane mixture prior to TRESI-HDX analysis. If pre-incubated with Cyt *c*, MS would be able to confirm of peroxidase activity by monitoring the oxidation of CL (+16 Da product). The kinetic TRESI-MS workflow[142] could also be used to study the kinetics of Cyt c-CL peroxidase catalysis alongside conformational analysis by millisecond HDX-MS. In a kinetic TRESI-HDX experiment, H_2O_2 could be introduced with D_2O via the outer

capillary to better understand the sequence of conformational transformations required for conversion of Cyt *c* into its peroxidase-active state. Kinetic TRESI-HDX analysis could also be used to link structural transitions to functional intermediates in the mitochondrial-induced apoptotic pathway. Nevertheless, such HDX-MS experiments would best be performed using higher ionic strength buffers that more analogous to physiological conditions. For TRESI-HDX-MS, increasing the salt concentration would require refinement of the setup and integration of an online desalting apparatus prior to MS analysis.

For the residual toxin work, a strategy to evaluate protein-level recovery is needed to further assess accuracy of the measurements obtained by targeted nanoLC-MS. This may include development of a recombinant protein or concatenated construct such as a QConCAT[186] similar to the work reported by Health Canada for LC-MS quantitation of flu antigens.[187] Coupling a protein construct with a digestion control standard would provide better control of sample preparation upstream of MS analysis.

For more routine implementation, the residual toxin method should be transferred to a conventional flow LC and triple quad MS. A higher flow targeted (MRM) method would improve throughput and robustness, but require bridging to ensure adequate sensitivity for DNT detection. Shifting from nanoLC to conventional flow LC would also mitigate limitations of the current trap-elute nanoLC setup for hydrophilic TCT. Different column chemistries and LC configurations could be evaluated for refined TCT detection while maintaining the ability to resolve ACT and DNT tryptic peptides. This work could be performed in parallel with establishment of a suitable TCT analogue as a standard for absolute quantitation. A higher throughput residual toxin LC-MS method would further facilitate improved characterization of acellular Pertussis vaccines.

Chapter 6: References

1. Griffiths, J. A Brief History of Mass Spectrometry. *Anal. Chem.* **80**, 5678–5683 (2008).
2. Thomson Joseph John. Bakerian Lecture:—Rays of positive electricity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **89**, 1–20 (1913).
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. Graves, P. R. & Haystead, T. A. J. Molecular Biologist's Guide to Proteomics. *Microbiol Mol Biol Rev* **66**, 39–63 (2002).
6. Suzuki, T., Jones, D. J. & Heaney, L. M. Mass spectrometry in medicine: a technology for the future? *Future Sci OA* **3**, FSO213 (2017).
7. Xia, Y. & McLuckey, S. A. Evolution of Instrumentation for the Study of Gas-Phase Ion/Ion Chemistry via Mass Spectrometry. *J Am Chem Soc* **19**, 173–189 (2008).
8. Whitehouse, C. M., Dreyer, R. N., Yamashita, M. & Fenn, J. B. Electrospray interface for liquid chromatographs and mass spectrometers. *Anal. Chem.* **57**, 675–679 (1985).
9. Taylor, G. I. & McEwan, A. D. The stability of a horizontal fluid interface in a vertical electric field. *Journal of Fluid Mechanics* **22**, 1–15 (1965).
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* **2012**, (2012).
11. Smith, J. N., Flagan, R. C. & Beauchamp, J. L. Droplet Evaporation and Discharge Dynamics in Electrospray Ionization. *J. Phys. Chem. A* **106**, 9957–9967 (2002).
12. Gomez, A. & Tang, K. Charge and fission of droplets in electrostatic sprays. *Physics of Fluids* **6**, 404–414 (1994).
13. Taflin, D. C., Ward, T. L. & Davis, E. J. Electrified droplet fission and the Rayleigh limit. *Langmuir* **5**, 376–384 (1989).
14. Dole, M. *et al.* Molecular Beams of Macroions. *J. Chem. Phys.* **49**, 2240–2249 (1968).
15. Iribarne, J. V. & Thomson, B. A. On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **64**, 2287–2294 (1976).
16. Fernandez de la Mora, J. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Analytica Chimica Acta* **406**, 93–104 (2000).
17. Konermann, L., Ahadi, E., Rodriguez, A. D. & Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **85**, 2–9 (2013).

18. Cech, N. B. & Enke, C. G. Relating electrospray ionization response to nonpolar character of small peptides. *Anal. Chem.* **72**, 2717–2723 (2000).

19. Allen, M. H. & Vestal, M. L. Design and performance of a novel electrospray interface. *J Am Soc Mass Spectrom* **3**, 18–26 (1992).

20. Wilm, M. & Mann, M. Analytical Properties of the Nanoelectrospray Ion Source. *Anal. Chem.* **68**, 1–8 (1996).

21. 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).

22. Wood, T. D. *et al.* Miniaturization of Electrospray Ionization Mass Spectrometry. *Applied Spectroscopy Reviews* **38**, 187–244 (2003).

23. Wilson, S. R., Vehus, T., Berg, H. S. & Lundanes, E. Nano-LC in proteomics: recent advances and approaches. *Bioanalysis* **7**, 1799–1815 (2015).

24. Scigelova, M., Hornshaw, M., Giannakopulos, A. & Makarov, A. Fourier Transform Mass Spectrometry. *Mol Cell Proteomics* **10**, (2011).

25. Ferguson, P. L. *et al.* Hydrogen/Deuterium Scrambling during Quadrupole Time-of-Flight MS/MS Analysis of a Zinc-Binding Protein Domain. *Anal. Chem.* **79**, 153–160 (2007).

26. Eliuk, S. & Makarov, A. Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annual Rev. Anal. Chem.* **8**, 61–80 (2015).

27. Scheltema, R. A. *et al.* The Q Exactive HF, a Benchtop Mass Spectrometer with a Pre-filter, High-performance Quadrupole and an Ultra-high-field Orbitrap Analyzer. *Mol Cell Proteomics* **13**, 3698–3708 (2014).

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

29. Curry, S. Structural Biology: A Century-long Journey into an Unseen World. *Interdiscip Sci Rev* **40**, 308–328 (2015).

30. Mann, M. & Jensen, O. *Mann, M. & Jensen, O.N. Proteomic analysis of post-translational modifications. Nat. Biotechnol. 21, 255-261*. **21**, (2003).

31. Frueh, D. P., Goodrich, A., Mishra, S. & Nichols, S. NMR methods for structural studies of large monomeric and multimeric proteins. *Curr Opin Struct Biol* **23**, 734–739 (2013).

32. Englander, S. W. & Kallenbach, N. R. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **16**, 521–655 (1983).

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

34. Smith, D. L., Deng, Y. & Zhang, Z. Probing the Non-covalent Structure of Proteins by Amide Hydrogen Exchange and Mass Spectrometry. *Journal of Mass Spectrometry* **32**, 135–146 (1997).

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

36. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **17**, 75–86 (1993).

37. Persson, F. & Halle, B. How amide hydrogens exchange in native proteins. *PNAS* **112**, 10383–10388 (2015).

38. Harrison, R. A. & Engen, J. R. Conformational insight into multi-protein signaling assemblies by hydrogen–deuterium exchange mass spectrometry. *Current Opinion in Structural Biology* **41**, 187–193 (2016).

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

40. Deng, B., Lento, C. & Wilson, D. J. Hydrogen deuterium exchange mass spectrometry in biopharmaceutical discovery and development – A review. *Analytica Chimica Acta* **940**, 8–20 (2016).

41. Wilson, D. J. & Konermann, L. A Capillary Mixer with Adjustable Reaction Chamber Volume for Millisecond Time-Resolved Studies by Electrospray Mass Spectrometry. *Anal. Chem.* **75**, 6408–6414 (2003).

42. Lento, C. & J. Wilson, D. Unravelling the mysteries of sub-second biochemical processes using time-resolved mass spectrometry. *Analyst* **142**, 1640–1653 (2017).

43. Zechel, D. L., Konermann, L., Withers, S. G. & Douglas, D. J. Pre-Steady State Kinetic Analysis of an Enzymatic Reaction Monitored by Time-Resolved Electrospray Ionization Mass Spectrometry. *Biochemistry* **37**, 7664–7669 (1998).

44. Rob, T. & Wilson, D. J. A Versatile Microfluidic Chip for Millisecond Time-Scale Kinetic Studies by Electrospray Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **20**, 124–130 (2009).

45. Lento, C., Ferraro, M., Wilson, D. & Audette, G. F. HDX-MS and deletion analysis of the type 4 secretion system protein TraF from the Escherichia coli F plasmid. *FEBS Lett.* **590**, 376–386 (2016).

46. Konermann, L. & Douglas, D. J. Acid-Induced Unfolding of Cytochrome c at Different Methanol Concentrations: Electrospray Ionization Mass Spectrometry Specifically Monitors Changes in the Tertiary Structure. *Biochemistry* **36**, 12296–12302 (1997).

47. Smith, D. P. *et al.* Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. *Eur J Mass Spectrom (Chichester)* **15**, 113–130 (2009).

48. Chen, Y.-L., Collings, B. A. & Douglas, D. J. Collision cross sections of myoglobin and cytochrome c ions with Ne, Ar, and Kr. *Journal of the American Society for Mass Spectrometry* **8**, 681–687 (1997).

49. Ewing, M. A., Glover, M. S. & Clemmer, D. E. Hybrid ion mobility and mass spectrometry as a separation tool. *Journal of Chromatography A* **1439**, 3–25 (2016).

50. Ben-Nissan, G. & Sharon, M. The application of ion-mobility mass spectrometry for structure/function investigation of protein complexes. *Curr Opin Chem Biol* **42**, 25–33 (2018).

51. Kinnunen, P. K. J., Kõiv, A., Lehtonen, J. Y. A., Rytömaa, M. & Mustonen, P. Lipid dynamics and peripheral interactions of proteins with membrane surfaces. *Chemistry and Physics of Lipids* **73**, 181–207 (1994).

52. Arumugam, S., Chwastek, G. & Schwille, P. Protein–membrane interactions: the virtue of minimal systems in systems biology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* **3**, 269–280 (2011).

53. Seddon, A. M., Curnow, P. & Booth, P. J. Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1666**, 105–117 (2004).

54. Fernández, C. & Wüthrich, K. NMR solution structure determination of membrane proteins reconstituted in detergent micelles. *FEBS Letters* **555**, 144–150 (2003).

55. Sanders, C. R. & Landis, G. C. Reconstitution of Membrane Proteins into Lipid-Rich Bilayered Mixed Micelles for NMR Studies. *Biochemistry* **34**, 4030–4040 (1995).

56. Whiles, J. A., Deems, R., Vold, R. R. & Dennis, E. A. Bicelles in structure–function studies of membrane-associated proteins. *Bioorganic Chemistry* **30**, 431–442 (2002).

57. Hopper, J. T. S. *et al.* New vehicles allow detergent-free mass spectrometry of membrane protein complexes. *Nat Methods* **10**, (2013).

58. Beaugrand, M. *et al.* Magnetically Oriented Bicelles with Monoalkylphosphocholines: Versatile Membrane Mimetics for Nuclear Magnetic Resonance Applications. *Langmuir* **32**, 13244–13251 (2016).

59. Montenegro, F. A., Cantero, J. R. & Barrera, N. P. Combining Mass Spectrometry and X-Ray Crystallography for Analyzing Native-Like Membrane Protein Lipid Complexes. *Front Physiol* **8**, (2017).

60. Cross, T. A., Sharma, M., Yi, M. & Zhou, H.-X. Influence of solubilizing environments on membrane protein structures. *Trends in Biochemical Sciences* **36**, 117–125 (2011).

61. Harayama, T. & Riezman, H. Understanding the diversity of membrane lipid composition. *Nat. Rev. Mol. Cell Biol.* **19**, 281–296 (2018).

62. Reading, E. *et al.* Interrogating Membrane Protein Conformational Dynamics within Native Lipid Compositions. *Angewandte Chemie International Edition* **56**, 15654–15657 (2017).

63. Kagan, V. E., Chu, C. T., Tyurina, Y. Y., Cheikhi, A. & Bayir, H. Cardiolipin asymmetry, oxidation and signaling. *Chem Phys Lipids* **179**, 64–69 (2014).

64. Hood, L. & Rowen, L. The Human Genome Project: big science transforms biology and medicine. *Genome Medicine* **5**, 79 (2013).

65. Yang, K. & Han, X. Lipidomics: Techniques, Applications, and Outcomes Related to Biomedical Sciences. *Trends in Biochemical Sciences* **41**, 954–969 (2016).

66. Evans, T. G. Considerations for the use of transcriptomics in identifying the 'genes that matter' for environmental adaptation. *Journal of Experimental Biology* **218**, 1925–1935 (2015).

67. Newgard, C. B. Metabolomics and Metabolic Diseases: Where Do We Stand? *Cell Metabolism* **25**, 43–56 (2017).

68. McDonald, W. H. & Yates, J. R. Shotgun Proteomics and Biomarker Discovery. *Dis Markers* **18**, 99–105 (2002).

69. Wolters, D. A., Washburn, M. P. & Yates, J. R. An Automated Multidimensional Protein Identification Technology for Shotgun Proteomics. *Anal. Chem.* **73**, 5683–5690 (2001).

70. Meissner, F. & Mann, M. Quantitative shotgun proteomics: considerations for a high-quality workflow in immunology. *Nat. Immunol.* **15**, 112–117 (2014).

71. Siuti, N. & Kelleher, N. L. Decoding protein modifications using top-down mass spectrometry. *Nat Methods, Nature methods, Nature methods., PLoS medicine* **4**, 817–821 (2007).

72. Meiring, H. D., Heeft, E. van der, Hove, G. J. ten & Jong, A. P. J. M. de. Nanoscale LC–MS(n): technical design and applications to peptide and protein analysis. *Journal of Separation Science* **25**, 557–568 (2002).

73. Gaspari, M. & Cuda, G. Nano LC–MS/MS: A Robust Setup for Proteomic Analysis. *Nanoproteomics* 115–126 (2011).

74. Vandermarliere, E., Mueller, M. & Martens, L. Getting intimate with trypsin, the leading protease in proteomics. *Mass Spectrometry Reviews* **32**, 453–465 (2013).

75. Olsen, J. V. & Mann, M. Status of Large-scale Analysis of Post-translational Modifications by Mass Spectrometry. *Mol Cell Proteomics* **12**, 3444–3452 (2013).

76. Giansanti, P., Tsiatsiani, L., Low, T. Y. & Heck, A. J. R. Six alternative proteases for mass spectrometry–based proteomics beyond trypsin. *Nature Protocols* **11**, 993–1006 (2016).

77. Saveliev, S. *et al.* Trypsin/Lys-C protease mix for enhanced protein mass spectrometry analysis. *Nature Methods* **10**, 1134 (2013).

78. Tabb, D. L. The SEQUEST Family Tree. *J Am Soc Mass Spectrom* **26**, 1814–1819 (2015).

79. Dorfer, V. *et al.* MS Amanda, a Universal Identification Algorithm Optimized for High Accuracy Tandem Mass Spectra. *J Proteome Res* **13**, 3679–3684 (2014).

80. Armengaud, J. Microbiology and proteomics, getting the best of both worlds! *Environ. Microbiol.* **15**, 12–23 (2013).

81. Ma, B. & Johnson, R. De Novo Sequencing and Homology Searching‡‡. *Mol Cell Proteomics* **11**, (2012).

82. Geyer, P. E., Holdt, L. M., Teupser, D. & Mann, M. Revisiting biomarker discovery by plasma proteomics. *Mol Syst Biol* **13**, (2017).

83. Legrain, P. *et al.* The Human Proteome Project: Current State and Future Direction. *Mol Cell Proteomics* **10**, (2011).

84. Chandramouli, K. & Qian, P.-Y. Proteomics: Challenges, Techniques and Possibilities to Overcome Biological Sample Complexity. *Hum Genomics Proteomics* **2009**, (2009).

85. Husson, G. *et al.* Dual Data-Independent Acquisition Approach Combining Global HCP Profiling and Absolute Quantification of Key Impurities during Bioprocess Development. *Anal. Chem.* **90**, 1241–1247 (2018).

86. Gillet, L. C. *et al.* Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Molecular & Cellular Proteomics* **11**, O111.016717 (2012).

87. The, M. & Käll, L. Integrated Identification and Quantification Error Probabilities for Shotgun Proteomics. *Molecular & Cellular Proteomics* **18**, 561–570 (2019).

88. Bourmaud, A., Gallien, S. & Domon, B. Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: Principle and applications. *Proteomics* **16**, 2146–2159 (2016).

89. Frese, C. K. *et al.* Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. *J. Proteome Res.* **10**, 2377–2388 (2011).

90. Lange, V., Picotti, P., Domon, B. & Aebersold, R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* **4**, 222 (2008).

91. Ronsein, G. E. *et al.* Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. *J Proteomics* **113**, 388–399 (2015).

92. Domon, B. & Aebersold, R. Options and considerations when selecting a quantitative proteomics strategy. *Nat Biotech* **28**, 710–721 (2010).

93. Osinalde, N., Aloria, K., Omaetxebarria, M. J. & Kratchmarova, I. Targeted mass spectrometry: An emerging powerful approach to unblock the bottleneck in phosphoproteomics. *Journal of Chromatography B* **1055–1056**, 29–38 (2017).

94. Rauniyar, N. Parallel Reaction Monitoring: A Targeted Experiment Performed Using High Resolution and High Mass Accuracy Mass Spectrometry. *Int J Mol Sci* **16**, 28566–28581 (2015).

95. Gygi, S. P. *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999 (1999).

96. Kirkpatrick, D. S., Gerber, S. A. & Gygi, S. P. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* **35**, 265–273 (2005).

97. Ow, Y.-L. P., Green, D. R., Hao, Z. & Mak, T. W. Cytochrome c: functions beyond respiration. *Nat. Rev. Mol. Cell Biol.* **9**, 532–542 (2008).

98. Liu, X., Kim, C. N., Yang, J., Jemmerson, R. & Wang, X. Induction of Apoptotic Program in Cell-Free Extracts: Requirement for dATP and Cytochrome c. *Cell* **86**, 147–157 (1996).

99. Alvarez-Paggi, D. *et al.* Multifunctional Cytochrome c: Learning New Tricks from an Old Dog. *Chem. Rev.* **117**, 13382–13460 (2017).

100. Orrenius, S. Mitochondrial regulation of apoptotic cell death. *Toxicology Letters* **149**, 19–23 (2004).

101. Iverson, S. L. & Orrenius, S. The cardiolipin–cytochrome c interaction and the mitochondrial regulation of apoptosis. *Archives of Biochemistry and Biophysics* **423**, 37–46 (2004).

102. Ascenzi, P., Polticelli, F., Marino, M., Santucci, R. & Coletta, M. Cardiolipin drives cytochrome c proapoptotic and antiapoptotic actions. *IUBMB Life* **63**, 160–165 (2011).

103. Ardail, D. *et al.* Mitochondrial contact sites. Lipid composition and dynamics. *J. Biol. Chem.* **265**, 18797–18802 (1990).

104. Lewis, R. N. A. H. & McElhaney, R. N. The physicochemical properties of cardiolipin bilayers and cardiolipin-containing lipid membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1788**, 2069–2079 (2009).

105. Orrenius, S. & Zhivotovsky, B. Cardiolipin oxidation sets cytochrome C free. *Nature Chemical Biology* **1**, 188–189 (2005).

106. Ott, M., Robertson, J. D., Gogvadze, V., Zhivotovsky, B. & Orrenius, S. Cytochrome c release from mitochondria proceeds by a two-step process. *Proc Natl Acad Sci U S A* **99**, 1259–1263 (2002).

107. Li, P. *et al.* Cytochrome c and dATP-Dependent Formation of Apaf-1/Caspase-9 Complex Initiates an Apoptotic Protease Cascade. *Cell* **91**, 479–489 (1997).

108. Kagan, V. E. *et al.* Oxidative lipidomics of apoptosis: redox catalytic interactions of cytochrome c with cardiolipin and phosphatidylserine. *Free Radical Biology and Medicine* **37**, 1963–1985 (2004).

109. Kagan, V. E. *et al.* Cytochrome c acts as a cardiolipin oxygenase required for release of proapoptotic factors. *Nature Chemical Biology* **1**, 223–232 (2005).

110. Belikova, N. A. *et al.* Peroxidase Activity and Structural Transitions of Cytochrome c Bound to Cardiolipin-Containing Membranes. *Biochemistry* **45**, 4998–5009 (2006).

111. Mandal, A. *et al.* Structural Changes and Proapoptotic Peroxidase Activity of Cardiolipin-Bound Mitochondrial Cytochrome c. *Biophysical Journal* **109**, 1873–1884 (2015).

112. Kapralov, A. A. *et al.* The Hierarchy of Structural Transitions Induced in Cytochrome c by Anionic Phospholipids Determines Its Peroxidase Activation and Selective Peroxidation during Apoptosis in Cells. *Biochemistry* **46**, 14232–14244 (2007).

113. Muenzner, J., Toffey, J. R., Hong, Y. & Pletneva, E. V. Becoming a peroxidase: cardiolipin-induced unfolding of cytochrome c. *J Phys Chem B* **117**, 12878–12886 (2013).

114. Muenzner, J. & Pletneva, E. V. Structural Transformations of Cytochrome c upon Interaction with Cardiolipin. *Chem Phys Lipids* **179**, 57–63 (2014).

115. Balakrishnan, G., Hu, Y. & Spiro, T. G. His26 Protonation in Cytochrome c Triggers Microsecond β-sheet Formation and Heme Exposure: Implications for Apoptosis. *J Am Chem Soc* **134**, 19061–19069 (2012).

116. Heimburg, T. & Marsh, D. Investigation of secondary and tertiary structural changes of cytochrome c in complexes with anionic lipids using amide hydrogen exchange measurements: an FTIR study. *Biophys J* **65**, 2408–2417 (1993).

117. Salamon, Z. & Tollin, G. Surface plasmon resonance studies of complex formation between cytochrome c and bovine cytochrome c oxidase incorporated into a supported planar lipid bilayer. I. Binding of cytochrome c to cardiolipin/phosphatidylcholine membranes in the absence of oxidase. *Biophys J* **71**, 848–857 (1996).

118. Basova, L. V. *et al.* Cardiolipin Switch in Mitochondria: Shutting off the Reduction of Cytochrome c and Turning on the Peroxidase Activity. *Biochemistry* **46**, 3423–3434 (2007).

119. Abe, M., Niibayashi, R., Koubori, S., Moriyama, I. & Miyoshi, H. Molecular Mechanisms for the Induction of Peroxidase Activity of the Cytochrome c–Cardiolipin Complex. *Biochemistry* **50**, 8383–8391 (2011).

120. Capdevila, D. A. *et al.* Specific methionine oxidation of cytochrome c in complexes with zwitterionic lipids by hydrogen peroxide: potential implications for apoptosis. *Chem. Sci.* **6**, 705–713 (2014).

121. Rytömaa, M. & Kinnunen, P. K. Evidence for two distinct acidic phospholipid-binding sites in cytochrome c. *J. Biol. Chem.* **269**, 1770–1774 (1994).

122. Tuominen, E. K. J., Wallace, C. J. A. & Kinnunen, P. K. J. Phospholipid-Cytochrome c Interaction Evidence for the Extended Lipid Anchorage. *J. Biol. Chem.* **277**, 8822–8826 (2002).

123. Kalanxhi, E. & Wallace, C. J. A. Cytochrome c impaled: investigation of the extended lipid anchorage of a soluble protein to mitochondrial membrane models. *Biochem J* **407**, 179–187 (2007).

124. Sinibaldi, F. *et al.* Extended cardiolipin anchorage to cytochrome c: a model for protein–mitochondrial membrane binding. *J Biol Inorg Chem* **15**, 689–700 (2010).

125. Hanske, J. *et al.* Conformational properties of cardiolipin-bound cytochrome c. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 125–130 (2012).

126. Hong, Y., Muenzner, J., Grimm, S. K. & Pletneva, E. V. Origin of the Conformational Heterogeneity of Cardiolipin-bound Cytochrome c. *J Am Chem Soc* **134**, 18713–18723 (2012).

127. Hannibal, L. *et al.* Alternative Conformations of Cytochrome c: Structure, Function, and Detection. *Biochemistry* **55**, 407–428 (2016).

128. O'Brien, E. S., Nucci, N. V., Fuglestad, B., Tommos, C. & Wand, A. J. Defining the Apoptotic Trigger. *J Biol Chem* **290**, 30879–30887 (2015).

129. Gorbenko, G. P. Structure of cytochrome c complexes with phospholipids as revealed by resonance energy transfer. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1420**, 1–13 (1999).

130. Brown, K. A. *et al.* Distinct Dynamic Modes Enable the Engagement of Dissimilar Ligands in a Promiscuous Atypical RNA Recognition Motif. *Biochemistry* **55**, 7141–7150 (2016).

131. Artimo, P. *et al.* ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* **40**, W597-603 (2012).

132. Strohalm, M., Hassman, M., Košata, B. & Kodíček, M. mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Communications in Mass Spectrometry* **22**, 905–908 (2008).

133. Rey, M. *et al.* Mass Spec Studio for Integrative Structural Biology. *Structure* **22**, 1538–1548 (2014).

134. Kierdaszuk, B. *et al.* Fluorescence of tyrosine and tryptophan in proteins using one- and two-photon excitation. *Photochem. Photobiol.* **61**, 319–324 (1995).

135. Moon, C. P. & Fleming, K. G. Using Tryptophan Fluorescence to Measure the Stability of Membrane Proteins Folded in Liposomes. in *Methods in Enzymology* **492**, 189–211 (Elsevier, 2011).

136. Ladokhin, A. S., Jayasinghe, S. & White, S. H. How to measure and analyze tryptophan fluorescence in membranes properly, and why bother? *Analytical biochemistry* **285**, 235–245 (2000).

137. Sanchez, K. M., Gable, J. E., Schlamadinger, D. E. & Kim, J. E. Effects of tryptophan microenvironment, soluble domain, and vesicle size on the thermodynamics of membrane protein folding: lessons from the transmembrane protein OmpA. *Biochemistry* **47**, 12844–12852 (2008).

138. Vadas, O. & Burke, J. E. Probing the dynamic regulation of peripheral membrane proteins using hydrogen deuterium exchange–MS (HDX–MS). *Biochemical Society Transactions* **43**, 773–786 (2015).

139. Hebling, C. M. *et al.* Conformational Analysis of Membrane Proteins in Phospholipid Bilayer Nanodiscs by Hydrogen Exchange Mass Spectrometry. *Anal. Chem.* **82**, 5415–5419 (2010).

140. Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S.-J. & Robinson, C. V. Ion mobility–mass spectrometry analysis of large protein complexes. *Nature Protocols* **3**, 1139–1152 (2008).

141. Clemmer, D. E., Hudgins, R. R. & Jarrold, M. F. Naked Protein Conformations: Cytochrome c in the Gas Phase. *J. Am. Chem. Soc.* **117**, 10141–10142 (1995).

142. Liuni, P., Deng, B. & J. Wilson, D. Comparing equilibrium and kinetic protein unfolding using time-resolved electrospray-coupled ion mobility mass spectrometry. *Analyst* **140**, 6973–6979 (2015).

143. May, J. C. *et al.* Conformational landscapes of ubiquitin, cytochrome c, and myoglobin: Uniform field ion mobility measurements in helium and nitrogen drift gas. *International Journal of Mass Spectrometry* **427**, 79–90 (2018).

144. Konermann, L., Collings, B. A. & Douglas, D. J. Cytochrome c Folding Kinetics Studied by Time-Resolved Electrospray Ionization Mass Spectrometry. *Biochemistry* **36**, 5554–5559 (1997).

145. Hersleth, H.-P., Ryde, U., Rydberg, P., Görbitz, C. H. & Andersson, K. K. Structures of the high-valent metal-ion haem–oxygen intermediates in peroxidases, oxygenases and catalases. *Journal of Inorganic Biochemistry* **100**, 460–476 (2006).

146. Birk, A. V., Chao, W. M., Liu, S., Soong, Y. & Szeto, H. H. Disruption of cytochrome c heme coordination is responsible for mitochondrial injury during ischemia. *Biochim. Biophys. Acta* **1847**, 1075–1084 (2015).

147. Cerella, C. *et al.* Multiple Mechanisms for Hydrogen Peroxide–Induced Apoptosis. *Annals of the New York Academy of Sciences* **1171**, 559–563 (2009).

148. Giorgio, M., Trinei, M., Migliaccio, E. & Pelicci, P. G. Hydrogen peroxide: a metabolic by-product or a common mediator of ageing signals? *Nat. Rev. Mol. Cell Biol.* **8**, 722–728 (2007).

149. Yin, V., Shaw, G. S. & Konermann, L. Cytochrome c as a Peroxidase: Activation of the Precatalytic Native State by H2O2-Induced Covalent Modifications. *J. Am. Chem. Soc.* **139**, 15701–15709 (2017).

150. Capdevila, D. A. *et al.* Specific methionine oxidation of cytochrome c in complexes with zwitterionic lipids by hydrogen peroxide: potential implications for apoptosis †Electronic supplementary information (ESI) available: CV, RR, SERR, UV-vis, HPLC and MS data. See DOI: 10.1039/c4sc02181a. *Chem Sci* **6**, 705–713 (2015).

151. Yin, V., Mian, S. H. & Konermann, L. Lysine carbonylation is a previously unrecognized contributor to peroxidase activation of cytochrome c by chloramine-T. *Chem. Sci.* (2019).

152. Galassie, A. C. & Link, A. J. Proteomic contributions to our understanding of vaccine and immune responses. *Proteomics Clin Appl* **9**, 972–989 (2015).

153. Adamczyk-Poplawska, M., Markowicz, S. & Jagusztyn-Krynicka, E. K. Proteomics for development of vaccine. *Journal of Proteomics* **74**, 2596–2616 (2011).

154. Bracewell, D. G., Francis, R. & Smales, C. M. The future of host cell protein (HCP) identification during process development and manufacturing linked to a risk-based management for their control. *Biotechnol. Bioeng.* **112**, 1727–1737 (2015).

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

156. Prakash, K. & Chen, W. Analytical Methods for the Measurement of Host Cell Proteins and Other Process-Related Impurities. in *ACS Symposium Series* (eds. Schiel, J. E., Davis, D. L. & Borisov, O. V.) **1201**, 387–404 (American Chemical Society, 2015).

157. Clark, T. A. Changing Pertussis Epidemiology: Everything Old is New Again. *J Infect Dis* **209**, 978–981 (2014).

158. Greco, D. *et al.* A Controlled Trial of Two Acellular Vaccines and One Whole-Cell Vaccine against Pertussis. *New England Journal of Medicine* **334**, 341–349 (1996).

159. Cherry, J. D. Pertussis: Challenges Today and for the Future. *PLOS Pathogens* **9**, e1003418 (2013).

160. Li, Z. *et al.* Purification design and practice for pertactin, the third component of acellular pertussis vaccine, from Bordetella pertussis. *Vaccine* **34**, 4032–4039 (2016).

161. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nature Methods* **6**, 359–362 (2009).

162. Zougman, A., Selby, P. J. & Banks, R. E. Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* **14**, 1006–1000 (2014).

163. HaileMariam, M. *et al.* S-Trap, an Ultrafast Sample-Preparation Approach for Shotgun Proteomics. *J. Proteome Res.* **17**, 2917–2924 (2018).

164. Parkhill, J. *et al.* Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nature Genetics* **35**, 32–40 (2003).

165. Yu, Y.-Q., Gilar, M., Lee, P. J., Bouvier, E. S. P. & Gebler, J. C. Enzyme-friendly, mass spectrometry-compatible surfactant for in-solution enzymatic digestion of proteins. *Anal. Chem.* **75**, 6023–6028 (2003).

166. James, D. A., Fradin, M. J., Pedyczak, A. & Carpick, B. W. Detection of Residual Proteins UL5 and UL29 Using a Targeted Proteomics Approach in HSV529, a Replication-Deficient HSV-2 Vaccine Candidate. *Journal of Pharmaceutical Sciences* **107**, 3022–3031 (2018).

167. Mellacheruvu, D. *et al.* The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. *Nature Methods* **10**, 730–736 (2013).

168. European Pharmacopoeia. in *Current edition European monograph for adsorbed pertussis vaccines (Acellular component)* 1356 (2013).

169. Hewlett, E. L., Urban, M. A., Manclark, C. R. & Wolff, J. Extracytoplasmic adenylate cyclase of Bordetella pertussis. *Proc Natl Acad Sci U S A* **73**, 1926–1930 (1976).

170. Carbonetti, N. H. Pertussis toxin and adenylate cyclase toxin: key virulence factors of Bordetella pertussis and cell biology tools. *Future Microbiol* **5**, 455–469 (2010).

171. Dadaglio, G. *et al.* Antigen Targeting to CD11b+ Dendritic Cells in Association with TLR4/Toll/IL-1R Domain–Containing Adapter-Inducing IFN-β Signaling Promotes Strong CD8+ T Cell Responses. *The Journal of Immunology* 1302974 (2014).

172. Carbonetti, N. H., Artamonova, G. V., Andreasen, C. & Bushar, N. Pertussis Toxin and Adenylate Cyclase Toxin Provide a One-Two Punch for Establishment of Bordetella pertussis Infection of the Respiratory Tract. *Infect Immun* **73**, 2698–2703 (2005).

173. Guiso, N. Bordetella Adenylate Cyclase-Hemolysin Toxins. *Toxins (Basel)* **9**, (2017).

174. Recommendations to assure the quality, safety and efficacy of acellular pertussis vaccines-WHO Technical Report Series No. 979 (2013).

175. Gentry-Weeks, C. R. *et al.* Dermonecrotic toxin and tracheal cytotoxin, putative virulence factors of Bordetella avium. *Infect. Immun.* **56**, 1698–1707 (1988).

176. Melvin, J. A., Scheller, E. V., Miller, J. F. & Cotter, P. A. Bordetella pertussis pathogenesis: current and future challenges. *Nat. Rev. Microbiol.* **12**, 274–288 (2014).

177. Locht, C. Molecular aspects of Bordetella pertussis pathogenesis. *Int. Microbiol.* **2**, 137–144 (1999).

178. Goldman, W. E., Klapper, D. G. & Baseman, J. B. Detection, isolation, and analysis of a released Bordetella pertussis product toxic to cultured tracheal cells. *Infect. Immun.* **36**, 782–794 (1982).

179. Ogawa, C., Liu, Y.-J. & Kobayashi, K. S. Muramyl dipeptide and its derivatives: peptide adjuvant in immunological disorders and cancer therapy. *Curr Bioact Compd* **7**, 180–197 (2011).

180. Luker, K. E., Collier, J. L., Kolodziej, E. W., Marshall, G. R. & Goldman, W. E. Bordetella pertussis tracheal cytotoxin and other muramyl peptides: distinct structure-activity relationships for respiratory epithelial cytopathology. *Proc Natl Acad Sci U S A* **90**, 2365–2369 (1993).

181. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966–968 (2010).

182. Hilpert, K., Winkler, D. F. H. & Hancock, R. E. W. Peptide arrays on cellulose support: SPOT synthesis, a time and cost efficient method for synthesis of large numbers of peptides in a parallel and addressable fashion. *Nat Protoc* **2**, 1333–1349 (2007).

183. Martin, S. A., Rosenthal, R. S. & Biemann, K. Fast atom bombardment mass spectrometry and tandem mass spectrometry of biologically active peptidoglycan monomers from Neisseria gonorrhoeae. *J. Biol. Chem.* **262**, 7514–7522 (1987).

184. Zhao, P. *et al.* Combining High-energy C-trap Dissociation and Electron Transfer Dissociation for Protein O-GlcNAc Modification Site Assignment. *J Proteome Res* **10**, 4088–4104 (2011).

185. Carr, S. A. *et al.* Targeted Peptide Measurements in Biology and Medicine: Best Practices for Mass Spectrometry-based Assay Development Using a Fit-for-Purpose Approach. *Mol Cell Proteomics* **13**, 907–917 (2014).

186. Pratt, J. M. *et al.* Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nature Protocols* **1**, 1029–1043 (2006).

187. Smith, D. G. S., Gingras, G., Aubin, Y. & Cyr, T. D. Design and expression of a QconCAT protein to validate Hi3 protein quantification of influenza vaccine antigens. *J Proteomics* **146**, 133–140 (2016).

Appendix

Appendix A - Curve Fitting Analysis for Global TRESI-HDX Results

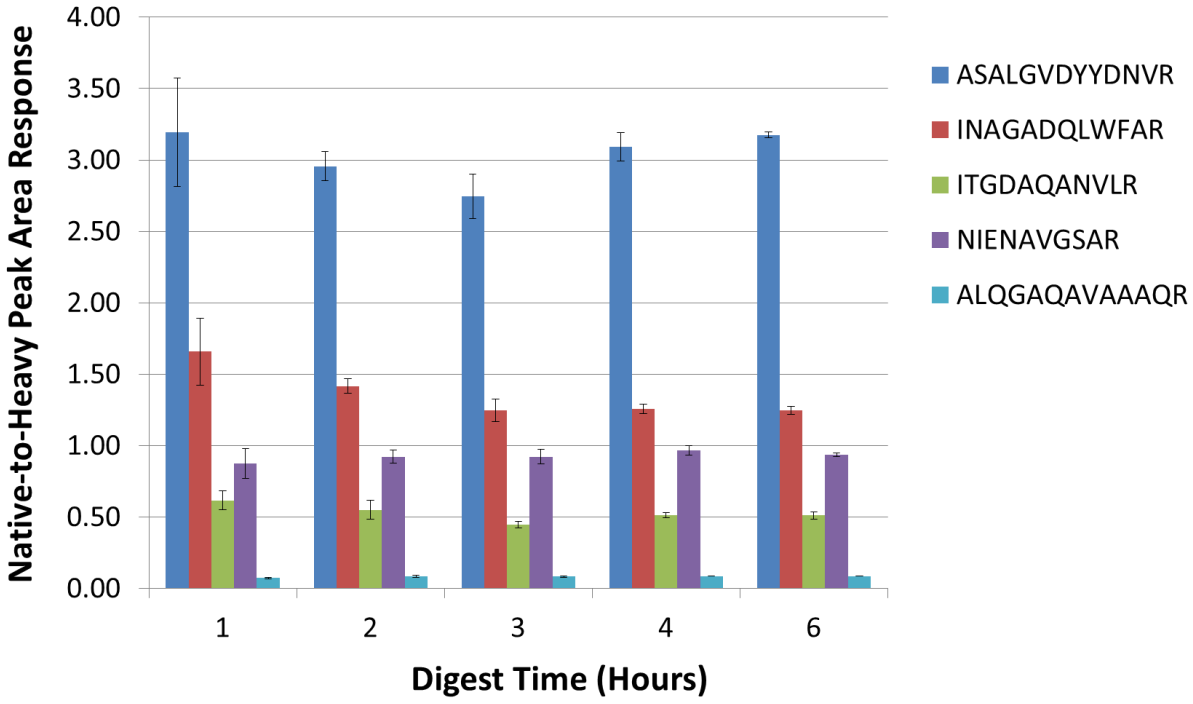
Fit to exponential Rise to Maximum, Single, 2 Parameter

f=a*(1-exp(-b*x))

Parameter	Cyt *c* only	Cyt *c* & DMPC	Cyt *c* & CL
a	85.74	86.80	53.07
Std Error a	2.89	1.65	2.41
b	2.17	5.15	3.29
Std Error b	0.20	0.53	0.54
R^2	0.99	0.99	0.98

[Constraints]
b>2
b<10
[Options]
tolerance=1e-10
1iterations=200

Appendix B– ACT quantitation in unpurified culture supernatant using Qtag heavy peptide sequences.



Protein	Sequence	Qtag Peptide Standard Concentration (fmol/µL)	Average Native-to-Heavy Area Ratio	%CV peptide (n=9)	Calculated Peptide Concentration in Sample (fmol/µL)	Calculated Protein Concentration	%CV Protein
ACT	ITGDAQANVLR	20.0	0.979	5.2%	44.4	59.9 fmol/µL	18.8%
	NIENAVGSAR		1.424	4.7%	64.6		
	INAGADQLWFAR		1.560	5.3%	70.8		
DNT	NDDLVSIAATYDR	5.0	5.70E-03	20.4%	64.6	64.6 amol/µL	N/Ap