Development of liquid extraction surface analysis mass spectrometry for protein analysis in microbial colonies

---

by

Klaudia Izabela Kocurek

A thesis submitted to the University of Birmingham for the degree of
DOCTOR OF PHILOSOPHY

School of Biosciences
College of Life and Environmental Sciences
University of Birmingham

October 2018

ABSTRACT

Liquid extraction surface analysis (LESA) is an ambient mass spectrometry (MS) technique applied for the analysis of proteins from a range of biological substrates. The work described in this thesis focuses on the development of LESA MS-based approaches for the analysis of proteins from microbial biofilms directly on solid nutrient media. Both microbiological research applications as well as clinical applications for the identification of pathogens were considered during development. The reproducibility of LESA sampling on model surfaces and bacterial colonies was investigated. Complementary imaging methods were used to characterise the effects of LESA sampling on the surface of colonies. Sampling protocols were optimised to access proteins from Gram-positive bacteria. Identification of over 40 proteins by LESA MS was subsequently demonstrated for Gram-positive and Gram-negative clinical isolates; *de novo* sequencing of a novel protein from an unknown strain was also achieved. Characteristic protein masses were used to differentiate closely related species of streptococci. The capabilities of LESA MS for protein identification and characterisation were expanded by the coupling of high-field asymmetric waveform ion mobility spectrometry (FAIMS) into the workflow, overcoming several limitations encountered in the earlier stages of the project. The final challenge involved the sampling of yeast colonies, not ordinarily amenable to lysis by solvents compatible with LESA due to the presence of a thick cell wall. A device was constructed to allow the lysis of yeast colonies directly on agar plates by application of high voltage electric pulses to the colonies prior to LESA sampling. The extraction and identification of proteins in three yeast species was subsequently demonstrated.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Professors Helen J. Cooper, Robin C. May and Josephine Bunch, for their guidance, support and encouragement throughout the project. The funding for this work was provided by EPSRC and the National Physical Laboratory.

I would also like to thank all the members of the Cooper and Bunch groups for all their help, company and cake, and for putting up with me when I was stressed out. I thank Leanne Stones for teaching me how to grow microbes, Rian Griffiths for teaching me how to perform LESA, and Andy Tanner for the engineering work on the electric lysis device. I thank Jana Havlikova for all the jolly discussions of mass spec and microbiology. To my friends, thank you for the laughs and the tea and the banter. And to my parents, thank you for giving me the chance to do what I love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AC – alternating current

AGC – automatic gain control

AUI – advanced user interface

CAD – collisionally activated dissociation

CAP – cold atmospheric plasma

CF – compensation field

CHCA – α-cyano-4-hydroxycinnamic acid

CI – chemical ionisation

CID – collision induced dissociation

CPCD – coupled physical and chemical dynamics

DC – direct current

DESI – desorption electrospray ionisation

DF – dispersion field

DHB – 2,5-dihydroxybenzoic acid

DT-IMS – drift time ion mobility spectrometry

EI – electron impact ionisation

ESI – electrospray ionisation

FAB – fast atom bombardment

FAIMS – high field asymmetric waveform ion mobility spectrometry

FET – field-effect transistor

FT-ICR – Fourier transform ion cyclotron resonance

HCD – higher energy collision dissociation

IMS – ion mobility spectrometry

LB – lysogeny broth

LBA – lysogeny broth agar

LC – liquid chromatography

LESA – liquid extraction surface analysis

LMJ-SSP – liquid microjunction surface sampling probe

LPS – lipopolysaccharide

MAI – matrix assisted ionisation

MALDI – matrix-assisted laser desorption ionisation

MS – mass spectrometry

nanoDESI – nanospray desorption electrospray ionisation

nanoESI – nanoelectrospray ionisation

PD – plasma desorption

PMF – peptide mass fingerprint

PSM – phenol soluble modulin

PTM – post-translational modification

RF – radio frequency

SALDI – surface-assisted laser desorption ionisation

SIMS – secondary ion mass spectrometry

TLC – thin layer chromatography

TOF – time-of-flight

TS – thermal spray

UV – ultraviolet

YPD – yeast extract peptone dextrose

LIST OF PUBLICATIONS

Peer-reviewed journal articles

K.I. Kocurek, J. Bunch, R.C. May and H.J. Cooper. Liquid extraction surface analysis mass spectrometry of yeast colonies lysed by application of electric current. Manuscript in preparation.

K.I. Kocurek, J. Bunch, R.C. May and H.J. Cooper. Application of high-field asymmetric waveform ion mobility separation to LESA mass spectrometry of bacteria. Manuscript in preparation.

K.I. Kocurek, R.L. Griffiths and H.J. Cooper. Ambient ionization mass spectrometry for in situ analysis of intact proteins. J. Mass Spectrom. 2018, 53: 565–578.

R.L. Griffiths, K.I. Kocurek and H.J. Cooper. Ambient surface mass spectrometry-ion mobility spectrometry of intact proteins. Curr. Opin. Chem. Biol. 2018, 42: 67–75.

K.I. Kocurek, L. Stones, J. Bunch, R.C. May and H.J. Cooper. Top-down LESA mass spectrometry protein analysis of Gram-positive and Gram-negative bacteria. JASMS 2017, 28: 2066–2077.

Presentations

Talks

K.I. Kocurek, J. Bunch, R.C. May and H.J. Cooper. Intact protein analysis of bacterial colonies by liquid extraction surface analysis mass spectrometry. Spring SciX 2018 (invited).

K.I. Kocurek, J. Bunch, R.C. May and H.J. Cooper. Liquid extraction surface analysis mass spectrometry: a novel tool for in situ analysis of intact bacterial proteins. Microbiology Society Annual Conference 2018.

K.I. Kocurek, J. Bunch, R.C. May and H.J. Cooper. Liquid extraction surface analysis-high field asymmetric waveform ion mobility mass spectrometry of bacterial colonies. BMSS Ambient Ionisation Special Interest Group Meeting 2018.

K.I. Kocurek, J. Bunch, R.C. May and H.J. Cooper. Liquid extraction surface analysis mass spectrometry for protein analysis directly from clinical isolates of Staphylococcus aureus and Pseudomonas aeruginosa. ASMS Annual Conference 2017.

Posters

K.I. Kocurek, J. Bunch, R.C. May and H.J. Cooper. Liquid extraction surface analysis mass spectrometry of bacterial colonies: expanding capabilities with high-field asymmetric waveform ion mobility spectrometry. ASMS Annual Conference 2018.

J. Havlikova, K.I. Kocurek, W. van Schaik, R.C. May, I.B. Styles and H.J. Cooper. Development of liquid extraction surface analysis mass spectrometry for identification of ESKAPE pathogens. ASMS Annual Conference 2018.

R.L. Griffiths, E.K. Sisley, A.L. Simmonds, K.I. Kocurek, J. Hughes, J. Havlikova, I.B. Styles and H.J. Cooper. Combining ion mobility spectrometry with ambient mass spectrometry: Toward spatial profiling of protein conformations. ASMS Annual Conference 2018.

K.I. Kocurek, R.C. May, H.J. Cooper and J. Bunch. Assessing the performance of liquid extraction surface analysis mass spectrometry of bacterial colonies. BMSS Annual Meeting 2017.

K.I. Kocurek, J. Bunch, R.C. May and H.J. Cooper. Top-down protein identification directly from bacterial colonies by liquid extraction surface analysis mass spectrometry. BMSS Annual Meeting 2016.

K.I. Kocurek, J. Bunch, R.C. May and H.J. Cooper. Liquid extraction surface analysis mass spectrometry for protein analysis directly from *Escherichia coli* and *Staphylococcus epidermidis*. ASMS Annual Conference 2016.

Placements

Experiments described in Chapter 3, encompassing six months of work, were carried out as part of a placement at the National Physical Laboratory.

CHAPTER 1

General introduction

1.1 Introduction to mass spectrometry

Mass spectrometry is an analytical technique which is used to determine the chemical composition of a sample based on the masses of its components.[1] The components of the sample are ionised and subsequently separated according to their mass-to-charge ratio (m/z).

The general layout of a mass spectrometer consists of an ion source, a mass analyser and a detector.

1.1.1 Ion sources

1.1.1.1 Overview

In order to be detected in a mass spectrometer, the analytes contained in the sample first need to be transferred to the gas phase and ionised. The ionisation step is crucial due to the basic operation of the instrument as the manipulation of analytes within a mass spectrometer is achieved via the application of radio-frequency (RF) voltages, generating oscillating electric fields.

In early mass spectrometric experiments, the two steps required for the formation of gas-phase ions, vaporisation and ionisation, were separated and sequential (vaporisation preceding ionisation).[2] Analytes would be subjected to heating in order to generate a vapour and subsequently ionised by collision with energetic electrons in a process termed electron (impact) ionisation (EI).[3] The vaporisation step limited analysis to relatively low molecular weight, volatile compounds. The process led to extensive fragmentation of larger analytes, both during transition to gas phase and upon ionisation due to the high energy (most

commonly 70 eV) of the electron beam (this was termed hard ionisation). Chemical ionisation (CI) partly circumvented this issue. Rather than affecting the analyte directly, the energetic electron beam was directed at a readily ionisable reagent which would then pass on its charge to the analyte in a less disruptive manner; it is thus an example of a soft ionisation technique.[4] Atmospheric pressure chemical ionisation (APCI), introduced in 1973,[5, 6] relied on the chemical ionisation of a vaporised analyte, similarly to CI. Unlike CI, which was performed in vacuum, APCI was carried out under atmospheric pressure. Analytes were delivered to the source in a continuous flow of heated solvent (making it suitable for coupling to liquid chromatography), which was nebulised by a flow of neutral gas. Molecules of gas and solvent were ionised by corona discharge; these ions would then transfer charge to the analytes similarly to CI. An alternative, termed atmospheric pressure photoionisation (APPI), which replaced the corona discharge with a UV laser (photon energy in the 10 eV range), was developed more recently.[7, 8] Thermal decomposition of larger analytes remained, however, unresolved.

Secondary ion mass spectrometry (SIMS), relying on the desorption and ionisation of solid samples through ion bombardment, was developed as a means of directly analysing solid surfaces; the desorption and ionisation steps were coupled in time and space.[9] Similarly, fast atom bombardment (FAB) allowed the ionisation of non-volatile, thermally unstable compounds.[10] For FAB and a variant of SIMS termed liquid SIMS (L-SIMS), the analytes would first be mixed with a matrix compound on the support surface. Analytes were subsequently transferred to the gas phase and simultaneously ionised directly from the solid surface by collision with energetic neutral atoms (FAB) or ions/ion clusters (SIMS) at very short timescales (secondary ion emission achieved in $<10^{-12}$ s). Thus, thermal decomposition of the analytes was avoided and primarily singly charged intact ions were generated, classing this

approach as a soft ionisation technique, though some fragmentation could still occur upon ionisation. Whilst fragments could be used for structural elucidation, the occurrence of in-source fragmentation limited the molecular weight of intact analytes detectable by the two techniques. Intact mass analysis of peptides by FAB up to 4.5 kDa was reported;[11, 12] an upper limit of approximately 25 kDa was determined by use of a magnetic sector instrument, though sensitivity for the molecular ion at high molecular weights was poor.[13] Whilst SIMS has traditionally been used to probe very small analytes (including their isotopic compositions), the recent introduction of large cluster ion beams (C_{60}^+, Ar_n^+), combined with fine control over the kinetic energy of the system, enabled the analysis of intact molecular ions of proteins,[14] including insulin and cytochrome c at 5.8 kDa and 12.3 kDa, respectively.[15]

Similarly, thermal spray (TS) ionisation[16] developed for coupling to liquid chromatography and plasma desorption (PD) ionisation[17-19] available on time-of-flight instruments reported upper molecular weight limits of 8 kDa[20] and 45 kDa,[21] respectively. In all three cases, the molecular weight limit was determined by the fragmentation of the analytes inherent to the ionisation processes, and the corresponding decrease in sensitivity for the molecular ion.[22]

The analysis of larger species was thus dependent upon the generation of intact gas-phase ions without in-source fragmentation; that is, the energy supplied to the analytes should not significantly exceed the energy required for ionisation. Two such soft ionisation techniques, matrix-assisted laser desorption ionisation (MALDI)[23] and electrospray ionisation (ESI)[24] were developed in the 1980s and allowed for the first time the analysis of intact large biomolecules, including proteins, which are the focus of this work.[25] Other techniques compatible with large analytes, such as surface-assisted laser desorption ionisation (SALDI) of frozen protein solutions on solid supports (in a non-imaging capacity),[26] matrix assisted ionisation (MAI), where the analytes are coated with a matrix but ionisation occurs without

laser irradiation,[27] and a number of ambient ionisation techniques (summarised in section 1.1.1.4) have since been developed.

ESI, in the form of chip-based nanoelectrospray ionisation (described in section 1.1.1.3), was used for the majority of this work (Chapters 4-6); MALDI (described below) was used for work described in Chapter 3.

1.1.1.2 MALDI

MALDI (**Figure 1.1**) is primarily a vacuum technique developed as an improvement to infrared and ultraviolet laser desorption mass spectrometry, two methodologies which proved notoriously difficult to standardise across different instruments.[23] Unlike the approaches it superseded, MALDI requires the deposition of a matrix compound onto the target sampling surface. The coated sample is then irradiated with a laser beam, which interacts with the matrix molecules, desorbs the analytes and directs them into the mass spectrometer.

Multiple models have been proposed to explain the mechanism of MALDI.[28-33] Over time, the hypotheses converged on the so-called two-step framework,[29] postulating that the ionisation process consists of two separate events following analyte desorption: the formation of primary ions or ion clusters and their subsequent reaction with neutral molecules in the gas plume which produces the secondary ions detected in the mass spectrometer. The process tends to generate singly charged species which, along with the pulsed nature of signal generation, makes this ionisation technique particularly suitable for use with time-of-flight (TOF) mass analysers (described in section 1.1.2.1). Its use in conjunction with other mass analysers is, however, limited by the requirement for a high mass range.

Whilst the general framework of MALDI is now well understood, further refinements to the model focused on the mechanism of formation of primary ions, the details of which remain

contested by several groups.[34] The two major models currently proposed for the formation of primary ions are the coupled physical and chemical dynamics (CPCD) model[33] and the thermal proton transfer model, based on the polar fluid model.[35] The major source of discrepancy between the two models revolves around the predicted ion-to-neutral ratio, or ion yield, and its agreement with observations for each model; of the two, the CPCD model predicts higher ion yields.[36] Whilst the agreement of predicted ion yield and laser fluence appears to favour the CPCD model,[36] the direct relationship of ion yield and temperature strongly indicates the involvement of thermal factors in the mechanism.[37] The extent to which thermal processes influence ion yields, as well as the methods of measurement of the ion yield for matrices and analytes, were subsequently contested.[38-42] Neither model currently appears to account for all phenomena observed in MALDI, necessitating further research.

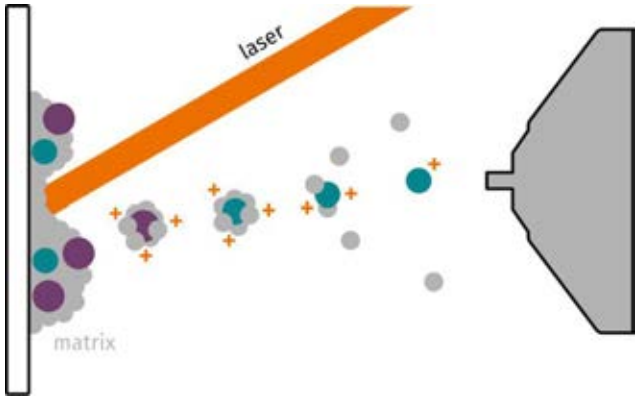


Figure 1.1. The general principle of MALDI. Sample co-crystallised with matrix is irradiated by a laser beam. Analyte and matrix are desorbed and ionised upon irradiation, following which the gas-phase ions are transferred into the mass spectrometer.

It was demonstrated that the sample preparation, ionisation and analysis parameters could be tuned to generate a reliable signal from intact proteins.[43] Protein ions in excess of 10 kDa were observed, predicting a near-unlimited theoretical mass range in the MALDI-TOF configuration. Refinement of this approach led to the widespread adoption of MALDI-TOF MS as a tool for the identification of proteins by peptide mass fingerprinting.[44] More recently,

advances in sample preparation and detection technology allowed the profiling of large (≥50 kDa) proteins directly from biological samples[45-47] as well as non-covalently bound protein complexes.[48-50] Thus far, protein aggregates of up to 400 kDa have been reliably ionised and detected as calibrants on commercially available instruments.[51] By incorporation of superconducting tunnel junction cryodetection, dimers of horse spleen holoferritin were observed at 1.63 MDa.[52] Achievement of such results requires, however, careful sample preparation and access to a TOF instrument; moreover, ionisation occurs in vacuum where the conformation of proteins may be disturbed. ESI (described below) is capable of achieving similar outcomes without mass analyser limitations and in atmospheric pressure.

1.1.1.3 ESI

The mechanism of ESI relies on the application of an electric potential between the inlet of the mass spectrometer and the sample (**Figure 1.2**).[24] Analytes are placed in solution, which is subsequently fed by use of a syringe through an emitter directed at the inlet of the mass spectrometer. A difference in electric potential is applied between the solution and the inlet of the mass spectrometer. The difference in electric potential, aided by backpressure applied by the syringe pump, draws the solution out of the emitter. The shape of the liquid film formed at the tip of the capillary by surface tension is distorted by the electric field, forming a Taylor cone.[53] Microscopic, charged droplets separate from the tip of the cone and subsequently travel towards the inlet of the mass spectrometer.

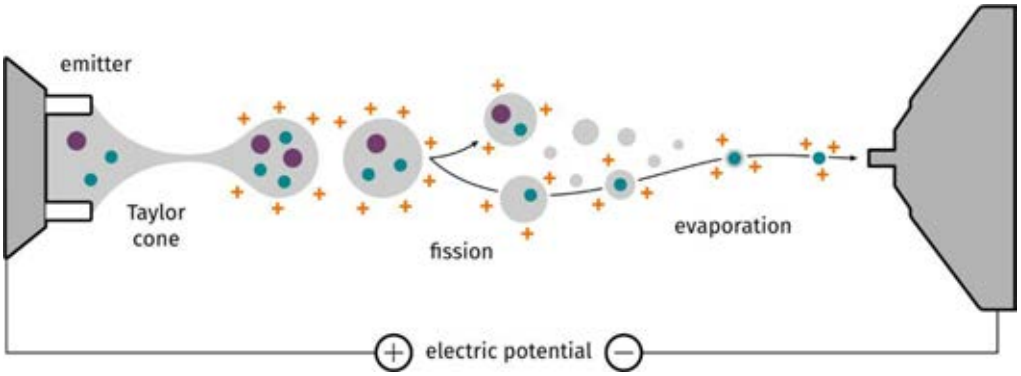


Figure 1.2. The charged residue model of ESI.

According to the most widely accepted model applied to large molecules including natively folded proteins, termed the charged residue model,[54-57] the droplets evaporate through their interaction with the surrounding air and shrink during the flight towards the inlet of the mass spectrometer. The charge accumulated on the surface of the droplet eventually becomes too great for the size of the droplet (the Rayleigh limit is reached), which causes fission of the droplet. The cycle of evaporation and fission repeats until the solvent evaporates completely and the charge carried by the droplet is passed directly onto the analyte within. The process thus generates multiply charged species.[58]

Two other models of ESI, the ion evaporation model and the chain ejection model, have also been proposed. The ion evaporation model is applied to low molecular weight analytes.[59] Solvent evaporation proceeds as in the charged residue model, until the Rayleigh limit for the droplet is reached. An ion carrying the excess charge is subsequently emitted from the droplet. Molecular dynamics simulations have shown that a solvent bridge extends from the parent droplet to the ejected ion; upon rupture, the remnants of the solvent bridge form a solvation shell around the ion,[60] which is subsequently lost upon entry into the mass spectrometer and collision with background gas.[61]

The chain ejection model is applied to disordered polymers, including fully denatured proteins.[62, 63] In an acidic solvent environment, globular proteins unfold. The hydrophobic, nonpolar core of the protein is exposed to solvent, which gives rise to unfavourable interactions with the polar solvent molecules in the interior of the droplet. Thus, the hydrophobic protein chain is rapidly removed to the surface of the droplet. Ejection of the protein chain into the gas phase proceeds sequentially, starting with one terminus of the protein chain protruding out of the droplet and continuing through a series of extended intermediates until the entire chain is expelled. This has recently been confirmed by experimental ion mobility observations.[64]

ESI provides several advantages over MALDI. No vacuum is required for its operation, which renders it compatible with a wide variety of instruments to which it can be externally coupled.[65] The continuous manner of sample delivery, contrasted with the pulsed MALDI signal, further enables its coupling to separation methods such as liquid chromatography, thus making the technique more flexible.[65, 66] The multiple charging of analytes allows for a variety of fragmentation methods to be applied to targeted precursors, facilitating their identification by tandem mass spectrometry (see section 1.2.2); it also enables the analysis of macromolecules in mass analysers of a limited *m/z* range (see section 1.1.2). Sample preparation is straightforward and consists, in principle, of dissolving the analyte of choice in a volatile solvent. This latter step can be performed in a variety of ways and may be easily automated by use of a suitable platform; multiple methodologies were consequently devised to directly couple the sample extraction and electrospray ionisation steps. This contributed to the development of an array of ambient ionisation techniques (described in section 1.1.1.4).

A miniaturised variant of ESI, termed nanoelectrospray ionisation (nanoESI) was subsequently developed.[55, 67] In the classical nanoESI apparatus, a glass capillary pulled into

a fine needle with an orifice diameter not exceeding 10 μm is used. The capillary is loaded with several microlitres of sample. A potential difference is applied through a conductive coating on the capillary, or through a fine wire inserted directly into the solution. Unlike ESI, in which the flow rate is controlled mechanically by a syringe feeding liquid into the ESI nozzle, the flow rate of nanoESI is driven by the voltage applied between the capillary and the inlet of the mass spectrometer, and does not exceed 50 nl/min. Sensitivity is increased by a combination of the low flowrate (resulting in a reduced sample consumption per experiment) and the reduced size of the nanoelectrospray plume, which reaches the inlet of the transfer capillary in its entirety; thus, the approach is ideal for samples which are only available in small quantities.

1.1.1.4 Ambient ionisation techniques

The term "ambient ionisation technique," as defined in the latest available edition of the IUPAC recommendations for terminology related to mass spectrometry,[68] refers to any technique permitting the "desorption of molecules and formation of ions outside the mass spectrometer directly from samples in their native environment with no or minimal sample preparation".[69-79] This definition has been a subject of considerable debate since its introduction and was most recently challenged by Venter and coworkers, who propose to redefine minimal sample preparation as "proximal and in real time with the ionization and analysis step".[80, 81]

The lack of extensive prior sample preparation distinguishes ambient ionisation techniques from MALDI and other similar approaches. Upwards of 40 ambient ionisation techniques have been described to date.[82, 83] According to Monge et al, these can be broadly categorised according to the mechanism of ion desorption into solid-liquid extraction, plasma, thermal/mechanical two-step, laser, acoustic and multimodal techniques, in addition to other,

unique approaches which defy categorisation.[82] A number of ambient mass spectrometry methods, particularly in the solid-liquid extraction class, have proven ideal for mass spectrometry imaging applications (described in section 1.1.1.5); these include desorption electrospray ionisation (DESI)[84] and a suite of liquid microjunction-based techniques, including nanospray desorption electrospray ionisation (nanoDESI),[85] the liquid microjunction surface sampling probe (LMJ-SSP)[86, 87] and liquid extraction surface analysis (LESA).[88] These four techniques, selected based on their suitability for the analysis of proteins and the history of their use for the analysis of bacteria, are described below. DESI and LESA were applied in this work, LESA being its main focus.

1.1.1.4.1 DESI

DESI was one of the first ambient ionisation techniques which evolved from ESI.[84] It requires samples to be deposited on a hard, non-conductive surface such as a glass slide. An electrospray jet is then directed at high pressure at the sample surface (**Figure 1.3**). Microdroplets of solvent impact the sample surface and desorb gaseous ions of analyte through a combination of pneumatic and electrostatic forces. Ions are collected by an atmospheric pressure ion transfer capillary, which delivers them to the inlet of the mass spectrometer for analysis.[89]

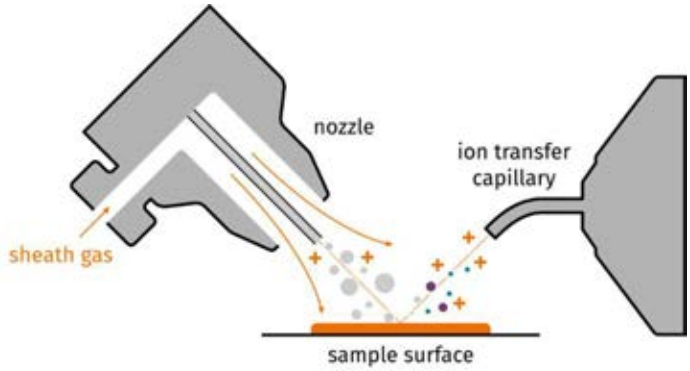


Figure 1.3. The mechanism of DESI.

By careful adjustment of the nozzle geometry and solvent flow rate under optimum conditions, sampling spot diameters of <50 μm are achievable, making this approach useful for mass spectrometry imaging applications.[90] The standard technique allows the desorption and detection of purified protein standards up to approximately the mass of bovine serum albumin (66 kDa) under optimum conditions, although the instrumental response decreases with increasing analyte sizes due to the formation of analyte and analyte-contaminant clusters.[91] The approach initially suffered from a restricted effective mass range in complex matrices as compared to liquid microjunction-based techniques, favouring its application for the analysis of small molecules and lipids. The incorporation of a wash step combined with the optimisation of source geometry enabled the detection of proteins in the 10-15 kDa range directly from biological samples.[92] Native mass spectrometry of purified protein complexes up to 800 kDa was also demonstrated by use of a modified ion source, where the ion transfer capillary was removed and the entire assembly was shifted directly to the inlet of the mass spectrometer.[93]

1.1.1.4.2 NanoDESI

NanoDESI (mechanistically unrelated to DESI) is an alternative technique which relies on the formation of a solvent junction between the sampling apparatus and the sample surface.[85] The apparatus consists of two capillaries held at an angle with respect to each other (**Figure**

1.4). A small gap is left between the capillaries. Solvent is continuously fed through the first capillary, at the end of which it spills onto the sample in a controlled manner, forming the liquid microjunction. It is aspirated into the second capillary by application of an electric potential between the capillary and the inlet of the mass spectrometer; the electric potential also drives the formation of a Taylor cone at the end of the second capillary and the subsequent ionisation of analytes by nanoESI.

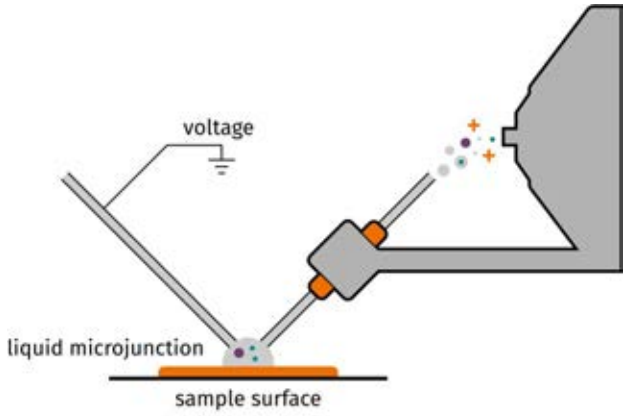


Figure 1.4. Schematic representation of a nanoDESI apparatus.

The main application of the technique is mass spectrometry imaging of biological samples, primarily thin tissue sections and microbial colonies.[94-101] Analysis of proteins up to 15 kDa in molecular weight has also been demonstrated.[94] By adjusting the size of the gap between the two capillaries, the size of the liquid junction can be decreased to approximately 10 μm in diameter as measured on a rhodamine standard grid, yielding the best spatial resolution for imaging available among liquid sampling techniques.[97] Pixel diameters of 20 μm or better were measured on tissue sections.[96, 97] Whilst such resolution is not trivial to achieve on uneven surfaces such as microbial colonies, the recent integration of a shear probe providing automatic height adjustment has solved this issue.[102, 103]

1.1.1.4.3 LMJ-SSP

The liquid microjunction surface sampling probe (LMJ-SSP) is another continuous flow design which has its roots in an automated thin layer chromatography (TLC) plate-reading system. [86] The probe consists of two coaxial capillaries, measuring approximately 600 μm in outer diameter (**Figure 1.5**). The outer capillary contains solvent flowing down towards the sample surface at a rate between 10 and 60 μl/min; the inner capillary withdraws the solvent and delivers it to a pneumatically aided sprayer attachment for delivery by electrospray ionisation. Unlike nanoDESI, the design is not self-aspirating and therefore the two flow rates are controlled independently to adjust the size and depth of the microjunction at the tip of the two capillaries.

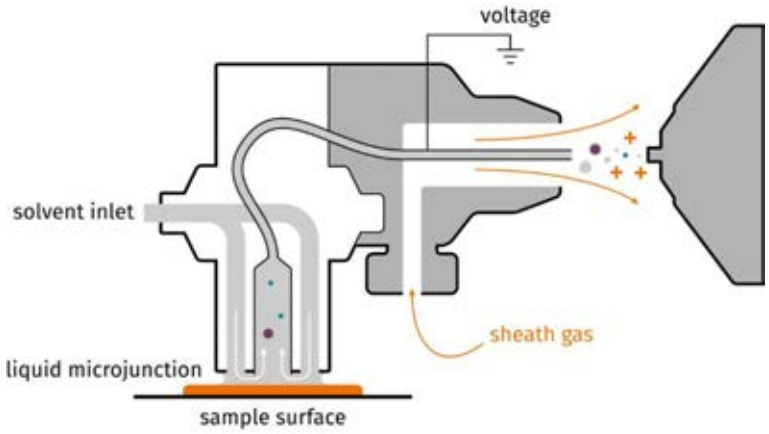


Figure 1.5. Schematic representation of the LMJ-SSP.

The dynamics of the fluid in the liquid microjunction affect extraction efficiency, introducing variability into the system; this can be partially controlled by altering the geometry of the probe.[104] As for nanoDESI and other continuous flow designs, extraction efficiency can be increased by dwelling at a single sampling location for extended periods of time. Sensitivity is limited by the continuous dilution of analytes during sampling. As an additional challenge to protein extraction in the current commercial iteration of the apparatus (Prosolia Flowprobe), acetonitrile-based solvent systems optimised for protein analysis are

incompatible with the polyimide coating of the capillaries, which swells upon exposure to acetonitrile and impairs solvent flow.[105, 106] Despite this limitation, proteins up to 15 kDa have thus far been observed.[106, 107]

1.1.1.4.4 LESA

Liquid extraction surface analysis (LESA) is a technique based on the formation of a liquid microjunction on the sample surface. Developed by Kertesz and Van Berkel in 2010, it was designed primarily for spot sampling, addressing the primary disadvantage associated with continuous sampling approaches, which is the continuous dilution of analytes during sampling, decreasing sensitivity for low abundance species.[88] The description of LESA operation principles and performance supplied here consists entirely of information available prior to the start of this project unless otherwise noted.

The technique makes use of a commercially available Advion TriVersa NanoMate robotic pipette system (**Figure 1.6**). The sampling process starts with the withdrawal of a small volume of solvent, either from a dedicated reservoir or a well plate (**Figure 1.7**). The volume of solvent used is limited to 0.5-10 µl by the calibration of the system but typically ranges from 1-3 µl. The solvent droplet is deposited on the sample surface and left in contact with the surface for several seconds, allowing for the diffusion of analytes; dwell times of 1-3 seconds are used for volatile solvent systems but can be increased to 60 seconds or longer for native solvent systems, contributing to increased extraction efficiency. The droplet is then re-aspirated and introduced into the mass spectrometer via nanoelectrospray ionisation.

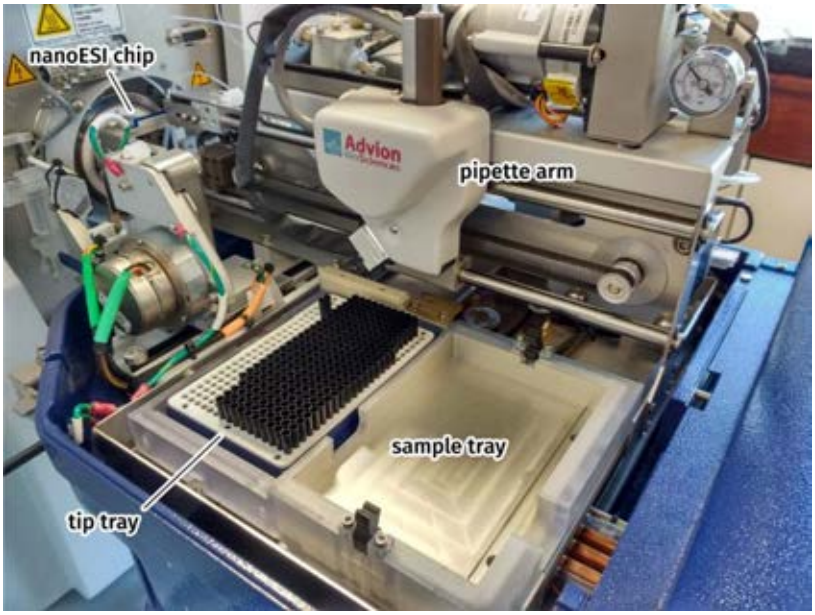


Figure 1.6. The TriVersa NanoMate robotic pipette system.

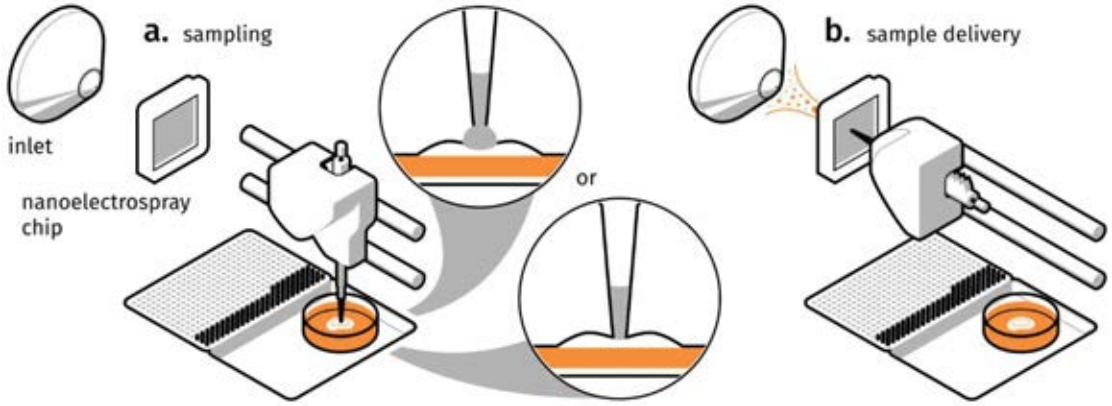


Figure 1.7. LESA carried out by use of a TriVersa NanoMate (Advion). **a**. Sampling. Analyte extraction can be achieved by the formation of a liquid microjunction on the surface of the sample (top inset) or by driving the extraction pipette tip into contact with the sample surface (bottom inset). **b**. Delivery of the extract into the mass spectrometer. The robotic pipette of the NanoMate swivels and engages with the nanoelectrospray chip for sample delivery. Voltage is applied to the chip.

Whilst this setup is incapable of performing raster mode imaging (see section 1.1.1.5), it offers significantly increased sensitivity as there is no continuous solvent flow diluting the analytes.[108] The relatively large surface area samples (see below) may also contribute to increased sensitivity.[109] As with other solvent-based techniques, it allows both the analysis of multiple classes of compounds at the same time and the tailoring of the system to a specific

class.[69] The sensitivity of the technique specifically with regards to small molecule analytes has recently been exploited for the purposes of drug localisation[110, 111] and quantification in imaging experiments.[112] Conversely, LESA was also successfully applied to very large analytes extracted under native-like conditions. Proteins up to 64 kDa (haemoglobin tetramer) were detected directly from biological substrates,[113, 114] whilst the GroEL complex (800 kDa) was the largest protein complex detected from a purified standard spotted onto a glass slide.[115] Detection of trimeric AmtB (140 kDa), a membrane protein encased in detergent micelles, was also accomplished;[115] crucially, whilst the capability of DESI to analyse such large assemblies and membrane proteins has also been shown, it required substantial modifications to the apparatus, eliminating the transfer capillary and moving the desorption/ionisation process directly to the inlet of the mass spectrometer. LESA, by contrast, requires no such modifications.

Spatial resolution is poorer than that achievable by DESI, nanoDESI or the LMJ-SSP due to the relatively large volumes of solvent forming the microjunction, as well as the limited ability of the NanoMate robot to move the pipette arm in small increments, restricted to 1 mm in the *x*, *y* plane by software and to 0.2 mm on the *z* axis by the precision of the mandrel stepper motor. The exact effective resolution was investigated using MALDI and optical imaging to analyse LESA-sampled tissue sections.[116] Sampling was performed in two modes: the standard (non-contact) mode and contact mode. For standard, non-contact sampling, the droplet of solvent was suspended from the pipette tip, forming a visible liquid microjunction at the surface of the sample. For the alternative, contact sampling mode, first applied to bacterial colonies,[117] the sampling height was decreased until the pipette tip came into contact with the sample surface prior to solvent deposition. Thus, the size of the liquid junction could theoretically be limited by the dimensions of the pipette tip (750 μm outside

diameter, 400 μm internal diameter) rather than by uncontrolled solvent spread. The spot diameter for the standard sampling routine averaged 1158 μm for 0.5 μl of solvent dispensed, as determined by optical imaging. The diameter of the sampling spot, as determined by optical imaging, decreased to approximately 690 μm in line with expectations. The average contact sampling area determined by MALDI imaging was 0.41 mm^2. MALDI imaging, however, also revealed a sampling precision issue in the spacing between the individual sampling spots (measured centre to centre), which varied between 0.8 and 1.3 mm.

In addition to the potential increase in spatial resolution, the contact LESA approach was shown to be particularly helpful on hydrophilic surfaces, including bacterial colonies (in which case it was the only LESA approach capable of protein detection), as well as for sampling using highly volatile solvent systems.[118] The extent of solvent spread within three-dimensional samples probed by the contact approach (as it the case for bacterial colonies) and thus the volume sampled in such manner is not currently known.

A summary of the techniques described in the previous four sections is provided in **Table 1.1**.

	DESI	NanoDESI	LMJ-SSP	LESA
spatial resolution (μm)	50	10-20	600	690
protein extraction	15 kDa (biological substrate) 800 kDa (purified standard, modified source)	15 kDa (biological substrate)	15 kDa (biological substrate)	64 kDa (biological substrate) 800 kDa (purified standard)

Table 1.1. A summary of the four ambient ionisation techniques in terms of protein extraction and imaging.

1.1.1.5 Imaging mass spectrometry

Many ionisation techniques, including MALDI and FAB/SIMS, as well as most ambient ionisation techniques, can be used in a mass spectrometry imaging workflow to study samples deposited on a surface. Mass spectrometry imaging is performed by dividing the sample surface into a grid of locations (or pixels) and acquiring a mass spectrum in each location.[119] The spatial distribution of analytes across the sample can thus be determined.

Imaging can be performed in two modes, spot mode and raster mode. Spot mode involves discrete sampling and imaging of each location. The alternative raster mode involves the continuous movement of the sampling apparatus across the sample surface (or the analogous movement of the sample stage underneath the sampling apparatus) such that continuous line scans, rather than individual pixels, are sampled at a time.[120, 121] Individual pixels are subsequently resolved from the line scans by correlating the velocity of the sampling apparatus and the mass spectrum acquisition times. DESI and nanoDESI are, by design, capable of a raster mode approach as the flow of solvent is not interrupted between adjacent sampling locations. The registration of sampling position and data acquisition is straightforward as desorbed analytes reach the inlet of the mass spectrometer almost instantaneously. By contrast, Flowprobe raster imaging required significant optimisation and software development due to the presence of a long transfer capillary between the probe end and the electrospray attachment, which desynchronised sampling and corresponding data acquisition.[106]

FAB and (L-)SIMS provide the highest achievable imaging resolution in mass spectrometry (sub-100 nm for nanoSIMS)[122] at the cost of low signal intensity (the primary ion beam only interacts with and ionises a low number of analytes at a time). Depth imaging is possible through controlled ablation of sample layers by use of the primary beam or a dedicated,

secondary ion gun, yielding a sub-10 nm depth profiling resolution.[123, 124] Resolution on organic samples can be further increased by conducting experiments on frozen samples in cryogenic conditions (cryo-SIMS), thus providing chemical and structural stabilisation of the samples and allowing for sub-cellular imaging.[125, 126] The two techniques are not well suited to the analysis of proteins due to the relatively high energies involved, which leads to concomitant fragmentation of larger analytes; imaging of intact proteins has not been reported. Protein imaging has been extensively performed by MALDI MS, primarily from tissue sections.[45, 47, 127-131] For ambient imaging analysis of intact proteins, liquid microjunction-based techniques have been dominant. NanoDESI has been shown to enable imaging of proteins up to 15 kDa from tissue sections;[94] the Flowprobe achieved similar results.[106] LESA was applied for protein imaging from tissue in a workflow including high-field asymmetric waveform ion mobility spectrometry (FAIMS; see section 1.2.2);[132] it was also used on its own to produce the first reported images of natively folded proteins directly from biological substrates.[113, 114] Thus far, DESI is the only non-liquid microjunction-based ambient technique which has been successfully used for protein imaging.[92]

1.1.2 Mass analysis

1.1.2.1 Overview

Mass analysers separate ions according to their mass-to-charge ratios (*m/z*). The three types used in this work were the time-of-flight (TOF) mass analyser, the linear quadrupole mass analyser and the Orbitrap. The Orbitrap (as part of the LTQ-Orbitrap Elite and Velos mass spectrometers) was used for the majority of the work described (see Chapters 3 to 6). The TOF mass analyser (as part of the Synapt G2-Si mass spectrometer) was used for MALDI and DESI work described in Chapter 3. The linear quadrupole forms part of the LTQ-Orbitrap

Elite and Velos systems, as well as the Synapt G2-Si TOF mass spectrometer, as an ion guide and mass analyser.

1.1.2.2 The time-of-flight mass analyser

A schematic diagram of a time-of-flight (TOF) mass analyser is provided in **Figure 1.8a**. Ions are pulsed from the ionisation source into an acceleration region (delineated by the inlet of the instrument at the front end and an extraction grid at the other end) to which an electric field is applied.[133] The electric field imparts kinetic energy on the ions and accelerates them through the extraction grid into a field-free flight region or flight tube. The time of flight through the flight region is directly related to the mass-to-charge ratio of a given ion; ions of the lowest *m/z* have the highest velocity and thus reach the detector first.

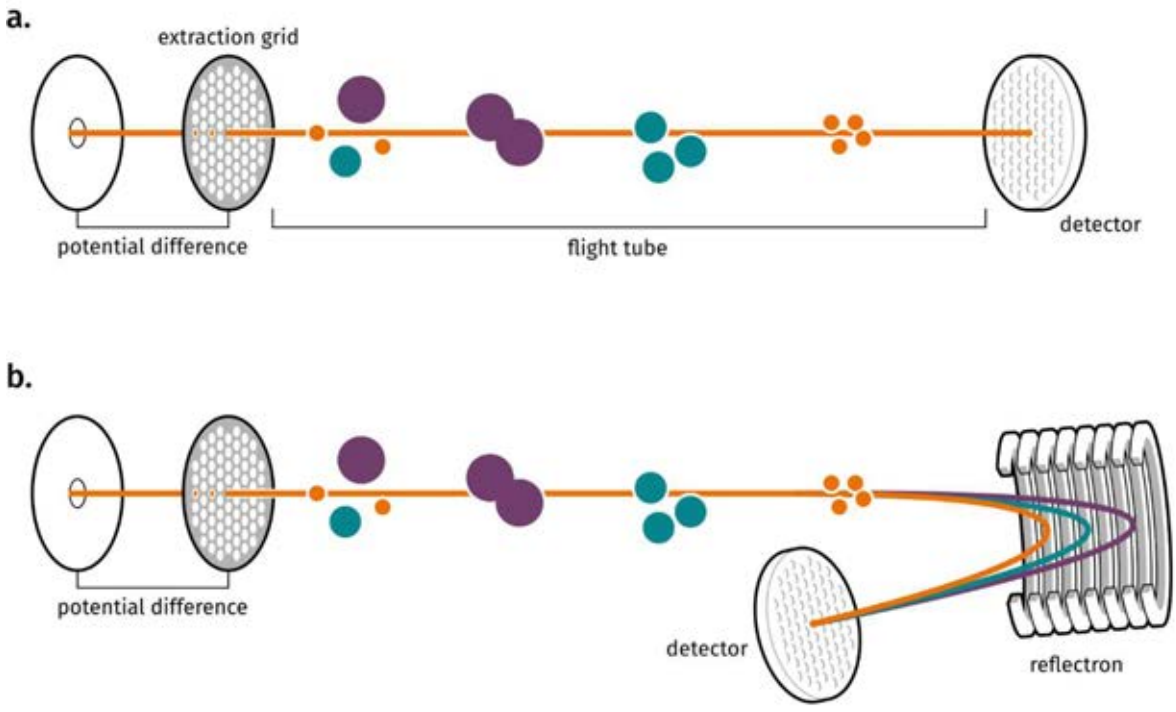


Figure 1.8. Schematic representation of a TOF instrument. **a**. Linear TOF. **b**. Reflectron TOF.

Time-of-flight mass analysis relies on several fundamental principles. First, if a range of ions of different masses (m) have the same kinetic energy (E_k), ions of lower mass will have a higher velocity v (**Equation 1.1**):[1, 134]

$$m = \frac{2E_k}{v^2} \tag{1.1}$$

Velocity can be expressed as a function of distance d and time t, yielding **Equation 1.2**:

$$m = \frac{2E_k}{d^2}t^2 \tag{1.2}$$

Kinetic energy is imparted onto the ions by a homogeneous electrical field in the acceleration region of the instrument. Thus, expressing E_k as a function of the acceleration potential U and charge z yields the quadratic relationship between the mass-to-charge ratio of a given ion and its flight time (**Equation 1.3**):

$$\frac{m}{z} = \left(\frac{2eU}{d^2}\right)t^2 \tag{1.3}$$

The mass resolution is affected by two factors. First, whilst in theory all ions of the same *m/z* are imparted with the same kinetic energy in the acceleration region, the energy is in fact distributed across a narrow range of values, resulting from the differences in the time of formation of the ions, their initial position in the acceleration region, and the variability in initial kinetic energies prior to acceleration.[135] This variation introduces variation in the flight times and thus broadens the peak corresponding to a given *m/z*. In order to increase the mass resolution, particularly for low molecular weight analytes, two alterations can be made to the instrument. First, delayed extraction can be implemented.[135, 136] Upon introduction of the ions into the acceleration region, the application of the accelerating potential is delayed so that the ions are spatially distributed within the acceleration region according to their initial

kinetic energies. Thus, once the accelerating potential is applied, ions of a lower kinetic energy (and thus positioned closer to the source) are accelerated at a greater potential than ions of a higher initial kinetic energy located closer to the extraction grid, which only experience the potential for a short time before leaving the acceleration region. Thus, ions of the same mass-to-charge ratio are focused prior to entering the flight region, which decreases the spread of their flight times and increases resolution. Secondly, a reflectron can be implemented (**Figure 1.8b**).[137] The reflectron uses a constant electric field to deflect the path of incoming ions towards the detector. For ions of the same *m/z*, those with a lower energy are deflected earlier than more energetic ions which penetrate deeper into the reflectron field. There is a point on the deflected trajectory, termed the time-of-flight focus, in which the flight times of the lower energy and the higher energy ions of the same *m/z* converge; thus, the distribution of ion arrival times is narrowed, the width of the resultant peak is decreased and thus the mass resolution increases. Additionally, since mass resolution is directly related to the length of the flight path, the reflectron further increases mass resolution by extending the effective length of the flight region. The Waters Synapt G2-Si instrument used in this work (see Chapter 3) is a hybrid quadrupole-TOF instrument fitted with a dual-stage reflectron-ion mirror configuration (**Figure 1.9**).

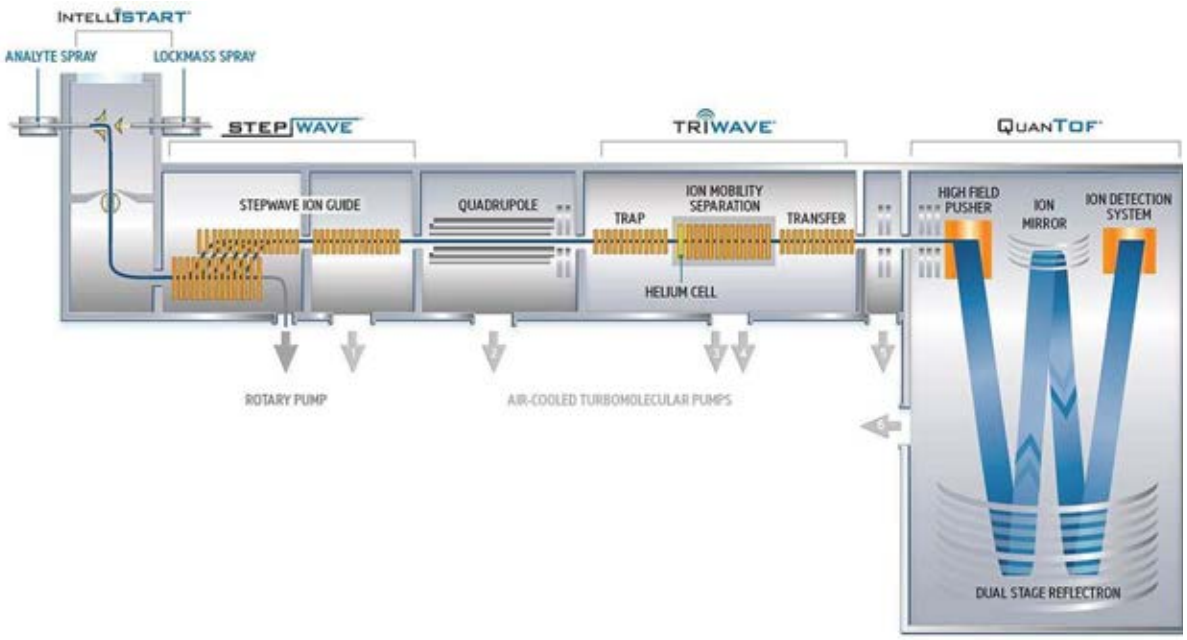


Figure 1.9. Manufacturer schematic of a Waters Synapt G2-Si instrument as used in this work. Figure adapted from http://proteomique.ipbs.fr/front-page/waters-synapt-g2-si/.

1.1.2.3 The quadrupole mass filter/analyser

Linear quadrupoles have been in use as mass analysers since the 1950s[138, 139] and now form an integral part of multiple hybrid mass spectrometer designs, including several instruments used in this work (Orbitrap Velos and Elite, Waters Synapt G2-Si). A quadrupole confines ions within the space between four rod-shaped electrodes operated in opposing pairs (**Figure 1.10**).[138, 139]

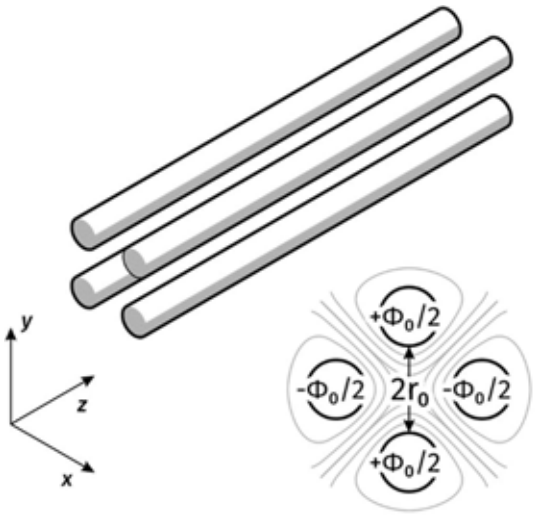


Figure 1.10. Quadrupole in a cylindrical rod configuration, commonly used as an ion guide and mass filter.

The potential applied to the quadrupole rods, Φ_0 is described as **Equation 1.4**:

$$\Phi_0 = U + V\cos\Omega t \tag{1.4}$$

where U is a direct current (DC) voltage, V is a radio frequency (RF) voltage, Ω is the driving frequency and t is time; thus, the potential has both a static DC component and a time-dependent RF component. The motion of the ions within the quadrupole can be mathematically modelled in the form of Mathieu equations. Thus, the ion motion in the x-axis of the quadrupole is described by **Equations 1.5** and **1.6**:

$$a_x = \frac{8eU}{mr_0^2\Omega^2} \tag{1.5}$$

$$q_x = -\frac{4eV}{mr_0^2\Omega^2} \tag{1.6}$$

where a and q are the DC and RF voltage, respectively, m is the mass of the ion and r_0 is the inscribed radius of the ion trap. Similarly, the motion along the y-axis is described by **Equations 1.7** and **1.8**:

$$a_y = \frac{8eU}{mr_0^2\Omega^2} \tag{1.7}$$

$$q_y = -\frac{4eV}{mr_0^2\Omega^2} \qquad 1.8$$

For each axis, there exists a set of a and q values (and hence, a set of DC and RF voltages) which provide bounded, or stable, solutions to the equations (**Figure 1.11**). In RF-only operation (where no DC is applied), such an arrangement can serve as an ion guide; in devices designed as such, the electrodes are usually simple cylindrical rods which adequately approximate the ideal hyperbolic shape whilst being cheap and easy to machine to a high precision. At higher DC voltages, however, the device can be used as a mass analyser: by ramping up the DC and RF voltages proportionally such that $\frac{a}{q}$ remains constant, ions of successive masses can be selectively brought into the narrow window of stability.

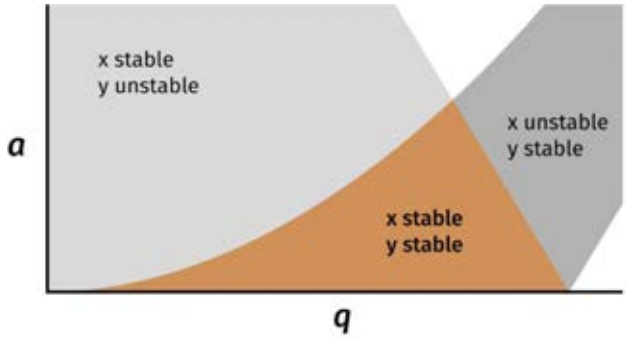


Figure 1.11. Stability diagram for the linear quadrupole.

1.1.2.4 The linear ion trap

A modern linear ion trap shares the basic mechanism of operation with quadrupole ion guides, confining ions in the x and y directions by the application of a quadrupolar field.[140] **Figure 1.12** illustrates the architecture of a linear ion trap as found in Orbitrap Velos and Elite instruments. Each rod of hyperbolic cross-section is split into three sections. The voltages of the front and end sections can be independently varied, functioning as trapping plates. Once the ions enter the trap, the potential at both end sections is ramped up, leaving the centre section at a lower potential. Thus, ions are also confined in the z direction.

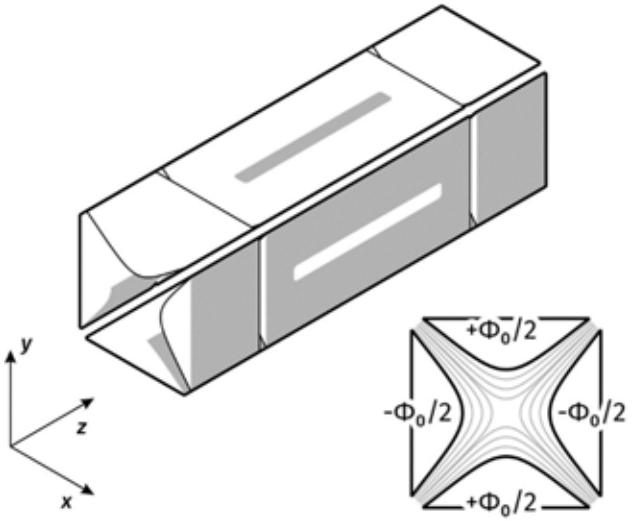


Figure 1.12. Schematic representation of the linear ion trap.

Mass analysis in the linear ion trap is achieved by resonant ejection. An RF voltage is applied to the four rods and linearly ramped. Ions confined in the trap become destabilised in the *x* and *y* directions simultaneously and are ejected through slits placed in the middle sections of the rods. Ejection happens sequentially, in order of increasing *m/z*. Furthermore, by applying a broadband RF voltage omitting the frequency corresponding to an *m/z* of interest, ions of interest can be isolated within the ion trap for further manipulation, including fragmentation.

1.1.2.5 The Orbitrap

The Orbitrap mass analyser was introduced in 2000 by Alexander Makarov and coworkers.[141, 142] The design consists of a central spindle electrode and an outer barrel electrode split in half by a narrow slit (**Figure 1.13**). The ions are injected into the analyser at a trajectory offset from the long axis of the central spindle. In positive ion mode, the potential at the spindle electrode is negative, which attracts the incoming ions and causes their trajectory to bend around the spindle. The electric field within the analyser is subsequently manipulated through the application of a DC ramp to the spindle, trapping the ions within the analyser and establishing their final, tight orbital trajectory. The radial motion of the ions is

randomised by a number of starting conditions. The frequency of the back and forth oscillation along the z-axis, ω, described by **Equation 1.9**

$$\omega = \left(\frac{kz}{m}\right)^{\frac{1}{2}} \qquad 1.9$$

where k is the curvature of the field, is independent of starting conditions and directly related to the mass-to-charge ratio of a given ion. Thus, the originally injected ions split into coherent packets, each one oscillating with a specific frequency determined by its m/z. The measurement of these frequencies as the ions move from one end of the Orbitrap to the other provides the mass analysis capability.

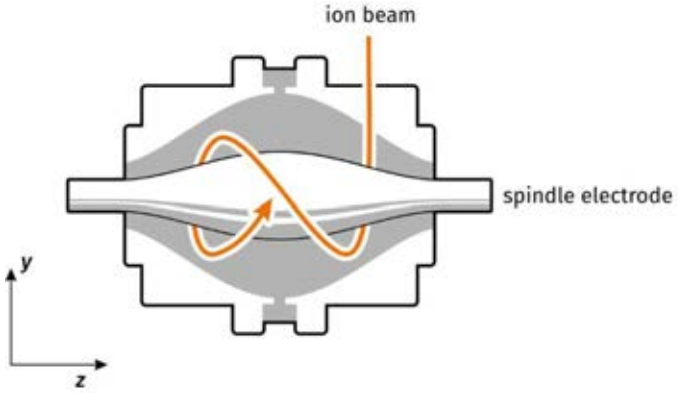


Figure 1.13. Schematic representation of an Orbitrap mass analyser.

A vital element of the orbitrap assembly is the C-trap. The C-trap is a C-shaped quadrupole ion trap consisting of rods of a square cross-section.[142] Ions are ejected from the trap through a slit in the electrode on the inner side of the C-shape. The geometry of the trap focuses the ejected ions into a narrow beam, which can be directly injected into the orbitrap for analysis. The C-trap also serves as an ion guide to transfer ions to the HCD cell (see section 1.3.2) and return fragment ions to the orbitrap for mass analysis.

Schematic diagrams of the Orbitrap Velos and Elite instruments used in this work are provided in **Figure 1.14**.

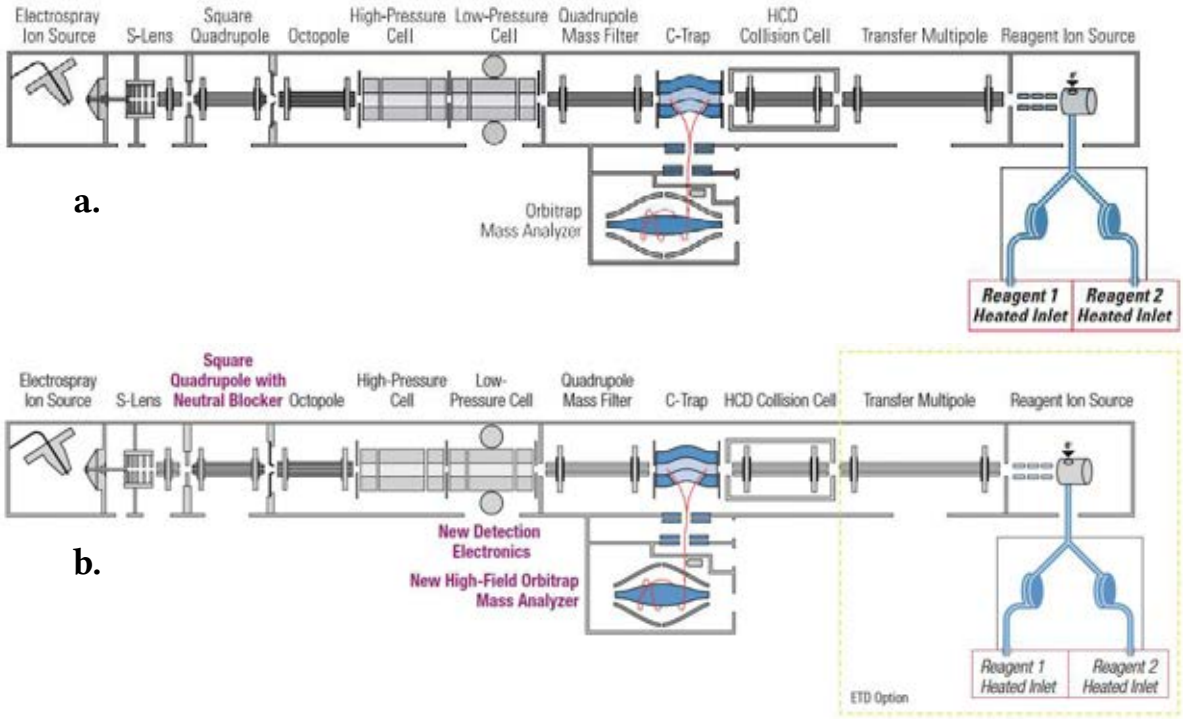


Figure 1.14. Manufacturer schematics of the LTQ Orbitrap Velos and Elite instruments. **a**. Orbitrap Velos. **b**. Orbitrap Elite. The overall layout of the instruments is identical; changes introduced to the Elite are marked in red type. Figures adapted from http://proteomique.ipbs.fr/front-page/thermo-ltq-orbitrap-velos-etd/.

1.1.3 Detection

Acquisition of mass spectrometric data requires the ions to be detected, and their mass-to-charge ratios recorded, following the mass analysis step. Broadly, ion detection can be achieved either directly through impact, or indirectly by detection of transient electrical disturbances induced by the ions without their neutralisation.

In time-of-flight and ion trap instruments, ions are detected directly by their collision with the surface of a flat dynode, where they generate a pulse of secondary electrons. The signal is amplified by an electron multiplier.[1] Microchannel plate detectors, where each channel functions as an individual electron multiplier, are also commonly used.[143] The detected ion current is digitised at set time intervals, thus binning the data during acquisition.

In Orbitrap mass analysers, the frequency of motion for the ions is measured by the induction of an image current on the detector plates.[144] The two halves of the outer barrel electrode fulfil the role of detector plates.[141] The oscillating movement of the ions between the two electrodes induces an image current which records the frequency of this oscillation as a sine wave. All frequencies are measured simultaneously and thus the raw readout, or transient, is comprised of many such sine waves summed together in the time domain. The transient is subsequently separated into the individual frequencies by Fourier transform, yielding a frequency spectrum. Since every frequency corresponds directly to an *m/z* value as shown by **Equation 1.9**, the frequency spectrum can be converted and calibrated to a mass spectrum.

1.2 Ion mobility spectrometry

1.2.1 Overview

Ion mobility spectrometry (IMS) encompasses techniques designed to separate ions according to their size and shape.[145] It is commonly coupled to mass spectrometry to provide the capability for structural elucidation; it can also function as a sample separation mechanism prior to mass spectrometric analysis. Ion mobility spectrometry relies on measuring the mobility of gas-phase ions through an ion mobility chamber in the presence of an inert gas.[146] Collisions between the analyte ion and the gas atoms/molecules determine the mobility of the ion through the device. Analytes of a larger cross-sectional area will experience more frequent collisions; thus, by measuring the mobility of the ion in the drift gas, the general features of its structure can be inferred.[147]

High field asymmetric waveform ion mobility spectrometry (FAIMS) was used in this work (Chapter 5) as a separation technique. It is described below and contrasted with drift tube ion mobility spectrometry (DT-IMS).

1.2.2 Drift tube ion mobility spectrometry

Traditional drift tube or drift time ion mobility spectrometry (DT-IMS), the earliest and most common form of ion mobility spectrometry,[148] measures the time it takes for a given analyte to traverse the length of a drift tube filled with buffer gas in the presence of an electric field.[147] This can be achieved at reduced or atmospheric pressure. The electric field provides the motive force for the ions; their velocity is directly proportional to the strength of the field. Ion mobility, described by the ion mobility constant K, is defined by **Equation 10**:[145]

$$K = \left(\frac{3q}{16N}\right)\left(\frac{2\pi}{kT}\right)^{\frac{1}{2}}\left(\frac{m+M}{mM}\right)^{\frac{1}{2}}\left(\frac{1}{\Omega}\right) \qquad 1.10$$

where q is the charge of the ion, N is the density of the buffer gas, k is Boltzmann's constant, T is temperature, m is the mass of the buffer gas, M is the mass of the ion and Ω is the collisional cross-section of the ion. Crucially, the collision cross section provides information on the ion's size and shape. DT-IMS is the only ion mobility spectrometry technique where this value can be calculated directly from experimental data; all other ion mobility approaches require calibration using known standards. DT-IMS has therefore been extensively used for structural elucidation, including studies of protein conformers and conformational changes.[148, 149]

1.2.3 High field asymmetric waveform ion mobility spectrometry

High field asymmetric waveform ion mobility spectrometry (FAIMS) (also termed differential mobility spectrometry, DMS) is an ion mobility spectrometry technique which separates ions

based on differences in their mobility in high and low electric fields.[150] Ions traverse the device between two parallel electrodes in the presence of a carrier gas (**Figure 1.15**). (Electrodes may have either planar or cylindrical geometry). An asymmetric waveform is applied to one of the electrodes, giving rise to an oscillating electric field which consists of a high field of short duration (known as the dispersion field (DF)) and a low field of long duration; the net electric field in one cycle is zero. The differential mobility of ions in the high and low electric fields results in deflection from their trajectories through the device and leads ultimately to their neutralisation at one of the electrode surfaces; thus, whilst in DT-IMS the electric field provides the motive force and the gas provides the separation, the reverse is true for FAIMS. A DC compensation voltage, giving rise to the compensation field (CF), is superimposed on the asymmetric waveform to selectively correct the trajectories of a subset of ions of particular differential mobility and transmit them through the device.

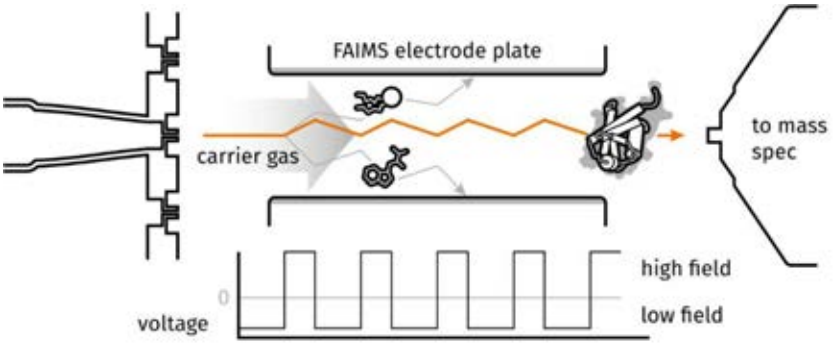


Figure 1.15. The general mechanism of FAIMS.

The mobility of a given ion in high and low electric fields is determined by its size and shape in addition to a number of other properties. Mobility in low electric fields is dependent on the temperature of the gas; in high electric fields, it is additionally dependent on the velocity of the ion, and thus the strength of the field. No model has thus far been proposed to account for all factors influencing the mobility of individual ions and thus FAIMS cannot currently be used for calculation of collision cross sections. The behaviour of ions in an electric field has

been classified into three types: type A, where ion mobility increases with increasing electric field, type C where ion mobility decreases with increasing electric field, and type B, where ion mobility initially increases, but subsequently decreases with increasing electric field (the classification only holds over the range of electric fields applicable to FAIMS).[151] The composition of the carrier gas was found to influence ion mobility in electric fields.[152, 153] It was also found that, with increasing mass (≥100-350 Da), ions tend towards type C mobility;[152] small molecules exhibit type A behaviour. Type B behaviour was observed for several ions of approximately 150 Da (e.g. H^+lysine, 147 Da).[154] Dipole alignment, which would render mobility dependent on the specific collision cross section in the direction orthogonal to dipole orientation rather than on the average rotational cross section, was proposed to account for the uncharacteristic type B mobility of large protein ions (approximately ≥20 kDa).[155-157] Two general models were proposed to account for the listed observations, the clusterisation model (whereby ion-neutral clusters form in the low field portion of the dispersion cycle and rapidly dissociate during the high field portion, amplifying the difference in mobilities)[158, 159] and the rigid sphere scattering model.[160] The effects described by both models are likely to provide significant contributions to ion mobility.[161]

Whilst FAIMS is not currently used for the calculation of collision cross sections, progress has been made to predict other structure-related parameters based on FAIMS data. Notably, the derivation of pK_a/pK_b of drug compounds from their differential mobility data was demonstrated;[162] more recently, machine learning was used to derive further parameters of importance in drug design, such as solubility, polar surface area and cell permeability.[163]

The continuous transmission of ions through a FAIMS device, in contrast to the ion packets produced by conventional ion mobility techniques, render it particularly useful as an ion

filter; Chapter 5 describes the results of coupling FAIMS to LESA mass spectrometry in that capacity.

1.3 Top-down protein analysis

1.3.1 Advantages and considerations

Whilst individual proteins have been extensively studied by mass spectrometric techniques in their intact form, global protein analysis by mass spectrometry has traditionally been carried out by a set of approaches collectively termed bottom-up proteomics.[164, 165] Bottom-up proteomics involves the extraction and enzymatic digestion of proteins of interest prior to their mass spectrometric analysis. Their separation was originally achieved at the intact protein level by two-dimensional electrophoresis, followed by enzymatic protein digestion;[166] a more efficient approach using liquid chromatography for on-line separation at the peptide level was later developed and is now prevalent.[167] Tandem mass spectrometry (MS/MS) is used for the identification of characteristic peptides. Individual peptides are separated, isolated and fragmented to yield sequence information. By matching peptide fragment masses against a database comprised of the sequences of a given organism's complete proteome subjected to *in silico* digestion, the individual proteins originally present in the sample can be identified. The bottom-up LC-MS/MS approach is now considered routine for many applications, such as proteome profiling, protein quantification, as well as the detection and quantification of post-translational modifications (e.g. glycosylation, phosphorylation).[164]

Highly sophisticated instrumentation is not required; a mass range of up to m/z 2000 is sufficient for the analysis of tryptic peptides. Furthermore, due to the technique's maturity, protocols for use with many commercially available systems are are widely accessible in the

literature. In broad terms, the extracted protein mixture is first prepared (this can be accomplished in a number of ways depending on the sample type); pre-fractionation and enrichment for specific protein or modification classes can be done at this stage. Tryptic digestion is then performed. The digested peptides are loaded onto an LC column and separated by on-line liquid chromatography. Eluted peptides are delivered to the mass spectrometer, commonly by ESI, and analysed by MS/MS. At the conclusion of the experiment, MS/MS data are imported into dedicated database-searching software (e.g. Proteome Discoverer) and automatically assigned to putative proteins. Statistical analysis is automatically performed to score the hits and provide a measure of certainty for the identifications.

An alternative, top-down approach involves the analysis of proteins by tandem mass spectrometry of their intact form.[168-170] Extracted proteins are introduced into the mass spectrometer without prior digestion. A separation step (LC) can be included; such an approach is commonly termed top-down proteomics, being analogous to bottom-up proteomics described above. Following the determination of the intact mass by a survey scan, the protein ions are isolated and fragmented within the mass spectrometer and the masses of the fragments are measured (see section 1.3.2). Thus, a list of characteristic peptide masses can be compiled. The process can be performed automatically, using a similar protocol to that used for bottom-up proteomics but tuned to accept larger analytes. The combination of accurate intact mass and fragment masses affords the capacity for protein identification, as well as the identification and localisation of post-translational modifications. Identification of both proteins and their modifications is accomplished by the comparison of experimental data against databases comprised of the intact masses of candidate proteins and theoretical fragment lists generated from the intact protein sequences. Data analysis is done

automatically by use of algorithms such as ProSight[171] or MS-Align+.[172] ProSight (as online-based ProSightPTM 2.0 as well as standalone software ProSightPC versions 3.0 to 4.1 alpha) was used in the work described in Chapters 4, 5 and 6. By contrast, protein identification by bottom-up proteomics relies on the automatic identification of tryptic peptides by intact mass and fragmentation patterns, commonly carried out by database searching software such as Mascot.[173] The intact mass cannot, however, be inferred; information on the existence of sequence deviations or post-translational modifications located along the sequence outside of the detected tryptic peptides is likewise lost.

Top-down protein analysis is, however, technically more challenging than bottom-up analysis. Intact proteins have a higher mass than peptides and thus require instruments capable of achieving a high mass range. Whilst TOF mass analysers have a theoretically unlimited mass range and a high mass accuracy, their sensitivity decreases if a wide *m/z* range is monitored.[174] Fourier transform mass spectrometers such as the Orbitrap, employed for their high mass resolution and mass accuracy, suffer a decrease in mass resolution with increasing *m/z*.[175] High resolution and mass accuracy is, however, necessary for the accurate assignment of high intact masses. The issue can be partially circumvented by use of ionisation sources which generate multiply charged ions (such as ESI and its derivatives), which allows for the analysis of intact proteins within the $<m/z$ 2000 range commonly used for peptides, easily accessible to instruments such as the Orbitrap. Large analytes, however, attain a wider range of charge states, which splits their abundance into a number of peaks. Thus, their analysis requires high sensitivity. Software availability is poorer than that for bottom-up proteomics due to the comparative novelty of the approach and the lack of its routine use; wider implementation is limited by the prohibitive costs of specialised instrumentation. The development of dedicated new technologies for top-down proteomics, such as compatible

high-performance LC columns or high-resolution, high-mass range instruments such as the Exactive EMR (Extended Mass Range, up to *m/z* 20000)[176] Orbitrap instrument, may eventually contribute to the wider adoption of the approach.

Top-down protein analysis can offer more detailed information on the position and connectivity of post-translational modifications, information which may be lost upon proteolytic cleavage.[177] It also offers the possibility of investigating the structure and conformation of intact proteins by native mass spectrometry.[178-180] By extracting proteins into native solvent systems, the tertiary and quaternary protein structure can be retained in solution and transferred into the gas phase, most commonly by ESI. Thus, the stoichiometry of protein complexes, stoichiometry of ligand binding as well as binding affinities can be measured, whilst the coupling of ion mobility spectrometry provides collision cross sections.[180] Native top-down mass spectrometry is by far the least routine approach described here, requiring highly specialised instrumentation and expertise to perform successfully. Separation is not employed; the majority of studies are carried out on purified proteins, although detection of natively folded proteins directly from biological substrates has also been demonstrated.[113, 114] Its instrumentation requirements are similar to those for standard top-down mass spectrometry; the high mass range requirement commonly limits the size of protein complexes which can be analysed, although dedicated developments have recently been made to enable the analysis of megadalton complexes at high resolution.[181] Other elements of the instrument, including the frequency of the quadrupoles and the vacuum pressures, also require consideration to retain native-like structure in the gas phase, preserve binding and deliver intact complexes to the mass analyser.

This project focused on top-down protein analysis by LESA MS without prior sample preparation in order to minimise the analysis time. Top-down proteomics (LC-MS/MS) was included in Chapter 5 strictly to provide a comparison to the method being developed.

1.3.2 Tandem mass spectrometry

Tandem mass spectrometry is a method involving multiple stages of mass analysis, employed for the structural analysis of ions.[1] It involves the selection of a precursor ion, followed by its fragmentation within the mass spectrometer and the subsequent analysis of fragment masses (MS/MS or MS^2). Fragments of the original precursor ion can be further fragmented in MS^n experiments where the architecture of the instrument allows it.[182]

Throughout this work, collision-induced dissociation (CID, also known as collisionally activated dissociation, CAD)[183, 184] and (in a few cases) higher energy collision dissociation, originally termed higher-energy C-trap dissociation (HCD),[185] were used to fragment intact proteins for identification.[186] In CID, precursors are accelerated into a collision chamber containing an inert collision gas, commonly helium. In ion trap-TOF or ion trap-orbitrap hybrid instruments, the ion trap functions as the collision chamber. Collisions between the precursors and the gas molecules results in the conversion of some of the ion's kinetic energy into internal vibrational energy, leading to the cleavage of the weakest bond in the ion along the lowest energy pathway.[187] In peptides and proteins, the peptide bond breaks first, generating *b* and *y* fragments (**Figure 1.16**).[188] The mobile proton model is commonly used to explain the fragmentation pattern observed in CID of peptides.[189, 190] Following collisional activation, a proton is mobilised from one of the basic sites along the peptide to the peptide backbone at either side of the peptide bond (either the oxygen of the carbonyl group or the nitrogen of the amide group). This initiates fragmentation at the peptide bond.

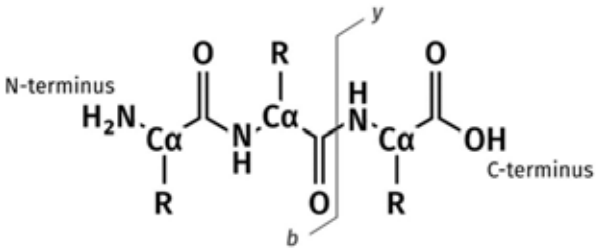


Figure 1.16. The formation of *b* and *y* ions through cleavage of the peptide bond in CID/HCD fragmentation.

Higher-energy C-trap dissociation is a variant of the technique currently unique to orbitrap instruments as used in this work.[185] In the Orbitrap Elite, fragmentation occurs in a dedicated collision cell located behind the C-trap and fragment mass analysis is achieved in the orbitrap. This configuration bypasses the low-mass cut-off characteristic of tandem mass spectrometry as performed in an ion trap, allowing the analysis of smaller fragment ions.[185]

1.4 Microbial colonies as a sample type

1.4.1 Overview

Microorganisms, first independently observed and documented by van Leeuwenhoek and Hooke in the 1650s,[191, 192] are single-celled organisms which are ubiquitously found in all ecological niches on the planet. Whilst bacteria constitute the best-understood microbes, the term also encompasses archaea as well as unicellular fungi and yeasts, thus representing all three kingdoms of life. The work described in this thesis focused on LESA mass spectrometric analysis of Gram-positive and Gram-negative bacteria (Chapters 3 to 5) as well as yeasts (Chapter 6). The structure of microbial cells and microbial communities (termed biofilms or colonies) influences the effectiveness of sampling approaches. An overview of cell and colony structure with respect to the three sample types enumerated above is provided here, followed by a summary of microbiological mass spectrometry.

1.4.2 The structure of microbial cells and colonies

1.4.2.1 Gram-negative bacteria: Escherichia coli

The three classes of microorganisms key to this work were Gram-negative and Gram-positive bacteria, and yeasts. The most relevant differences between these classes of organisms concern the outer structure of the cells, which affects the response to solvent extraction.[193]

The cell envelope of Gram-negative bacteria will be discussed based on the well-studied example of *Escherichia coli*; another Gram-negative bacterium, *Pseudomonas aeruginosa*, was also used throughout the project. The cell envelope of Gram-negative bacteria is comprised of three parts: the outer membrane, the peptidoglycan layer and the inner membrane (**Figure 1.17**).[194] The protein-rich space between the two membranes is termed the periplasm.

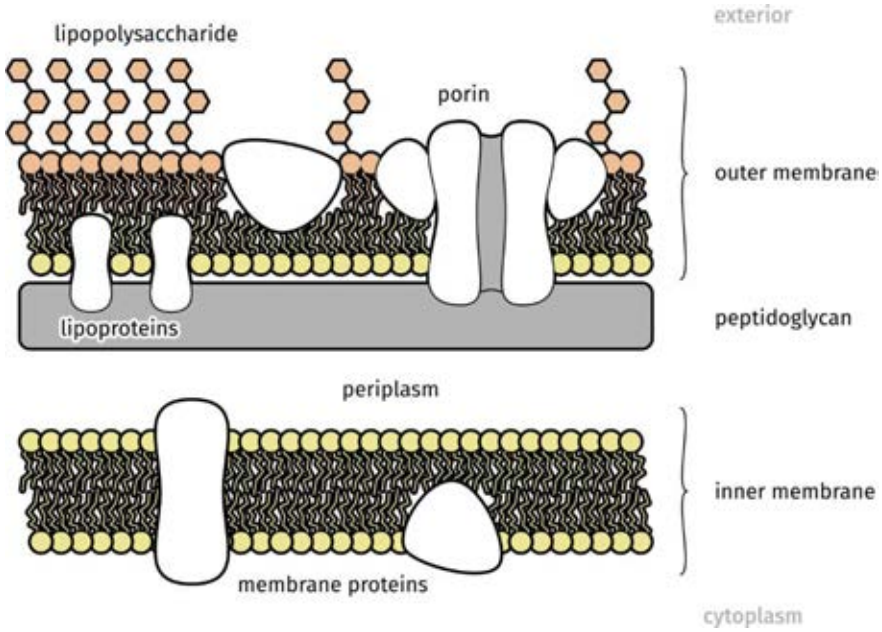


Figure 1.17. The Gram-negative cell envelope. Based on the structure of *Escherichia coli*.

The outer membrane is the most characteristic part of this structure, as it distinguishes Gram-negative bacteria from Gram-positive species. It is a lipid bilayer containing phospholipids in its inner leaflet and primarily lipopolysaccharide (LPS) in the outer leaflet. Proteins are also present. Two of the most abundant classes are transport channels in the outer membrane

allowing the diffusion of select compounds, such as nutrients, into the bacterium, and lipoproteins which anchor the outer membrane to the peptidoglycan layer below.

The peptidoglycan layer is responsible for maintaining the shape and structural rigidity of the bacterial cell.[195] Peptidoglycan is a polymer comprised of alternating N-acetyl muramic acid and N-acetyl glucosamine units (**Figure 1.18**), crosslinked by peptides three to five amino acids long; in *E. coli*, these are pentapeptides. The crosslinked rings of peptidoglycan run parallel to the short axis of the bacterium. The inner membrane, located below, is a phospholipid bilayer, similar to that found in eukaryotes. It serves as an anchoring point for proteins homologous to those found in eukaryotic mitochondria and also performs functions related to protein and lipid synthesis.

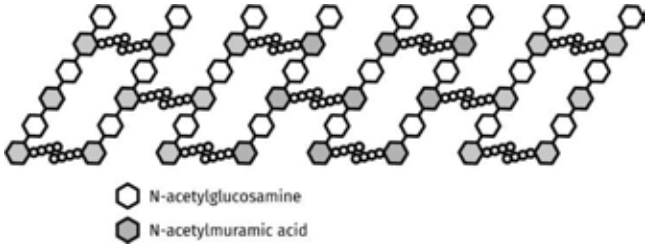


Figure 1.18. The general structure of peptidoglycan. Alternating units of N-acetylglucosamine and N-acetylmuramic acid form linear peptidoglycan chains linked into a mesh by peptide crosslinks.

1.4.2.2 Gram-positive bacteria: Staphylococcus aureus

The structure of a Gram-positive cell envelope, based on the structure of *Staphylococcus aureus*, is shown in **Figure 1.19**.[193] Gram-positive bacteria do not possess the outer membrane found in Gram-negative species. Instead, they have an extremely thick (up to 100 nm) peptidoglycan cell wall as the outermost layer of their structure. The structure of the peptidoglycan layer is similar to that found in Gram-negative species, with the notable exception of its thickness. In *S. aureus*, the structure of the crosslinks is branched, in contrast to the linear structure of *E. coli* crosslinks.[193] Two classes of teichoic acids are found within

the peptidoglycan layer; wall teichoic acids are associated with the peptidoglycan layers themselves, while lipoteichoic acids are anchored to the inner membrane below the peptidoglycan and thus span the entire thickness of the peptidoglycan wall. Teichoic acids contribute to the establishment of net negative charge throughout the surface of Gram-positive organisms; the role of this arrangement is currently poorly understood.

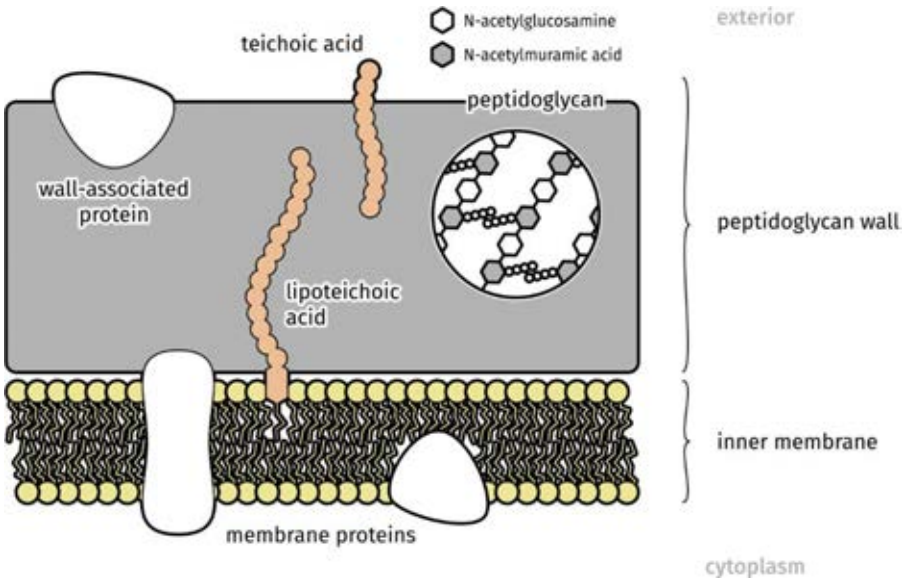


Figure 1.19. The Gram-positive cell envelope. Based on the structure of *Staphylococcus aureus*.

Multiple proteins responsible for the interaction of Gram-positive bacteria with their environment are found embedded in the peptidoglycan wall. In *S. aureus*, the most important are proteins responsible for adhesion to the extracellular matrix of the host, in the event of colonisation or infection. Proteins can be loosely associated with the surface of the peptidoglycan wall, or can be covalently attached to the crosslinking peptides.[193]

1.4.2.3 Yeasts: Saccharomyces cerevisiae

An overview of the structure of a yeast cell envelope, based on the example of *Saccharomyces cerevisiae*, or baker's yeast, is shown in **Figure 1.20**. Yeasts proved to be by far the most challenging sample type for LESA in this project. Their cells are surrounded by a complex

cell wall composed of polysaccharides and glycoproteins. In the case of the model organism *S. cerevisiae*, the cell wall contains glucans (β(1⟶3)-glucan and β(1⟶6)-glucan), mannoproteins and chitin.[196, 197] The structure is extremely resilient whilst retaining selective permeability and is thus not amenable to chemical lysis. To exemplify the severity of this challenge, the development of new lysis protocols for MALDI identification of clinically relevant yeasts was required, as the standard approaches did not provide sufficient extraction efficiency for fingerprint proteins to yield a satisfactory identification rate.[198]

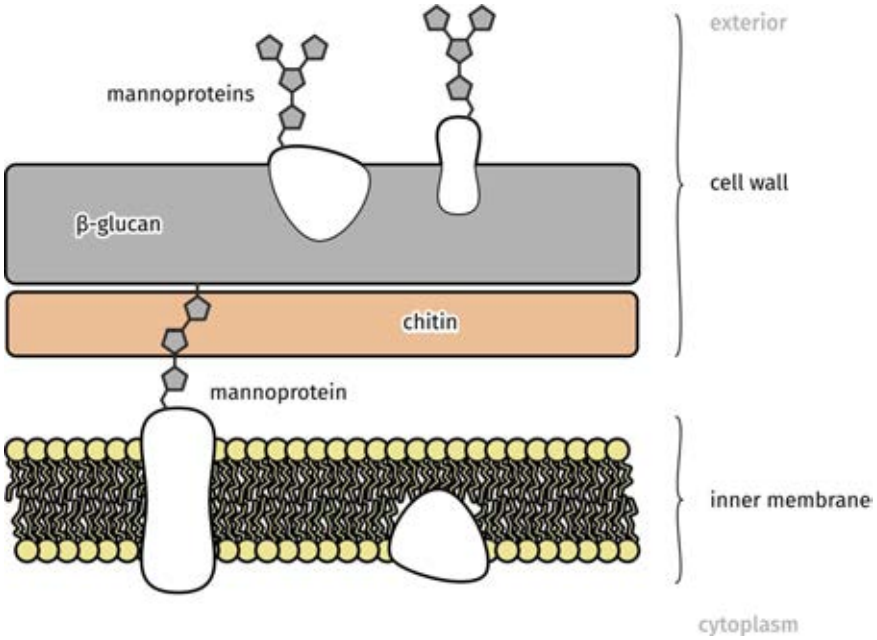


Figure 1.20. The yeast cell envelope. Based on the structure of *Saccharomyces cerevisiae.*

1.4.2.4 A summary of colony structure

A separate question concerns the structure of entire microbial colonies. Biofilms and colonies are not amorphous aggregates of cells, but organised structures composed of adherent cells held together by an extracellular matrix, termed the glycocalyx, comprised of a network of exopolysaccharides.[199, 200] The matrix encases the microbial colony and safeguards its structural integrity. It also contains an organised architecture of water channels which allow communication, metabolic exchange and a degree of cooperation between individual cells.

Populations of microorganisms forming part of a biofilm thus function differently to cells in the planktonic phase, exhibiting marked phenotypic changes caused by the expression of factors triggered, in turn, by adhesion to substrate surfaces and neighbouring cells.[201]

The most relevant property of microbial biofilms to this work is their increased, intrinsic resistance to antimicrobial agents.[202-205] Multiple mechanisms have been proposed to account for this, including impeded penetration of solutes into the biofilm matrix, the presence of dormant cells and the activation of a general stress response within the biofilm, which decreases the permeability of individual cells. It is likely that these mechanisms also determine the susceptibility of microbial colonies to chemical lysis, thus directly affecting their behaviour during LESA sampling. It is the most likely factor which necessitates the use of a contact approach for LESA sampling of colonies on agar; it has been proposed that the matrix needs to be mechanically disturbed in order to allow access to cytosolic proteins.[117]

1.4.3 Identification and characterisation of microorganisms by mass spectrometry

1.4.3.1 Seminal studies

The roots of microbiological mass spectrometry lie in the extensive use of pyrolysis coupled to gas-liquid chromatography[206-215] for the differentiation of clinically relevant microbial strains. The mass spectrometric analysis of the products generated from the original constituents of bacterial cells was, however, inconclusive as the detected compounds were very similar regardless of the species under investigation.[216] This failure was caused by the harsh conditions of sample preparation, which were sufficiently vigorous to break characteristic macromolecules down to their basic components. Nevertheless, pyrolysis-mass spectrometry (frequently aided by an additional gas chromatography separation step) continued to be developed alongside pyrolysis-gas-liquid-chromatography, ultimately yielding information useful for the characterisation of microbes.[216-218]

The first study to characterise intact bacterial cells at a sufficient level of detail to allow for their identification based on mass spectra was conducted by Fenselau and coworkers in 1975.[219] Sample introduction, whilst still technically relying on pyrolysis, was performed in a different manner—lyophilised cells were fed directly into the EI ion source, which was heated to over 300 °C. This treatment retained the structural integrity of phospholipids and quinones. Bacterial species could subsequently be differentiated according to the masses of intact phospholipids and quinones detected in their mass spectra.

Alongside pyrolysis, many other techniques, such as fast atom bombardment[220] and direct probe mass spectrometry,[221, 222] were applied to whole bacteria in order to extract information on metabolites and cell wall components. These approaches were sufficient for the analysis of carbohydrates and fatty acids; proteins, however, remained largely elusive until the development of MALDI and ESI in the 1980s.

1.4.3.2 Microbial protein analysis by MALDI MS

The capability of MALDI MS for protein analysis unlocked the potential for the rapid identification of bacteria based on protein mass spectra.[223-228] Any such procedure demands standardisation in sample preparation in order to ensure high repeatability and reproducibility of results, necessary for the high statistical confidence of identification which is required in a clinical setting. To that end, colony material picked up from agar plates was deposited onto the target plate and covered with a set volume of pre-made matrix solution. The resulting fingerprint mass spectra were consistent enough to enable the construction of spectral databases and the confident matching of any new strains against previously identified microorganisms.

Mass spectra generated for the purpose of identification of bacteria constitute peptide mass fingerprints (PMFs). These mass spectra are typically acquired in the 2000-20000 *m/z* range,

which corresponds to proteins of up to 20 kDa. The range of proteins encompassed within this mass range spans primarily ribosomal proteins, which constitute up to 70% of the dry weight of bacteria; some of the most abundant small housekeeping proteins may also be included. An example PMF spectrum acquired from *Escherichia coli*, as compared to the reference peaks derived from the proprietary Bruker Biotyper databases, is shown in **Figure 1.21**.[229]

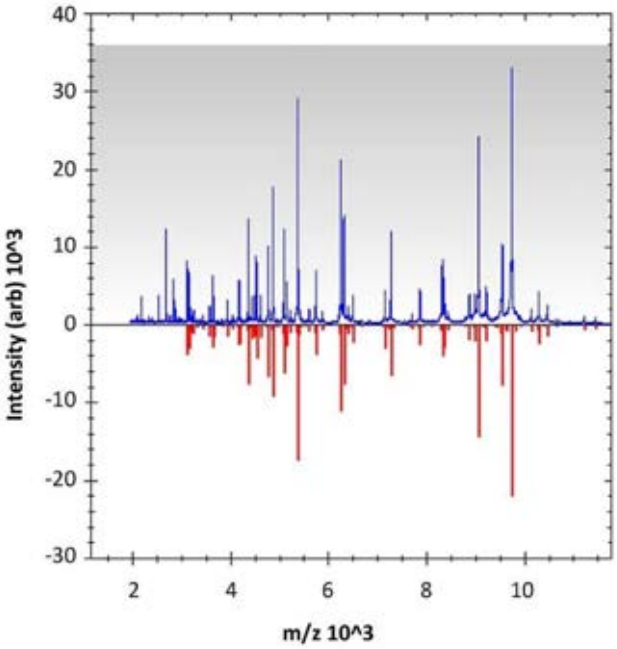


Figure 1.21. An example peptide mass fingerprint derived from a clinical isolate of *E. coli*. The experimentally acquired mass spectrum is shown in blue. The reference mass spectrum derived from the Bruker Biotyper databases is shown in red. Figure adapted from Figure S1 in Egli et al. 2015;[229] reproduced under a Creative Commons Attribution Licence.

Numerous targeted studies, reviews and evaluations have been performed since,[230-243] leading to the development of commercial TOF-based platforms for microbial identification down to the species and, less commonly, strain level; two of these, BioMérieux VITEK-MS and Bruker Biotyper were approved by the US FDA for diagnostic use,[243, 244] prompting their gradual introduction by healthcare systems in other countries, including the UK.

Whilst the technique is also suited for microbial characterisation and imaging of samples directly on growth media, the protocols for the analysis of proteins from microbial colonies on agar have not yet been as strictly standardised as those applied for the imaging of tissue sections. The poor reproducibility for experiments carried out on agar-bound biofilms can only be partly explained by the chemical and morphological differences between the samples; the variations in equipment and a rigid adherence to preferred sample preparation methods all contribute to the fact that no two groups currently working on the characterisation of bacterial biofilms by MALDI-MS follow the same protocol.[245-249] Issues with sample adhesion to sampling surfaces (both target plates and glass slides) have also been noted.[248]

1.4.3.3 Microbial analysis by ambient ionisation techniques

The introduction of ambient ionisation techniques allowed, for the first time, the mass spectrometric analysis of intact microbial colonies directly on the substrate. The applications of DESI and LESA to microbial colonies are described below; the applications of two techniques not used in this work but also relevant to the analysis of microbial colonies, nano-desorption electrospray ionisation (nanoDESI) and the liquid microjunction surface sampling probe (LMJ-SSP), are also briefly summarised.

The microbiological applications of DESI take advantage primarily of the imaging capabilities of the technique. It has been extensively used to analyse imprints of microbial colonies deposited onto filter paper or membranes,[250-252] providing information on secreted compounds including small molecules as well as cyclic and linear lipopeptides. Imprinting was developed to circumvent the requirement for a hard surface supporting the sampled material; hydrated agar media are both elastic and highly conductive, preventing the analysis of colonies directly in the petri dish.[252] Alternative approaches involved the removal of bacteria from agar plates, or their growth in liquid culture, followed by their deposition onto

clean glass slides prior to analysis (the spatial information in this case is lost),[253-255] incubation of colonies on membranes which could be lifted off the agar media[256] or on wood or cardboard insets,[257] transverse cryo-sectioning of colonies and subsequent analysis of sections on glass slides,[250] freeze-drying of colonies on agar media,[257] or drying down the colonies, along with the agar substrate, to form a thin, hard layer suitable for sampling.[258]

The main limitation of DESI as applied to microorganisms stems from its incompatibility with sampling fresh colonies on growth media. The imprinting process distorts the positions of analytes on the membrane compared to their original location in and around the colonies, reducing the effective resolution of imaging on such samples. An attempt was made to replicate the analysis of intact, dried colonies on an agar substrate by DESI; the experiments are described in Chapter 3.

Similarly to DESI, nanoDESI has primarily been employed for imaging applications.[69] Despite the similarity in terminology, it is a liquid microjunction sampling technique and thus unrelated to DESI by its mechanism.[95, 100]

Watrous et al. were the first to apply nanoDESI to microbial colonies growing directly in petri dishes for metabolomic analysis.[259] The technique was used for profiling across the surfaces of a variety of colonies exhibiting a range of morphologies. Multiple studies were since conducted on a range of organisms, focusing on metabolites, glycolipids and other relatively small analytes.[95, 100, 260-262] The capability of the technique to continuously extract analytes from a single location for extended periods of time enabled more in-depth metabolic analyses, culminating in the demonstration of a data-independent metabolomics approach.[261]

The microbiological applications exposed several practical weaknesses of nanoDESI. It was discovered that the effective resolution on the rough surfaces of microbial colonies drops

from 10 μm (as achieved on perfectly flat thin tissue sections)[97] to approximately 1 mm; the decrease resulted from the necessity of using a larger droplet size to render the system less sensitive to changes in sample height.[100] A modification has recently been implemented to control the distance from the sample by the integration of a shear force probe into the system;[102] the probe would mechanically detect the height of the sample surface and feed the information back to the computer in a closed feedback loop, thus enabling constant-distance imaging. The extracellular matrix which comes into contact with the solvent is highly hydrophilic and tends to destabilise the liquid microjunction, leading to the loss of solvent into the colony. The low, continuous flow through the system and the self-aspirating design both render the setup sensitive to such losses. In addition, the solvent flow may on occasion dislodge and lift colony material—this can easily occur on uneven colony surfaces should the capillaries collide with the sample. Sufficient volume of aspirated debris will block the transfer capillary, necessitating its cleaning or replacement.[263]

Finally, the liquid microjunction surface sampling probe (LMJ-SSP), commercialised as the Flowprobe, was employed for the metabolomic analysis of over 30 microbial isolates, including bacteria, yeasts, filamentous fungi and marine organisms.[263] Compounds secreted by fungal colonies were also analysed by droplet-LMJ-SSP.[264, 265]

LESA is thus far the only ambient ionisation technique which has proven capability for the extraction of proteins from microbial colonies. The first study performed in our group was conducted on samples of *Escherichia coli* K-12.[117] Cytosolic proteins were extracted directly from colonies growing on agar media with no prior sample preparation. Crucially, detection could not be achieved with the standard LESA protocol which only probes the top surface of the sample, but rather with the contact LESA approach in which the pipette tip is brought into contact with the colony, touching or piercing the surface layer. Six proteins were

identified in the *E. coli* samples by top-down tandem mass spectrometry, by matching the resulting fragmentation mass spectra against a database of *E. coli* proteoforms. One more protein was identified by the addition of high-field asymmetric waveform ion mobility spectrometry (FAIMS) into the protocol.[266] The work described in this thesis is a direct continuation of these studies.

1.5 Cell lysis by application of electricity

1.5.1 Introduction to electroporation

In the case of certain microorganisms, such as yeasts or encapsulated Gram-positive bacteria, chemical lysis by use of solvents may prove challenging. Thus, it was necessary to investigate other mechanisms of lysis which may be incorporated into a LESA workflow. One such method involved the application of high voltage direct current pulses to the microbial colonies. The design, construction and performance of a home-built electric cell lysis apparatus is described in Chapter 6. Construction of the apparatus was done by Andrew Tanner in the University of Birmingham Biosciences workshop.

The application of high voltage direct current pulses to various kinds of cells, including bacteria and yeasts, was extensively studied by Sale and Hamilton in the 1960s.[267-269] It was found that, by subjecting suspensions of microbial cells to an electric field of up to 30 kV/cm (limited by the ionisation of air above the sample), the viability of the organisms was drastically reduced. Different organisms presented different sensitivities to the electric field; notably, yeasts (including *Saccharomyces cerevisiae* and *Candida utilis*) required the weakest electric field (5 kV/cm) to be rendered entirely unviable. The lethal effect was traced to the electrical breakdown of the cell membranes.[268] The process of rendering membranes

permeable by application of an electric field, whether in a reversible or irreversible manner, was termed electroporation.

Electroporation has been a staple biochemical and biological technique for over 40 years. Reversible electroporation is one of the major methods by which cells are permeabilised to artificially introduced foreign DNA, which allows the transformation and modification of microorganisms and cell lines.[270] In a clinical setting, irreversible electroporation has been adopted as an alternative to thermal ablation for multiple therapies, including the selective elimination of tumour cells without thermal collateral damage.[270-277] Selective ablation of cardiac cells was shown to aid in the management of cardiac arrhythmias.[278] More recently, the introduction of chemotherapy agents into targeted cancer cells by electroporation (termed electrochemotherapy) was described.[279, 280]

The mechanism of electroporation, both reversible and irreversible, is described in detail below. For the purposes of this work, irreversible electroporation was desired.

1.5.2 Principle of operation

Lipid bilayers, which are the foundation of cell membranes, are intrinsically stable structures.[281] This behaviour is due to the structure of individual lipids and their resulting interactions. A typical membrane lipid contains a hydrophilic headgroup and hydrophobic fatty acid tails (**Figure 1.22a**). Exposed to water, the molecules spontaneously arrange themselves into a bilayer, the hydrophobic tails buried within the interior of the sheet and the hydrophilic headgroups pointing outwards into the aqueous environment (**Figure 1.22b**). The structure is flexible and semi-fluid, allowing for limited migration of lipids within each leaflet of the bilayer; migration of lipids between leaflets is rarely observed. Under harsher conditions, such as increased temperature, the fluidity of the bilayer increases to the point

where the formation of transient, aqueous pores can be detected. The pores allow the limited passage of water molecules and ions through the bilayer and spontaneously reseal.

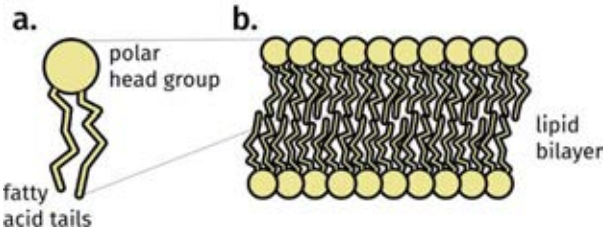


Figure 1.22. The structure of a lipid bilayer. **a**. A single lipid molecule. **b**. A lipid bilayer.

Aqueous pore formation is the most widely accepted mechanism by which electroporation might occur.[281] It postulates that the application of an electric field reduces the energy required to induce the formation of pores in the bilayer, thus increasing the number of pores formed and extending their lifetime. Molecular dynamics simulations have shown that the formation of a pore is preceded by the induction of a 'single-file defect' or water wire—a series of water molecules which spans the entire thickness of the membrane (**Figure 1.23a**).[282, 283] Hydrophilic headgroups of the lipids subsequently reorient themselves to face the water molecules, lining the nascent pore and stabilising the expanding water channel (**Figure 1.23b**). Whilst the presence of hydrophilic headgroups contributes to pore stabilisation, it is not necessary for the induction of pore formation as identical behaviour has been observed in entirely nonpolar octane bilayers.[284] In an electric field, water molecules become highly oriented according to their dipole moment; it has been suggested that this rearrangement significantly induces the penetration of water molecules into the bilayer and thus contributes to the electroporation phenomenon. Upon removal of the electric field, nanometre-scale pores reseal within a period of time determined by various groups to range from nanoseconds to seconds.[283, 285, 286] Upon application of a sufficiently strong electric field, the formation of irreversibly expanding aqueous pores is observed, leading to the disintegration of the membrane. The transition from sealable to expanding pore formation is

determined by the critical size of the membrane defect, which varies according to the type of membrane or cell and is currently poorly understood.[287]

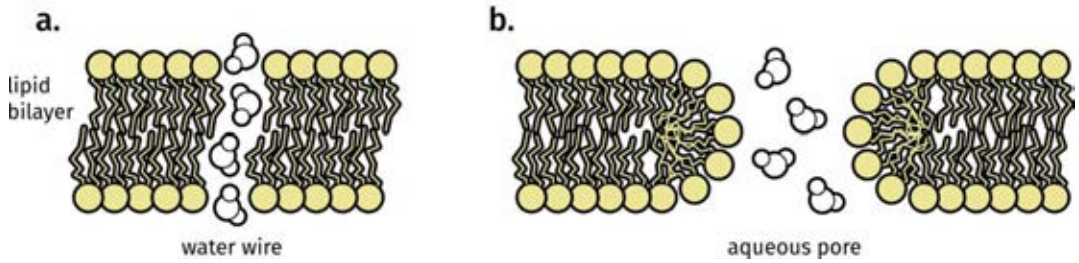


Figure 1.23. The formation of pores in a lipid bilayer. **a**. The penetration of a water wire across the bilayer. **b**. Rearrangement of the polar head groups towards the water wire, forming a quasi-stable aqueous pore.

1.5.3 Components of the apparatus

1.5.3.1 General layout

A simplified schematic of a generic electroporator is seen in **Figure 1.24.** The generic electroporator consists of a variable high voltage power supply, a bank of capacitors which accumulate and store the charge for the pulse, a switch or relay which releases the pulse, and a vessel in which the sample is placed with integrated electrodes which transfer the energy of the electric pulse into the sample.[288] The characteristics of the individual components of the device, listed in the order in which they contribute to the generation of a voltage pulse, are described below.

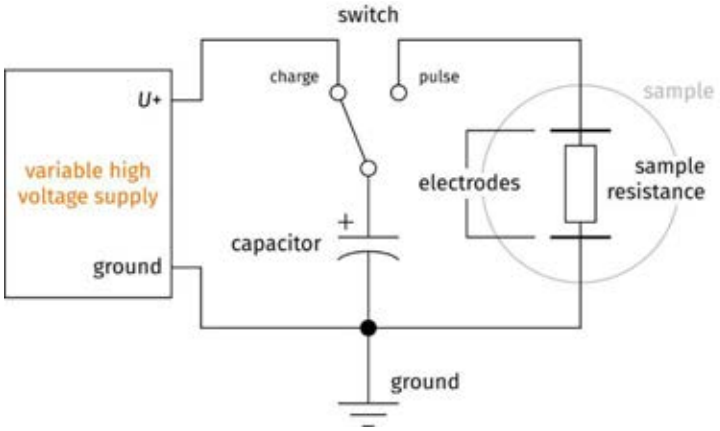


Figure 1.24. Schematic representation of a generic electroporator. The sample is represented by a resistor.

1.5.3.2 The power supply

The power supply receives 230 V AC from a standard UK wall socket and converts it to a high voltage direct current. This commonly involves the use of a transformer and a rectifier.[289]

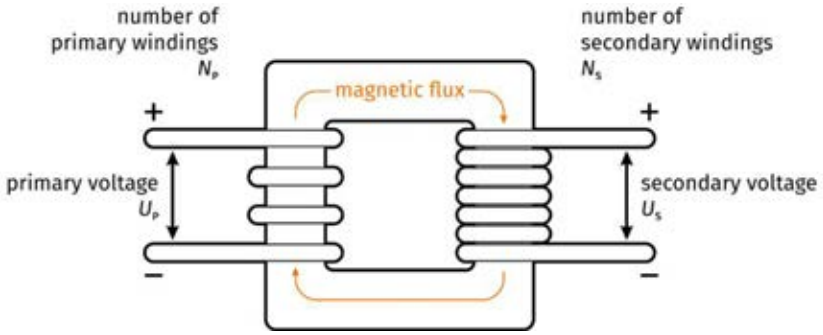


Figure 1.25. Schematic representation of a transformer.

The transformer (**Figure 1.25**) is responsible for controlling the output voltage.[290] It consists of a core and two sets of windings, the primary (input) winding and the secondary (output) winding. Voltage applied to the primary winding generates a magnetic field, which induces current in the secondary winding. Following from Faraday's law of induction, the ratio of the primary and secondary voltages is proportional to the ratio of turns in the primary and secondary windings (**Equation 1.11**). In step-up transformers as required here, the number

of turns in the secondary winding is greater than the number of turns in the primary winding, generating a higher output voltage.

$$\frac{U_P}{U_S} = \frac{N_P}{N_S} \qquad 1.11$$

The rectifier is responsible for the conversion of alternating current, as supplied from the mains, into direct current. This process is termed rectification.[291] The simplest form of a rectifier is a diode. A diode is a circuit component which has a low resistance when current flows through it in one direction and a high resistance in the other direction.[289] Upon application of an alternating current to a diode, the current flowing in the high-resistance direction of the diode is blocked, resulting in the generation of a pulsating direct current in the low-resistance direction (half-wave rectification; **Figure 1.26a**). A bridge rectifier is a more complex arrangement which provides full-wave rectification. It consists of four diodes connected as shown in **Figure 1.26b**.

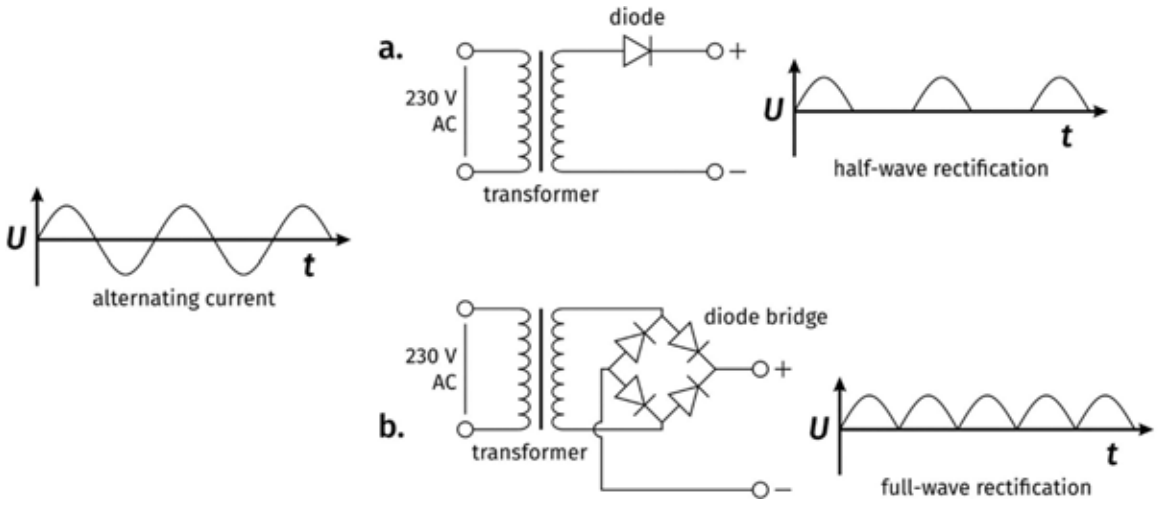


Figure 1.26. Rectification of alternating current to direct current. **a.** Half-wave rectification using a diode. **b.** Full-wave rectification using a diode bridge.

1.5.3.3 Charge storage and release

The power supply alone is not capable of generating precise pulses required for electroporation. Thus, the energy delivered by the power supply is accumulated and then controllably released as a single burst. The accumulation of energy is accomplished by use of capacitors.

A capacitor is an electrical component which stores potential energy in an electric field.[289] The simplest type of capacitor, the thin-film capacitor, is constructed from two thin conductive plates separated by a dielectric material.[292] As voltage is applied across the capacitor, a net positive charge develops on one plate of the capacitor and negative charge accumulates on the other plate. The storage capacity, or capacitance, is measured in farads (F).

The discharge of a capacitor, accomplished by use of a relay or spark gap switch, produces an exponential decay pulse (**Figure 1.27a**); that is, an instantaneous pulse of high voltage followed by a gradually fading tail of lower voltage.[288] The length, or time constant, of such a pulse is conventionally measured from its peak value to 37% of the peak value. The time constant associated with the discharge of capacitors commonly used in simple electroporators using exponential decay pulses (capacitance up to 500 μF) usually falls in the range of milliseconds. Whilst the low voltage tail is desirable in certain applications, for example providing a driving force for DNA to enter permeabilised cells in transformation experiments, it also damages cell viability which makes it unsuitable for work with more fragile or precious cell types.[288]

A more complex but at the same time more controllable setup involves the generation of a square wave pulse (**Figure 1.27b**).[288] In the case of an ideal square wave, the maximum desired voltage is reached almost instantaneously and maintained for the duration of the

pulse, following which it drops to baseline. In practice, since the current is supplied by a bank of capacitors, a slight drop in voltage over the duration of the pulse can still be observed. The duration of square wave pulses capable of lysing bacterial cells falls in the range of microseconds and can be tightly controlled in the absence of a low voltage tail. The capacitance required to generate such pulses is in the range below 1 μF.

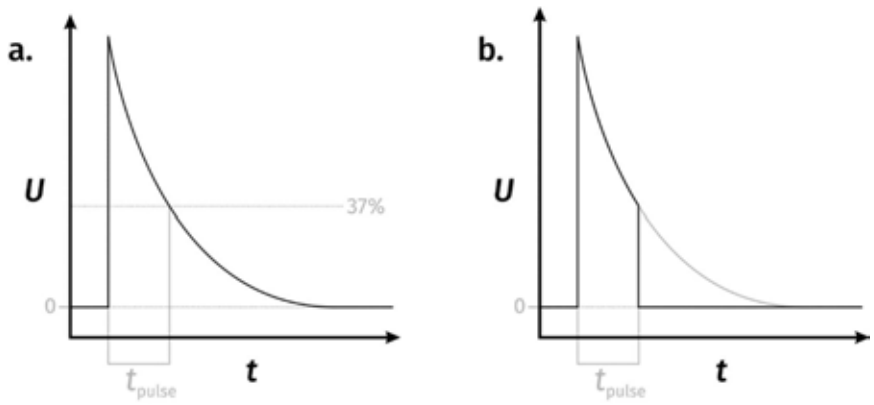


Figure 1.27. Shapes of pulses generated by capacitor discharge. **a.** Exponential decay. The time constant (t_{pulse}) is measured at the time it takes for the voltage (U) to reach 37% of its maximum value. **b.** Square wave.

The difficulties in constructing such a setup all concern the need for a fast switching mechanism to generate the square wave pulse. A standard mechanical relay or spark gap is insufficient for such purposes as the switching time exceeds the duration of the desired pulse. Thus, it is necessary to use a fast high voltage field-effect transistor switch instead.

Transistors are electrical components which switch the flow of current. Of particular relevance to this work is the field-effect transistor, or FET (**Figure 1.28**).[289] The salient components of a FET are the source, gate and drain. The gate accepts an input voltage. If the voltage applied to the gate exceeds the threshold, the channel between the source and the drain is rendered conductive and thus charge flows across the transistor from the source to the drain. Since there are no mechanical parts in a FET switch, its operation is orders of magnitude faster than that of a standard relay.

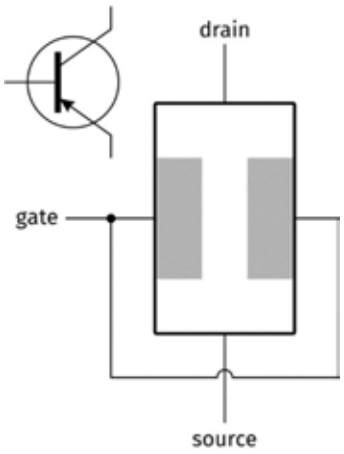


Figure 1.28. Schematic representation of a field-effect transistor. The corresponding electrical diagram symbol for a transistor is shown in the upper left.

1.5.3.4 Pulse delivery

The electric pulse is delivered by a pair of electrodes. A number of geometries can be used; by far the simplest and most common, however, is the planar electrode setup.[288] Planar electrodes are flat pieces of conducting material arranged in parallel to one another. The strength of an electric field is measured in volts per centimetre. Thus, stronger electric fields can be generated by decreasing the distance between the electrodes, enabling the lysis of less susceptible organisms without any invasive modifications to the circuitry of the apparatus.

1.6 Structure and aims of the project

This work presented in this thesis focused on the development of LESA mass spectrometry for the analysis of proteins directly from microbial colonies. LESA was chosen due to its demonstrated capability to extract intact cytosolic proteins directly from bacterial colonies grown on agar media, an achievement which has not hitherto been replicated by any other ambient ionisation technique due to the requirement for colony and cell lysis (achieved mechanically by contact LESA); additionally, the sensitivity of LESA was considered an advantage for the analysis of low abundance proteins not accessible by other techniques. No harsh ionisation techniques were considered due to their incompatibility with protein

analysis. Imaging performance was not prioritised due to it being non-essential for the identification and characterisation of relatively uniform bacterial biofilms; the spatial resolution of LESA could, however, be potentially sufficient for the detection of secreted peptides in media surrounding the colonies, which was previously demonstrated by nanoDESI and Flowprobe on freshly prepared colonies. Whilst DESI was also shown to be suitable for the imaging of bacteria in that capacity, it necessitated the drying of samples; furthermore, its capacity for protein analysis had not yet been realised at the beginning of this project.

Results of MALDI-TOF-MS identification of bacteria in a clinical setting were considered a benchmark for bacterial identification by LESA; whilst the speed of operation of MALDI-TOF-MS in conjunction with spectral fingerprint-based identification would not be matched by a LESA based approach, LESA offered the capability of *in situ* sampling of live colonies and potential development avenues towards direct sampling of clinically relevant surfaces (contaminated equipment, patient skin etc) which could not be performed by MALDI. A similar depth of identification (species level, ideally strain level) to MALDI-TOF-MS was expected to be achievable LESA MS due to the similar mass range of proteins currently accessible by both techniques, encompassing characteristic ribosomal proteins commonly used for identification purposes (up to approximately 20 kDa).

The aims of the project were as follows:

- to develop a better understanding of the LESA process and measure its reproducibility (Chapter 3),
- to optimise sampling protocols for the top-down identification of proteins in a wider range of microorganisms (Chapter 4),

- to investigate the potential of LESA as a tool for the identification of bacterial species based on LESA protein mass spectra (Chapter 4),
- to improve the performance of LESA mass spectrometry on bacterial colonies by including FAIMS in the workflow (Chapter 5),
- to develop a plate-based electroporation platform compatible with LESA sampling for the identification of proteins in yeasts (Chapter 6).

CHAPTER 2

Materials and methods

2.1 Materials

2.1.1 Chemicals

Acetonitrile, HPLC-grade water, methanol and ethanol were purchased from J.T.Baker, Deventer, Netherlands. α-Cyano-4-hydroxycinnamic acid (α-CHCA) and dihydroxybenzoic acid (DHB) matrices, as well as formic acid, rhodamine B and raffinose, were purchased from Sigma Aldrich, Gillingham, UK. Glycerol was purchased from Fisher Scientific, Loughborough, UK.

Pre-mixed, powdered lysogeny broth (LB) (Miller composition) and yeast extract peptone dextrose (YPD) broth were purchased from Sigma Aldrich, Gillingham. Bacteriological agar was purchased from VWR International, Leuven, Belgium. Blood agar plates were sourced from surplus stock of the teaching laboratories at the School of Biosciences, University of Birmingham, UK.

2.1.2 Strains

Escherichia coli K-12, *Cryptococcus neoformans* H99 and *Candida glabrata* (in-house strain) were provided by Robin C. May from the collection of the Institute of Microbial Infection (IMI), University of Birmingham, UK; *E. coli* BL21 mCherry was also sourced from the same collection. *Pseudomonas aeruginosa* PS1054 and *Staphylococcus aureus* MSSA476 were provided by Mark A. Webber from the library of clinical isolates of the Queen Elizabeth Hospital Birmingham. The in-house strain of *Staphylococcus* sp. was sourced from the collection stored at the teaching laboratories in the School of Biosciences, University of

Birmingham, UK. *Streptococcus pneumoniae* D39 was provided by Tim Mitchell and Andrea Mitchell at the IMI, University of Birmingham. *S. oralis* ATCC 35037 and *S. gordonii* ATCC 35105 were provided by Rachel Sammons at the School of Dentistry, University of Birmingham.

2.2 Methods

2.2.1 Culturing of bacteria and yeast

2.2.1.1 Preparation of media and plates

Lysogeny broth agar (LBA) was prepared by dissolving 20 g l^{-1} of the pre-mixed lysogeny broth base and 15 g l^{-1} of agar in purified (18 MΩ) water, and sterilised by autoclaving (121 ℃, 0.5 bar). For YPD agar, the LB base was replaced by 50 g l^{-1} YPD broth, the other steps remaining the same. Bottles of media were stored at room temperature.

Agar plates were prepared by gently melting a previously prepared bottle of media in a microwave. The agar was subsequently poured into empty petri dishes by the flame of a Bunsen burner, to a thickness of 3-5 mm, and left to set. For a consistent agar height where noted, a set volume of 5 ml liquefied media was dispensed. Plates were sealed with parafilm and stored inverted at 4 ℃ until use.

LBA media for use in the work described in Chapter 3 were prepared as outlined above, with the addition of 10 mg/100 ml of rhodamine B.

2.2.1.2 Preparation of glycerol stocks

A 50% (v/v) solution of glycerol in purified (18 MΩ) water was sterile-filtered through a 0.2 µm syringe filter and stored at room temperature.

A 5 ml LB culture was started in a 50 ml sterile plastic tube to provide sufficient aeration. A single colony of a chosen bacterium was picked up from a donor agar plate by use of a sterile plastic loop or pipette tip and stirred into the liquid media held close to the flame of a Bunsen burner; the top of the tube was left slightly unscrewed. The tube was incubated at 37 ℃ with shaking (200 rpm) overnight for a maximum of 18 hours, before the onset of stationary phase where possible. For yeast, YPD broth was used instead of LB and the temperature of incubation was lowered to 25 ℃ for *C. neoformans* and 30 ℃ for *S. cerevisiae* and *C. glabrata.*

Following incubation, 400 µl of the overnight culture were transferred in a sterile fashion into a cryovial (Nalgene) and topped up with 600 µl of the 50% glycerol solution. The cryovial was gently inverted several times until the contents became homogeneous, and subsequently stored at -80 ℃.

2.2.1.3 Culture of samples for analysis by mass spectrometry

Two approaches were used for the inoculation of cultures on solid agar media. For overnight preparation, agar plates were inoculated directly from glycerol stocks by scraping the surface of the stock with a 1 µl inoculation loop or sterile pipette tip by the flame of a Bunsen burner and subsequently touching a spot on the agar plate where the growth of a colony was desired. Alternatively, for the growth of even, round colonies (following conclusions derived from work described in Chapter 3), an overnight liquid culture was prepared instead as described in section 2.2.1.2. Following incubation, a 1 µl (bacteria) or 10 µl (yeast) inoculation loop was used to withdraw an aliquot of liquid culture and spot it onto the agar plate. Plates were inverted and incubated in the dark at 37 (bacteria) or 30 ℃ (yeast) for 12-48 hours, until the size of the colonies was suitable for sampling (approx. 5 mm in diameter).

2.2.2 LESA sampling

The solvent system used for liquid extraction consisted of acetonitrile, water and formic acid (Sigma-Aldrich, Gillingham, UK); the relative proportions were initially 39.5:59.5:1 for Gram-negative species (this was changed to 40:60:1 prior to the introduction of BSL2 organisms into the study), and either 40:60:1 or 50:45:5 for Gram-positive species as noted.

A grid based on the LESA universal adapter plate, marking the sampling positions programmed into the advanced user interface (AUI) of the ChipSoft 8.3.3 software, was placed at the bottom of the sample tray, enabling the selection of a suitable sampling position on the bacterial colony. Agar plates containing colonies of interest were placed in the sample tray of a TriVersa NanoMate robotic pipette system (Advion, Ithaca, NY, USA), adjacent to half a 96-well PCR plate holding the solvent. 3 µl of solvent were aspirated from a well of the PCR plate. The pipette tip was moved above the colony to the position specified by the advanced user interface and lowered to a height of approximately -10 mm, bringing it into contact with the colony surface; the exact distance would be adjusted for each individual agar plate and regularly reassessed to account for the gradual changes in the height of the solid agar medium. On contact, 2 µl of the solvent was dispensed onto the surface of the colony encompassed by the pipette tip. Contact was maintained for 3-15 s (specified in the main text), following which the solution was reaspirated into the pipette tip.

Sampling of glass slides and rhodamine B-containing agar media for work described Chapter 3 was carried out in a similar fashion, albeit using the standard LESA interface in ChipSoft rather than the advanced user interface. Glass slides were placed directly on top of a universal adapter plate provided with the TriVersa NanoMate. The sampling height was increased accordingly. The volume of solvent dispensed was reduced to 1 µl. The liquid microjunction was maintained for 3 s.

For specific alterations and optimisation of the protocol, see Chapters 3 and 4.

2.2.3 LESA mass spectrometry

All MS experiments, with the exception of the work described in Chapter 3, were performed by use of an Orbitrap Elite instrument (Thermo Fisher Scientific, Bremen, Germany) at a resolution of 120,000 at *m/z* 400, acquiring and coadding ten microscans per scan. The inlet capillary temperature was set to 250 ℃. Samples were introduced into the instrument by use of the TriVersa NanoMate integrated, chip-based nanoelectrospray system at a pressure of 0.3 psi and a voltage of 1.75 kV. Full scan mass spectra in the *m/z* range from 600 to 2000 were acquired for a minimum of 5 minutes. Precursor ions were selected for fragmentation with an isolation window of 2, 3, or 5 *m/z* as appropriate. Collision-induced dissociation (CID) was performed in the ion trap with use of helium gas at a normalized collision energy of 35%; the fragments were detected in the Orbitrap. Where CID was inadequate to provide sufficient fragmentation, HCD was performed instead. The normalised collision energy for HCD was optimised for each protein individually in 5% increments; values between 20 and 35% were used. MS/MS spectra were recorded for 5 minutes. Each MS/MS scan comprised 30 coadded microscans. Automatic gain control targets were set to 1×10^6 charges for full scan mass spectra and 5×10^5 charges for MS/MS spectra.

For Chapter 3, the instrument used was an Orbitrap Velos (Thermo Fisher Scientific, Bremen, Germany). The general protocol for sampling and mass spectrometric analysis followed the protocol described above for the Orbitrap Elite; the resolution was decreased to *m/z* 60,000 at *m/z* 400 in line with the specification of the instrument. The *m/z* range of full scan mass spectra and the duration of data acquisition were varied as described in Chapter 3, section 3.2. MS/MS was not performed.

2.2.4 FAIMS

High-field asymmetric waveform ion mobility separation (FAIMS) was coupled to LESA MS by use of an external FAIMS device (Owlstone, UK). For all analyses, the chip temperature of the FAIMS device was set to 100 ℃ and the inlet capillary temperature was increased to 350 ℃.

Analyses on bacterial colonies were performed in 1D and static field modes. 1D sweeps were performed for each bacterial species at DF 270 Td. CF values were ramped from 0.0 to 4.0 Td over a period of 240 seconds. Data for static field analyses were collected for a minimum of one minute at each setting.

Protein peaks of interest were isolated in static field mode using an isolation width of 3 or 5 *m/z*. The larger windows were used to retain as much signal intensity as possible; due to the effects of ion mobility separation, co-isolation of unwanted peaks was not an issue. Selected precursor ions were fragmented using either CID (NCE 35%) or higher energy collision dissociation (HCD) (NCE ranging from 20 to 35%, varied in 5% increments).

2.2.5 Protein identification

Top-down protein identification was performed with ProSightPC software, versions 3.0 to 4.1 alpha (Thermo Fisher Scientific, Bremen, Germany). MS/MS spectra were deconvoluted by the THRASH algorithm at a signal-to-noise ratio of 3. These were then matched against custom databases constructed for each species from full proteome data available via UniProt. Each database was constructed as a standard top-down database, taking into account the cleavage of initial methionines and N-terminal acetylation; in databases constructed in ProSightPC 3.0 (Chapter 4), formylation was ignored due to software instability induced by the selection of this option. The subsequent upgrade to ProSightPC 4.1 alpha eliminated this issue. SNPs and all posttranslational modifications annotated within the database were

considered, with up to 13 features per sequence and a maximum mass of 70 kDa. Fixed modifications, including methionine oxidation, were not selected at this stage. For both strains of *E. coli*, the complete, annotated reference proteome of the K-12 strain was used (UniProt ID UP000000625, 4306 protein entries). For *S. aureus* MSSA476, two databases were generated: one based on the reference proteome of a representative strain (NCTC 8325, UniProt ID UP000008816, 2889 protein entries), making use of the more complete, reviewed annotation, and one specific to the MSSA476 strain (UniProt ID UP000002201, 2598 protein entries). For *P. aeruginosa* PS1054 the reference proteome (UniProt ID UP000002438, 5563 protein entries) was used exclusively. *S. pneumoniae* D39 was searched against the reference proteome of a closely related, non-pathogenic strain (UniProt ID UP000000586, 2030 protein entries). The *S. oralis* database was based on the only available reference proteome for this species (UniProt ID UP000005621, 2022 protein entries). Two *S. gordonii* databases were used, one based on the annotated chromosome (UniProt ID UP000001131, 2050 protein entries) and one based on the whole genome sequence (UniProt ID UP000069207, 2095 protein entries). A multispecies database was also constructed based on the pan-proteomes of *S. pneumoniae* ATCC BAA-255/R6 (UniProt ID UP000000586) and *S. gordonii* Challis (UniProt ID UP000001131; a total of 61,367 protein entries). For *S. cerevisiae*, the reference proteome of the ATCC 204508/S288c strain was used (UniProt ID UP000002311, 6049 protein entries). For *C. glabrata*, the only available reference proteome was used (UniProt ID UP000002428, 5200 protein entries). For *C. neoformans*, the UniProt proteome corresponding to the H99 strain was used (UniProt ID UP000010091, 7430 protein entries).

Tandem mass spectra of the unknown species of *Staphylococcus* were searched against the reference *S. epidermidis* proteome (UniProt ID UP000000531, 2492 protein entries) and a

concatenated multispecies database comprising all available nonredundant proteomes in the *Staphylococcus* genus (145,289 protein entries).

A broad absolute mass search was specified as a starting point for each investigated species. Protein hits were considered within 2 kDa of the measured intact mass. Disulfide bridges were taken into account. The Δm mode, accounting for post-translational modifications or amino acid substitutions not present in the database, was active. All post-translational modifications annotated within the database were taken into account. Fragment tolerance was set to ±15 ppm. Putative hits, as well as any post-translational modifications, were verified by narrowing down the search criteria and optimizing any remaining matches in ProSightPC's Sequence Gazer; this included the manual addition of post-translational modifications such as methylation or acetylation to the sequence within Sequence Gazer and subsequent automatic re-annotation of fragments. For the work described in Chapter 4, manual peak assignment with an acceptance threshold of ±5 ppm was subsequently carried out.

The protein sequence generated *de novo* from the unknown *Staphylococcus* species was searched by standard protein BLAST against all nonredundant protein sequences belonging to the *Staphylococcus* genus (taxonomic ID 1279), by use of the protein–protein BLAST algorithm with the default BLOSUM62 scoring matrix.

2.2.6 Liquid chromatography-tandem mass spectrometry (LC-MS/MS)

LC-MS/MS was performed on an Orbitrap Elite instrument (Thermo Fisher Scientific, Bremen, Germany) coupled to a Dionex UltiMate-3000 liquid chromatography system. All LC-MS/MS experiments were self-contained and performed separately from ion mobility spectrometry experiments; the techniques were not combined. A PepMap C4 column (300 Å pore size) was paired with a filter flanked by two IDxl 30 μm × 10 cm trapping columns. The

mobile phases used were 99.9% water, 0.1% formic acid (A) and 99.9% acetonitrile, 0.1% formic acid (B); 99.9% water, 0.1% formic acid was used as a loading buffer. A 30 minute gradient (3.2-60% B) was used, followed by a 10 minute wash step at 90% B. The injection volume was 5 µl. The column compartment temperature was 60 ℃. Analytes were eluted and delivered to the mass spectrometer via the TriVersa NanoMate nanoelectrospray coupler at a spray voltage of 1.7 kV.

Full scan acquisition was performed in the Orbitrap analyser at a resolution of 60,000 at *m/z* 400. Following a survey scan encompassing the *m/z* 600-2000 range, the top five peaks were selected for MS/MS analysis by CID (normalised collision energy 35%) in the Orbitrap, also at a resolution of 60,000 at *m/z* 400. Dynamic exclusion of 60 s was applied.

The optimisation of parameters is briefly discussed in Chapter 5, section 5.3.4.

2.2.7 MALDI MS

Colonies of *E. coli* K-12 were cultured as described above, starting with a liquid culture. Standard diameter (11 cm) LBA plates were used. Plates were transported in an ice box and stored at 4 ℃ until use.

24 hours prior to MALDI MS imaging analysis, a square of agar media containing a single colony was excised from the agar plate using a sterile scalpel, lifted off the plate with a thin, sterile spatula and deposited onto a glass slide. The sample was subsequently dried overnight at 40 ℃. The dried sample was coated with matrix solution consisting of a 1:1 mix of CHCA and DHB, 20 mg/ml in 80% MeOH. Matrix deposition was accomplished by used of a TM-Sprayer (HTX Imaging, Chapel Hill, NC, USA). The spraying parameters for the analysis described in Chapter 3, section 3.3.4.2 were as follows: nozzle temperature 65 ℃, 8 passes,

flow rate 0.12 ml/min, track velocity 1200 mm/min, track spacing 3 mm, drying time 30 s between passes.

The MALDI experiments were performed on a Synapt G2-Si MALDI instrument (Waters, Milford, MA) equipped with an Nd:YAG laser (355 nm) and operated in resolution mode. Calibration was performed in the *m/z* 100-1200 range by use of red phosphorus as calibrant. Glass slides carrying the dried samples were inserted into a glass slide holder and scanned into Waters HDImaging software. The area containing the colony was selected by use of the polygonal selection tool; a small proportion of plain agar media around the colony was included in the selection to provide background signal. The parameters used for the analysis described in Chapter 3, section 3.3.4.2 were as follows: pixel size 90 µm, 1 scan/pixel, 0.5 s/scan, laser energy 175 arbitrary units, laser repetition rate 500 Hz. Data were acquired in the *m/z* 100-1200 range.

2.2.8 DESI MS

Colonies for DESI analysis were prepared as described for MALDI prior to the matrix deposition step. 95% MeOH with raffinose (0.01 mg/ml) as a mass standard was used. Source geometry was optimised by use of rhodamine B-coated glass slides. The imaging parameters were as follows: pixel size 100 µm, stage movement rate 50 µm/s (2 s data acquisition/pixel). Electrospray voltage was 5 kV. Data were acquired in the *m/z* 100-1500 range by use of a Synapt G2-Si instrument (Waters, Milford, MA) operated in resolution mode.

CHAPTER 3

Determination of LESA sampling reproducibility and fundamentals

3.1 Introduction

The aim of the work presented in this chapter was to probe the fundamentals of the LESA process. A statistical analysis of the repeatability of LESA was performed on model surfaces (rhodamine B-coated glass slides and agar media containing rhodamine B). Rhodamine B, a red dye emitting fluorescent light at 627 nm, was used as a visible marker compound to monitor the solvent spread during the sampling process, as well as provide characteristic peaks for the measurement and comparison of absolute ion intensities across a number of mass spectra. Some of the sources of variability inherent to the LESA process were identified. Sampling of model surfaces was followed by basic statistical analysis and general observations pertaining to the sampling of bacterial colonies exemplified here by *Escherichia coli* K-12. A number of growth and storage conditions (listed in section 3.2.1.3) were selected for sample preparation, with the ultimate aim of detecting resultant differences in mass spectra; subsequent analysis was, however, diverted towards a simple statistical summary due to a number of challenges encountered during data acquisition and described below. These challenges, and the improvements made to the sample preparation and LESA sampling protocols in response, are described.

Attempts to further the understanding of the LESA process through characterisation of LESA sampling spots and the disturbances introduced in the sampled colonies by use of complementary techniques are described in the latter sections of the chapter. Confocal imaging was used, followed by chemical imaging by MALDI and DESI MS.

3.2 Experimental

3.2.1 Measurement of repeatability

3.2.1.1 Sampling of glass slides

Glass slides coated with rhodamine B were sourced from existing stock prepared by Gurdak et al.[293] Sampling was carried out as described in Chapter 2, section 2.2.2; alterations to the protocol are described below.

Sampling was carried out by LESA in both contact (pipette tip touching the glass slide) and surface sampling (pipette tip suspended approximately 0.2 mm above the glass slide) modes in a grid-like fashion, the number of sampling spots was determined by the spread of solvent and the dimensions of the glass slides (75 × 25 mm). The solvent system used consisted of 40:60:1 acetonitrile-water-formic acid, the same as used for sampling colonies of *E. coli* K-12. The volume of solvent initially aspirated was 3 µl. 1 µl was dispensed for both sampling modes; 1.2 µl was reaspirated to ensure complete withdrawal of solvent from the surface.

Fifty mass spectra, corresponding to 50 contact sampling attempts on a single rhodamine-coated glass slide, were acquired as described in Chapter 2, section 2.2.3 by use of an Orbitrap Velos. The AGC was turned off to allow for the direct comparison of ion intensities between mass spectra; the optimum injection time was verified by prior use of AGC and set to 0.1 ms. Data were acquired in the *m/z* 400-500 range. Data acquisition was set to 30 seconds; in order to normalise the number of scans in each mass spectrum, as well as eliminate the sporadic artifacts caused by higher current at the beginning of spectrum acquisition, only the last 18 scans in each analysis were taken into consideration. For surface sampling, the sampling height was increased by 0.2 mm compared to contact sampling; the pipette tip was thus suspended above the glass slide so that a visible solvent microjunction was formed. Fifteen

mass spectra, corresponding to 15 surface sampling attempts, were acquired from a replica glass slide using the same parameters. The number of repeats performed in surface sampling mode was reduced due to the greater solvent spread observed on the glass slide, yielding larger sampling spots.

3.2.1.2 Sampling of agar media

LBA media containing rhodamine B were prepared as described Chapter 2, section 2.2.1.1. Sampling of agar media was carried out as described in section 3.2.1.1., with alterations as follows: Data were acquired in the *m/z* 400-1200 range. The ion injection time was set to 10 ms.

3.2.1.3 Sampling of bacterial colonies

Colonies of *Escherichia coli* K-12 were prepared as described in Chapter 2, section 2.2.1.3, directly from glycerol stocks. One colony was inoculated per plate. Sampling was carried out as described in Chapter 2, section 2.2.2. A photocopy of the LESA universal adapter plate, corresponding to a grid of 1536 unique locations programmable for sampling within the ChipSoft software, was placed on the sampling tray of the TriVersa NanoMate underneath the agar plate containing the target colony. The grid was then used to select sampling locations on the colony. Data acquisition was carried out as described in Chapter 2, section 2.2.3 by use of an Orbitrap Elite.

Five growth/storage condition sets were prepared consisting of three colonies each. These were as follows:

- 3 days incubation at 20 ℃, 4 days storage at 4 ℃
- 6 days incubation at 20 ℃, 1 day storage at 4 ℃
- 7 days incubation at 20 ℃, 1 day storage at 4 ℃

- 7 days incubation at 20 ℃
- 24 hours incubation at 37 ℃.

3.1.1 Optical imaging and microscopy

Optical imaging of rhodamine B-coated glass slides following LESA sampling was performed by use of a DinoLite Premier microscope camera. Optical and confocal microscopy images of LESA sampling spots were acquired by use of an Olympus LEXT confocal microscope.

3.2.2 MALDI

Colonies of *E. coli* K-12 were generated as described in Chapter 2, section 2.2.1.3. Following incubation, all plates were transported in a cooled container and stored at 4 ℃ until use. Sample preparation for MALDI is outlined in Chapter 2, section 2.2.7. Imaging parameters are listed in Chapter 2, section 2.2.7.

3.2.3 DESI

As for MALDI, colonies of *E. coli K-12* were generated as described in Chapter 2, section 2.2.1.3. Following incubation, all plates were transported in a cooled container and stored at 4 ℃ until use. DESI mass spectrometry imaging was carried out as described in Chapter 2, section 2.2.8.

3.3 Results and discussion

3.3.1 Sampling repeatability on glass slides

The absolute intensity plots for sampling of rhodamine B (*m/z* 443.23, $[M+H]^+$; molecular structure and representative LESA mass spectrum shown in **Figure 3.1a**) from glass slides in surface and contact modes are shown in **Figure 3.1b** and **3.1c**, respectively. For contact mode, the pipette tip was lowered into contact with the glass slide during sampling. The spread of solvent being dispensed should therefore be limited to the diameter of the pipette tip. For

surface sampling mode, the pipette tip was suspended approximately 0.2 mm above the surface of the glass slide, establishing a visible liquid microjunction during sampling. No normalisation to the total ion current was performed. The number of data points (50 for contact sampling mode, 15 for surface sampling mode) was determined by the number of sampling spots which could be accommodated on a single glass slide without risk of solvent overflowing onto adjacent sampling spots.

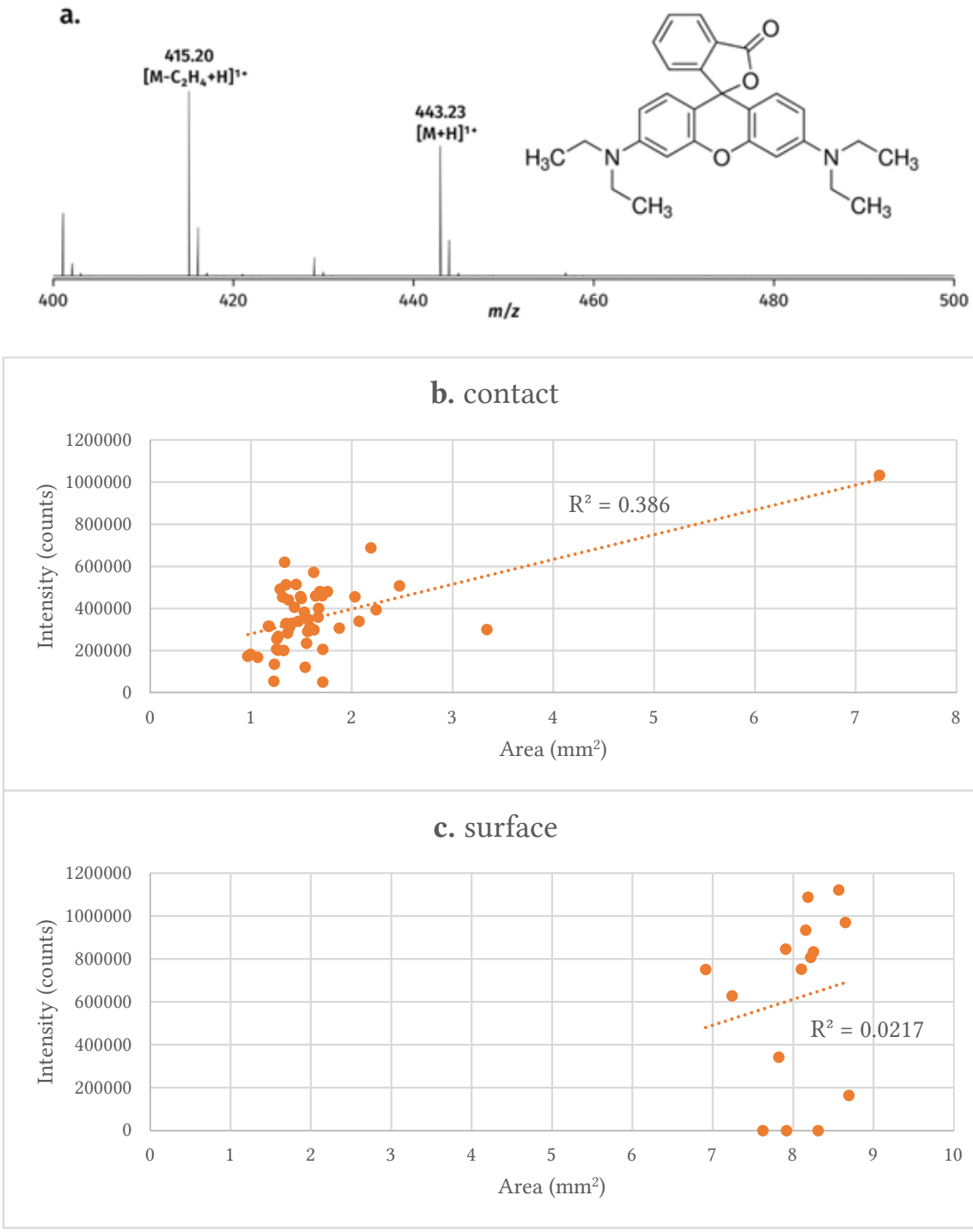


Figure 3.1. Absolute intensities and corresponding sampling spot areas for sampling of rhodamine B-coated glass slides. **a**. Representative LESA mass spectrum of rhodamine B on glass. **b**. Absolute intensities of the rhodamine peak (*m/z* 443.23) and sampling spot areas measured in contact mode. **c**. Absolute intensities of the rhodamine peak and sampling spot areas measured in surface sampling mode.

The mean measured intensity of the rhodamine B peak was 358784 counts for contact sampling and 616041 counts for surface sampling. It was noted that, in all acquired mass spectra, the peak corresponding to the expected molecular ion (*m/z* 443.23) was lower in intensity than the *m/z* 415.20, determined by intact mass to correspond to the $[M\text{-}C_2H_4+H]^{1+}$ fragment ion. As no fragmentation energy was applied, this observation was putatively explained by the natural breakdown of rhodamine on the glass slides, caused by extended storage (over 1 year) prior to the experiment. The intensity of the rhodamine B peak was correlated with the area of each sampling spot calculated by use of ImageJ for both contact and surface sampling modes; the calculation was made by linear regression. The R^2 value for contact sampling mode was 0.386, indicating a putative, weak correlation; for surface sampling, the R^2 value was 0.0217, indicating no correlation. The mean measured sampling area was 1.8 mm^2 with a standard deviation (SD) of 0.9 mm^2 for contact sampling and 8.0 mm^2 with a standard deviation of 0.5 mm^2 for surface sampling (**Figure 3.2**); for both sampling modes, the volume of solvent dispensed was the same (1 µl). For surface sampling, the pipette was suspended approximately 0.2 mm above the surface of the glass slide. The mean intensity for surface sampling was thus approximately double that of contact sampling, whilst the mean sampling area for surface sampling was approximately quadruple that of contact sampling. An anomalous sampling spot at location 17 (7.5 mm^2; shown as the largest sampling spot in **Figure 3.2a**), putatively a result of an imprecise vertical positioning of the pipette tip and the subsequent spillage of solvent, was associated with the highest recorded intensity of the rhodamine peak in the contact sampling series.

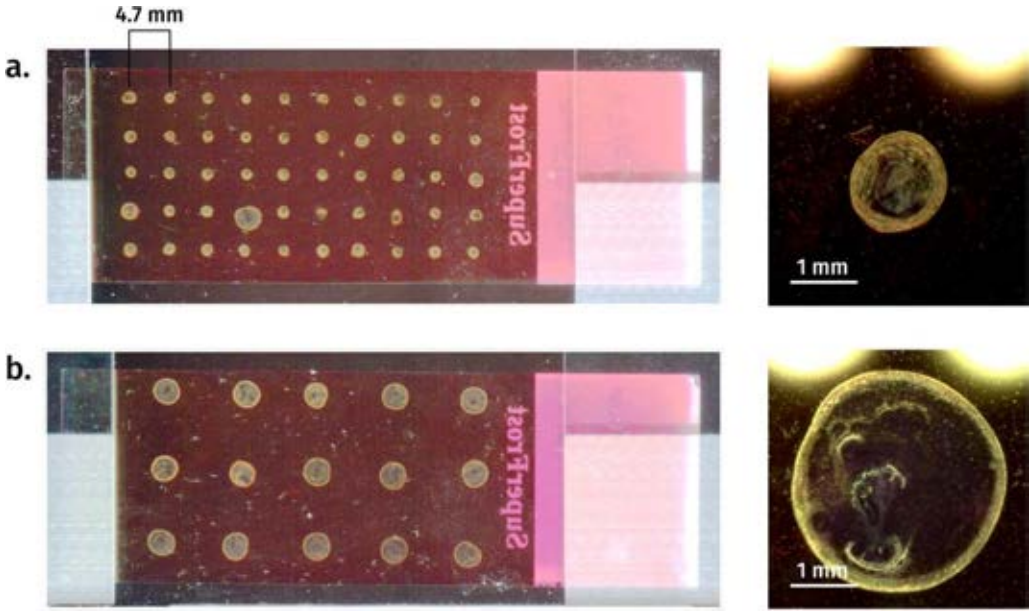


Figure 3.2. Optical images of the rhodamine B-coated glass slides used to determine sampling repeatability. **a**. Contact mode; representative sampling spot surface area 1.8 mm^2. Anomalous sampling spot described in main text (surface area 7.5 mm^2) is seen in row 4, column 4. An example measurement of the inter-sample distance is shown for the first two sampling spots. **b**. Surface sampling mode; representative sampling spot surface area 8.2 mm^2.

Sampling repeatability was calculated as described before,[294] by expressing the standard deviation of the absolute intensities of the rhodamine B peak as a percentage of the mean. The values obtained were 47.4% for contact sampling and 65.8% for surface sampling on glass slides. These results indicate significant variability in the LESA process. Part of the variability observed for surface sampling could be explained by the collapse of the nanoelectrospray (that is, a rapid decrease of the nanoelectrospray current to zero, accompanied by a reduction of the observed signal to the level of noise), observed for three sampling attempts. The behaviour may have been caused by incomplete reaspiration of the solvent into the pipette tip following sampling, leading to the premature depletion of solvent available for nanoelectrospray delivery, or alternatively leading to the formation of air bubbles or an air gap within the pipette tip which would subsequently destabilise the nanoelectrospray. Aspiration of particles from the surface of the glass slide, which would subsequently obstruct

the aperture of the nanoelectrospray nozzle, could have also contributed to nanoelectrospray destabilisation. Finally, the volume of solvent dispensed (1 µl) was halved compared to the volume routinely used for the sampling of bacteria. Whilst this reduced the spread of solvent on the glass slides, thus allowing the acquisition of a greater number of data points from a single slide, it also amplified the variability in the actual solvent volumes dispensed and aspirated due to the inherent limitations of the TriVersa NanoMate's pressure system.

Inconsistencies were detected in the arrangement of sampling spots on the glass slide. Consistent with prior literature,[116] the inter-sample distances measured from centre to centre of contact sampling spots (as shown in **Figure 3.2a**) were found to vary between 3.8 and 5.1 mm (expected distance 4.5 mm), indicating an imprecision in the *x* and *y* axis movements of the pipette arm. The mean measured distance was 4.4 mm in the *x* axis and 4.5 mm in the *y* axis, suggesting that the variation is random and converges to expected values over a large number of samples. Similarly, the variation in sample spot size was indicative of an imprecision in the *z* axis movements of the mandrel, variation in the volume of aspirated solvent, or a combination of both.

3.3.2 Sampling repeatability on agar plates

The repeatability of sampling on agar media containing rhodamine B was subsequently investigated in both contact and surface sampling modes. The area of the sampling spots was not investigated as the sampling of agar media did not result in visible, measurable changes to the media under visible or UV light. The intensities of three characteristic *m/z* ranges, *m/z* 443-446, *m/z* 652-662 and *m/z* 1100-1150 were measured (**Figure 3.3a**). The range *m/z* 443-446 encompassed the entire isotopic distribution of rhodamine B (*m/z* 443.23). The ranges *m/z* 652-662 and *m/z* 1100-1150 both contained highly abundant, characteristic peaks derived from agar media without interference from peaks derived from solvent. The most abundant

peaks in these ranges were *m/z* 652.41, 659.35, 1100.59, 1124.65 and 1146.63, all singly charged. For each range, the summed intensity within the range was plotted separately. The results are displayed in **Figure 3.3. Figure 3.3a** indicates the location and with of the three m/z ranges investigated with respect to a representative agar mass spectrum. **Figure 3.3b** and **3.3.c** display the recorded absolute intensities in the three *m/z* ranges for contact sampling and for surface sampling, respectively.

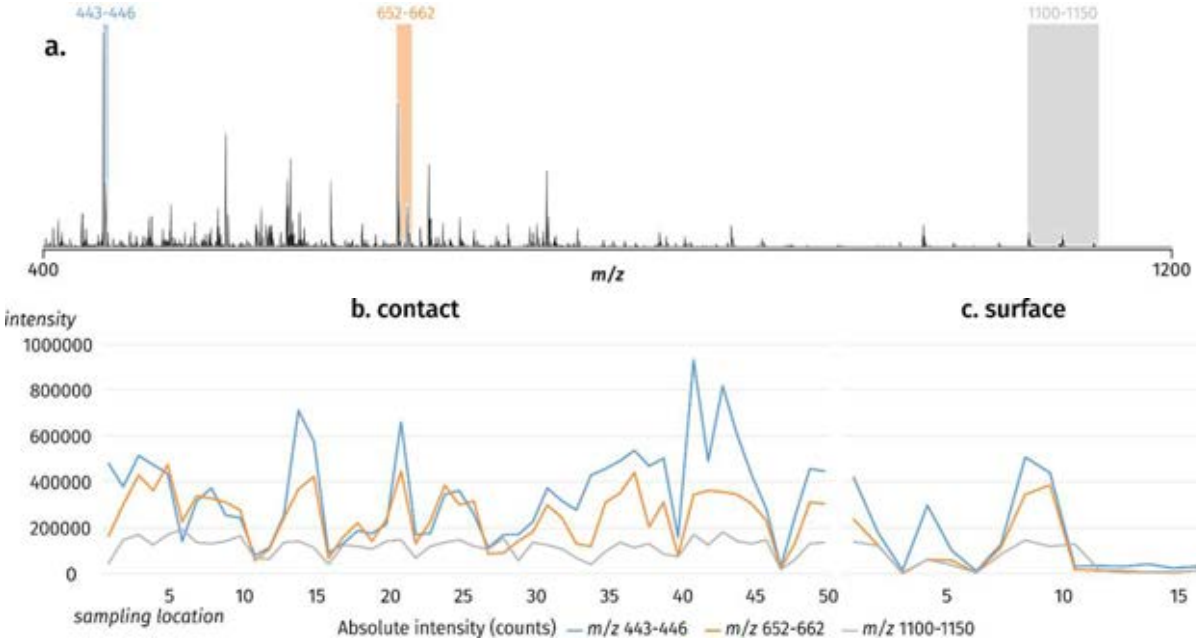


Figure 3.3. Absolute intensities of three *m/z* regions in mass spectra acquired from agar media containing rhodamine B. **a**. A representative mass spectrum of rhodamine-spiked agar media. The three *m/z* ranges chosen for analysis are shown. **b**. Contact mode. **c**. Surface sampling mode. The *x*-axis (sampling location) defines the arbitrary number assigned to each sampling spot based on chronological order of data acquisition.

The absolute repeatability was 45.0% for contact sampling and 115.5% for surface sampling, calculated as the mean of absolute repeatabilities calculated for each *m/z* range as described in section 3.3.1. Relative repeatability was also calculated as described previously[294] to estimate the influence of the ion source instability on the repeatability of the LESA process. Briefly, mass spectra were normalised to the total ion current recorded across all three evaluated *m/z* ranges and the relative intensities between the three ranges were subsequently

compared to return the relative repeatability value. The values were 25.4% for contact sampling and 39.8% for surface sampling. For comparison, the relative repeatability values reported in the literature for DESI approximated 5%.[293] Nanoelectrospray stability in LESA may be affected by the aspiration of debris or solid sample fragments into the pipette tip and the subsequent blockage of the nanoelectrospray nozzle; this occurrence is particularly common in the sampling of bacteria. It can also be induced by the loss of solvent to a hydrophilic sample surface during sampling. The values indicated that the instability of the nanoelectrospray significantly contributed to, but did not account for all observed variability in the LESA process. As suggested by data collected from glass slides and discussed above, part of the unexplained variability may be tentatively attributed to the imprecise operation of the TriVersa NanoMate itself.

3.3.3 Sampling of *E. coli*

Sampling locations on colonies of *E. coli* K-12 were divided into centre and edge locations; the classification was based on a common morphological feature of all *E. coli* colonies, the presence of a central depression depicted in **Figure 3.4**. Sampling locations were assigned to the edge group if any of the sampling area encompassed by the pipette tip at the given location fell outside of the central concave area. Three example analyses of *E. coli* colonies are shown in **Figure 3.5**. Photographs of all colonies used for the assessment of reproducibility are included in **Appendix 2**.

Figure 3.4. Photograph of a colony exemplifying the characteristic elevation profile with a central depression.

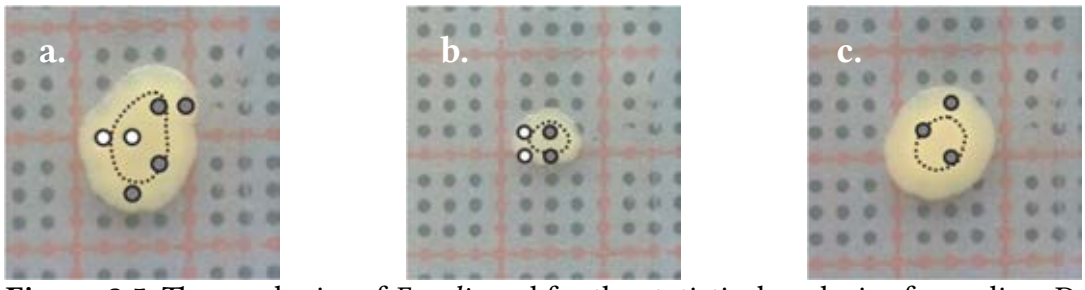


Figure 3.5. Three colonies of *E. coli* used for the statistical analysis of sampling. Dots delineate selected sampling locations. Grey indicates that a mass spectrum was acquired from a given location; white indicates no successful mass spectrum acquisition. The dotted line indicates the boundary between the centre and the edge region. **a**. Colony incubated at 20 ℃ for 7 days. 14 sampling attempts in total (7 centre, 7 edge); 1 protein mass spectrum acquired. **b**. Colony incubated at 37 ℃ for 24 hours. 10 sampling attempts in total (7 centre, 3 edge); 1 protein mass spectrum acquired. **c**. Colony incubated at 20 ℃ for 7 days followed by 1 day of storage at 4 ℃. 8 sampling attempts in total (3 centre, 5 edge); 5 protein mass spectra acquired.

The outcomes of 124 LESA sampling attempts on colonies of *E. coli* K-12 incubated and stored under a range of conditions are summarised in **Table 3.1**. Success was defined as an acquisition of a mass spectrum over five minutes without nanoelectrospray destabilisation and loss of signal during acquisition. Successfully acquired mass spectra were inspected for the presence of protein peaks. The percentage success rate, calculated by dividing the number of sampling attempts resulting in the acquisition of a mass spectrum over the total number of sampling attempts made, was calculated separately for each growth and storage condition

set; the success rates for sampling in edge regions versus centre regions were also calculated separately to systematically investigate in which area of the colony LESA sampling may be more effective. Calculations were made for both the total success rate (for sampling attempts resulting in the acquisition of any mass spectrum) and the protein acquisition success rate (for sampling attempts resulting in the acquisition of mass spectra containing peaks attributable to protein). AGC was active and thus the intensities of observed peaks were not compared.

Growth and storage conditions	Total sampling attempts	Sampling attempts in edge regions	Total mass spectra/protein mass spectra acquired from edge regions	Sampling attempts in centre regions	Total mass spectra/protein mass spectra acquired from centre regions	Total mass spectra/total protein mass spectra	Mass spectrum/protein mass spectrum success rate (%)
3d 20°C 4d 4°C	13	7	4/4	6	3/2	7/6	54/46
6d 20°C 1d 4°C	26	19	9/3	7	4/2	13/5	50/19
7d 20°C 1d 4°C	28	16	6/3	12	10/6	16/9	57/32
7d 20°C	33	17	8/4	16	5/3	13/7	39/21
24h 37°C	24	13	6/2	11	7/2	13/4	54/17
Total	**124**	72	33/16	52	29/15	**62/31**	**50/25**

Edge success (%):	**46/22**
Centre success (%):	**56/29**

Table 3.1. Summary of results generated by LESA sampling of colonies of *E. coli* K-12. For a fully descriptive list of growth and storage conditions used, see section 3.2.1.3.

Overall, the acquisition of mass spectra was successful following 50% of the sampling attempts; 25% of the sampling attempts yielded mass spectra containing a detectable protein signal. Sampling was more successful for mass spectrum acquisition (56% vs 46%) within the central depression of the colonies (29% vs 22% for acquisition of mass spectra containing peaks attributable to protein); the differences were not statistically significant at the .05 level.

Any putative effect favouring sampling at the centre of the colonies could be rationalised by the relatively uniform height distribution of colony matter within the central area. Edge locations ranged from sampling locations falling marginally outside of the central concave area to locations directly bordering agar media and thus subject to a steeper height gradient, which complicated the selection of a suitable sampling height and increased the risk of aspirating solid colony matter. Thus, nanoelectrospray destabilisation could be expected to occur more readily following sampling at edge locations.

Whilst the statistical analysis was not carried further, the protocol for sampling microbial colonies was improved in direct response to the practical challenges encountered during the experiment. The rigid utilisation of the universal adapter plate grid limited the choice of suitable sampling locations and severely reduced the number of sampling locations available on smaller colonies; thus, for all future experiments, the target sampling location was fixed a single, central position and the agar plate itself was moved to bring the desired sampling location on the colony in line with the predetermined sampling position programmed in the system. Thus, all of the available colony area was rendered accessible for sampling. The selection of sampling height was found to be critical for the reliable acquisition of mass spectra; irregular colony topology necessitated frequent manual adjustments to sampling height and thus decreased reproducibility. Both the variation of growth conditions and the variability inherent to bacterial inoculation and incubation led to high variability of colony size and topology, complicating the selection of sampling height and thus contributing to reduced sampling success. The size and shape of colonies for LESA sampling was subsequently rendered consistent by first inoculating a liquid culture, incubated overnight, followed by the dispensation of identical aliquots of the liquid culture onto solid media. The difference was most prominent for yeast colonies grown for work described in Chapter 6

(**Figure 3.6**). Colonies inoculated directly from glycerol stocks exhibited an irregular margin and an irregular, raised profile (**Figure 3.6a**) which negatively impacted correct sampling height selection and necessitated adjustments for resampling the same colony. Colonies inoculated from liquid cultures exhibited a uniform, flat surface which allowed repeated, consistent sampling of the same colony at the same parameters (**Figure 3.6b**). To estimate the effectiveness of the changes introduced to the protocol, an overall success rate for protein mass spectrum acquisition was calculated at a later stage for the experiment carried out on *E. coli* K-12 as described in Chapter 4, section 4.3.7. Sampling was carried out in the centre of colonies. Sampling height was carefully adjusted to obtain contact of the pipette tip with the colony without penetration. The calculated value of the success rate for protein mass spectrum acquisition was 49%, indicating a twofold improvement compared to the result of the initial experiment described in this section (25%). For *Pseudomonas aeruginosa* and *Staphylococcus aureus*, two clinical isolates analysed alongside *E. coli* (see Chapter 4), the values were 87% and 97%, respectively; this indicated that species-specific colony and cell structure may also affect LESA sampling, and that *E. coli* was a comparatively challenging model system.

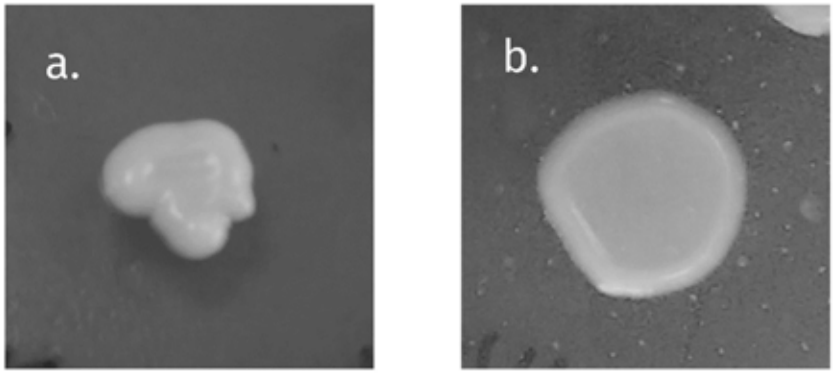


Figure 3.6. Topology of *Saccharomyces cerevisiae* colonies inoculated **a**. directly from glycerol stocks, **b**. from liquid culture. Uniform, lateral growth was promoted, yielding a consistent flat surface amenable to repeated, reproducible LESA sampling.

3.3.4 Investigation into the characteristics of LESA sampling spots

3.3.4.1 Optical and confocal imaging

Optical imaging and microscopy were used to investigate the topology of LESA sampling spots on colonies of *E. coli* K-12 and attempt to assess the disruption of bacterial colonies caused by LESA sampling. A representative image of a contact LESA sampling spot on the surface of an *E. coli* colony is shown in **Figure 3.7**. At higher magnification, outlines of individual, intact bacterial cells were observed at the centre of the sampling spot. The generated images were not significantly different to those observed on areas of the colony which were not subject to sampling; no evidence of solvent-induced lysis was detected within the LESA sampling spots. Topographical data derived from optical imaging was challenging to interpret. Chemical imaging methods were therefore proposed as a more suitable approach.

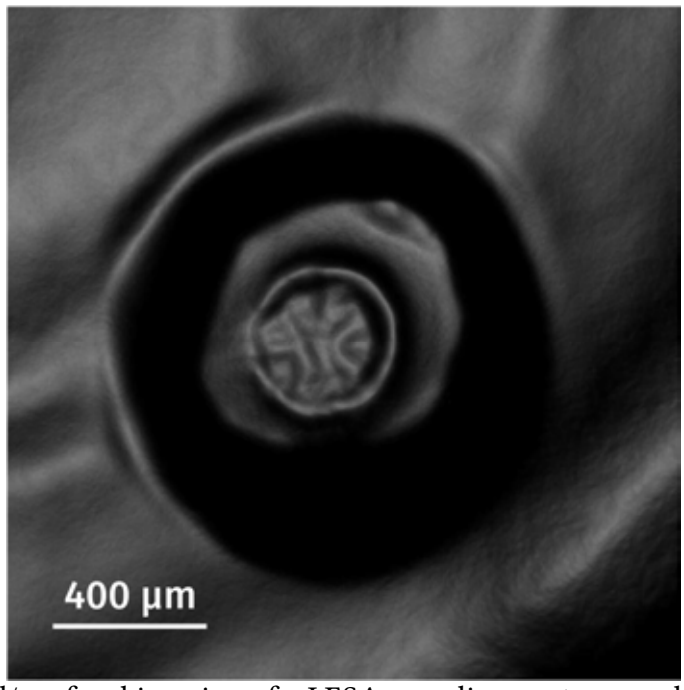


Figure 3.7. Optical/confocal imaging of a LESA sampling spot on a colony of *E. coli* K-12. Higher brightness indicates higher topology. 5× magnification of the sampling spot is shown. The dark ring putatively corresponds to the imprint of the pipette tip. The mismatch between the outer diameter of the observed imprint (~1100 µm) and the outer diameter of the pipette tip (750 µm) may have been introduced by solvent spread, partial drying of the colony in air in the time between LESA sampling and confocal microscopy, or both.

3.3.4.2 MALDI

Attempts were made to develop a protocol for the use of MALDI MS for the characterisation of LESA sampling spots on bacterial colonies directly on agar. (Due to issues with sample integrity, described in detail below, all MALDI data displayed here were acquired from plain bacterial colonies not sampled by LESA.) MALDI MS imaging of bacteria on nutrient media has been performed by multiple groups;[245-248] no consensus protocol has, however, been developed so far due to numerous challenges associated with sample preparation for bacterial on agar. Issues include insufficient matrix saturation, uneven matrix coverage, as well as sample flaking and disintegration under vacuum which may lead to instrument damage.[248] Following recommendations published by Dorrestein and coworkers,[248] colonies were excised from agar plates along with a square of agar media approximately 15 mm × 15 mm,

and deposited on glass slides, following which they were air dried at 40 ℃ for 18 hours prior to CHCA matrix deposition and MALDI analysis. Issues were, however, encountered at the transition of samples into the vacuum, which caused partial or complete sample disintegration (**Figure 3.8**). Due to the possibility that LESA sampling may further weaken the structure of the colonies, only intact colonies were used for optimisation. Following recommendations described in literature,[248] a 1:1 mix of CHCA and DHB matrices was used to increase adhesion between the sample and the glass slide; the colony thus prepared for MALDI MS analysis is shown in **Figure 3.9a**. The alteration in matrix composition increased sample durability sufficiently to allow the acquisition of an image, shown in **Figure 3.9b;** a mean mass spectrum generated from the image is shown in **Figure 3.9c**. Resultant mass spectra were of poor quality; the only difference recorded between the agar media and the colony matter was a reduction of the total ion current. No peaks were confidently assigned by their intact mass to any known *E. coli* metabolites. Due to the significant difficulties in sample preparation coupled with unsatisfactory results, the MALDI approach was not continued.

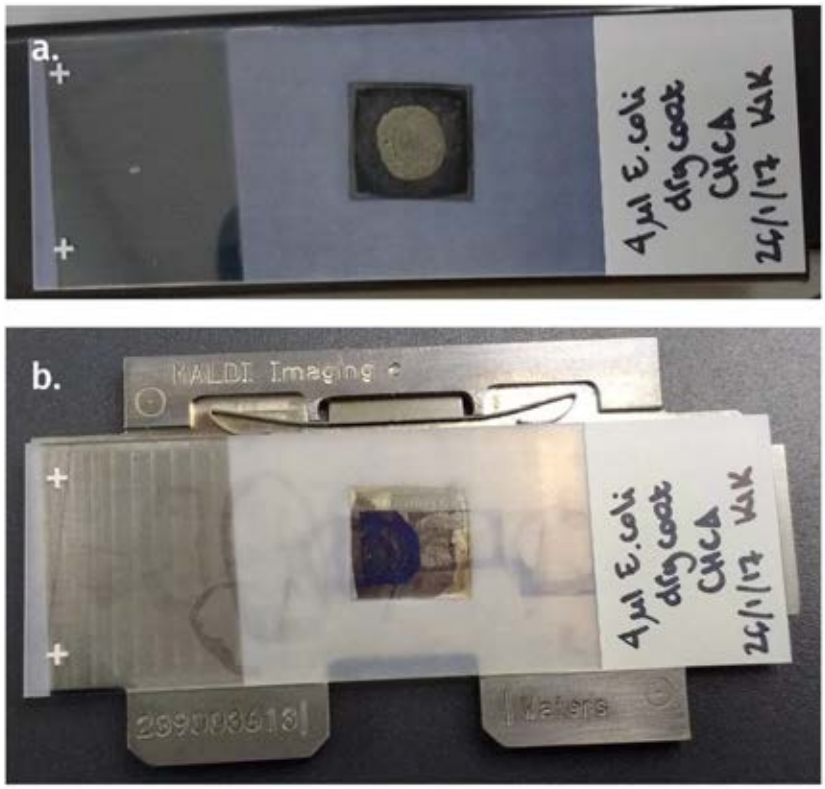


Figure 3.8. An example case of sample disintegration in MALDI. **a**. Colony of *E. coli* K-12 on a glass slide coated with CHCA alone, immediately prior to attempted MALDI analysis. **b**. The same glass slide following introduction into vacuum and subsequent withdrawal from the instrument. The surface of the slide holder is visible through the portion of the glass slide formerly occupied by the sample.

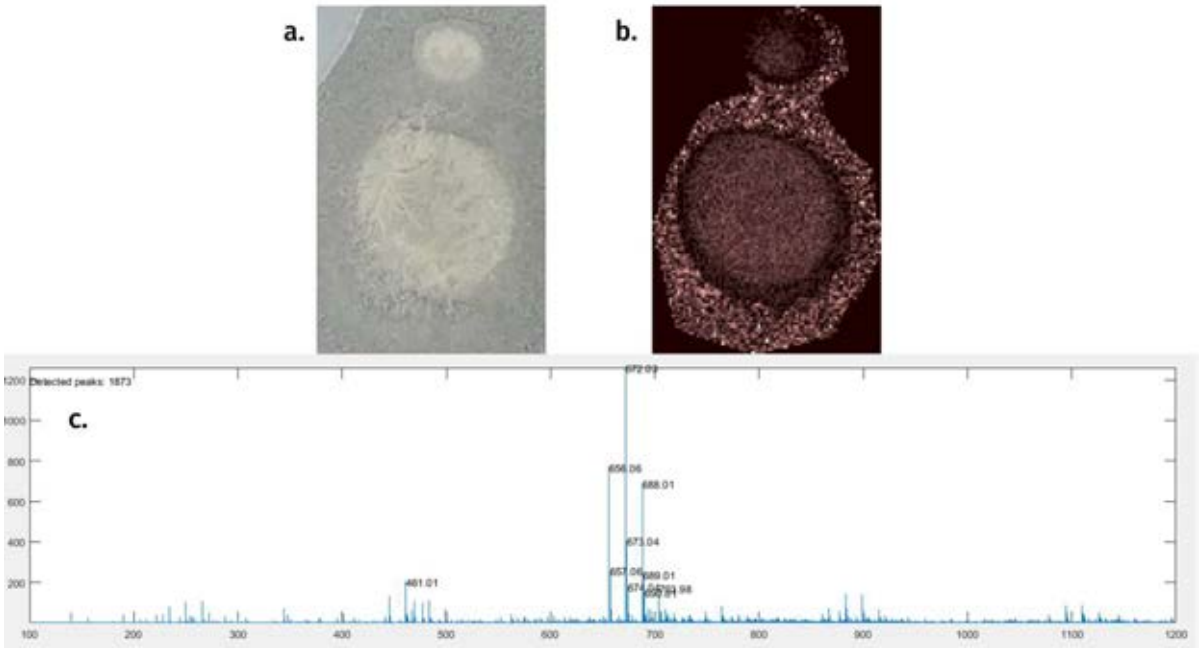


Figure 3.9. MALDI MS imaging of *E. coli* K-12. **a**. Colony coated with 1:1 DHB/CHCA immediately prior to MALDI MS analysis. **b**. Total ion current image generated from the colony (pixel size 90 μm). **c**. Mean mass spectrum generated from the image. No ions could be assigned to putative metabolites.

3.3.4.3 DESI

Sample preparation for DESI was carried out as for MALDI, by air-drying colonies excised along with underlying nutrient media. As described in Chapter 1, DESI requires a hard, non-conductive and uniform surface to maintain consistent signal. The topological variability of bacterial colonies is thus fundamentally incompatible with DESI; furthermore, any attempt to direct a jet of solvent at a hydrated colony would most likely result in the delocalisation of colony matter, formation of aerosolised bacterial cells and subsequent contamination of the ion source. Retaining structural integrity of the samples remained a challenge, in this instance due to the tendency of dry colony material to be displaced by the jet of solvent. Sample disintegration was, however, observed at a much smaller scale and did not pose a risk to the instrument; thus, colonies previously sampled by LESA, which could have been

structurally compromised by the process, could be sampled immediately following brief optimisation of DESI geometry and parameters.

A colony of *E. coli* K-12 previously subjected to LESA sampling was used as a test system to determine whether DESI would yield information on the disturbances introduced in the colony by LESA sampling, primarily concerning analyte extraction and displacement. A photograph of the intact colony prior to LESA sampling is shown in **Figure 3.10a**. A photograph of the colony following LESA sampling and partial drying is shown in **Figure 3.10b**; the sampling spot is clearly visible. A total ion current image generated by DESI is shown in **Figure 3.10c**. No changes attributable to sampling were initially detected. Irregular areas of low signal intensity, attributable to sample flaking, were detected in the data; these are indicated by white arrows in **Figure 3.10c**. One peak (*m/z* 257.1465, MW_{obs} = 117.0787 Da) was attributed to the $[2M + Na]^+$ ion of free valine (MW_{theor} = 117.0790; Δppm = -3.4) by a metabolite database search and subsequent manual verification of intact mass. Three other peaks with similar spatial distributions, highly abundant in the colony but absent in the agar background (*m/z* 118.0899, 140.0714, 156.0454), corresponding to $[M+H]^+$, $[M+Na]^+$ and $[M+K]^+$ ions of the same compound (MW_{obs} = 117.0826, 117.0822 and 117.0822, respectively) were also initially assigned to free valine (Δppm = 27.5) due to the absence of other candidate metabolites found in the database within ±0.01 Da (±85 ppm) of the monoisotopic mass of the peaks under investigation. (For comparison, the known standard raffinose peak had a calculated Δppm of 11.9.) ChemCalc (www.chemcalc.org/mf_finder) was used to calculate putative molecular formulae from the observed monoisotopic mass; the three candidate formulae with the lowest ppm differences were $C_5H_{10}FN_2$ (Δppm =5.1), $C_6H_{14}P$ (Δppm = 9.5) and $C_5H_{11}NO_2$ (valine, Δppm = 27.5). No other metabolites in the mass spectrum were identified. The mean mass spectrum acquired from the colony and single ion images

corresponding to valine peaks are shown in **Figure 3.11**. Sample flaking, not clearly visible by optical imaging, was observed in the single ion data as irregular patches of low signal intensity (compare **Figure 3.10c**). Upon application of total ion current normalisation, for one ion in the valine series (*m/z* 140, $[M+Na]^{1+}$), an outline of the LESA sampling spot was revealed, indicating extraction of the analyte by LESA (**Figure 3.12**). Corroboration with LESA MS data was not possible as routine data acquisition on the orbitrap instrument was carried out from *m/z* 150 onwards. No other ions of known or unknown identity yielded comparable images.

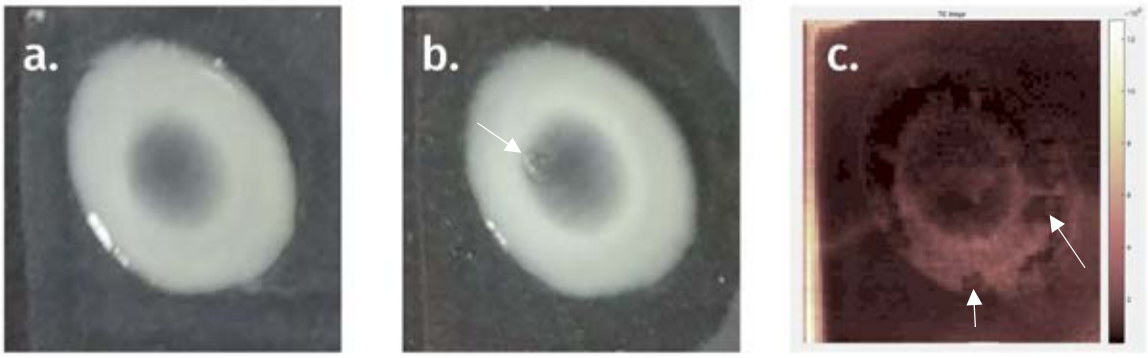


Figure 3.10. DESI mass spectrometry imaging sample preparation. **a**. A colony of *E. coli* K-12 immediately following excision from the petri dish. **b**. Colony following LESA sampling and partial drying; the LESA sampling spot is visible in the upper left quarter of the colony and marked with an arrow. **c**. Total ion current image of the colony (pixel size 100 μm); areas of sample flaking and disruption are indicated with arrows.

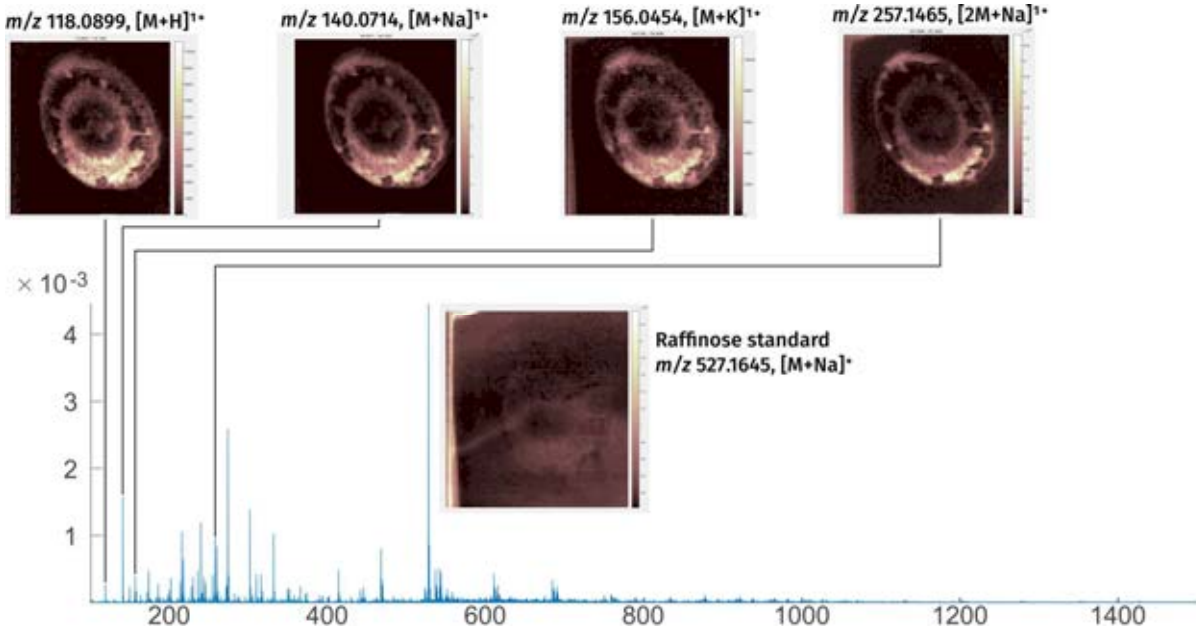


Figure 3.11. Mean mass spectrum derived from the image shown in Figure 3.8c. Single ion images corresponding to peaks putatively assigned to valine by their intact mass are shown in the upper row.

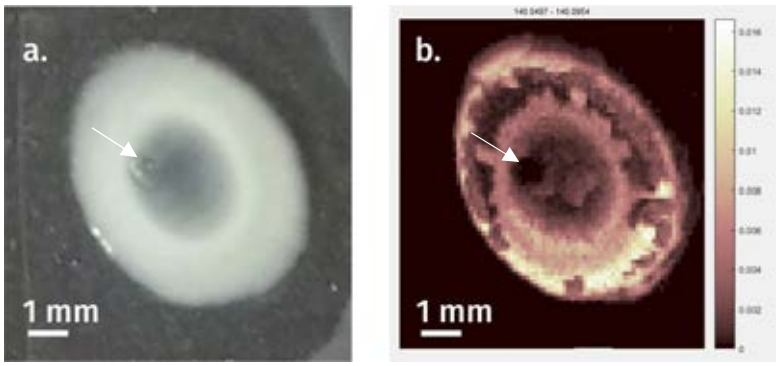


Figure 3.12. The outline of the LESA sampling spot revealed by DESI. **a**. Optical image of the colony prior to DESI sampling. **b**. Single ion image of the valine ion centred at *m/z* 140.1; total ion current normalisation was applied. The position of the LESA sampling spot is marked with arrows.

3.4 Conclusions

The repeatability of LESA sampling on model surfaces relevant to microbial sampling was quantified. An attempt to attribute sources of the observed variability was made. For sampling of glass slides, a weak correlation between surface area sampled and the intensity

of the reporter peaks was detected. A major source of variability stemming from the imprecision of the NanoMate movements was identified based on the inter-sample distances measured on the *x* and *y* axes, and the variability in solvent spread attributable to variation on the *z* axis and to variation in the volume of solvent aspirated/dispensed. For imaging purposes, this imprecision may introduce variability in the positioning and diameter of sampled areas (and thus pixels), affecting resolution. For bacterial sampling, variability in the sampling height may affect the undesirable aspiration of colony material and thus the reproducibility of sampling. For sampling of agar media, the contribution of source instability was quantified. The findings described here, coupled to findings reported elsewhere prior to this project,[116] suggest that there exists variability inherent to the LESA apparatus which affects the repeatability of mass spectrum acquisition. Whilst the variability detected in the positioning of the LESA pipette tip would be relevant to imaging experiments, it was not considered significant enough to affect experiments involving the acquisition of single mass spectra from bacterial colonies. The selection of appropriate sampling height was improved by alterations to the sample preparation protocol, yielding colonies of a consistent topology.

The reproducibility of LESA sampling on bacterial colonies, measured as a percentage success rate (where success was defined as the maintenance of a stable nanoelectrospray over the entire duration of mass spectrum acquisition, coupled to the detection of peaks attributable to protein in the resultant mass spectrum), was investigated. Colonies of *E. coli* yielded a 50% success rate for mass spectra in general, and 25% for mass spectra containing peaks attributable to protein. A more in-depth analysis was not attempted due to the large number of uncontrolled, unanticipated variables such as the morphology of the colonies, colony density, and height of the agar media, compounded by the variability introduced by manual sampling height adjustment. The results highlighted areas of improvement to the sample

preparation and LESA sampling protocols which were subsequently carried over to the remainder of the project. The sampling location grid was abandoned in favour of fully manual location targeting, which increased the number of locations available for sampling from any one colony and allowed for the selection of locations of optimum topology for sampling. The method of preparation for microbial samples was altered to yield more uniform sampling surfaces. The alterations resulted in significantly improved outcomes, subjectively detectable during ongoing work and quantified at a later stage in the project (see Chapter 4).

The attempts to characterise LESA sampling spots by other methods were unsuccessful. DESI constituted the most promising approach; the applicability of the information contained within the imaging data was, however, limited and not comparable to methods used before for the analysis of LESA sampling on tissue sections.[116] Optical/confocal imaging data of the LESA sampling spots proved difficult to interpret; as part of future work, confocal imaging could, however, potentially be used to investigate both the topological changes occurring on the surface of LESA sampled colonies and (in conjunction with fluorescent solvents) the depth and volume of the colony accessed with each sampling, crucial for imaging applications.

Following the conclusion of the work described in this chapter, the approach developed for protein analysis in *E. coli* was translated to the analysis of clinical isolates, described in the next chapter.

CHAPTER 4

Top-down identification of proteins in bacterial colonies

The work presented in this chapter has been published in K.I. Kocurek, L. Stones, J. Bunch, R.C. May and H.J. Cooper. Top-down LESA mass spectrometry protein analysis of Gram-positive and Gram-negative bacteria. *JASMS* 2017, 28: 2066–2077. A copy of the paper is provided in **Appendix 1**.

4.1 Introduction

The main advantage of LESA over other ambient ionisation techniques is its proven capability for the extraction of proteins, including low abundance species. As described in the general introduction in Chapter 1, this aspect of the technique was previously exploited for the extraction and analysis of proteins in colonies of *Escherichia coli* K-12.[117, 266] The original study demonstrated that, by driving the pipette tip of the TriVersa NanoMate directly into contact with the colony surface and sampling with an acetonitrile-based solvent, it was possible to extract both periplasmic and cytosolic proteins directly from microbial colonies growing on agar media.

In order to conclusively show, however, that this approach is suitable for real-world clinical and research applications, it was necessary to further optimise the protocols for the sampling of a much wider variety of species, including clinically relevant strains. This expansion should be accompanied by a corresponding increase in the number of identified proteins, demonstrating greater potential for microbiological research applications. Finally, it was proposed that the differences in protein mass spectra of various microbial species could be

used for their unambiguous identification, potentially in a clinical setting. All of these aims were addressed by the work described in this chapter.

Protein identification by LESA MS directly from microbial colonies is described for eight bacterial strains: *Escherichia coli* K-12, *Escherichia coli* BL21 mCherry, *Pseudomonas aeruginosa* PS1054, *Staphylococcus aureus* MSSA476, *Streptococcus pneumoniae* D39, *Streptococcus oralis* ATCC 35037 *Streptococcus gordonii* ATCC 35105 and an unknown species of *Staphylococcus*. The capabilities of LESA MS for the differentiation of bacterial species are explored using the three streptococci as a model system. The three species are representative of the viridans group, which are notoriously difficult to distinguish by commercially available MALDI-TOF identification systems due to the high similarity of their fingerprint mass spectra.[239, 295, 296] The impact of varying growth media is discussed, as is the influence of storage conditions on the composition of acquired mass spectra.

4.2 Experimental section

4.2.1 Colony culture

Colonies of *Escherichia coli* K-12, *Escherichia coli* BL21 mCherry, *Pseudomonas aeruginosa* PS1054, *Staphylococcus aureus* MSSA476 and the unknown species of *Staphylococcus* were inoculated directly from glycerol stocks onto LB agar plates. *Streptococcus pneumoniae* D39, *Streptococcus oralis* ATCC 35037, and *Streptococcus gordonii* ATCC 35105 were re-streaked from donor plates and cultured on horse blood agar either in open air or under semi-anaerobic conditions induced by placing the plates in a candle jar. All strains were incubated as described in Chapter 2, section 2.2.1.3. Culturing conditions were altered to investigate the influence of cold storage on mass spectra; see below.

4.2.2 Sampling

The solvent system used for liquid extraction consisted of acetonitrile, water and formic acid; the relative proportions were initially 39.5:59.5:1 for Gram-negative species (this was changed to 40:60:1 prior to the introduction of BSL2 organisms into the study), and either 40:60:1 or 50:45:5 for Gram-positive species. The dwell (solvent contact) time was 3 seconds for *E. coli* and 15 seconds for all other species. Optimisation of sampling for Gram-positive bacteria is described in section 4.3.2.

4.2.3 Mass spectrometry, MS/MS and protein identification

Data acquisition, in both full scan and tandem modes, was carried out as described in Chapter 2, section 2.2.3. Protein identification was achieved by use of ProSightPC software, versions 3.0 to 4.1 alpha (Chapter 2, section 2.2.5). *De novo* sequencing was carried out manually.

4.2.4 Investigation of the influence of cold storage

Colonies of *E. coli* K-12, *P. aeruginosa* PS1054 and *S. aureus* MSSA476 were inoculated from a glycerol stock onto three replica LB agar plates each to yield a total of 9 colonies per plate, divided into three sectors (**Figure 4.1**). Each plate was first incubated for 24 hours at 37 ℃, following which one mass spectrum per colony in the first sector was taken. The second set of mass spectra was acquired from the second sector of the plate following storage of the plates at 4 ℃ for 4 days; the plates were then returned to 37 ℃ for an additional 24 hours following which the final set of mass spectra was taken from the colonies in the last sector. Data acquisition was maintained for 5 minutes following the establishment of a stable nanoelectrospray.

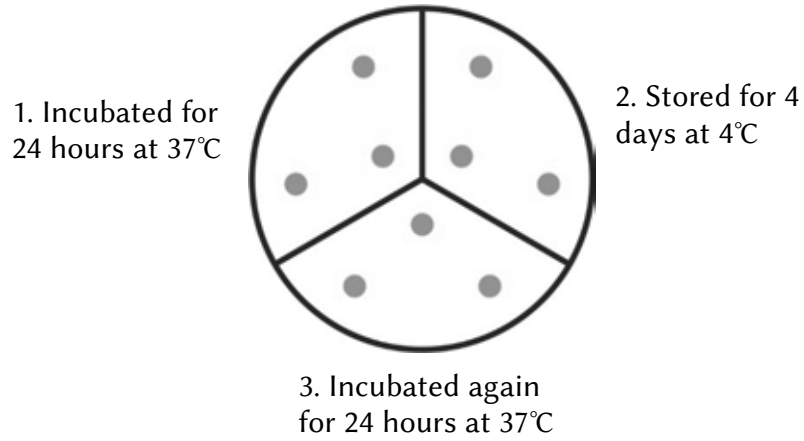


Figure 4.1. Schematic representation of a plate used to monitor the influence of sequential temperature changes on mass spectra. Three replica plates were prepared.

4.3 Results and discussion

4.3.1 Overview of mass spectra

Figure 4.2 shows representative mass spectra generated from the total of seven known bacterial species and strains investigated during this project: *E. coli* K-12, *E. coli* BL21 mCherry, *P. aeruginosa* PS1054, *S. aureus* MSSA476, *S. pneumoniae* D39, *S. oralis* ATCC 35037, and *S. gordonii* ATCC 35105. All colonies were sampled immediately following incubation at 37 ℃. The solvents used were based on acetonitrile, with a composition of 40:60:1 acetonitrile-water-formic acid for Gram-negative species and 50:45:5 for Gram-positive species (solvent optimisation is described in section 4.3.2). **Table 4.1** lists the proteins identified in each species; details are provided in section 4.3.5.

Table 4.1. List of proteins identified by LESA MS/MS (as described in Chapter 2, sections 2.2.3 and 2.2.5) across all sampled species included in this chapter. Names in bold highlight proteins which have not been previously detected in their intact form by any technique. Asterisks mark proteins identified in the previous study on *E. coli* K-12 using LESA MS.

m/z	**charge**	**MW_{obs}**	**ppm**	**ID**	**UniProt accession**	**% cov**	**conditions**	**modifications**
E. coli BL21								
959.1288	+6	5748.73	-0.8	YmdF	P56614	66	Incubation: 24h, 37℃ Storage: 4 days, room temp.	-Met
978.9706	+6	5867.78	-2.0	YciG	P21361	36		-Met
1101.2759	+7	7701.88	-1.5	YahO	P75694	91		-signal peptide
1189.5904	+7	8320.08	-1.3	UPF0337 protein YjbJ*	P68206	80		
1254.2614	+7	8772.78	-2.2	YdfK	P76154	31		
1494.9477	+7	10457.58	+1.02 Da	YbgS	P0AAV6	45		-signal peptide; putative disulfide 71-76 and deamidation
E. coli K-12								
923.0049	+10	9219.98	-2.0	HU-β*	P0ACF4	43	Incubation: 24h, 37℃ Storage: 2 days, room temp.	
953.9263	+10	9529.93	0.2	HU-α*	P0ACF0	40	Incubation: 24h, 37℃ Storage: 15 days, 4℃	
963.5065	+16	15399.99	+1.02 Da	H-NS*	P0ACF8	30	Incubation: 24h, 37℃ Storage: 2 days, room temp.	-Met
1019.5584	+5	5092.76	-1.9	SRA	P68191	11		
1039.4342	+7	7268.99	0.8	L29	P0A7M6	41	Incubation: 24h, 37℃ Storage: 15 days, 4℃	
1088.8935	+6	6527.32	-1.1	BhsA*	P0AB40	35		-signal peptide
1094.2255	+8	8745.75	-0.6	YnaE	P76073	62		
1212.1269	+6	7266.72	-0.4	CspC	P0A9Y6	37	Incubation: 24h, 37℃ Storage: 2 days, room temp.	-Met
1356.6814	+3	4067.02	0.3	CydX	P56100	46	Incubation: 24h, 37℃ Storage: 18 days, 4℃	fMet
P. aeruginosa PS1054								
672.2674	+9	6041.34	-1.8	L33	Q9HTN9	31	Incubation and storage: 4 days, room temp.	
723.8214	+10	7228.14	-2.0	L35	Q9I0A1	41		-Met
739.5860	+6	4431.47	-1.2	L36	Q9HWF6	45		
759.9734	+11	8348.63	-1.4	S21	Q9I5V8	29		-Met
826.5565	+11	9081.04	-0.3	HU-β	P05384	51	Incubation: 48h, 37℃ Sampled fresh	
951.8583	+8	7606.81	0.0	UPF0337 protein PA4738	Q9HV61	58		
956.3376	+6	5731.98	0.1	PA0039	Q9I793	24	Incubation: 24h, 37℃ Sampled fresh	-signal peptide, 4-42 disulfide
958.5127	+16	15320.09	3.5	PA5178	Q9HU11	27	Incubation: 48h, 37℃ Sampled fresh	-Met
983.5923	+10	9825.85	-1.6	Peptidylprolyl isomerase	Q9HWK5	45	Incubation: 24h, 37℃	-Met

990.3747	+8	7914.94	-1.4	L31	Q9HUD0	31	Storage: 4 days, room temp.	
996.4895	+14	13936.75	-0.8	Azurin	P00282	14		-signal peptide, disulfide
1029.1389	+7	7196.92	-0.6	L29	Q9HWE3	29	Incubation: 48h, 37℃ Sampled fresh	
1090.1381	+5	5445.65	0.4	PA2146	Q9I1W9	67		-Met
1223.5075	+7	8557.50	-0.7	PA4739	Q9HV60	61		-signal peptide (1-32)
1245.9733	+6	7469.80	-0.2	CspA	P95459	31		-Met
S. aureus MSSA476								
733.7894	+3	2198.35	0.6	Phenol soluble modulin α4	P0C826	100	Incubation: 24h, 37℃ Sampled fresh	fMet
759.6682	+4	3034.64	0.8	δ-hemolysin	Q9I5V8	100	Incubation: 48h, 37℃ Sampled fresh	fMet
921.1190	+6	5520.67	2.0	Uncharacterised protein SAOUHSC_01729	Q2FXV2	96	Incubation: 24h, 37℃ Sampled fresh	
1136.8509	+4	4521.39	2.0	Uncharacterised protein SAOUHSC_01135	Q2FZA4	82		fMet, sodium adduct
1148.2490	+6	6883.45	1.7	UPF0337 protein SAOUHSC_00845	Q2FZY9	89		
S. pneumoniae D39								
930.0688	+7	6503.43	0.3	UPF0337 protein spr1385	Q8CYJ5	85	Incubation: 24h, 37℃ Sampled fresh	-Met
984.6548	+7	6884.53	0.4	UPF0337 protein spr1626	Q8CYD7	76		
S. oralis								
963.0736	+7	6734.46	-2.0	UPF0337 protein HMPREF9950_0438	F9Q496	82	Incubation: 24h, 37℃ Sampled fresh	-Met
S. gordonii								
980.2427	+7	6854.65	-0.6	UPF0337 protein AWH02_03040	A0A0F5MKW3	71	Incubation: 48h, 37℃, semi-anaerobic Sampled fresh	-Met

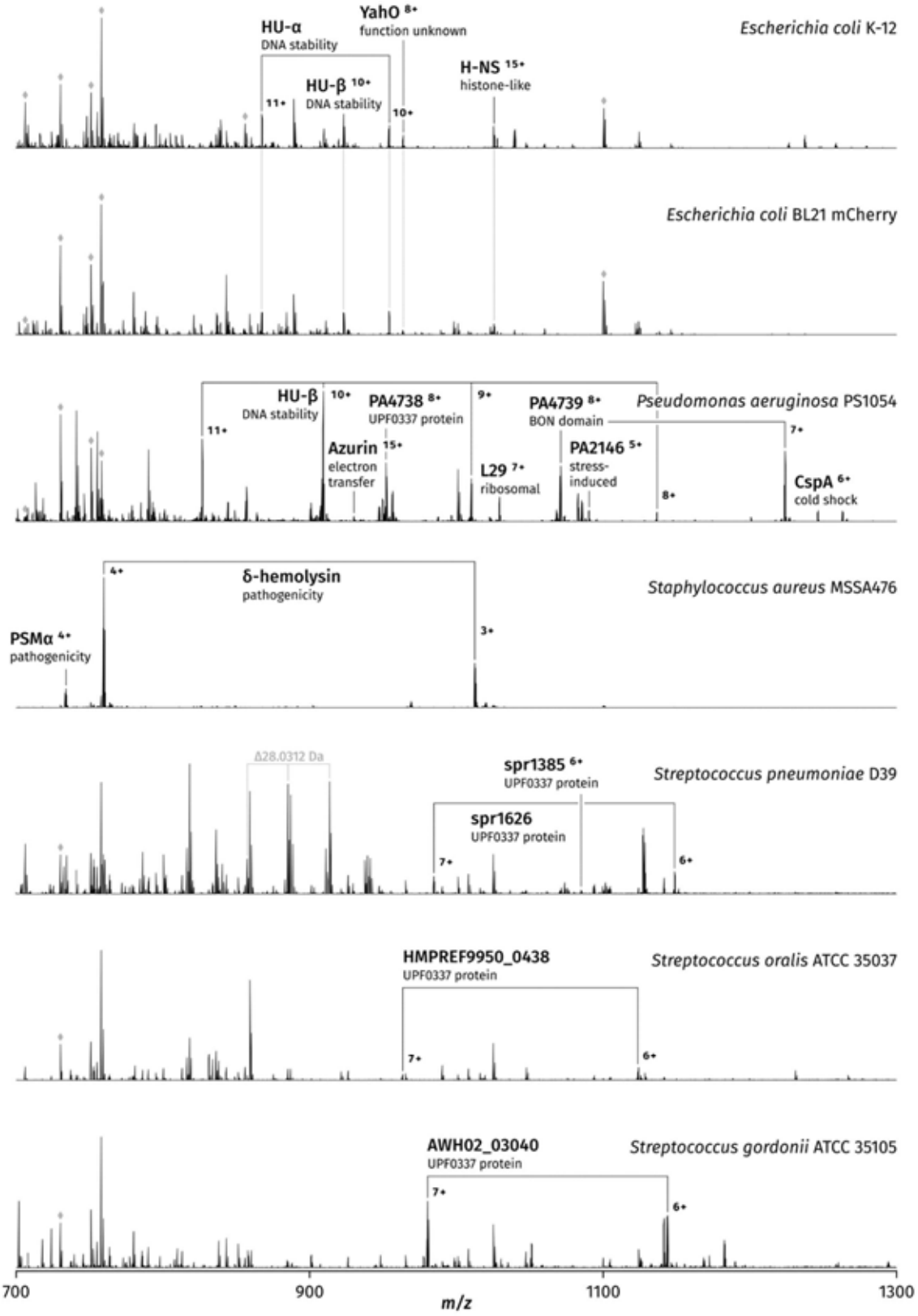


Figure 4.2. Comparison of LESA mass spectra acquired from the seven known strains of bacteria included in the study. All mass spectra were acquired immediately following

incubation at 37 ℃; the 40:60:1 acetonitrile-water-formic acid solvent composition was used for Gram-negative bacteria, whereas Gram-positive species were probed with the 50:45:5 solvent. The streptococci were incubated in semi-anaerobic conditions as required for optimum growth. The remaining strains were incubated in open air. Prominent background peaks are marked with diamonds. An unidentified compound characterised by a 28.0312 Da repeating unit is labelled in *S. pneumoniae* (see section 4.3.4).

Mass spectra generated from freshly cultured colonies of *E. coli* K-12 were remarkably similar to those shown in the prior study,[117] consistent with the reproducibility of the method demonstrated in Chapter 3. The lower end of the mass range (approximately *m/z* 600-800) contained the majority of singly charged peaks; these corresponded chiefly to background signal (**Figure 4.3**). (The removal of background interference by use of FAIMS is described in Chapter 5). From *m/z* 600 onwards, a total of 44 multiply charged protein peaks were observed in the representative mass spectrum, corresponding to 14 unique protein masses (**Table 4.2**). Identification was performed based on intact mass and MS/MS data, putatively assigned by ProSightPC software and verified manually; unless otherwise noted, the lack of listed identification indicates that a given protein was not selected for MS/MS analysis (see section 4.3.5 for details). Among these, peaks corresponding to the DNA-binding proteins HU-α and HU-β, each occurring in charge states 9+ to 12+ (*m/z* 1060.23, 954.02, 867.39 and 795.10 respectively for HU-α; *m/z* 1025.67, 923.11, 839.28 and 769.42 for HU-β), were reliably and consistently seen across all mass spectra acquired from this strain; charge state 13+ (*m/z* 734.10) was also observed for HU-α in the highest quality mass spectra at a lower abundance. By subjecting the colonies to altered growth and storage conditions (described in detail in section 4.3.5), the expression of additional multiple stress response proteins was observed; these were YdfK and YnaE, YbgS, BhsA and CspC (for details see section 4.3.5). The number of detected proteins was indisputably much lower than that generated by other methods, such as top-down LC/MS which yields upwards of 150 intact proteins from a colony extract.[297] The LESA MS approach is, however, an order of magnitude faster, offering usable

results within the span of minutes. It was therefore proposed that, for applications where the number of proteins detected is crucial whilst the time of analysis is not, it may be beneficial to couple an additional separation step into the LESA protocol. An exploration of hyphenated separation methods and the benefits they offer in terms of analytes detected, focusing on the coupling of high-field asymmetric waveform ion mobility spectrometry (FAIMS) to LESA MS, will be described in Chapter 5.

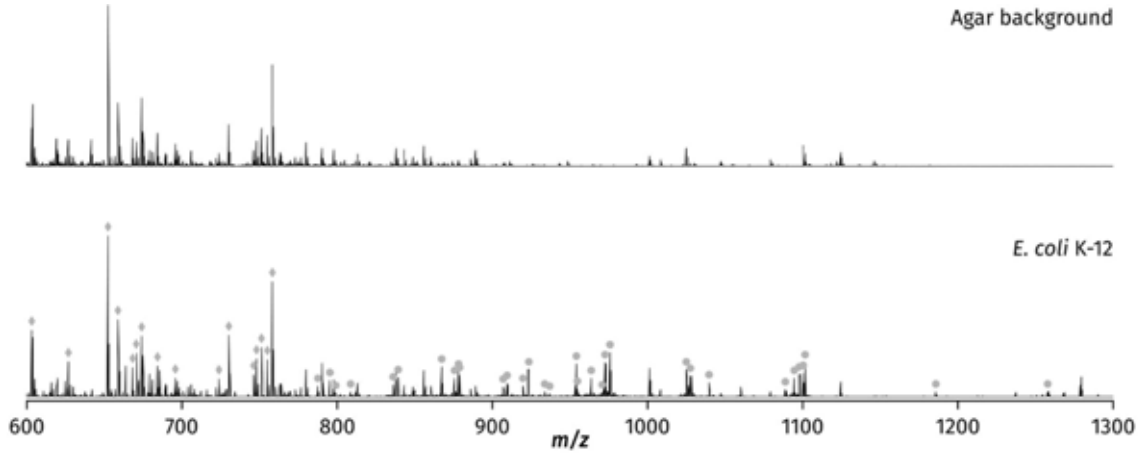


Figure 4.3. Agar background vs *E. coli*. Protein peaks, absent in the agar background spectrum, are marked with grey dots. The highest background signal is observed at the low mass range (*m/z* 600-800); the most prominent background peaks, which could not be assigned to any compounds naturally occurring in *E. coli*, have been marked with grey diamonds.

Table 4.2. List of protein peaks detected in the representative mass spectrum of *E. coli* K-12 shown in Figure 4.2. All proteins were identified on the basis of MS/MS data generated from a minimum of one charge state; all other charge states were assigned by accurate intact mass.

m/z	Charge state	Observed mass	Identity
624.0714	7	4361.4489	
626.0669	10	6250.5962	
632.1464	10	6311.3912	
637.6032	8	5092.7674	SRA
656.5846	12	7866.9279	
683.3051	13	8869.8717	
692.5345	13	8989.8539	

693.9608	9	6236.5817	
702.2725	9	6311.3870	
707.6190	13	9185.9524	
710.2398	13	9220.0228	HU-β
716.0896	11	7865.9056	
727.9152	6	4361.4475	
728.5452	7	5092.7655	SRA
734.0180	13	9529.1394	HU-α
740.3274	12	8871.8415	
766.5032	12	9185.9511	
769.4224	12	9220.9815	HU-β
771.2962	10	7702.8892	YahO
782.3314	8	6250.5930	
787.5980	10	7865.9072	
795.1052	12	9529.1751	HU-α
801.8318	8	6406.5962	
808.6701	9	7268.9654	L29
836.1846	11	9186.9506	
839.2789	11	9220.9879	HU-β
849.9681	6	5093.7649	SRA
867.3868	11	9530.1748	HU-α
875.1089	9	7866.9146	
893.9455	7	6250.5676	
906.9472	17	15400.9787	H-NS
909.6304	8	7268.9850	L29
916.3800	7	6407.6091	
919.6414	10	9186.3412	
923.1055	10	9220.9822	HU-β
925.1998	10	9241.9252	
954.0245	10	9530.1722	HU-α
963.6313	16	15401.9844	H-NS
973.5038	16	15559.9444	
1025.5659	9	9221.0276	HU-β
1027.8729	15	15402.9844	H-NS
1039.5760	7	7269.9811	L29
1060.1399	9	9532.1936	HU-α
1069.1070	6	6408.5983	

Mass spectra of *E. coli* K-12 and *E. coli* BL21 were highly similar, consistent with the high genotypic similarity of these strains (see **Figure 4.2**). The previously mentioned DNA-

binding proteins HU-α and HU-β, as well as the histone-like protein HN-S and the putative periplasmic protein YahO, were consistently detected in both strains and occurred at similar abundances; multiple unidentified low-abundance protein peaks were also common to both strains. It was hypothesised that *E. coli* BL21 mCherry could be distinguished from K-12 by the presence of the overexpressed mCherry protein which gives the colonies a characteristic pink hue; the 28.8 kDa protein could not, however, be assigned to any of the peaks found in the mass spectra of this strain.

Colonies of *P. aeruginosa* were by far the most amenable to LESA sampling, reliably generating mass spectra rich in protein peaks. Five prominent peaks (*m/z* 857.16 10+, 952.29 9+, 1071.32 8+, 1224.08 7+ and 1427.93 6+) corresponding to the stress response protein PA4739 (see section 4.3.5) were consistently observed in every mass spectrum. The total number of protein peaks detected in a typical mass spectrum acquired under optimum colony growth conditions was 57, corresponding to 20 unique protein masses (excluding adducts and modifications) listed in **Table 4.3**, the highest result among all probed bacterial species.

Table 4.3. List of protein peaks detected in the representative mass spectrum of *P. aeruginosa* shown in Figure 4.2.

m/z	Charge	Observed mass	Identity
668.0450	10	6670.3772	
672.2691	9	6041.3564	L33
681.7147	8	5445.6594	PA2146
699.6258	13	9082.0408	HU-β
705.8390	9	6343.4855	
723.4144	9	6501.6641	
732.6566	12	8779.7919	
739.7541	6	4432.4809	L36
742.1603	9	6670.3772	
756.1778	8	6041.3642	L33
757.8449	12	9082.0515	HU-β
759.9963	10	7589.8902	

768.1962	9	6904.7003	
778.9583	7	5445.6572	PA2146
779.0540	11	8558.5140	PA4739
793.9429	8	6343.4850	
799.1732	11	8779.8252	
800.6661	9	7196.9294	L29
808.5355	12	9690.3387	
826.5556	11	9081.0316	HU-β
834.9295	8	6671.3778	
837.5406	11	9201.8666	
844.3275	9	7589.8820	
856.8582	10	8558.5092	PA4739
864.2194	8	6905.6970	
882.1274	11	9692.3214	
900.6217	8	7196.9154	L29
907.2180	7	6343.4751	
908.6153	6	5445.6481	PA2146
909.2117	10	9082.0442	HU-β
930.1933	15	13937.7904	Azurin
934.1407	11	10264.4677	
934.9821	8	7471.7986	CspA
936.6415	15	14034.5134	
949.7409	8	7589.8690	
951.9525	9	8558.5070	PA4739
956.5046	6	5732.9839	PA0039
958.6952	16	15323.0068	PA5178
978.6643	12	11731.8843	
987.5362	7	6905.7025	
996.5543	14	13937.6583	Azurin
1010.1226	9	9082.0379	HU-β
1022.5388	15	15322.9729	PA5178
1029.2829	7	7197.9294	L29
1058.0842	6	6342.4615	
1068.2662	7	7470.8125	CspA
1070.6985	8	8557.5298	PA4739
1085.2756	7	7589.8783	
1090.1374	5	5445.6506	PA2146
1136.2616	8	9082.0346	HU-β
1147.6058	5	5732.9926	PA0039
1162.5724	12	13938.7815	Azurin
1200.8302	6	7198.9375	L29
1223.6505	7	8558.5026	PA4739
1246.1387	6	7470.7885	CspA

1265.9804	6	7589.8387	
1427.5908	6	8559.5011	PA4739

Mass spectra of *S. aureus* were entirely dominated by peaks corresponding to toxic secreted peptides, primarily δ-haemolysin (*m/z* 759.94 4+, 1012.92 3+, 1518.87 2+). The highest numbers of protein peaks, 16 corresponding to 11 unique protein masses (**Table 4.4**), were detected by sampling with a modified solvent composition (see below).

Table 4.4. List of protein peaks detected in the representative mass spectrum of *S. aureus* shown in Figure 4.2.

m/z	**Charge**	**Observed mass**	**Identity**
733.7769	3	2198.3089	PSMα4
759.6541	4	3034.5873	Haemolysin
775.4143	5	3872.0351	SAOUHSC_01729
849.1192	3	2544.3358	
861.5475	8	6884.3218	SAOUHSC_00845
878.7943	3	2633.3611	
921.2673	6	5521.5601	
969.0165	4	3872.0369	CAG03344
984.3377	7	6883.3130	SAOUHSC_00845
1012.5364	3	3034.5874	Haemolysin
1037.8406	3	3110.5000	
1039.8427	3	3116.5063	
1051.5626	3	3151.6660	
1131.5834	4	4522.3045	SAOUHSC_01135
1148.3934	6	6884.3167	SAOUHSC_00845
1291.6890	3	3872.0452	CAG03344

4.3.2 Optimisation of sampling—*S. aureus*

Initial sampling of the Gram-positive *S. aureus* with the same composition of solvent as that used for Gram-negative species (40:60:1 acetonitrile-water-formic acid) proved challenging. Whilst it was possible to obtain a stable electrospray, only about half of all sampling attempts yielded a useable mass spectrum. On closer inspection of the nanoelectrospray chip, this

appeared to be caused by excessive build-up of colony material. Whilst data could still be generated with sufficient persistence, more reliable sampling was crucial for time-sensitive experiments.

It was also noted that for this particular bacterium, the signal corresponding to peptides and proteins was unusually time-dependent; the initial three minutes of nanoelectrospray invariably displayed a prevalence of singly charged ions (**Figure 4.4**). Maintaining the sample delivery and data acquisition in full scan mode for extended periods of time led to an increase in the abundance of protein peaks, which were observed almost exclusively after 10 minutes of sample delivery and further, whilst smaller, singly charged analytes were depleted. These effects may be related to the differential diffusion of analytes in the sampling solvent, which may lead to the formation of an analyte gradient within the pipette tip. Thus, analytes less readily diffusible in the sampling solvent would remain closer to the opening of the pipette tip and thus be the first to be delivered to the mass spectrometer.

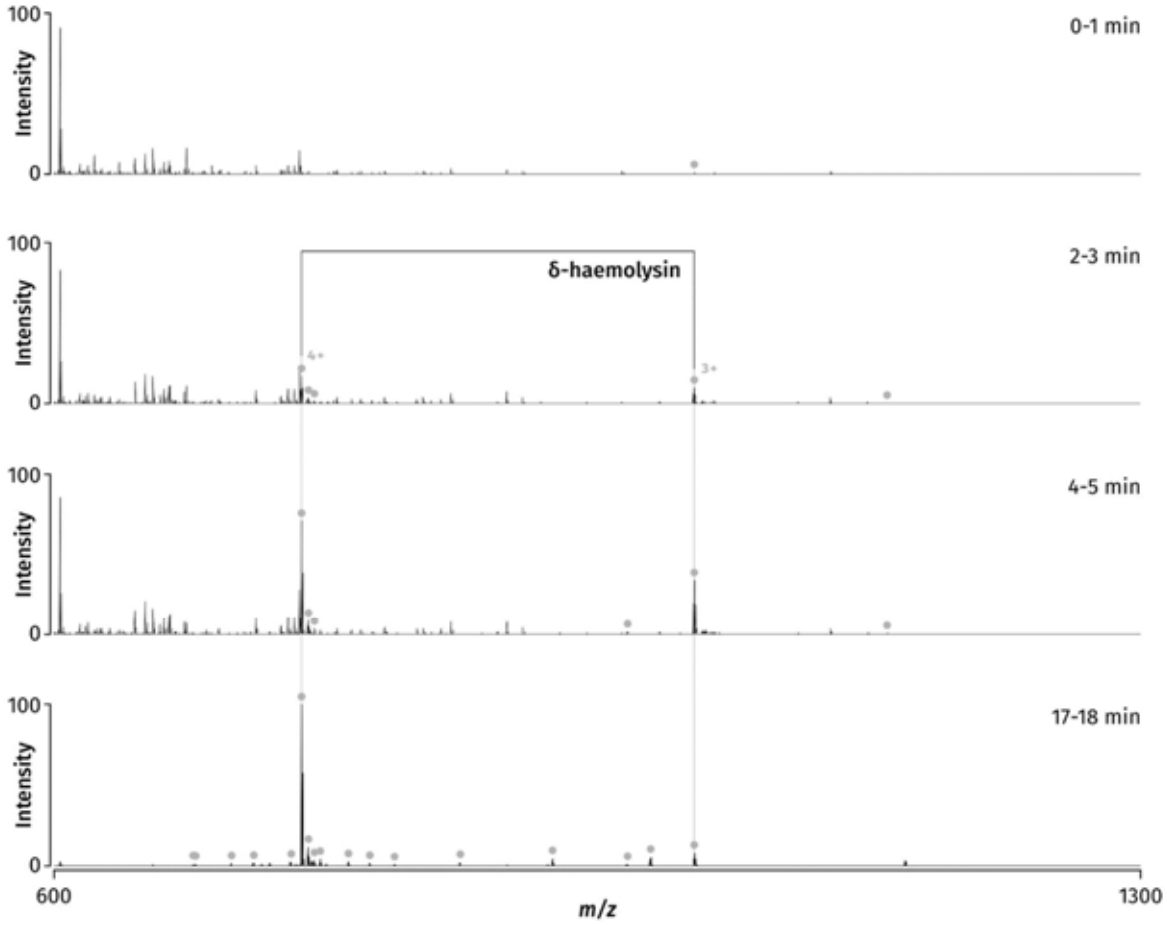


Figure 4.4. The progression of mass spectrum acquisition for *S. aureus*. Data were summed across a window of one minute within the chromatogram to generate each mass spectrum shown. Protein peaks are marked in grey. All mass spectra have been normalised to the same peak (haemolysin^{4+} in the bottom mass spectrum).

Figure 4.5 shows the major results of solvent optimisation performed on the same day on a single plate of *S. aureus*. Focus was placed on the rapid emergence of protein/peptide signal and repeatability. All mass spectra were summed across the initial five minutes of data acquisition. Following data acquisition, mass spectra were manually inspected for the presence of protein peaks. Three replicates were performed at each set of conditions to account for known sampling variability. The dwell time was increased to 15 seconds for all sampling attempts, with the aim of boosting the extraction efficiency for proteins. The proportions of both acetonitrile and formic acid were subsequently increased. As seen in the top mass spectrum, the solvent used for Gram-negative species was incapable of yielding any

protein signal within the first minutes of sample delivery. By increasing the proportion of formic acid to 10% (40:50:10 acetonitrile-water-formic acid), the haemolysin signal could be generated instantaneously. The repeatability of the approach was, however, problematic—whilst the first acquired spectrum (shown here) was extremely promising, even briefly showing indications of large (up to 10 kDa), most likely cytosolic proteins, these were not observed subsequently. Subsequent sampling attempts using this solvent composition generated mass spectra consisting primarily of singly charged background peaks; the detection and abundance of protein peaks, as well as the relative abundance of the most prominent background peaks, were irreproducible. A 50:45:5 composition of solvent, commonly used for extraction of bacterial colonies for LC-MS analysis, offered increased extraction of peptides and proteins without compromising repeatability. The δ-haemolysin peaks appeared reliably in the first minute of sample delivery, alongside peaks corresponding to toxic peptides belonging to the phenol soluble modulin α family. One peak corresponding to a UPF0337 stress response protein in the 6+ charge state (*m/z* 1148.72) was also reproducibly detected. This solvent was therefore retained for further work.

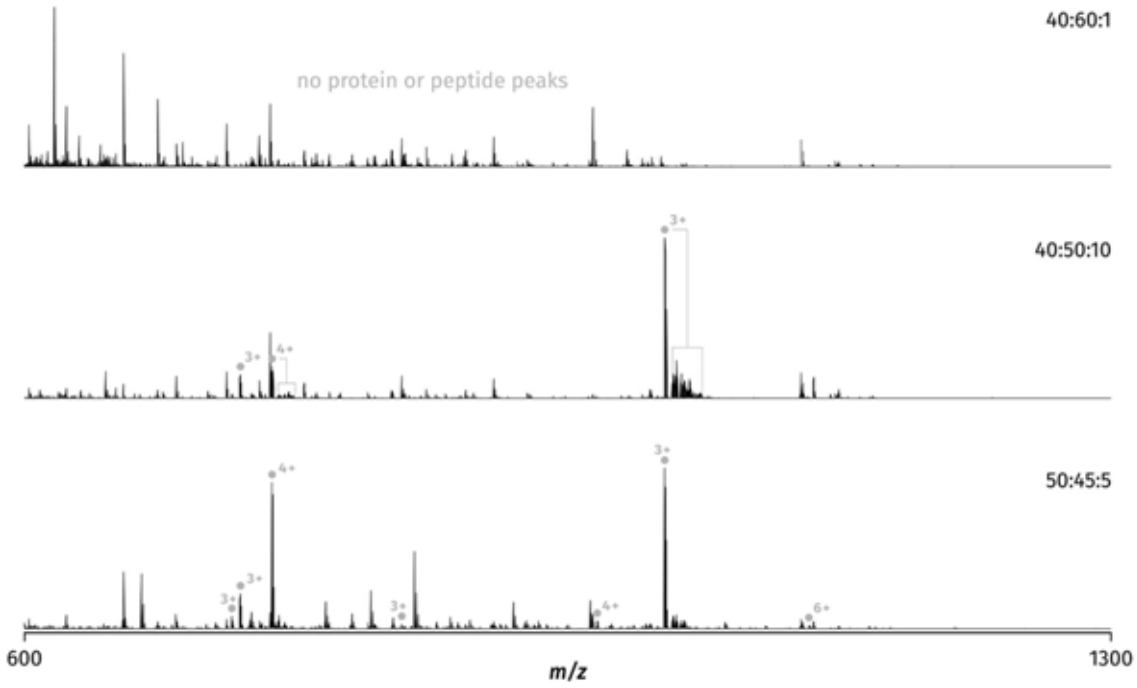


Figure 4.5. Optimisation of solvent composition on a single plate of freshly incubated *S. aureus*. The labels on the right hand side of the mass spectra indicate the relative proportions of acetonitrile, water and formic acid, respectively. Protein and peptide peaks detected, along with their charge states, are indicated in grey.

Whilst the reproducibility and reliability of the sampling process was significantly improved through this optimisation step, the routine observation of proteins larger than approximately 7 kDa (with the exception of the aforementioned UPF0337 stress response protein) remained a challenge. Considering the harsher sampling conditions which should induce cell lysis, it was hypothesised that it is the overabundance of small, secreted peptides and/or their highly preferential extraction which prevents the detection of larger proteins. Another attempt to reveal additional protein peaks by use of FAIMS is described in Chapter 5.

4.3.3 Optimisation of sampling—yeast

LESA sampling and protein identification was also extended to yeasts, the most impermeable class of microorganisms, represented here by *Cryptococcus neoformans* H99 and an in-house

strain of *Candida glabrata*. *C. glabrata* proved to be by far the most challenging sample encountered during the project.

According to literature, in their natural state below 37 ℃, colonies of *C. glabrata* may sprout a dense layer of aerial hyphae used for reproduction.[298-300] These, however, would impede LESA sampling by instantly blocking the nanoelectrospray nozzles and potentially the pipette tip itself. Therefore, colonies destined for sampling were initially grown at a sub-optimal 37 ℃ in order to discourage the growth of hyphae and obtain a smooth colony surface amenable to sampling. Test colonies were subsequently grown at 30 ℃ to document the generation of the hyphal form of *Candida*; no growth of hyphae was, however, observed up to 48 hours after inoculation, indicating that factors necessary to induce the transition into the hyphal form were absent during the time necessary to generate colonies sufficiently large for sampling. All subsequent batches of *Candida* colonies were therefore incubated at 30 ℃ as described in Chapter 2, section 2.2.1.3.

Sampling of the colonies yielded no protein signal whatsoever, even using the harshest solvents capable of maintaining a stable nanoelectrospray (**Figure 4.6**). Mass spectra were consistently dominated by peaks corresponding to the underlying YPD growth media, indicating that no analytes were extracted from the interior of the cells themselves. Lysis was not achieved even by the preparative treatment recommended for the clinical analysis of yeast samples by MALDI-TOF MS: A colony was scraped off the agar plate, placed in an Eppendorf tube and suspended in a 70% solution of formic acid. Following microcentrifugation at 13,000 rpm for 20 minutes, the pellet was resuspended in 100% acetonitrile and centrifuged again at the same settings. Analysis of the supernatant by direct infusion revealed no protein; a compound with a repeating unit of 28.0315 Da (C2H4) was detected but could not be identified (also see section 4.3.4). Once a small aliquot of the

supernatant was diluted in a 1:10 ratio into the 40:60:1 sampling solvent, protein signal was eventually observed but all attempts at recreating this series of extraction steps solely by use of the NanoMate failed. One protein, the 40S ribosomal constituent S30 (MW_{obs} = 6958.94 Da), was identified from the only protein mass spectrum obtained. Whilst identification of yeasts could also be performed by investigating lipid profiles or small molecules potentially accessible without complete lysis of the cell wall (although this was not seen here), intact proteins constituted the main interest of the project for the purposes of not only identification, but also characterisation of yeasts as part of microbiological studies. The sampling of yeasts was therefore revisited and discussed in detail following the construction of a high-voltage cell-lysing apparatus, described in Chapter 6.

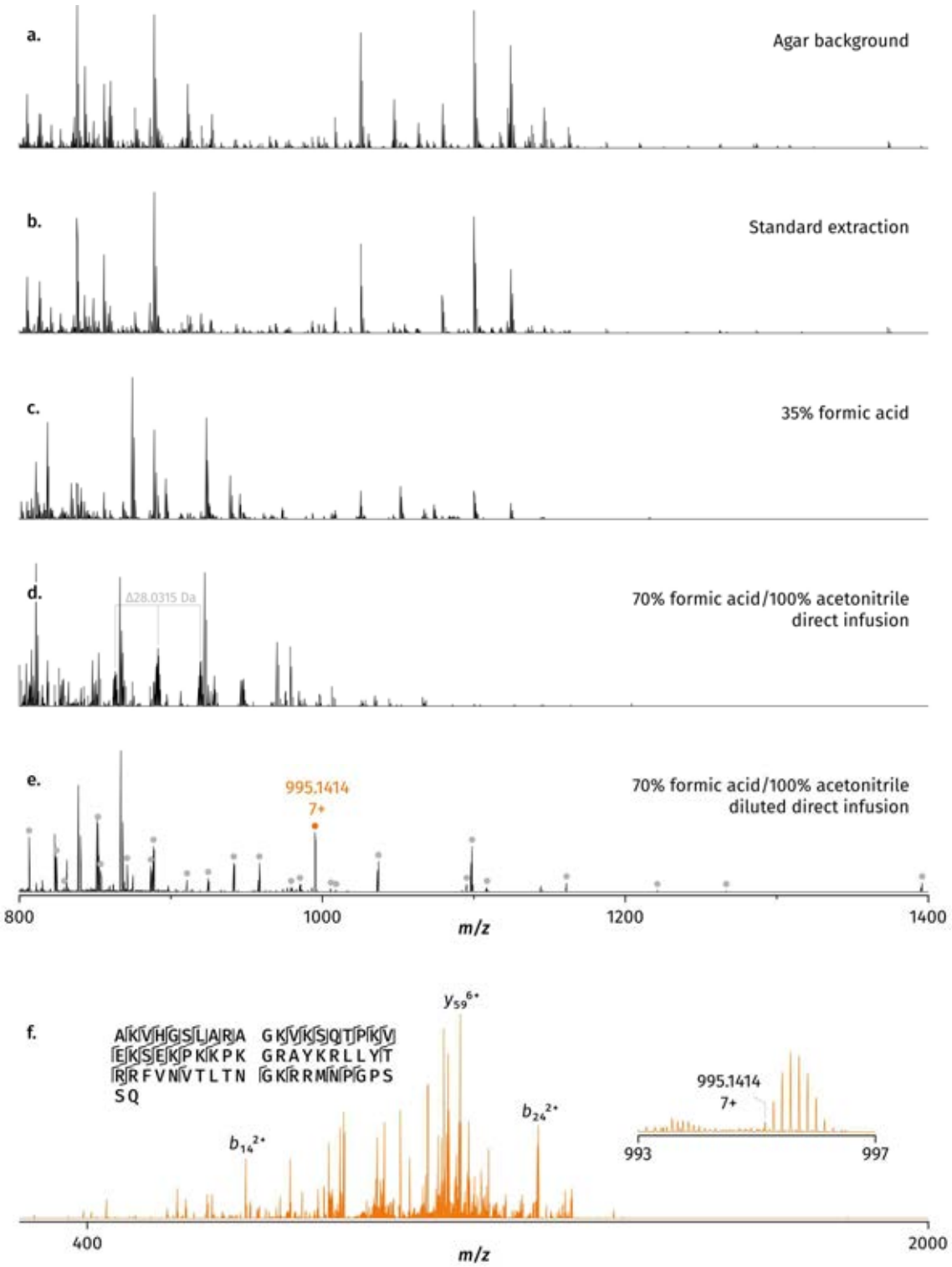


Figure 4.6. Mass spectra generated from colonies of *Candida glabrata*.

4.3.4 Streptococci and prospects for bacterial identification

A comparison of three mass spectra of *S. pneumoniae*, *S. gordonii* and *S. oralis*, acquired under identical conditions, is shown in the lower part of **Figure 4.2**. *S. mitis* was originally planned to participate in the comparison; colony growth was, however, very poor, regardless of the type or batch of growth media used and the strain was therefore excluded from further work. Mass spectra of *S. pneumoniae* were particularly characteristic due to the presence of a rich distribution of singly charged peaks in the *m/z* 700-1000 range. The thick capsule surrounding streptococcal cells was expected to hinder lysis even further compared to staphylococci and therefore reduce the efficiency of protein extraction. Accordingly, the intensity of the protein signal was much lower than was the case for all other investigated strains; both 1% and 5% formic acid-containing solvents were tested for completeness, displaying no marked differences between the two. Despite these challenges, a total of eight peaks attributable to protein were detected in *S. pneumoniae*, corresponding to five distinct protein masses (the *m/z* values and charge states are provided in **Table 4.5**); similarly, nine protein peaks corresponding to six protein masses were detected in *S. oralis*. *S. gordonii* gave rise to significantly richer mass spectra, comparable to those of the Gram-negative species described earlier, exhibiting 26 protein peaks and 15 protein masses detected. All protein peaks detected in representative mass spectra of the three streptococci are listed in **Table 4.5**. Only one protein with observed mass of 7981.44 Da (*m/z* 1141.21 7+; 998.69 8+ and 887.94 9+ also observed in *S. gordonii*) was common to all three species; all other detected protein masses were unique to the species they were found in, rendering the differentiation of the three species trivial. Among all observed proteins, four were selected for MS/MS identification (described in section 4.3.5); none of these have thus far been directly observed as intact proteins by any methods, mass spectrometric or otherwise (for details see the following

section). CID fragmentation data were searched against a comprehensive protein database comprising multiple strains of *S. pneumoniae*, *S. oralis*, *S. mitis*, *S. pseudopneumoniae*, *S. gordonii*, and a number of other, closely related species listed in Chapter 2, section 2.2.5. All four proteins were confidently assigned (expectation value less than 1×10^{-70}) to the correct species, reinforcing the idea that this process could be used for bacterial identification. Other classes of analytes were not focused on. A characteristic compound (*m/z* 859.5315, 887.5632, 913.5782, all singly charged) containing a 28.0312 Da repeating unit (C_2H_4) was detected in *S. pneumoniae* (see **Figure 4.2**). The compound was not found in mass spectra of other streptococci acquired on the same day or in the background, indicating that it was most likely produced by the colony. Its identity could not, however, be confirmed by intact mass alone.

Table 4.5. List of protein peaks detected in representative mass spectra of streptococci shown in Figure 4.2. The unidentified protein observed in all three species is labelled as 'Unknown common.'

m/z	Charge	Observed mass	Identity
S. oralis			
737.0612	6	4416.3235	
843.0606	8	6736.4266	HMPREF9950_0438
963.2094	7	6735.4149	HMPREF9950_0438
1123.5760	6	6735.4123	HMPREF9950_0438
1125.9569	6	6749.6977	
1141.6289	7	7984.3514	Unknown common
1188.2696	8	9498.0986	
1199.7982	7	8391.5365	
1358.0233	7	9499.1122	
S. pneumoniae			
839.9030	7	5872.2701	
984.6451	7	6885.4648	spr1626
986.7930	7	6900.5001	
1045.0943	6	6264.5221	
1085.0690	6	6504.3703	spr1385
1141.6293	7	7984.3542	Unknown common
1148.5856	6	6885.4699	spr1626
1151.2524	6	6901.4707	

S. gordonii			
737.2356	6	4417.3699	
828.1912	7	5790.2875	
848.9033	7	5935.2722	
855.3709	8	6834.9090	
857.8377	8	6854.6434	AWH02_03040
884.6827	5	4418.3771	
887.9449	9	7982.4386	Unknown common
900.9528	7	6299.6187	
944.9465	7	6607.5746	
977.5641	7	6835.8978	
980.2418	7	6854.6417	AWH02_03040
990.3881	6	5936.2849	
991.5584	7	6933.8579	
998.6869	8	7981.4370	Unknown common
1001.6590	7	7004.5621	
1034.9443	8	8271.4962	
1050.9431	6	6299.6149	
1125.8040	6	6748.7803	
1132.6064	8	9052.7930	
1141.2132	7	7981.4415	Unknown common
1143.4481	6	6854.6449	AWH02_03040
1155.3509	7	8080.4054	
1168.4347	6	7004.5645	
1182.6497	7	8271.4970	
1189.7899	8	9510.2610	
1294.1228	7	9051.8087	

One concern prior to the analysis of mass spectra from these species was the requirement for horse blood agar growth media. It was hypothesised that the haemoglobin signal from the horse blood may obscure peaks corresponding to bacteria. Accordingly, multiple peaks corresponding to the two chains of haemoglobin, as determined by intact mass, were observed at relatively low abundance in mass spectra acquired from plain blood agar plates stored at 4 ℃ (charge states 11+ to 20+). The number of observed haemoglobin charge states decreased to three (11+ to 13+) following incubation of plain blood agar plates at 37 ℃ for 24 h, conditions consistent with the growth of bacteria. The haemoglobin peaks observed in

mass spectra acquired from streptococci were limited to the 12+ charge state (*m/z* 1258.65 and 1259.90 for haemoglobin α and β, respectively) and did not interfere with detection of bacterial proteins (**Figure 4.7**). The age of the media may have contributed to the low abundance of haemoglobin—most of the plates used throughout the experiments were sourced from surplus stock and it is likely that, whilst the nutrient value remained sufficient for the growth of bacteria, a significant portion of haemoglobin had degraded.

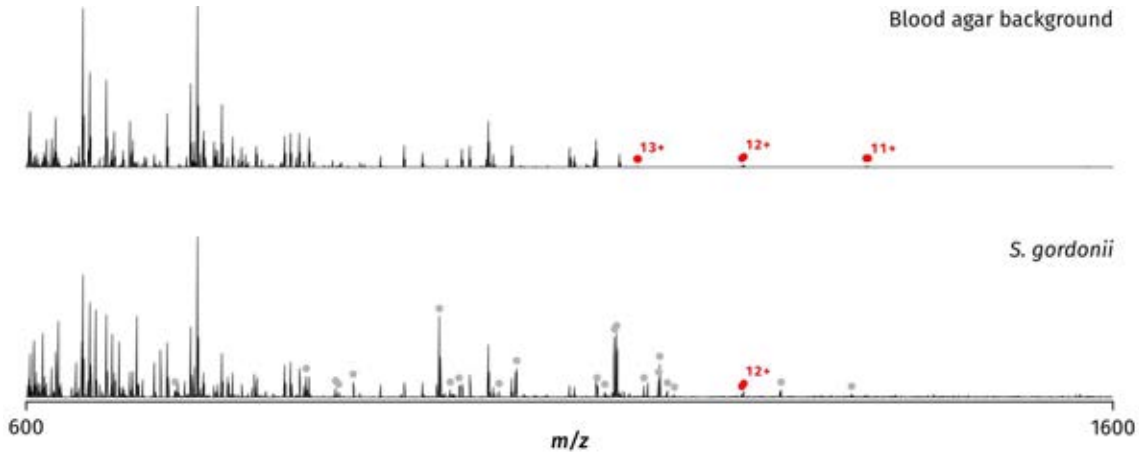


Figure 4.7. Background interference of blood agar peaks and mass spectra of *S. gordonii*. Haemoglobin peaks (both chains) are marked with red dots above their peaks; all other proteins are marked in grey.

4.3.5 Top-down protein identification

The proteins described above were identified by top-down MS/MS. Identification of each protein was accomplished based on a minimum of one tandem mass spectrum, deconvoluted by the THRASH algorithm, putatively assigned to a protein by use of ProSightPC software and confirmed by manual assignment of peaks in the original raw data against a theoretical peak list. Assignments were subject to a mass accuracy threshold of 5 ppm, derived from the calibration of the Orbitrap (where a 5 ppm difference between the expected and the measured masses of calibrants is considered the acceptable maximum by the calibration algorithm). The

accuracy of the assignment was not strictly gauged by the statistical values returned by the software (in the form an expectation value, *E* and a probability value, *P*), as the values were found to be highly reliant on the number of matching fragments, which could be identical for a large number of equally probable, putative hits. The assignment was instead considered accurate if the intact mass of the observed protein matched the intact mass of the candidate sequence (including software-derived and manually added post-translational modifications) to within 5 ppm, and if masses of abundant fragment ions in the tandem mass spectrum (derived from raw rather than deconvoluted data) matched the masses corresponding to fragments of the candidate sequence to within 5 ppm. Whilst such subjective assessment could potentially lead to misidentification, the data were generally of sufficient quality to provide unambiguous identification at the software stage, where the returned *E* and *P* values for the most likely hit were significantly lower (better) than the second most likely hit and in the region of 1×10^{-10} or lower (indicating a highly confident assignment); thus, manual peak assignment only served to validate the identification and provide additional sequence coverage through identification of low-abundance peaks missed by the software.

The variation of growth and storage conditions introduced for the purpose of measuring the reproducibility of protein mass spectra (Chapter 3), as well as the investigation into the effect of refrigeration and repeated incubation on protein mass spectra (section 4.3.7 below) allowed the identification of additional stress response proteins not observable in bacteria grown under optimal conditions. Thus, the proteins detected included proteins unique in mass and specific to a given species, as well as proteins specific to the conditions under which a given species was grown. **Table 4.1** lists all the proteins identified in this manner across all sampled species. The conditions under which each protein was isolated and fragmented are listed alongside other parameters; these, however, are not necessarily the only conditions under

which the given protein was observed. Notably, a number of identified proteins have not been previously observed intact by any other technique; these are highlighted in bold type. Tandem mass spectra and fragment assignments for all listed proteins are included in **Appendix 3.2**.

The sampling and analysis of *P. aeruginosa* was by far the most efficient among all probed species, thanks to the fast and reliable growth of colonies, few issues with nozzle blockage, as well as the largest range of proteins detected. Multiple homologs of proteins previously identified in *E. coli* were observed, including the DNA-binding protein HU-β (**Appendix 3.2.2.3**) as well as L29 (**Appendix 3.2.2.11**), a structural constituent of the ribosome. It was the stress response proteins, however, that proved arguably the most interesting. PA2146 (**Appendix 3.2.2.4**), belonging to the KGG motif family according to genomic data,[301] was identified as a close relative of *E. coli* YciG and YmdF, also identified here (**Appendix 3.2.1.1** and **3.2.1.4**, respectively). Even more notably, members of the UPF0337 family of stress response proteins were observed in all species investigated during this project; this included PA4738 detected in *P. aeruginosa* (**Appendix 3.2.2.1**), as well as YjbJ in *E. coli* (**Appendix 3.2.1.3**) and the uncharacterised protein SAOUHSC_00845 in *S. aureus* (**Appendix 3.2.3.3**).

In addition to PA2146, *P. aeruginosa* yielded multiple proteins which have never before been detected in their intact form by any other technique, but whose expression was predicted based on genomic data and homology searches. One notable example of such a protein was peptidylprolyl isomerase PpiC2 (**Appendix 3.2.2.2**), a putative enzyme sharing sequence homology with *E. coli* parvulin;[301] the measurement of its intact mass in *P. aeruginosa*, as well as its tandem mass spectrum, supplied evidence for the post-translational removal of the N-terminal methionine. Another predicted protein detected in *P. aeruginosa*, PA4739 (**Appendix 3.2.2.5**), contains the characteristic BON domain which indicates localisation to

either the periplasm or the outer membrane, and subsequent involvement in stress response.[302] According to the automatic annotation found in the UniProt database (accession number Q9HV60), the mature protein should be missing a 25-amino acid signal peptide cleaved from the N-terminus. Data gathered by LESA MS/MS, however, convincingly demonstrate that the cleavage site is instead located at position 32 (**Figure 4.8**). The tandem mass spectrum also contained fragments of a high mass precursor co-isolated with PA4739 which did not interfere with the identification of PA4739 by ProSightPC; identification of the high mass precursor was not successful based on this dataset. Further instances of compound tandem mass spectra containing fragments derived from more than one precursor were also recorded at *m/z* 908.62 6+ (PA2146, co-isolated with 909.21 10+ corresponding to HU-β) and *m/z* 742.16 9+ (50S ribosomal protein L32, co-isolated with *m/z* 739.75 6+ corresponding to 50S ribosomal protein L36). The resolution of compound mass spectra from *P. aeruginosa* by FAIMS is described in Chapter 5.

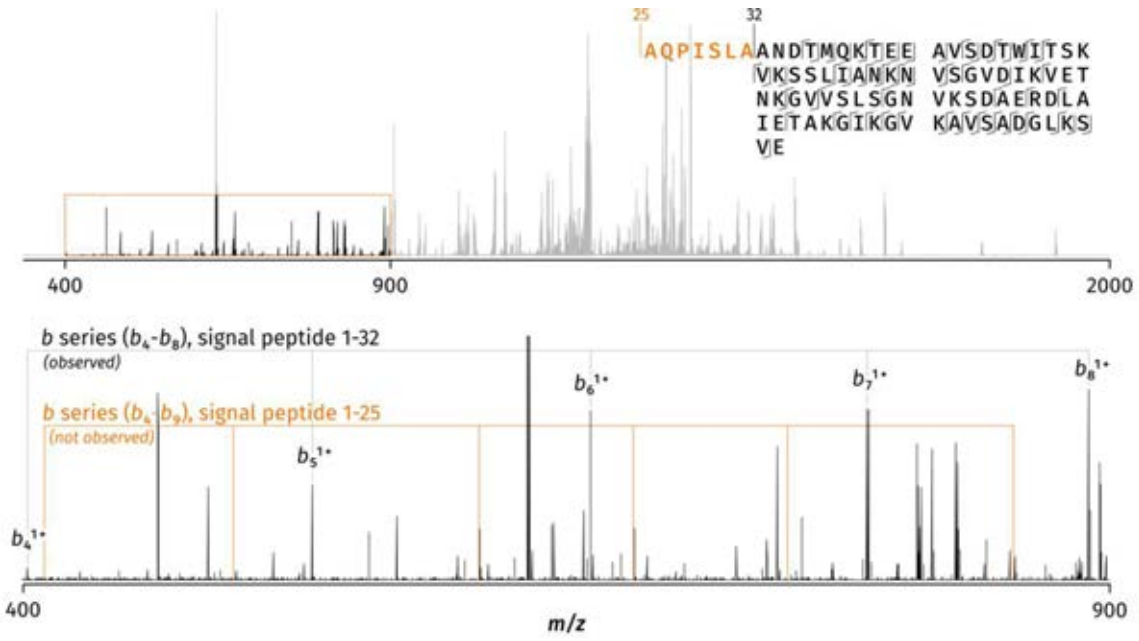


Figure 4.8. Tandem mass spectrometry data of *P. aeruginosa* PA4739. The masses of observed *b* fragments were consistent with the cleavage of a signal peptide at position 32, in contrast to the expected cleavage site at position 25. No *b* fragments consistent with a

shorter signal peptide sequence were observed. The observed intact mass (8557.50 Da) was also consistent with cleavage at position 32 (error -0.7 ppm).

Another protein missing its signal peptide, PA0039 (**Appendix 3.2.2.15**), generated a characteristic tandem mass spectrum (**Figure 4.9**) consistent with the presence of a disulfide bridge between the cysteines at positions 4 and 42, undetected by the automatic annotation algorithms of the UniProt database. No fragmentation was observed between these two positions, whilst the observed intact mass was 2.04 Da lower than predicted (MW_{obs} = 5731.98 Da, MW_{theor} = 5733.99 Da; Δppm = -351);[301] the theoretical mass of a disulfide-containing form of the protein matched the observed value to within 0.1 ppm. PA5178, sharing its BON domain with the previously described PA4739, generated particularly puzzling data (**Appendix 3.2.2.6**). A large C-terminal fragment of the protein (*m/z* 895.04 7+ and 1044.04 6+; tandem mass spectrum of the 6+ charge state as compared to the 16+ charge of the intact protein shown in **Figure 4.10**), along with a corresponding decrease in the abundance of the intact protein (*m/z* 958.82 16+), was repeatedly observed in two colonies exposed to the 37-4-37 ℃ temperature cycle described in section 4.3.7, although not in the third colony grown on the same plate and subjected to the same conditions. It was also detected and confirmed by tandem mass spectrometry in a colony of *P. aeruginosa* stored in the refrigerator for over three months; full scan mass spectra acquired from this particular colony were largely devoid of intact protein peaks, consistent with extensive protein breakdown. The site of cleavage was investigated more closely by matching the sequence of the detected fragment against the predicted secondary structure of the protein—curiously, it was traced to a disordered region located between the two major domains. PeptideCutter[303] identified a thermolysin cleavage site at the same location, indicating that the cleavage event may be catalysed by a yet unidentified protease, and implying functional relevance. Finally, two ribosomal proteins hitherto only predicted based on sequence homology with related species,[301] L35 (with its N-

terminal methionine removed; **Appendix 3.2.2.12**) and L36 (**Appendix 3.2.2.9**), were also detected and identified here in their intact form.

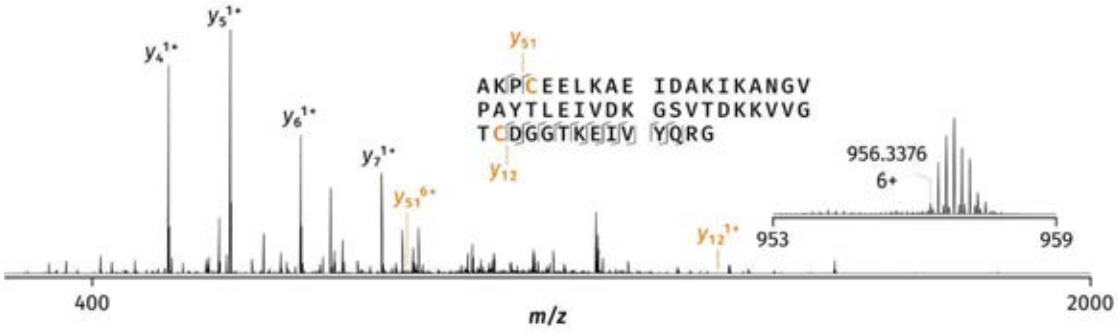


Figure 4.9. CID tandem mass spectrum of *P. aeruginosa* PA0039. No fragmentation was observed between the two cysteines (labelled in orange), indicating the presence of a disulfide bond. Corresponding *y* fragment ions are labelled in orange.

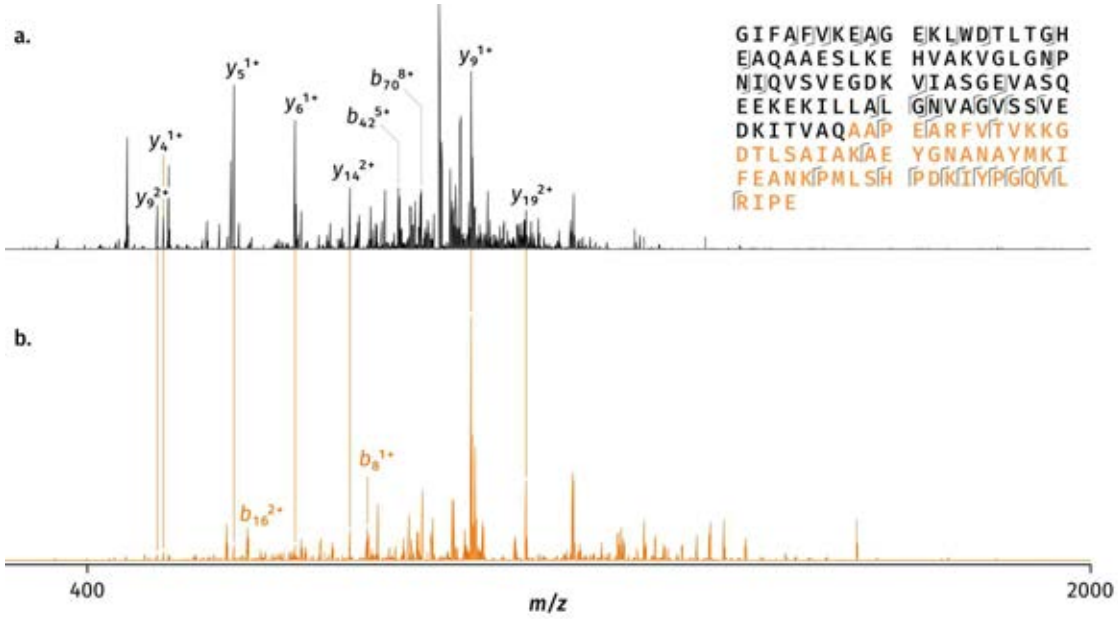


Figure 4.10. Tandem mass spectrometry data of *P. aeruginosa* PA5178. **a**. Tandem mass spectrum of the intact form of the protein. M_{obs} = 15320.09 Da (consistent with the cleavage of the initiator methionine, error 3.5 ppm). **b**. Tandem mass spectrum of the C-terminal fragment of the protein. M_{obs} = 6258.25, error 4.9 ppm. The most prominent common *y* fragments have been marked by orange lines. Representative *b* fragments unique to each protein form are labelled in their corresponding mass spectra.

In total, 16 proteins detected in *P. aeruginosa* were selected for tandem mass spectrometry; 15 of these were successfully identified using the approach described above. One protein, observed in charge states 6+ and 7+ at *m/z* 1237.45 and 1060.83, respectively which corresponded to a measured mass of 7419.73 Da, evaded identification by the current methodology. A single sequence tag, QTAVQ, was derived from the fragmentation mass spectrum by use of ProSightPC's Sequence Tag search mode; such a short signature is highly unusual and may be indicative of a high degree of modifications resistant to fragmentation by CID, or a non-canonical structure of the protein. It also suggests that the lack of specificity of the chosen *P. aeruginosa* database to the particular strain under investigation is not the reason for the lack of identification. Some functional significance of the protein can be gleaned from the pattern of its detection—as it was only observed in colonies which had been exposed to cold or room temperature storage conditions, suboptimal for the growth of *P. aeruginosa* which requires 37 ℃, it could be deduced that this strange protein may be involved in cold shock response. Other software, such as MSAlign+,[172] is available and may be better suited for the identification of uncharacterised, highly modified proteins; this, however, was not investigated further during the project.

Among the five proteins selected for tandem mass spectrometry in *S. aureus*, all were successfully identified. Three of these belonged to the phenol-soluble modulin family of small, secreted proteins which are characteristic to virulent staphylococci.[304] The most abundant and well-characterised by far was δ-haemolysin, a 26 amino acid toxic peptide suspected to form pore-like structures in erythrocyte membranes, subsequently causing their lysis.[305] The tandem mass spectrum generated by CID is displayed in **Figure 4.11** (see **Appendix 3.2.3.4** for fragment assignment). Full sequence coverage was obtained, confirming the presence of formylation on the N-terminal methionine (characteristic to phenol soluble modulins) as well

as an amino acid substitution (G-S, see inset in **Figure 4.11**) which differentiates the MSSA476 strain of *S. aureus* from the prototypical strain (NCTC 8325). A related peptide, phenol-soluble modulin α4, which binds to and lyses neutrophils,[306] was also identified (**Appendix 3.2.3.5**); whilst it has never been observed directly in this particular strain of *S. aureus*, identical gene products have been found and studied in MRSA. A third, less well-known member of this family of small toxic peptides, SAOUHSC_01135,[307] was also observed (**Appendix 3.2.3.1**); like δ-haemolysin, both were detected with the characteristic formylation of the N-terminal methionine,[304] as determined by their intact mass and fragmentation pattern. SAOUHSC_00845 (**Appendix 3.2.3.3**),[308] briefly mentioned above, was the representative member of the UPF0337 family of stress response proteins in this species, whilst the predicted protein SAOUHSC_01729 (**Appendix 3.2.3.2**),[307] observed here in an unmodified form, is hitherto entirely uncharacterized; identical gene products were located by a protein-protein BLAST search at default settings in numerous strains of *S. aureus*, as well as one strain of *S. haemolyticus*.

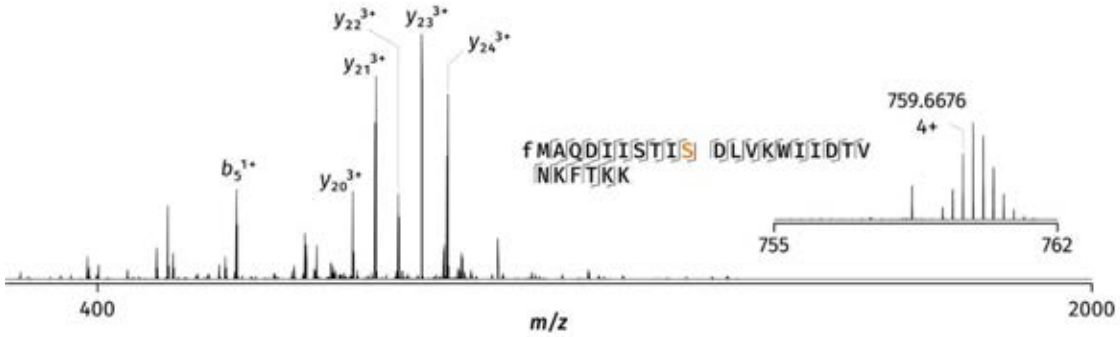


Figure 4.11. Tandem mass spectrometry data of δ-haemolysin isolated from *S. aureus* MSSA476. Complete sequence coverage was obtained. The serine residue unique to this strain is highlighted in orange.

Similarly to *P. aeruginosa*, 16 proteins were chosen for tandem mass spectrometric analysis in *E. coli* (including both K-12 and BL21 strains). Fifteen of those were successfully assigned; the remaining protein, a 30 kDa species only observed briefly in a single mass spectrum at a

low abundance, could not be identified, most likely due to the poor quality of the generated fragmentation spectrum. All proteins previously reported in *E. coli* by LESA MS[117] were detected here. Five proteins identified by LESA MS in *E. coli* K-12 in the previous study were also isolated and fragmented here, confirming their identity. Among the proteins not identified in the previous LESA MS study and presented here for the first time, YbgS (MW_{obs} = 10457.58 Da; **Appendix 3.2.1.15**), a protein inferred from its homology with other gene products found in related species, was possibly the most significant as it has never before been directly observed; the only information available on its function related to its genetic interactions with several partners.[309] Tandem mass spectrometry data (**Figure 4.12**) implied the cleavage of a signal peptide from the N-terminus, as well as the presence of a disulfide bond and possible deamidation, although the exact position of these two modifications could not be inferred.

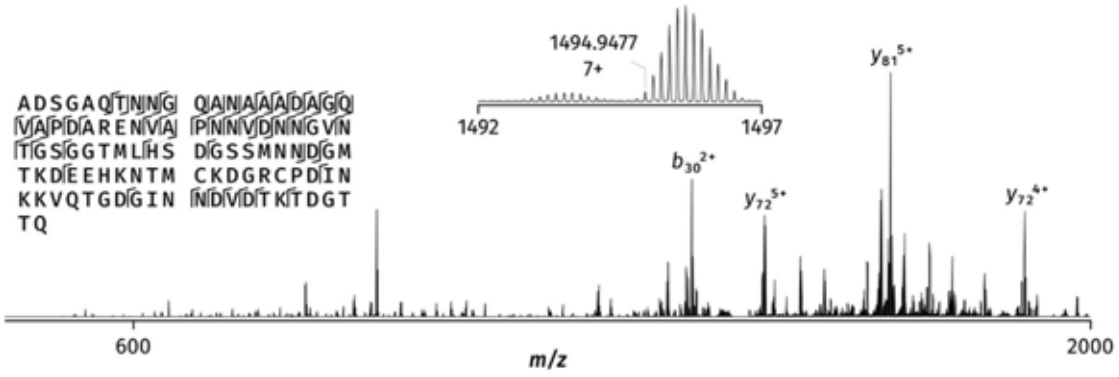


Figure 4.12. Tandem mass spectrum of YbgS.

The role of improvements in the software available for top-down protein analysis could clearly be seen when comparing the results obtained by use of ProSightPC 3.0 and 4.1 alpha in this study compared to the results generated by Randall et al. by use of ProSightPTM 2.0. Crucially, ProSighPC 3.0 first allowed the creation and maintenance of custom protein databases, thus expanding the range of species which could be analysed and allowing the use

of databases tailored to specific strains. Additional improvements included the Sequence Gazer, which allows the manual supplementation of post-translational modifications to candidate sequences to better match the intact mass and fragmentation patterns of putative hits and thus improve confidence in the assignments. The hypothetical cold shock protein YdfK (**Appendix 3.2.1.7**)[310] evaded identification by Randall et al.[117] using the functionality and databases provided by the web-based software ProSightPTM 2.0. ProSightPC 3.0 armed with a custom *E. coli* K-12 database generated from UniProt data was, however, capable of returning highly confident hits with the correct intact mass. Even more strikingly, the initial ProSightPC 3.0 database of *E. coli* K-12 was only able to identify the highly abundant protein YahO (**Appendix 3.2.1.2**) by use of the Biomarker search mode which takes into account artificially generated sub-sequences of the protein in addition to its full canonical sequence (ProSightPTM 2.0 returned no hits whatsoever); this was likely due to the presence of a 21 amino acid signal peptide at the N-terminus of the protein,[311] which was not correctly annotated. A database constructed less than a year later (2017), however, enabled the identification of this protein using the standard intact mass search with no issues.

Finally, four proteins were selected for fragmentation across the range of *Streptococcus* species described in detail in a previous section; two of these were detected in *S. pneumoniae*, one in *S. oralis* and one in *S. gordonii*. All four were successfully identified (**Figure 4.13**; see **Appendices 3.2.4**, **3.2.5** and **3.2.6** for fragment assignments); interestingly, all of them (including, unusually, both proteins from *S. pneumoniae*) belonged to the previously mentioned, largely uncharacterised family of UPF0337 stress response proteins. Proteins belonging to this family were detected in all bacterial species analysed thus far, irrespective of growth and sampling conditions. The ubiquitous detection of these particular proteins, as

well as their unique masses in each of the studied species, may make them especially suitable for use as identification markers.

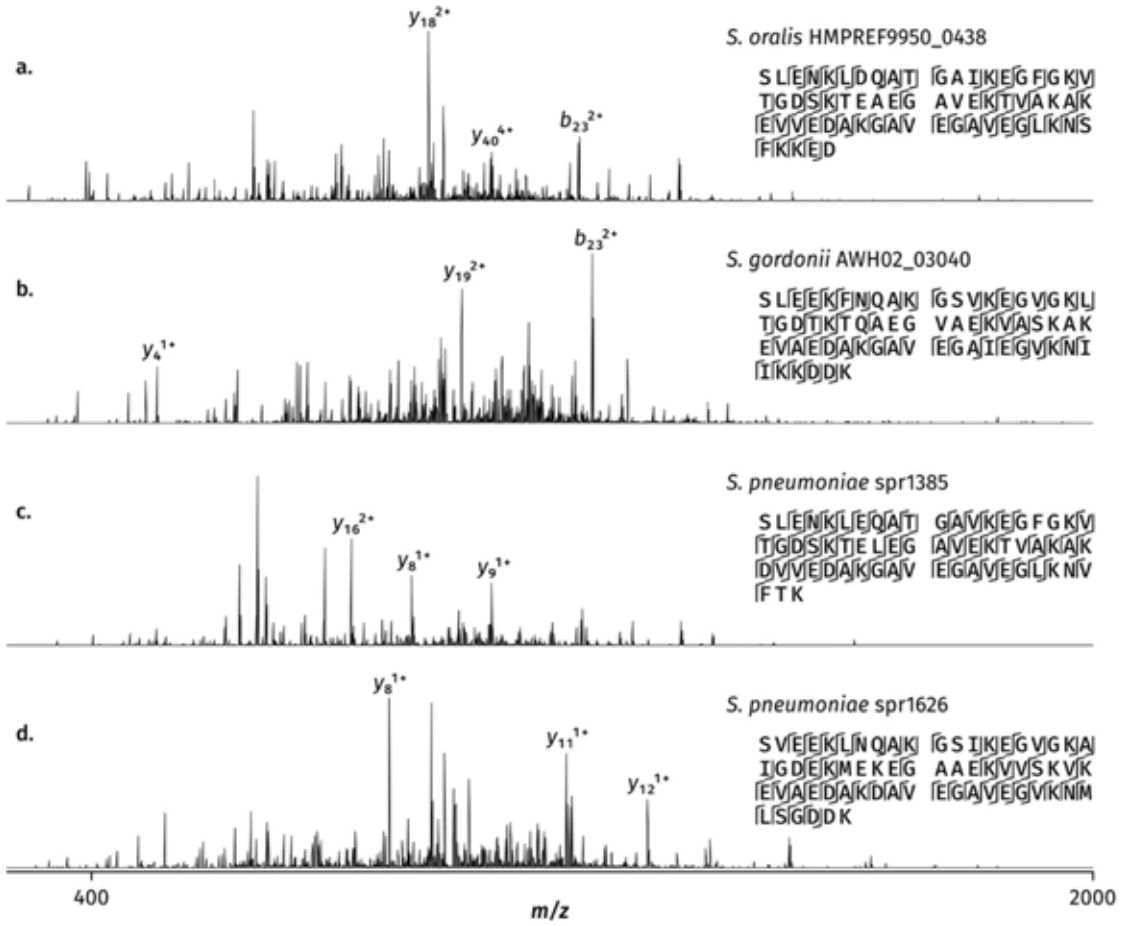


Figure 4.13. Tandem mass spectra of the four UPF0337 family stress response proteins isolated from streptococci. Note the high sequence similarity.

4.3.6 *De novo* sequencing of an unknown protein in *Staphylococcus* sp.

The majority of protein identification in this part of the project relied fully on the use of ProSightPC software to provide putative protein identification, later verified by manual peak assignment. This approach, however, breaks down when the protein under investigation contains multiple modifications not predicted by the databases on which ProSightPC relies, or if the protein is entirely absent from them. One such challenge was encountered while

investigating a sample of a bacterium believed to be *Staphylococcus epidermidis*. Mass spectra abundant in small secreted peptides, similar to those generated later from *S. aureus*, were acquired. Of particular interest was the characteristic distribution of peaks at m/z 1200-1300, dominated by peptides approximately 4 kDa in size. Acquired tandem mass spectra were matched using ProSightPC against a proteome database generated from the whole sequence genome of *Staphylococcus epidermidis* ATCC 35984 (the reference strain for this species). The masses or sequences of these peptides did not match any of the proteins in the database. The option to *de novo* sequence one of the peptides in order to determine the exact nature of the sample was therefore explored.

Whilst *de novo* sequencing of almost all peptides selected for tandem mass spectrometric analysis was attempted, the majority of mass spectra did not provide sufficient fragmentation to enable entirely manual analysis. The most abundant peptide in the distribution (*m/z* 1124.34), however, generated C-terminal fragments (y_3 to y_7, 1+) which clearly corresponded to the sequence of a putative hit returned by ProSightPC, although the intact mass was significantly different. This peptide became a focus for manual analysis, using raw data directly to search for fragments (**Figure 4.14**); see **Appendix 3.2.7** for the complete list of assigned fragments. The sequence of the putative hit was inserted into Protein Prospector and used to estimate the mass-to-charge ratio and the charge state of fragments expected to appear in the mass spectrum; any missing peaks were indicative of a sequence difference. Once a peak was identified which corresponded to the mass of a previously located fragment plus the mass of an amino acid (within a 5 ppm difference), the new sequence was noted and inserted into Protein Prospector to generate a new list of fragments taking into account the newly identified sequence deviation. This process was repeated until both the *b* and the *y* series of fragments were almost completely elucidated; two gaps in sequence coverage,

spanning three amino acids each, were identified, where the identity but not the order of amino acids could be determined. The final amino acid sequence, including all the permutations of the three-amino acid gaps, was used to conduct a BLAST search to identify any known homologs and thus determine the exact species of the sample; the closest match (80% identity) belonged to *S. capitis*, a commensal species found living on human skin, although such a score still indicates severe divergence of the sampled species from any bacteria sequenced before. It is therefore likely that the laboratory-strain *Staphylococcus* species became quite unique over the years, mutating and evolving from the originally purchased strain into a completely new species. The high mutation rate of phenol-soluble modulins observed in staphylococci[304] increases the likelihood of such evolution occurring over the span of decades.

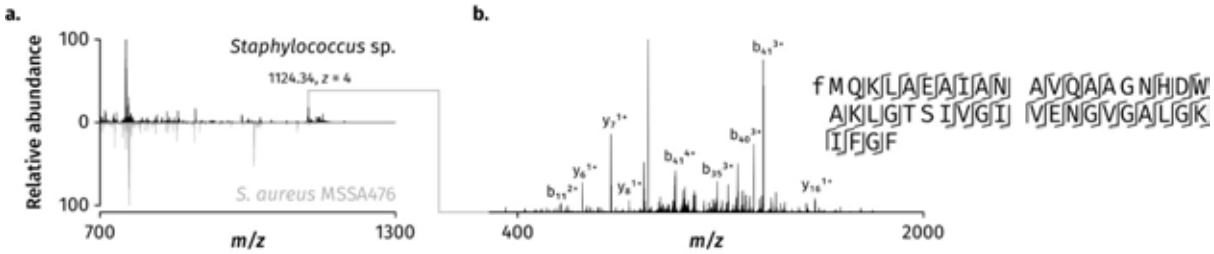


Figure 4.14. *De novo* sequencing of a novel protein from an unknown species of *Staphylococcus*. **a.** A direct comparison of mass spectra acquired from *Staphylococcus* sp. (top) and *S. aureus* MSSA476 (bottom); both spectra were generated with the 50:45:5 acetonitrile-water-formic acid solvent system. **b.** Tandem mass spectrum of ions centred at *m/z* 1124.34, corresponding to a putative novel phenol-soluble modulin β peptide.

4.3.7 Influence of cold storage and repeated incubation

Both in clinical and in research-focused settings, it is frequently the case that samples need to be stored for extended periods of time prior to analysis. This situation, however, can be reasonably expected to alter the composition of the sample, particularly where the sample is alive and capable of responding to the altered environmental conditions. Crucially, any changes to the mass spectra brought about by sample storage or handling need to be assessed

should LESA MS be ever deployed as a tool for the identification of bacteria, as these could potentially interfere with the identification algorithms. On the other hand, any detection of differences in mass spectra following differences in sample growth or storage would be indicative of the potential of LESA MS to detect and characterise the responses of living microbes to changes in their environment, opening avenues for its use as a microbiological characterisation tool.

It was previously noted that LESA mass spectra generated from colonies of *E. coli* stored at 4 ℃ for a number of days contained marked differences as compared to mass spectra collected from samples retrieved immediately from incubation.[117] In particular, a characteristic charge state envelope corresponding to an additional, highly abundant protein was observed; this was identified as BhsA, a multiple stress response factor. A more systematic investigation, encompassing *S. aureus* MSSA476 and *P. aeruginosa* PS1054 alongside *E. coli*, was therefore devised to identify any stress response proteins expressed and detected as a result of exposure to cold storage, as well as to determine whether a return to optimal temperatures could reverse phenotypic changes induced by storage conditions and recover mass spectra comparable to those recorded from bacteria sampled directly following incubation. Three representative bacterial species, *E. coli*, *P. aeruginosa* and *S. aureus* were included in the experiment. For each species, nine colonies grown on a single LBA plate were subjected to sequential changes in temperature: 37 ℃ for 24 h, 4 ℃ for 4 days, and finally 37 ℃ for 24 h. Following each period, three of the colonies were sampled by LESA MS. Each set of mass spectra was normalized. All experiments were performed in triplicate to ensure reproducibility and statistical significance.

Figure 4.15 shows representative results of this investigation, displaying one series of mass spectra per species. As previously mentioned, the appearance of the multiple stress response

protein BhsA following cold storage was expected in *E. coli* based on the previous study by Randall et al.[117] Accordingly, the protein was absent in mass spectra of *E. coli* acquired immediately following incubation but highly abundant peaks corresponding to the 7+ and 6+ charge states (*m/z* 933.90 and 1089.38, respectively; see **Appendix 3.2.1.13** for identification) were detected following cold storage. The signal was absent in colonies sampled following the second incubation step, indicating that the detection, and likely expression, of BhsA is strongly dependent on the recent exposure of the colonies to cold stress and does not persist once the growth conditions are brought back towards optimum, a behaviour largely consistent with the putative role of BhsA as a cold shock protein. A pair of proteins highly related to each other, YnaE (**Appendix 3.2.1.14**) and YdfK (**Appendix 3.2.1.7**), were also detected immediately following exposure to low temperatures; the expression of these proteins has previously been detected at transcript level in microbiological studies,[310] and the pattern of expression was determined to be consistent with that of putative cold shock proteins. Interestingly, peaks corresponding to this pair of proteins (*m/z* 1250.67 7+, 1094.46 8+, 972.85 9+ and 875.77 10+ for YnaE; *m/z* 1254.67 7+, 1097.71 8+, 975.86 9+, 878.37 10+, 798.61 11+ for YdfK) were also clearly observed in all mass spectra of *E. coli* acquired following the second incubation step, demonstrating that some stress-induced changes in bacterial phenotype can persist for multiple generations following the removal of a stressor.

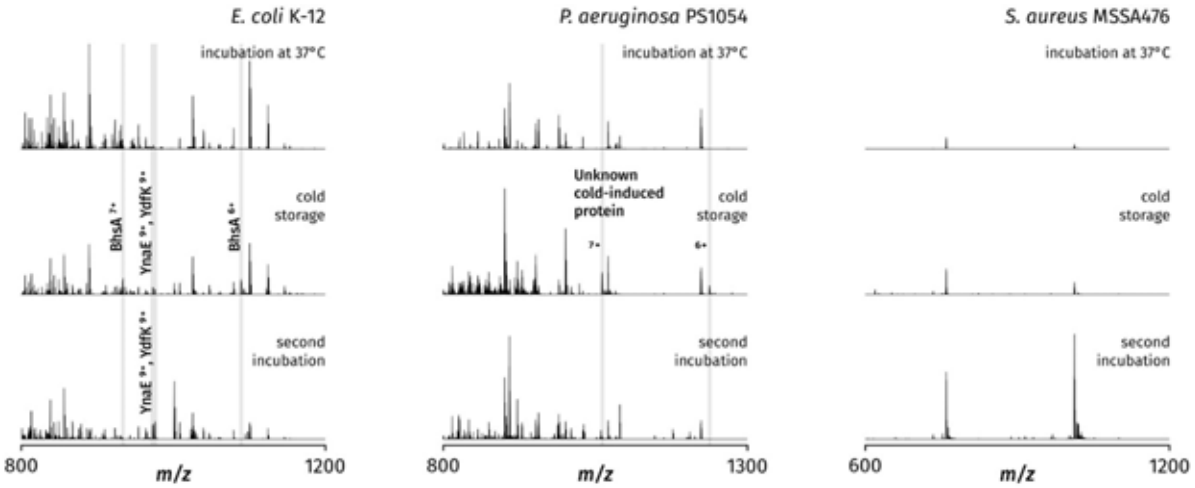


Figure 4.15. The investigation into the effects of cold storage on the mass spectra of three representative bacterial species: *E. coli* K-12, *P. aeruginosa* PS1054 and *S. aureus* MSSA476. Each set of mass spectra corresponds to three colonies sampled from the same agar plate; for each species, the first set of replicates is shown. The solvent system used was 40:60:1 for *E. coli* and *P. aeruginosa*, and 50:45:5 for *S. aureus*.

Changes similar to those observed in mass spectra of *E. coli* were also detected for *P. aeruginosa*. Whilst the majority of protein peaks remained consistent across all conditions, additional peaks (*m/z* 1237.45 6+ and 1060.83 7+) corresponding to an unknown product 7418.73 Da in mass were detected in mass spectra gathered immediately following the period of cold storage. This putative cold shock protein, briefly mentioned in section 4.3.5, could not be successfully assigned to any known *P. aeruginosa* gene product. Its clear appearance following cold storage, however, implies a function in stress response. It is possible that LESA could be more widely used in such manner, shedding light on the probable functions of uncharacterised proteins based on the patterns of their expression and detection following exposure to stressors. This could potentially include the response to the presence of antibiotic compounds, aiding in the deciphering of the mechanisms of resistance.

Whilst changes were also observed in the mass spectra of *S. aureus*, these were much different in nature compared to those detected in the other two species. The expression of cold shock proteins was not detected. δ-Haemolysin (*m/z* 759.92 4+, 1012.93 3+) consistently remained

the most abundant analyte in the mass spectra; however, normalisation of the mass spectra revealed a pattern in its abundance. Immediately following the first round of incubation, its signal was comparatively low but a marked increase was recorded over the course of the experiment, culminating after the second incubation step. This behaviour would be consistent with literature claiming that the highest expression of this toxic peptide is recorded during stationary phase.[312]

4.4 Conclusions

LESA MS was shown here to be a suitable, useful tool for the identification and characterisation of proteins in bacterial colonies. Its capability to investigate both Gram-positive and Gram-negative species was developed and demonstrated, unlocking a multitude of applications for the characterisation of bacteria, both in clinical and in research settings. Additionally, its capability to distinguish closely related pathogenic species was investigated, showing that single proteins detected by LESA MS may be used as markers for species identification. At this stage of development, the methodology could be employed to look at a number of clinically relevant systems as well as other growth substrates, further expanding the pool of potential applications and contributing to its eventual adoption as a tool for microbiological research. In order to validate the results, further clinically relevant species on a range of growth substrates should be investigated to ensure the consistent performance of LESA. It may be beneficial to also investigate the sampling of smaller colonies or contaminated surfaces to render the protocols more compatible with standard clinical procedures currently in use. A parallel project initially focusing on LESA MS analysis of the ESKAPE pathogens (*Enterococcus, Staphylococcus, Klebsiella, Acinetobacter, Pseudomonas, Enterobacter*), which currently constitute the main causes of nosocomial infections in the world, was established based on the results described here.

Three challenges were identified during the course of the work. First, the background signal generated by the solvent and the nutrient media dominated the low *m/z* range of mass spectra acquired. This was intrinsic to the process of LESA sampling on bacteria probed directly from agar plates. Second, multiple compound tandem mass spectra were acquired, particularly in *P. aeruginosa*, due to overlapping protein peaks. The problem was exacerbated in the case of low abundance proteins where larger isolation windows had to be used to generate sufficient signal for fragmentation. Finally, mass spectra of *S. aureus* were dominated by δ-haemolysin and did not provide satisfactory information on cytosolic proteins. A proposed solution to these challenges was the coupling of FAIMS to LESA MS, described in the next chapter.

CHAPTER 5

Application of ion mobility separation to LESA mass spectrometry of bacteria

5.1 Introduction

Several challenges pertaining to the sampling and analysis of bacteria were identified in the previous chapter and are briefly summarised here. First, multiple highly abundant peaks corresponding to components of the solvent system as well as the agar media were seen in the low *m/z* range, obstructing signal corresponding to putative small proteins. Second, in the higher *m/z* range, multiple compound tandem mass spectra were acquired where the isotopic distributions of protein precursor ions would overlap and prevent clean isolation. Third, in *S. aureus* in particular, the mass spectra were heavily skewed towards a specific subset of analytes (small toxic peptides).

This chapter describes the coupling and optimisation of high-field asymmetric waveform ion mobility spectrometry (FAIMS) to LESA MS to address the above challenges. FAIMS separation was expected to primarily improve signal-to-noise ratios, allowing for the observation of peaks which would otherwise remain undetectable in the mass spectra. Experiments were carried out both in one-dimensional (1D) sweep mode, where the compensation field (CF) was continuously scanned across a range of values whilst the dispersion field (DF) was kept constant, and in static field mode where both the CF and DF values were kept constant during data acquisition. Initial experiments were performed on *P. aeruginosa* PS1054; the FAIMS settings were optimised for maximum protein separation. The results of the optimisation were subsequently translated to *E. coli* K-12 and *S. aureus* MSSA476. Each species served as a test system for a specific challenge to be addressed by

FAIMS. For *E. coli*, FAIMS parameters were tuned to remove background signal interference and reveal additional protein peaks in the low *m/z* range. For *P. aeruginosa*, FAIMS was used to separate overlapping protein peaks and resolve composite tandem mass spectra. Finally, for *S. aureus*, FAIMS was used to reduce the abundance of δ-haemolysin in mass spectra and increase the abundance of other protein and peptide species. **Table 5.1** lists of proteins identified by LESA-FAIMS MS in these three species, including proteins not detected in the work described in Chapter 4. Supplementary tandem mass spectra and all fragment assignments are provided in **Appendix 3.3**.

Additionally, an attempt was made to quantify the capability for protein identification by LESA MS and LESA-FAIMS MS as compared to established methodologies. Top-down LC-MS experiments were carried out on both LESA extracts of *E. coli* colonies, and whole colony lysates.

5.2 Experimental

5.2.1 LESA

Colonies of *Escherichia coli* K-12, *Pseudomonas aeruginosa* PS1054 and *Staphylococcus aureus* MSSA476 were prepared as outlined in Chapter 2, section 2.2.1.3 directly from glycerol stocks. Sampling of each species was carried out according to the protocols optimised in Chapter 4 and summarised in Chapter 2, section 2.2.3.

5.2.2 FAIMS

Acquisition of LESA-FAIMS mass spectra was carried out as described in Chapter 2, section 2.2.5.

5.2.3 Protein identification

The acquisition and processing of top-down mass spectra was performed as described in Chapter 2, section 2.2.5 (acquisition) and 2.2.8 (processing); isolation windows of 5 *m/z* were used due to the lower risk of co-isolation of multiple analytes as a result of the separation provided by FAIMS. Top-down protein identification was carried out as described in Chapter 2, section 2.2.8.

5.2.4 LC-MS—sample preparation

Agar plates of *E. coli* K-12, each bearing three identical colonies, were prepared as described in Chapter 2, section 2.2.1.3. For each colony separately, three LESA extractions were performed, with no alterations made to the standard protocol; the extract was deposited in an empty well of the halved 96-well plate. The three extractions were pooled to generate a total volume of approximately 5 µl; some solvent was lost to the colony during the extraction process. The extract was diluted ten-fold with pure water to reduce the percentage content of acetonitrile and microcentrifuged at 12,000 × *g* for 5 minutes to pellet any colony debris that may have been introduced by the extraction pipette. The supernatant was transferred to a clean 96 well plate.

To generate a whole colony lysate, an established protocol was followed.[297] A colony was scraped off the agar plate in its entirety using a 10 µl inoculation loop and deposited in an Eppendorf tube with 500 µl of 70% ethanol. The colony matter was broken up using a sterile inoculation loop and vortexed until a loose suspension was formed. The tubes were sonicated for 3 minutes in an ultrasonic cleaner to further break apart the colony material. 100 µl of the suspension was transferred to a clean Eppendorf tube and made up to 1 ml by addition of a 50:45:5 acetonitrile-water-formic acid extraction solvent, following which it was subjected to further ultrasonication for 10 minutes. The suspension was spun down at 12,000 × *g* for 5

minutes to pellet colony debris. 40 µl of the supernatant was loaded into the 96 well LC plate for analysis.

Three biological replicates (one colony per replicate) were prepared for each approach. Every sample was run twice as technical replicates.

5.2.5 LC-MS—chromatography and data acquisition

LC-MS was carried out by use of an Ultimate 3000 nano-LC system (Thermo Fisher Scientific, Bremen, Germany) coupled to the LTQ-Orbitrap Elite. Optimisation of the chromatography and data acquisition settings is described in section 5.3.4 of this chapter. The final parameters are listed in Chapter 2, section 2.2.5.

Full scan data were collected in the *m/z* 600-2000 range by use of the orbitrap mass analyser at a resolution of 60,000 at *m/z* 400. Three microscans were coadded per scan. The full scan was followed by five data-dependent MS/MS scans corresponding to the top five most abundant peaks of charge state 4+ or higher; the abundance threshold was set to 5000 ion counts. The isolation width was set to 3 *m/z*. Fragmentation was carried out by CID at a normalised collision energy of 35%.

5.3 Results and discussion

5.3.1 Removal of background interference—*E. coli* K-12

In the absence of FAIMS, the low *m/z* region (600-800) of LESA mass spectra of *E. coli* is characterised by singly charged peaks corresponding to extraction solvent, metabolites, putative lipids and components of the agar media on which the colonies have been grown (see Chapter 4, section 4.3.1; for examples see **Figures 4.2** and **5.1**). Peaks corresponding to protein ions in this *m/z* region, while occasionally discernible, were largely obscured by the peaks corresponding to singly charged species, preventing the isolation of protein peaks and

their identification. It was proposed that these small molecules could be filtered out by FAIMS, leaving behind an undisturbed distribution of protein peaks in the same *m/z* range. Thus, smaller proteins and peptides, as well as higher charge states of larger species, should be observed.

A one-dimensional FAIMS experiment was performed at DF 270 Td to optimise compensation field values for the transmission of proteins of interest in the low *m/z* range. DF 270 Td was selected to maximise the separation capability of the device.[266] Representative mass spectra acquired during the one-dimensional analysis are shown in **Figure 5.1**. A mass spectrum acquired in the absence of FAIMS is shown in **Figure 5.1a**. At CF 0.0 to 1.2 Td, transmission of singly charged species was observed (**Figure 5.1b**). Transmission of protein ions was observed from CF 1.2 Td onwards with the emergence of peaks corresponding to an unidentified 18.1 kDa protein (*m/z* 865.40 21+, 908.77 20+, 956.50 19+, 1009.69 18+, 1068.90 17+, 1135.97 16+, 1211.30 15+, 1297.67 14+ and 1397.57 13+; MW_{obs} = 18155.29 Da; **Figure 5.1c**). A distribution of peaks of low abundance corresponding to the histone-like protein H-NS (*m/z* 856.67 18+, 906.94 17+, 963.56 16+, 1027.67 15+, 1100.93 14+, 1185.53 13+, 1284.58 12+ and 1401.36 11+; MW_{obs} = 15399.89 Da) was detected between CF 2.2 Td and CF 3.0 Td. Peaks corresponding to YahO (*m/z* 963.99 8+, 1101.42 7+ and 1284.82 6+; MW_{obs} = 7702.87 Da), one of the proteins consistently detected in every LESA mass spectrum acquired from *E. coli* K-12 (see Chapter 4, section 4.3.1), emerged at approximately CF 2.2 Td and persisted until CF 3.8 Td (**Figure 5.1d**); HU-α (*m/z* 734.10 13+, 795.12 12+, 867.30 11+, 953.92 10+, 1060.03 9+ and 1192.53 8+; MW_{obs} = 9529.17 Da) and HU-β (*m/z* 710.39 13+, 769.42 12+, 839.18 11+, 923.00 10+, 1025.45 9+ and 1153.63 8+; MW_{obs} = 9219.96 Da) exhibited a similar pattern of transmission. A correlation between the strength of the compensation field and the mass-to-charge ratio of transmitted ions was observed, exemplified by the distributions of HU-α and

HU-β charge states. Charge states 8+ and 9+ of both proteins emerged at CF 2.2 Td, followed by 10+ at CF 2.4 Td, 11+ at CF 2.6 Td, 12+ at CF 2.9 Td and finally 13+ at CF 3.0 Td. In the CF range of 3.0–3.5, singly charged peaks in the low *m/z* region were almost entirely removed and protein peaks were observed without interference of lower molecular weight analytes (**Figure 5.1e**). Between CF 3.5 Td and CF 4.0 Td, the protein signal rapidly decreased in intensity.

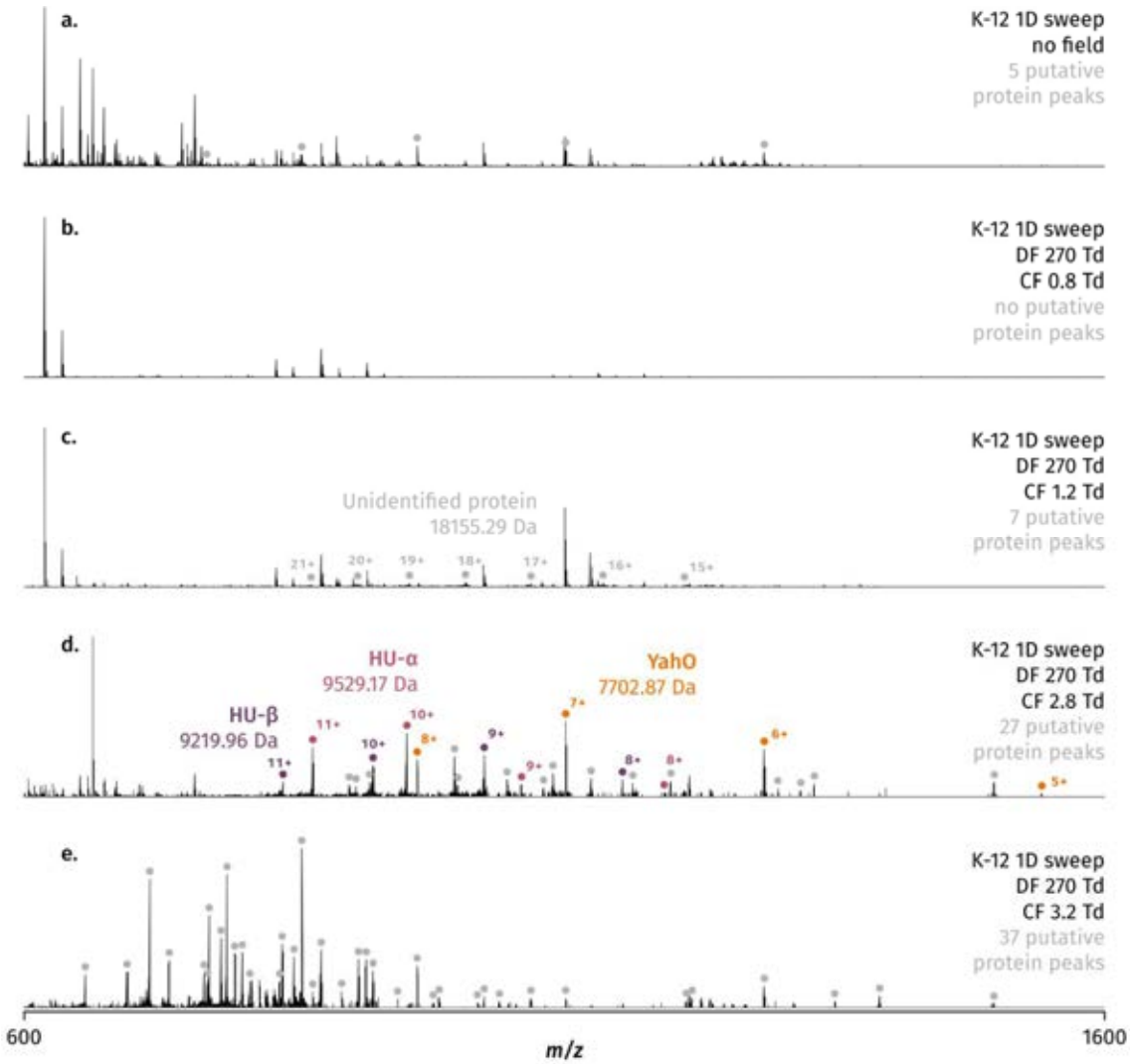


Figure 5.1. Representative LESA-FAIMS-MS single scan spectra (no averaging performed) of *E. coli* K-12 acquired during a one-dimensional FAIMS sweep experiment at DF 270 Td. The compensation field (CF) ramp ranged from CF 0.0 Td to CF 4.0 Td over 240 seconds. Putative protein peaks (charge states 4+ and above) are marked by dots.

Mass spectra in static FAIMS field mode were subsequently acquired in the range of CF 3.0 Td—3.5 Td, where the greatest reduction in the abundance of singly charged background peaks was observed. **Figure 5.2** shows a comparison of a LESA-FAIMS mass spectrum of *E. coli* K-12 acquired with no field applied versus two LESA-FAIMS mass spectra acquired in static field mode at DF 270 Td-CF 3.2 Td and DF 270 Td-CF 3.5 Td, respectively. For all mass spectra acquired in static mode, an increase in the number of detected proteins was observed compared to the number of proteins detected at identical settings during the one-dimensional FAIMS sweep experiment. The increase was attributed to the extended acquisition of data at a given set of FAIMS parameters in static mode (one minute, summed) compared to the one-dimensional FAIMS sweep experiment (single scans), and the corresponding increase in signal to noise ratio, revealing low abundance peaks. In the mass spectrum acquired without FAIMS, as seen in **Figure 5.2a**, only three protein peaks were observed in the *m/z* 600-800 region; the total number of peaks observed in the full displayed mass range (*m/z* 600-1600) was 23.

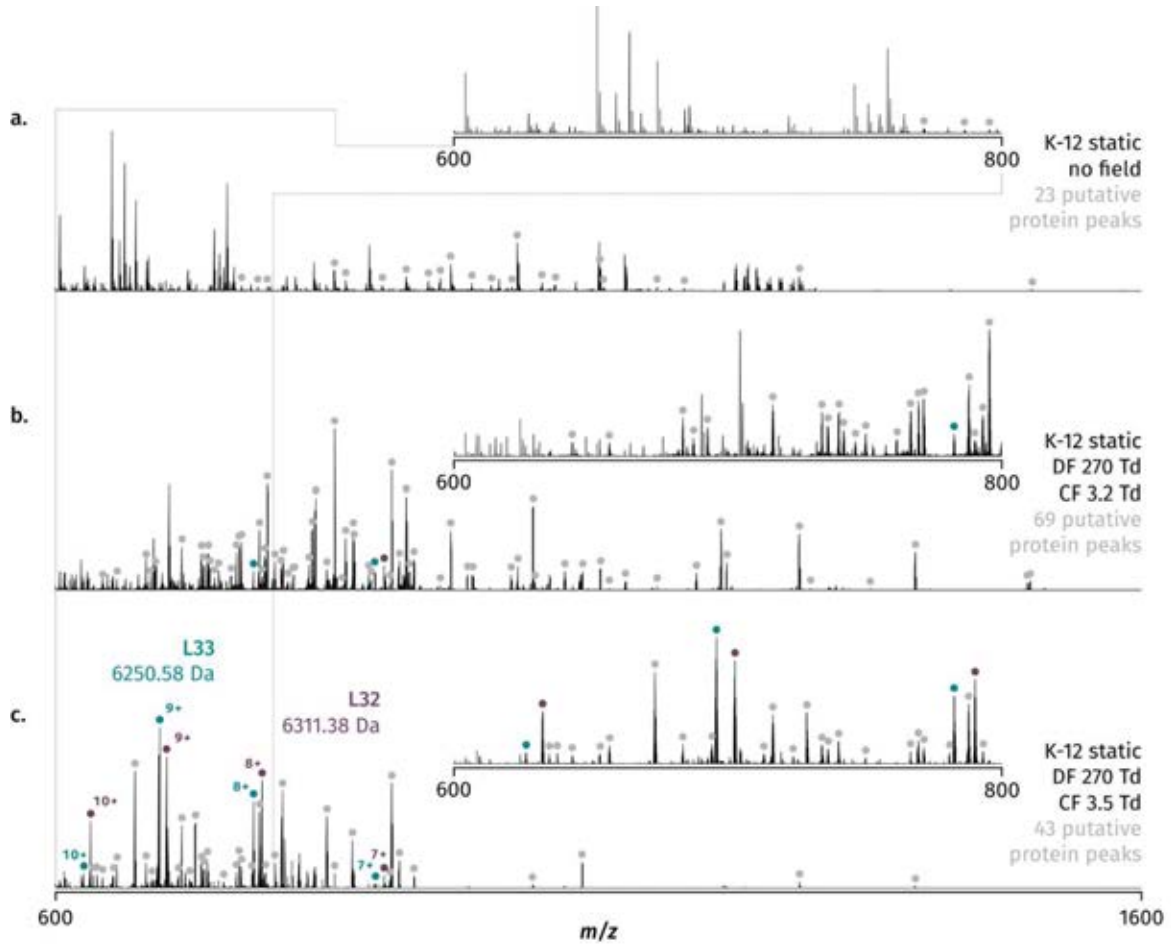


Figure 5.2. LESA-FAIMS mass spectra of *E. coli* K-12 acquired in static FAIMS mode. **a**. No FAIMS applied. **b**. DF 270 Td-CF 3.2 Td. **c**. DF 270 Td-CF 3.5 Td. Each mass spectrum was summed across 1 minute of data acquisition. The *m/z* 600-800 range of each mass spectrum is shown as an inset.

Figure 5.2b shows a mass spectrum acquired in static field mode at DF 270 Td-CF 3.2 Td. A total of 69 protein peaks were observed across the full mass range shown; 21 protein peaks were detected in the *m/z* 600-800 range. Singly charged background peaks were eliminated in the *m/z* 700-800 range but still constituted the majority of peaks in the *m/z* 600-700 range. Compared to the single scan displayed in Figure 5.1 for the same FAIMS settings, the number of protein peaks detected in static FAIMS mode was significantly higher (69 versus 27). This discrepancy could be attributed to the extended time of data acquisition in static FAIMS mode,

which allowed the summation of multiple scans and enabled the observation of low abundance proteins otherwise not evident in the data.

At CF 3.5 Td, 28 protein peaks were detected in the *m/z* 600-800 range (**Figure 5.2c**). Singly charged peaks were absent. A total of 43 protein peaks were observed across the full mass range shown. Four peaks (*m/z* 695.92 9+, 702.27 9+, 728.54 7+ and 916.37 7+) were selected for fragmentation and identification at these settings. **Table 5.1** lists all the proteins isolated and identified by use of FAIMS for *E. coli* and *P. aeruginosa* (see following section). The peak at *m/z* 695.92 (9+; MW_{obs} = 6250.58 Da) was identified as the ribosomal protein L33 (**Figure 5.3a**), not detected before without the use of FAIMS; the intact mass and fragmentation pattern were consistent with the cleavage of initiator methionine and the presence of methylalanine at the N-terminus, documented by prior literature (for fragment assignments, see **Appendix 3.3.1.2)**.[313] The peak at *m/z* 702.27 (9+; MW_{obs} = 6311.38 Da) was identified as L32 (**Figure 5.3b**), another structural constituent of the ribosome which was not previously observed without the use of FAIMS; the initiator methionine was not present (**Appendix 3.3.1.3**). The peak at *m/z* 728.54 (7+; MW_{obs} = 5092.75 Da) was confirmed as SRA, a sigma-associated factor previously identified without FAIMS (see Chapter 4, section 4.3.5; **Appendix 3.3.1.1**). Similarly, the peak at *m/z* 916.37 (7+; MW_{obs} = 6406.63 Da) was identified as ribosomal protein L30 (**Appendix 3.3.1.4**).

Table 5.1. A list of proteins identified on the basis of tandem mass spectra acquired by the LESA-FAIMS approach. All proteins were isolated at DF 270 Td. Proteins previously identified without the use of FAIMS are shown in grey.

Protein	MW_{obs}	ProSight E-score	CF (Td)	Notes
Escherichia coli K-12				
SRA	5092.75	8.2E-30	3.5	Ribosome-associated
L33	6250.58	2.2E-40	3.5	Ribosomal
L32	6311.38	1.7E-31	3.5	Ribosomal
L30	6406.63	1.2E-06	3.5	Ribosomal
RaiA	12647.50	7.8E-12	2.2	Ribosome-associated inhibitor
YgiW	11968.91	1.8E-49	3.0	Biofilm formation on agar
Pseudomonas aeruginosa PS1054				
L33	6041.32	3.3E-14	3.5	Ribosomal
L32	6670.37	1.3E-10	2.5	Ribosomal
CsrA	6904.68	5.3E-20	3.2	Carbon storage regulator
PA4739	8556.48	6.6E-14	3.2	Stress response; missing N-terminal sequence AQPISLA-
HU-β	9080.99	3.1E-14	3.2	DNA-binding
L11	14894.97	1.7E-08	1.0	Large ribosomal
Pilin	14983.82	3.9E-12	2.5	Pilus formation

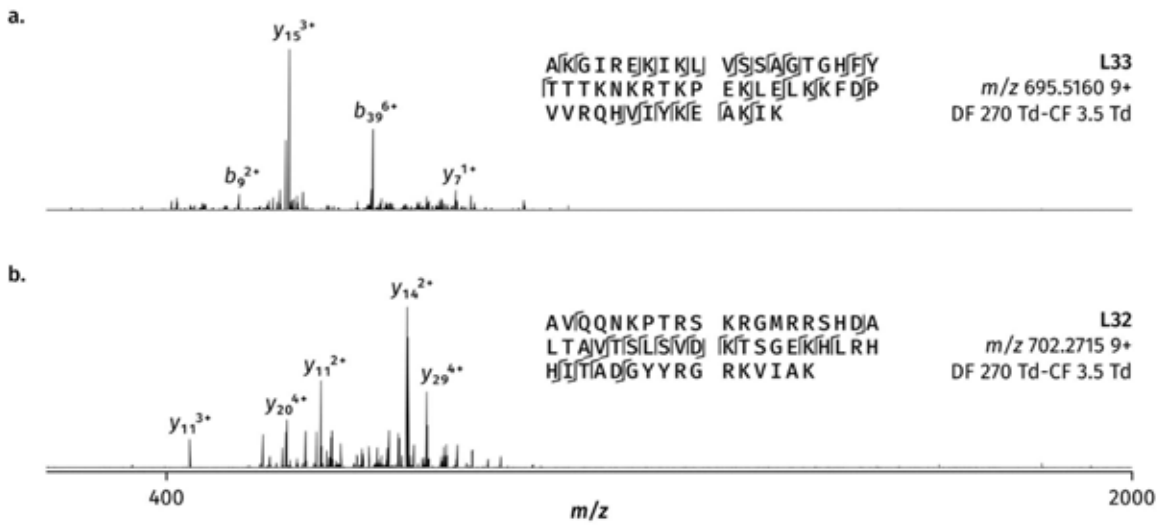


Figure 5.3. Tandem mass spectra of two ribosomal proteins detected in the low *m/z* range by use of FAIMS at DF 270 Td-CF 3.5 Td. **a**. L33 (MW_{obs} = 6250.58 Da); **b**. L32 (MW_{obs} = 6311.38 Da).

Two further peaks were isolated and identified in static field mode at other FAIMS settings.

The peak at *m/z* 1497.25 (8+; MW_{obs} = 11968.91 Da), isolated at DF 270 Td-CF 3.0 Td, was

identified as YgiW (**Appendix 3.3.1.6**), a periplasmic protein involved in biofilm formation on inanimate substrates (i.e. agar media); the intact mass confirmed the loss of a 20 amino acid signal peptide previously documented in the literature.[314] The peak at *m/z* 1054.88 (12+; MW_{obs} = 12647.50 Da), isolated at DF 270 Td—CF 2.2 Td, was identified as RaiA, a well-characterised ribosome-associated inhibitor which increases the accuracy of translation under stress conditions, such as low temperature and high growth density in stationary phase (**Appendix 3.3.1.5**).[315, 316] Its detection here is consistent with the high density of colonies used for LESA sampling.

5.3.2 Resolution of composite tandem mass spectra—*P. aeruginosa* PS1054

The coupling of FAIMS to LESA MS in the case of *P. aeruginosa* was spurred by the acquisition of multiple compound tandem mass spectra as described in Chapter 4. A particularly clear example of this issue could be seen in the data for the stress response protein PA4739 (see Chapter 4 section 4.3.5 and **Appendix 3.2.2.5** for the tandem mass spectrum). Whilst the tandem mass spectrum of the precursor isolated at *m/z* 1070.82 (charge state 8+) provided sufficient information to generate a satisfactory sequence coverage, numerous fragments were detected which could not possibly correspond to any theoretical fragments of the identified protein as they were significantly higher in mass than the precursor ion. An inspection of the isolation window (shown in **Figure 5.5a** below) revealed the presence of a distribution of very low abundance peaks carrying a high apparent charge (estimated as 14+), corresponding to an additional high-mass precursor. It was hypothesised that FAIMS could separate the overlapping precursor peaks and yield clean tandem mass spectra for subsequent identification.

As for *E. coli K-12*, a one-dimensional FAIMS (DF 270 Td, CF 0.0—4.0 Td) experiment was performed to elucidate the pattern of transmission of proteins of interest. Representative

mass spectra from the analysis are displayed in **Figure 5.4**. A single scan acquired in the absence of FAIMS is shown in **Figure 5.4a**. No protein peaks were observed between CF 0.0 and 1.0 Td. At a CF of 1.0 Td and above, a clear envelope of five charge states (13+ to 17+; *m/z* 1146.93, 1065.01, 994.07, 932.00 and 877.36, respectively) corresponding to L11, a large ribosomal protein (MW_{obs} = 14898.00 Da), was observed (**Figure 5.4b**; see below for details on identification).

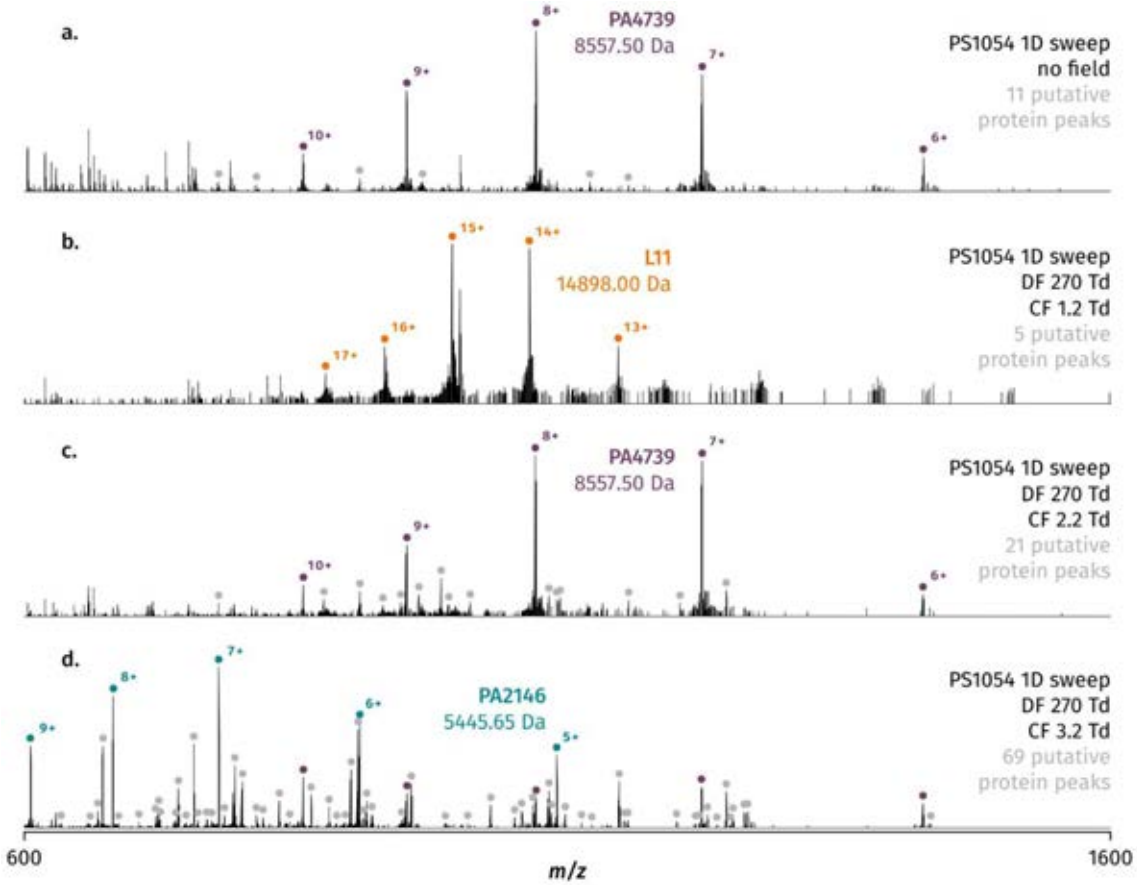


Figure 5.4. Representative LESA-FAIMS-MS single scan spectra of *P. aeruginosa* PS1054 acquired during a one-dimensional FAIMS sweep experiment at DF 270 Td. The compensation field (CF) ramp ranged from CF 0.0 Td to CF 4.0 Td over 240 seconds. Putative protein peaks (charge states 4+ and above) are marked by dots.

Increasing the strength of the compensation field led to the gradual disappearance of the L11 pattern and the subsequent emergence of a distribution of peaks (*m/z* 856.86 10+, 951.95 9+,

1070.82 8+, 1223.51 7+ and 1427.43 6+) corresponding to PA4739 (MW_{obs} = 8557.50 Da), a protein previously characterised by direct LESA mass spectrometry, at CF values between 1.2 and 3.2 Td (**Figure 5.4c**). At the highest compensation fields (**Figure 5.4d**), a large population of smaller proteins was observed, among which PA2146 (*m/z* 606.08 9+, 681.71 8+, 778.96 7+, 908.62 6+, 1090.1367 5+; MW_{obs} = 5445.65 Da) was the most abundant. The total transmission of proteins was highest at approximately CF 3.0-3.2 Td, offering good coverage of medium to small proteins, including putative ribosomal proteins.

The main goal of the experiment involved the resolution of co-isolating protein peaks for the purpose of fragmentation and unambiguous identification. Three instances of overlapping protein peaks were resolved by use of FAIMS at *m/z* 1071, *m/z* 864 and *m/z* 778.

As noted above, compound mass spectra of PA4739 containing fragments larger than the precursor ion were generated in the absence of FAIMS. The precursor of PA4739 isolated at *m/z* 1070.82 (8+) is shown in **Figure 5.5a**. At DF 270 Td-CF 2.5 Td, the co-isolating 14+ precursor ion was revealed at *m/z* 1071.35, MW_{obs} = 14984.81 Da (**Figure 5.5b**). Its fragmentation and subsequent automatic peak assignment (see **Appendix 3.3.2.7**) identified it as fimbrial protein (pilin), a structural protein involved in the formation of the *Pseudomonas* pilus and thus contributing to pathogenicity.[317] The cleavage of an N-terminal propeptide (six amino acids) and methylation of the N-terminal phenylalanine were confirmed by automatic fragment assignment (**Appendix 3.3.2.7**); the experimentally derived intact mass was, however, significantly different to the theoretical value (MW_{obs} = 14983.79 Da, MW_{calc} = 14870.71 Da; Δ111.07 Da). *b* and *y* ions of expected masses covering the entire sequence were identified by supplementary manual verification, suggesting that the mass difference may have been due to a labile post-translational modification which was lost upon fragmentation. No reports of such putative modifications were found in the literature. Whilst some of the

high molecular weight fragments first observed in the tandem mass spectrum of PA4739 (*m/z* 1127.42 12+, 1163.69 11+) could subsequently be assigned to pilin (y_{131}^{12+} and y_{123}^{11+}, respectively), fragments which could not be assigned to either protein (the fragment detected at *m/z* 1172.50, 11+ being the most abundant) remained unidentified. The observation of such fragments, as well as the discrepancy in mass-to-charge ratio of the pilin precursor and the high molecular weight precursor originally observed in the isolation window of PA4739 (as seen in **Figures 5.5a** and **5.5b**), suggest that a third, hitherto unidentified protein may be present in this *m/z* region.

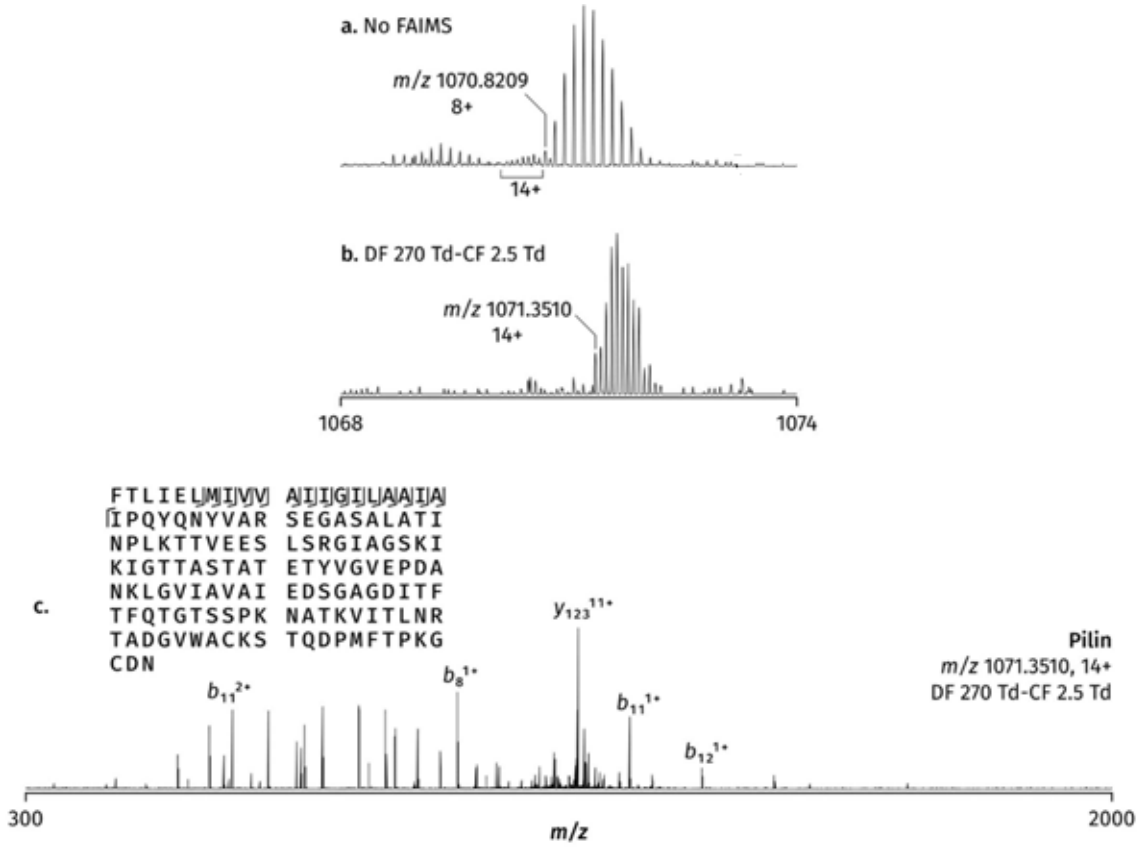


Figure 5.5. Isolation and identification of pilin in *P. aeruginosa*. **a**. No FAIMS applied; only the precursor of PA4739 (8+) is detected. **b**. DF 270 Td-CF 2.5 Td. The 14+ precursor (pilin) is isolated. The mass spectrum acquired without FAIMS is outlined in grey for comparison. **c**. Tandem mass spectrum of pilin (14+) isolated by application of FAIMS.

The second instance involved two proteins detected at *m/z* 864 (**Figure 5.6a**). At DF 270 Td-CF 3.5 Td, a clear peak corresponding to the 7+ charge state of the 50S ribosomal protein L33 (*m/z* 864.06, MW_{obs} = 6041.35 Da; also detected and identified without FAIMS, see Chapter 4 section 4.3.5) was observed (**Figure 5.6b**; see **Appendix 3.3.2.1** for fragment assignment). A low abundance precursor of charge state 8+ (*m/z* 864.09) could, however, be discerned underneath the 7+ distribution of L33; this peak was favoured at CF 3.2 Td and could thus be isolated (**Figure 5.6c**) and identified as CsrA (MW_{obs} = 6904.69 Da; **Appendix 3.3.2.3**), a translational regulator involved in the regulation of carbon storage, among numerous other functions (**Figure 5.6d**). CsrA was detected in multiple compound tandem mass spectra acquired previously without FAIMS, as identified by ProSighPC 3.0 at a low confidence level. Here, it was detected for the first time with sufficient clarity to be unambiguously identified.

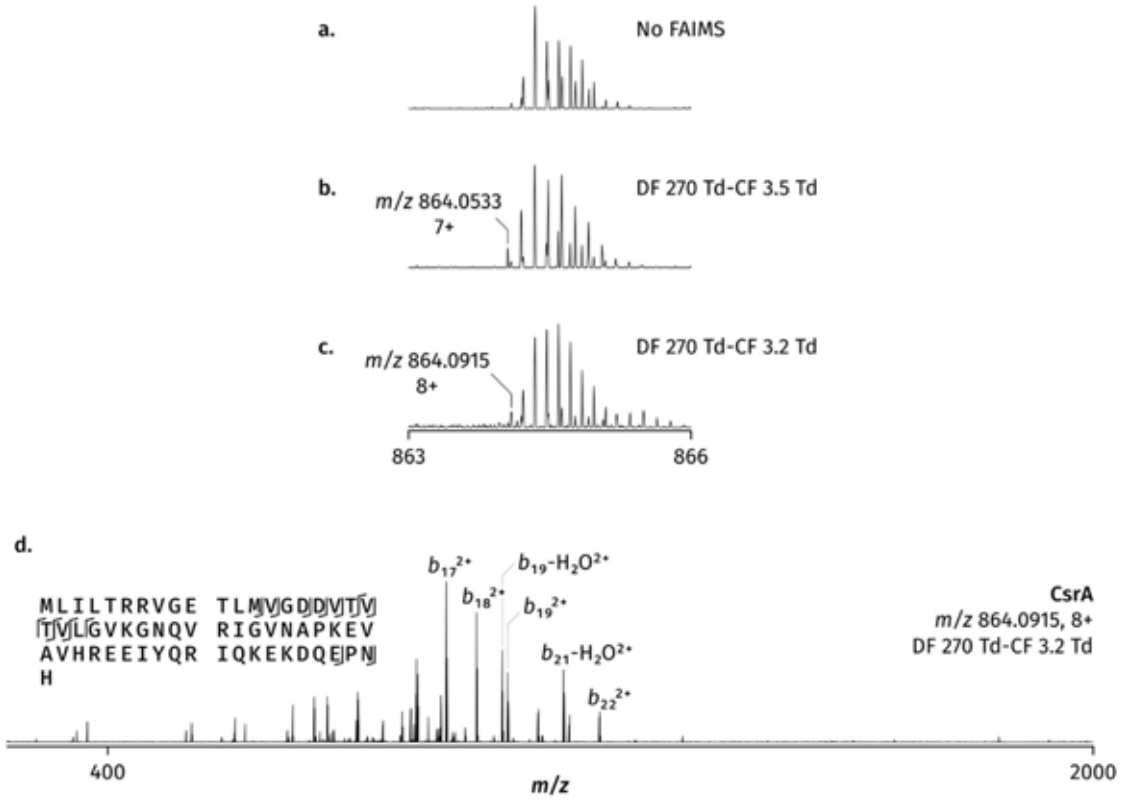


Figure 5.6. Isolation and identification of CsrA in *P. aeruginosa*. **a**. No FAIMS applied; both precursors are detected simultaneously. **b**. DF 270 Td-CF 3.5 Td. The 7+ precursor (L33) is more abundant; the 8+ precursor is still present within the isolation window. **c**. DF 270 Td-CF 3.2 Td. The 8+ precursor (CsrA) is isolated. Mass spectrum acquired without FAIMS is shown in grey for comparison. **d**. Tandem mass spectrum of CsrA (8+) isolated by application of FAIMS.

The third instance involved protein peaks detected at *m/z* 778, charge states 7+ and 11+ (**Figure 5.7**). Without FAIMS, both charge states were detected simultaneously (**Figure 5.7a**); the isotopic distribution of the 11+ charge state was entirely encompassed by the isotopic distribution of the 7+ charge state. By application of FAIMS at DF 270 Td-CF 3.5 Td, the 7+ peak (*m/z* 779.9576; MW_{obs} = 5445.65 Da), corresponding to PA2146 (see Chapter 4, section 4.3.5), was isolated on its own (**Figure 5.7b**). The 11+ peak (*m/z* 778.9604; MW_{obs} = 8557.48 Da) was isolated on its own at DF 270 Td-CF 3.2 Td (**Figure 5.7c**) and confirmed by intact mass and fragmentation pattern as PA4739 (see Chapter 4, section 4.3.5; fragment assignment for the 11+ charge state is supplied in **Appendix 3.3.2.4**). No fragments originating from the

PA2146 precursor were detected in the tandem mass spectrum of PA4739 acquired from the 11+ charge state using FAIMS, indicating complete separation of the two precursors. Whilst neither protein was identified here for the first time, isolation of these two specific charge states was not previously possible. When combined with the removal of low molecular weight background noise (see section 5.3.1), FAIMS could be used to target previously inaccessible, high charge states of proteins of interest, thus potentially yielding improved sequence coverage. Furthermore, the isolation of individual charge states may theoretically aid in the correct tuning of the collision energy, further improving sequence coverage, although this effect was not investigated here (all precursors were fragmented under the same conditions).

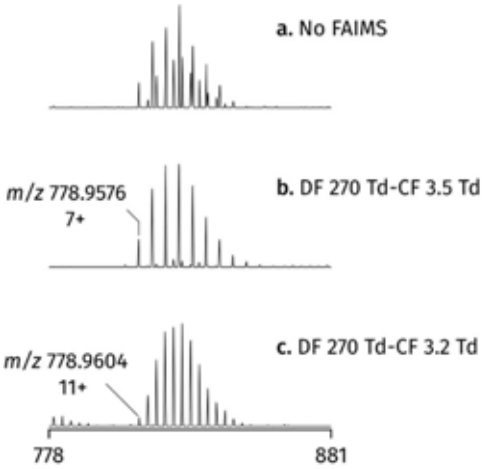


Figure 5.7. Isolation of two precursors at *m/z* 778. **a**. No FAIMS applied; both precursors are detected together. **b**. Isolation of the 7+ charge state of PA2146 at DF 270 Td-CF 3.5 Td. **c**. Isolation of the 11+ charge state of PA4739 at DF 270Td-CD 3.2 Td. The monoisotopic peaks of the two precursors are detected less than 0.01 *m/z* units apart.

In addition to proteins identified in compound mass spectra, further proteins previously not identified by LESA alone were identified here in static field mode. L11, the ribosomal protein mentioned above, was isolated at DF 270 Td-CF 1.0 Td (*m/z* 994.21, 15+; MW_{obs} = 14898.00 Da; **Figure 5.8**). This protein was detected at a very low abundance without FAIMS and was not previously identified by LESA. Fragmentation required the use of HCD (**Figure 5.8b**)

rather than CID (**Figure 5.8a**) due to the extremely poor fragmentation efficiency obtained by CID. Furthermore, whilst the observed intact mass did not match the predicted mass with satisfactory accuracy (ΔMW = 0.90 Da, 6.6 ppm by selection of the closest matching peak), manually calculated masses of the *y*-terminal fragments highlighted in **Figure 5.8** all matched their respective theoretical values to within 3 ppm (within 1 ppm for fragments up to y_{30}^{2+}; see **Appendix 3.3.2.6**); these observations were consistent with a probable amino acid substitution in the observed protein versus the theoretical sequence, proximal to the N-terminus.

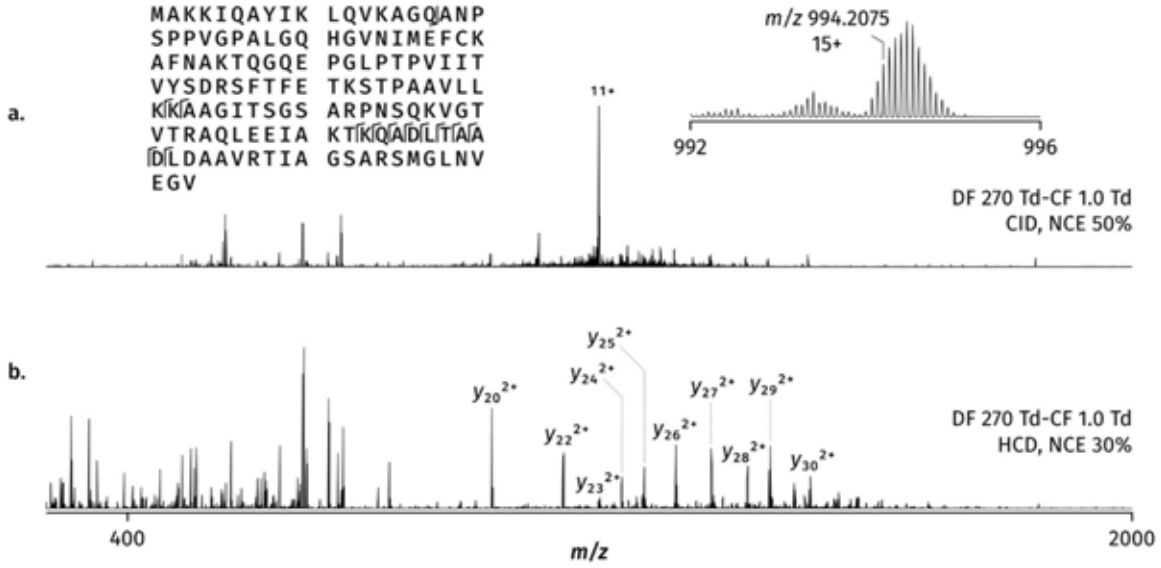


Figure 5.8. Tandem mass spectra of L11. **a**. MS/MS, CID, NCE 50%; **b**. MS/MS, HCD, NCE 30%.

In addition to pilin, isolated at DF 270 Td-CF 2.5 Td, another peak (m/z 953.92 7+; MW_{obs} = 6670.37 Da) was isolated at the same static field settings and identified as the ribosomal constituent L32. The observed intact mass was 14.01 Da higher than the theoretical mass and within 1.2 ppm of a theoretical mass consistent with the presence of methylation; fragment assignment confirmed the presence of monomethylation at the N-terminus of the protein, not present in the database (**Appendix 3.3.2.2**).

5.3.3 Reduction of dominant peaks—*S. aureus* MSSA476

The application of FAIMS to the analysis of *S. aureus* was proposed to eliminate the overwhelmingly dominant peaks of δ-haemolysin from the mass spectrum, allowing the observation of less abundant peptides and proteins.

When FAIMS was not applied, δ-haemolysin was consistently observed in two charge states, 4+ (*m/z* 759.67) and 3+ (*m/z* 1012.56); in several mass spectra, the 2+ charge state at *m/z* 1518.33 was also observed at low abundance. The 4+ charge state at *m/z* 759.67 was the most abundant and usually constituted the base peak of the mass spectrum. At DF 270 Td and CF sweeping from 0.0 to 4.0 Td, all three charge states of the toxin could be observed. Representative mass spectra from the one-dimensional FAIMS experiment are shown in **Figure 5.9**. A single mass spectrum acquired in the absence of FAIMS is shown in **Figure 5.9a**. The first peak corresponding to δ-haemolysin (*m/z* 1012.56 3+; MW_{obs} = 3034.64 Da) was detected at CF 0.5 Td. The peak remained detectable until CF 4.0 Td and constituted the base peak of the mass spectrum from CF 2.0 Td to 2.7 Td (**Figure 5.9b**); it was subsequently overtaken in abundance by the 4+ charge state (*m/z* 759.67; **Figure 5.9d**). The only exception occurred in one scan at CF 2.7 Td where the related peptide CAG44330 exceeded the abundance of both charge states of δ-haemolysin (*m/z* 969.03 4+ and 1291.71 3+, MW_{obs} = 3872.10 Da; **Figure 5.9c**). Peptides previously unseen by LESA alone (*m/z* 1142.65 4+, 1192.18 4+; MW_{obs} = 4566.58 and 4764.71 Da, respectively), most likely belonging to the phenol soluble modulin β class of secreted toxins as inferred from their intact masses, were also observed in one mass spectrum (see Chapter 4, section 4.3.6), as were nine peaks corresponding to six proteins of higher molecular weight (*m/z* 1070.95 6+, MW_{obs} = 6014.63 Da; *m/z* 1091.24 10+, 1212.16 9+ and 1363.93 8+, MW_{obs} = 10899.38 Da; *m/z* 1114.06 10+, MW_{obs} = 11130.56 Da; *m/z* 1164.47 7+, MW_{obs} = 8145.26 Da; *m/z* 1198.59 11+ and 1318.75 10+, MW_{obs}

= 13172.43 Da; *m/z* 1233.74 7+, MW_{obs} = 8624.09 Da); none, however, could be isolated and fragmented before their signal disappeared and the detection of the peaks in subsequently acquired mass spectra could not be reproduced. Thus, whilst the strong signal of δ-haemolysin could, to some extent, be reduced under the right FAIMS conditions, the original goal of its complete elimination was not achieved and the identification of previously undetected proteins was not accomplished.

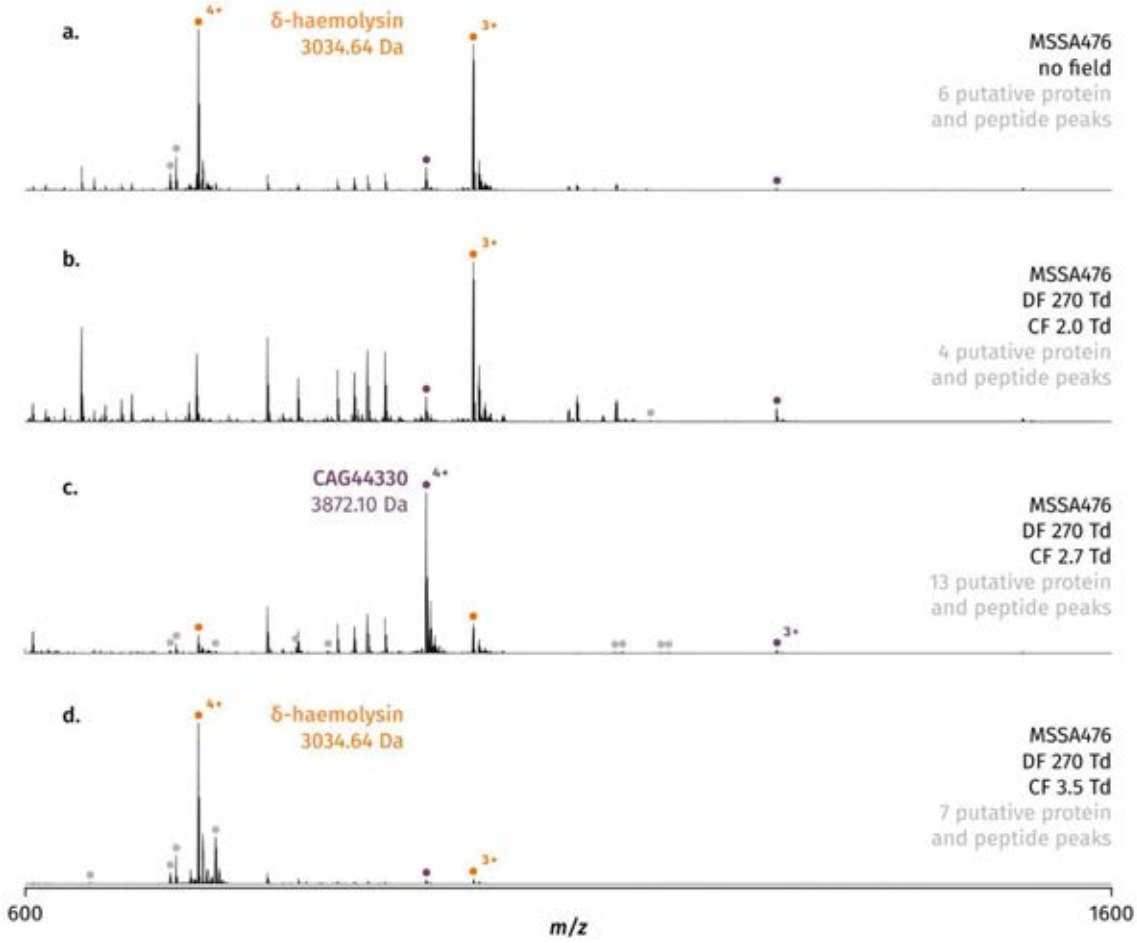


Figure 5.9. Representative LESA-FAIMS-MS single scan spectra of *S. aureus* MSSA476 acquired during a one-dimensional FAIMS sweep experiment at DF 270 Td. The compensation field (CF) ramp ranged from CF 0.0 Td to CF 4.0 Td over 240 seconds. Putative protein and peptide peaks (charge states 3+ and above) are marked by dots.

5.3.4 LC-MS

Top-down LC-MS/MS experiments were performed on *E. coli* K-12. The samples consisted of LESA extracts (using the standard LESA sampling parameters for *E. coli*), as well as colony lysates generated through a protocol adapted from literature.[297] The experiment was designed to compare the LC-MS/MS data not only to the separation provided by FAIMS, but also to approximate the protein fraction extracted by LESA versus the total protein available from lysed *E. coli* cells.

Optimisation of the liquid chromatography parameters was performed on standard samples of cytochrome c and myoglobin in <5% acetonitrile and 0.1% formic acid. A PepMap C4 column (300 Å pore size) was acquired and paired with a filter flanked by two IDxl 30 µm × 10 cm trapping columns. Parameters optimised for use with a C18 column were used as a start point. The 30 minute gradient (3.2%-60% B) followed by a wash step was retained. The maximum running pressure was reduced to 300 bar to account for the lower tolerance of the C4 column. The mobile phases used were 99.9% water, 0.1% formic acid (A) and 99.9% acetonitrile, 0.1% formic acid and were not altered from the start-point protocol.

Full scan acquisition was performed in the Orbitrap analyser at a resolution of 60,000 at *m/z* 400. The top five peaks were selected for fragmentation in the orbitrap, also at a resolution of 60,000 at *m/z* 400; the value was a compromise between the high resolution needed for successful intact protein identification by use of the ProSightPD node in Proteome Discoverer, and the data acquisition time. The minimum ion count for precursor selection was increased to 5000. **Figure 5.10** shows two sample chromatograms and corresponding mass spectra acquired at final settings following initial optimisation. The runs confirmed the correct functioning of the LC-MS system and the transmission of proteins, and the parameters were deemed sufficiently satisfactory to warrant further testing on more complex samples. Thus, the next batch consisted of LESA and whole colony lysate samples of *E. coli* K-12.

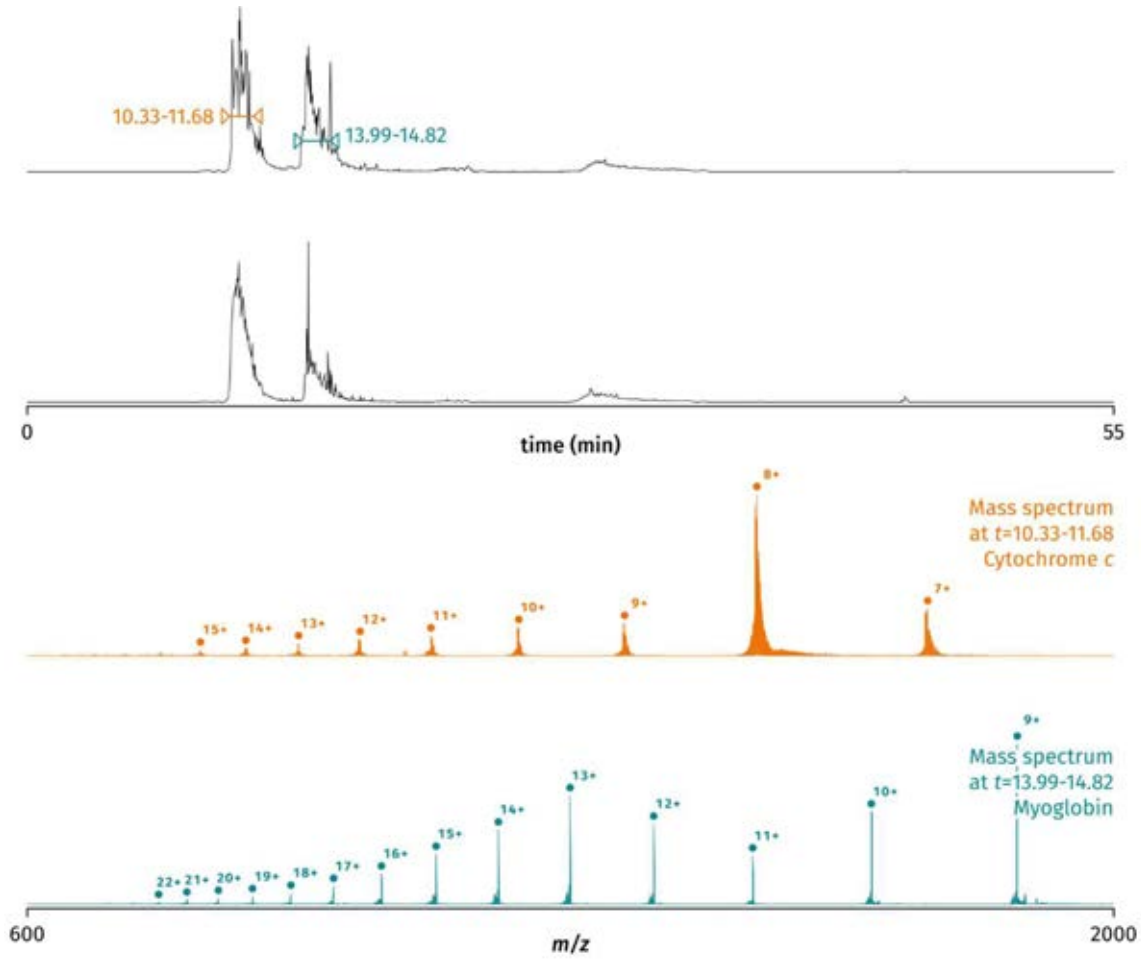


Figure 5.10. Preliminary LC-MS data generated from a standard mixture of cytochrome c and myoglobin.

Figure 5.11 shows a representative chromatogram and mass spectrum generated from LESA extracts of *E. coli* K-12; **Figure 5.12** shows the same data for whole-colony lysates. The comparative LC-MS/MS data from the *E. coli* LESA extracts yielded predominantly poor to acceptable quality tandem mass spectra of HU-α (ions centred at *m/z* 734.33 13+, 795.69 12+, 796.85 12+ acquired twice in the same LC-MS run, 795.60 12+ and 795.62 12+) and HU-β (*m/z* 771.00 12+ corresponding to the sodiated ion, acquired three times, 771.34 12+ sodiated ion, 769.75 12+, 769.67 12+ and 769.59 12+ acquired twice), at the expense of other components of the mixture, as well as BhsA, consistent with the storage of the samples at 4 ℃ (see Chapter 4 section 4.3.5; tandem mass spectra acquired from *m/z* 1089.56 6+ and 934.05 7+). This

observation was consistent with the high abundance of the HU complex in *Escherichia*, as well as with the previous observation of these proteins by LESA-MS. It was, however, also indicative of suboptimal precursor selection and exclusion settings as tandem mass spectra of the same proteins, frequently different peaks belonging to the same isotopic distribution, were acquired multiple times. No other proteins were detected in LC-MS data of LESA extracts.

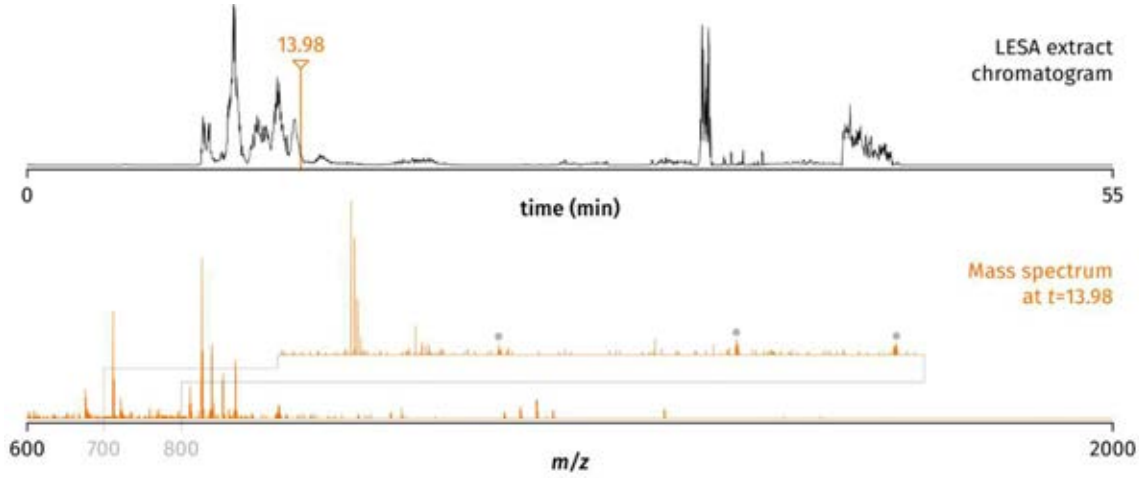


Figure 5.11. Chromatogram and the most protein-rich mass spectrum generated by LC-MS of a LESA extract of *E. coli* K-12. Protein peaks are marked in grey.

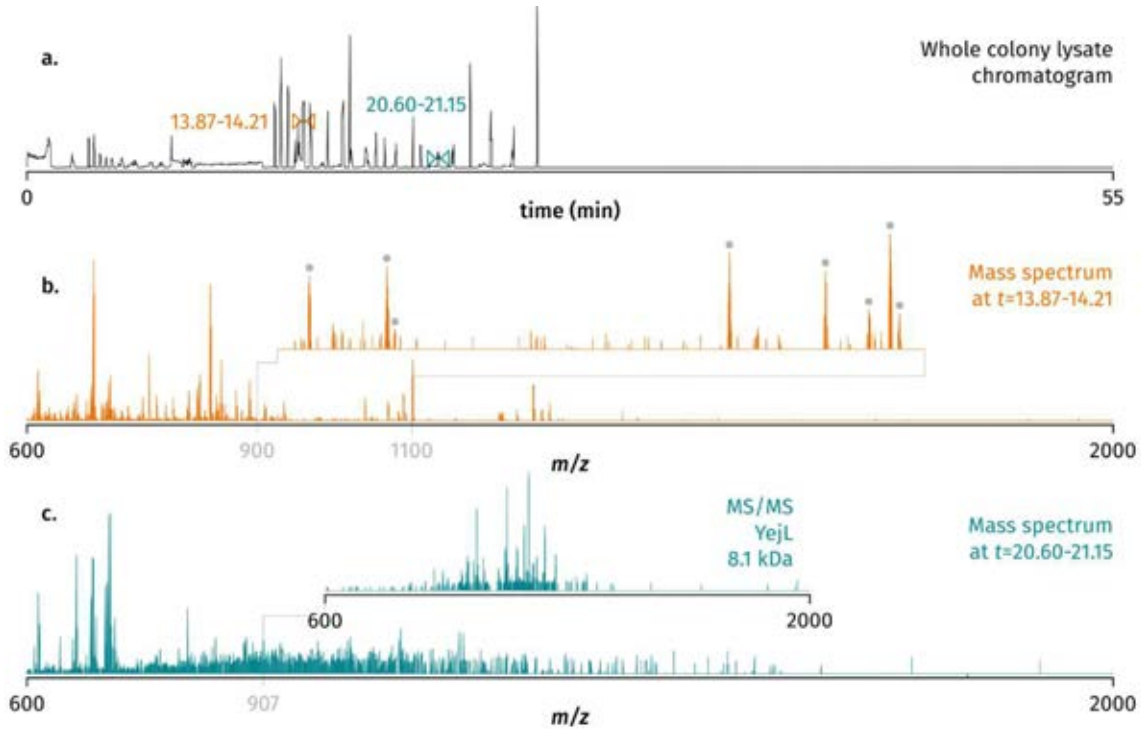


Figure 5.12. Whole colony lysate LC-MS. **a**. Chromatogram. **b**. Representative mass spectrum acquired at *t* = 13.87-14.21. The protein-containing region (m/z 900-1100) is enlarged in the inset; protein peaks are marked with dots. **c**. Representative mass spectrum acquired at *t* = 20.60-21.15. Tandem mass spectrum of YejL acquired during the run is shown as an inset.

Four peaks corresponding to a single protein (*m/z* 941.13 13+, 1019.39 12+, 1111.88 11+, 1222.87 10+, MW_{obs} = 12216.64 Da) were detected in full scan LC mass spectra generated from the offline whole-colony lysates, although none were selected for fragmentation. One stress response protein not detected by LESA, YejL (*m/z* 907.05 9+, MW_{obs} = 8154.36 Da), was identified with high confidence in the second replicate (**Figure 5.12c**); the ProSightPC identification parameters and peak assignments are supplied in **Appendix 3.3.1.7**. In addition, multiple tandem mass spectra of HU-α and HU-β in charge states 10+ to 13+ were acquired. Based on existing literature, however, the identification of over 150 proteins could be expected, indicating that extensive further optimisation of the data acquisition and processing parameters was required. This was corroborated by the detection of multiple,

clearly delineated protein peaks which were not selected for fragmentation by the software, indicating issues with the isolation process.

5.4 Conclusions

The coupling of FAIMS to the LESA MS workflow was shown here to resolve two of the three challenges outlined in section 5.1 as well as Chapter 4, section 4.4. Highly abundant background peaks in the *m/z* 600-800 region of *E. coli* mass spectra were successfully removed by use of FAIMS, revealing additional protein peaks. Three examples of compound mass spectra were resolved in *P. aeruginosa*.

The removal of δ-haemolysin from mass spectra of *S. aureus* was not successful. The wide band of transmission for δ-haemolysin (CF 0.5 Td to CF 4.0 Td), even at a high dispersion field (270 Td), precluded its complete elimination from mass spectra of this species, although a reduction of its abundance was briefly achieved. The solution to this challenge could be attempted by use of a more advanced FAIMS device, such as the novel cylindrical FAIMS introduced by Thermo Scientific shortly prior to the conclusion of this project. Alternatively, a different ion mobility technique such as drift tube IMS or travelling wave IMS could be used, where the size of the analyte influences mobility to a greater degree than it does for FAIMS; thus, small toxic peptides could potentially be separated from larger proteins of interest.

The data generated by use of LC-MS were of insufficient quality to provide a fair comparison with the performance of FAIMS. This outcome highlighted the complexity of the LC-MS based approach compared to FAIMS. Multiple parameters, including sample preparation, the choice of column, flow rates, mobile phases, gradient, as well as mass spectrometry settings including resolution, number of MS/MS scans and dynamic exclusion parameters require

extensive optimisation to increase the number of observed proteins and the identification capability. Optimisation of FAIMS in principle only requires the tuning of the two voltages corresponding to the DF and the CF; the number of proteins detected is, however, lower than that expected of a fully optimised LC-MS/MS protocol and the acquisition of tandem mass spectrometry data cannot be automated at this point. FAIMS separation is achieved at much shorter timescales (milliseconds) than the fastest available LC-MS protocols (minutes), allowing for more rapid data acquisition. Thus, the choice of separation technique for a particular application would depend on the priorities of a given experiment—at the current state of both technologies, LC-MS/MS based bacterial protein identification should yield a larger range of detected and identified proteins compared to FAIMS, at the cost of significant time investment both for optimisation and for the separation step itself.

CHAPTER 6

Lysis of microbial cells by application of electricity

6.1 Introduction

As described in Chapter 4, section 4.3.3, the identification of proteins from yeast by LESA proved highly challenging. This finding, as well as the potential progression of the project towards native LESA sampling of proteins directly from microbial colonies, prompted an investigation into alternative methods of cell lysis in addition to or instead of solvent-based approaches. Mechanical lysis, whilst technically the simplest and widely adopted method for lysis of cells in suspension,[318] could not be adequately reproduced for colonies sampled directly from agar due to the soft supporting material. Cold atmospheric plasma (CAP) lysis, which induces the formation of pores in membranes through the oxidation of phospholipids[319] and the destruction of bacterial cell walls through disintegration of peptidoglycan,[320] has been demonstrated as an effective method for cell ablation in medical applications[321-323] and could potentially be coupled directly to LESA. It would, however, be too technically challenging to recreate for this project without specialised engineering input. Electrical lysis, or electroporation, was chosen due to its relative technical simplicity, compatibility with the LESA process, the lack of additional chemical background (unlike for enzymatic lysis) and minimal thermal effects. The mechanism of electroporation is described in Chapter 1, section 1.4.1. Briefly, the application of an electric field to live cells induces the formation of pores in cellular membranes. Beyond an energy threshold specific to the given membrane, the disruption of the membrane becomes irreversible and cell lysis occurs. Whilst transformation/cloning applications require a finely controlled delivery of current to retain

cell viability, electroporation for the purpose of protein extraction would utilise higher voltage and current to induce complete, irreversible lysis.

As no commercial platform for the electroporation of microbial colonies directly on agar plates is currently available, a home-built apparatus was constructed. This chapter describes the optimisation of individual, major components of the apparatus and the final design of the device. The device was subsequently used to lyse yeast colonies for the purpose of top-down identification of proteins by LESA MS. Three species of yeast were investigated: *Saccharomyces cerevisiae* BY4741 (baker's yeast), *Candida glabrata* and *Cryptococcus neoformans* H99. *C. glabrata* and *C. neoformans* were previously used for LESA sampling in the absence of electrical lysis, with unsatisfactory results (see Chapter 4, section 4.3.3). *S. cerevisiae* was selected as a safe BSL1 organism which could be used to collect preliminary data in an unsecured workshop environment during the construction of the device. *S. cerevisiae* was also used for the optimisation of electroporation parameters, including pulse number, pulse length and voltage (section 6.3.3.1).

6.2 Experimental

6.2.1 Preparation of yeast samples

Colonies of *Saccharomyces cerevisiae* BY4741 (baker's yeast), *Candida glabrata* and *Cryptococcus neoformans* H99 were prepared as described in Chapter 2, section 2.2.1.3. Any deviations from the parameters described in Chapter 2, section 2.2.1.3 are highlighted in the text.

6.2.2 Electroporation

Electroporation was carried out by use of the home-built cell lysis apparatus. Sample agar plates were placed underneath the electrode assembly. Electrodes were lowered into the

colony of interest such that the electrode tips would not touch the agar media underneath the colony. A default voltage of 3000 V was used. Pulse width was set to 20 µs. Pulse delay was 200 µs. Pulse number was set to 1; the repetition number was set to 15. Repetition delay was 1 s. Capacitor discharge and voltage peak shape were monitored by use of an externally connected oscilloscope (Tektronix, Beaverton, USA).

The general principle of operation, including the selection of adjustable parameters, was identical for both the prototype device and the final iteration of the device with the exception of the safety key, which was absent in the prototype.

6.2.3 Optimisation of electroporation parameters

Colonies of *S. cerevisiae* grown at optimum temperature conditions (30 ℃) were used to test the optimum electroporation parameters for the liberation of protein from yeast colonies. Three settings were varied: pulse number, pulse length and voltage. For each variable to be investigated, three plates containing seven *S. cerevisiae* colonies each were prepared as described in Chapter 2, section 2.2.1.3. Each batch of three plates was consecutively subjected to electroporation, following which it was stored at 4 ℃ until LESA sampling. Electroporation and sampling were performed on the same day.

To investigate the effect of varying the number of pulses, 0, 2, 4, 5, 8, 10 and 15 pulses were tested. Pulse length (20 µs) and voltage (3000 V) were kept constant. To investigate the effect of varying the duration of the pulse, 0, 2, 4, 8, 12, 16 and 20 µs pulses were used. Number of pulses (10) and voltage (3000 V) were kept constant. To investigate the effect of varying the voltage, 0, 500, 1000, 1500, 2000, 2500 and 3000 V were used. Number of pulses (10) and pulse length (20 µs) were kept constant.

6.2.4 Sampling

Sampling of yeast colonies was carried out according to the protocols summarised in Chapter 2, section 2.2.3. The sampling tray of the TriVersa NanoMate was cooled to approximately 7 ℃, the lowest temperature attainable by the cooling system in ambient temperature. A 40:60:1 acetonitrile-water-formic acid solvent was used to minimise chemical disruption of cells while maintaining efficient protein extraction. FAIMS was used where noted in the text in static field mode (DF 270 Td, CF 3.0 Td) to remove the signal generated by cell wall debris and reveal protein peaks earlier during data acquisition.

6.2.5 Protein identification

Protein fragmentation and identification was carried out as described in Chapter 2, sections 2.2.3 and 2.2.5, respectively. The biomarker search was only used for truncated protein forms, where identification by the absolute mass search failed. All fragment assignments were generated automatically by ProSightPC 4.1 alpha using deconvoluted data and manually verified. The acceptance threshold for fragment assignment was increased to 15 ppm due to instrument calibration issues.

6.3 Results and discussion

6.3.1 Development of the electroporation platform

6.3.1.1 General layout of the apparatus

As outlined in Chapter 1, section 1.5.3.1, a generic electroporator consists of a high voltage power supply, a bank of capacitors, a switching mechanism and a vessel for delivering the pulses to the sample. The optimisation of each of these components is described below. For a description of the mechanism of operation and the parameters specific to each component, see Chapter 1, section 1.5.

The circuit design, the construction of the device (with the exception of parts of the electrode assembly) and software programming were carried out by Andrew Tanner at the University of Birmingham, Biosciences engineering workshop. The final build, shown in **Figure 6.1**, was encased in an aluminium box separating the user interface from the high-voltage components. The front panel contained a dual banana-plug port used to connect the electrode assembly to the device, a safety key switch, a capacitor charge button connected by a short lead, and the touchscreen interface. The rear of the device contained a power switch and two sockets used to connect an optional oscilloscope for the purpose of monitoring capacitor charge and pulse shape.

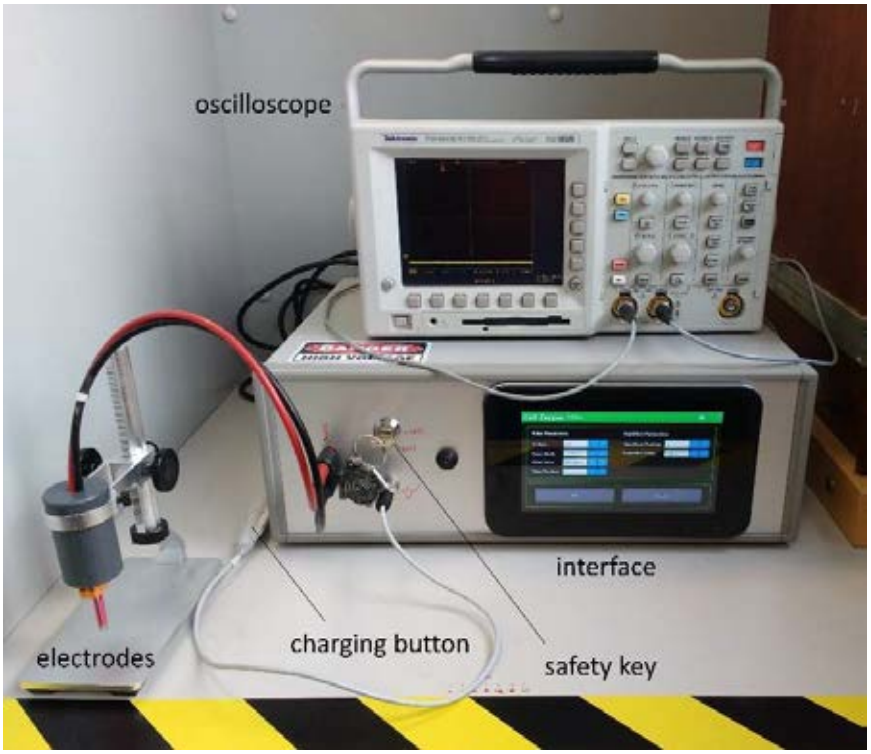


Figure 6.1. Photograph of the completed electroporator with an optional, attached oscilloscope.

6.3.1.2 High voltage power supply

The working output voltage for this design was set at 3 kV, with the aim of generating an electrical field of 15 kV/cm, sufficient for the complete lysis of yeast colonies, as well as *E.*

coli, as determined from literature.[267] The maximum achievable parameters were primarily limited by the poor availability of parts for higher voltage designs at a reasonable cost and in an acceptable time frame. A general purpose 15 kV high voltage supply (Spellman, USA) with an output limited to 3 kV was used here. A Bio-Rad 3000Xi 3 kV electrophoresis supply was also investigated as a potential voltage source but proved too complex to integrate into the apparatus due to the inclusion of multiple poorly documented, proprietary safety and inhibition circuits.

6.3.1.3 Charge storage

A bank of capacitors totalling 0.7 μF in capacitance was used in the first build, referred to hereinafter as the prototype. The capacitance was sufficient for the generation of approximately square wave pulses exhibiting a slight drop in voltage (**Figure 6.2**). All data described in section 6.3.2 were acquired using the 0.7 μF capacitor bank and the 41-03 fast high-voltage switch (see section 6.3.1.4).

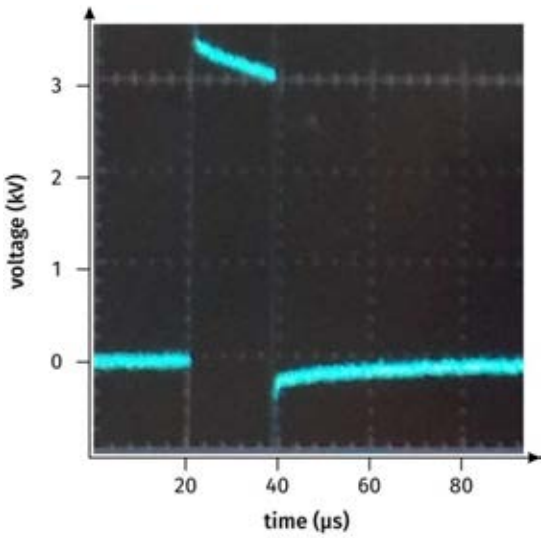


Figure 6.2. Photograph of an oscilloscope trace corresponding to a test pulse delivered without a load (no sample). The parameters used were 20 μs pulse length and 3000 V.

The capacitance was briefly increased to 2.5 μF in order to mitigate the voltage drop seen in the shape of the pulses generated by the 0.7 μF bank. The larger amount of stored energy

should take a longer time to discharge, thus slowing the decrease in voltage and generating pulses more closely approximating the ideal square wave. It would also mitigate any drop in maximum voltage which could occur between pulses due to capacitor depletion. In practice, however, it also led to the significant aggravation of overload conditions should a failure, such as a short circuit, occur somewhere in the apparatus. A short circuit subsequently occurred along the surface of a test agar plate during the first test of the new capacitor bank, which led to a severe electrical overload and consequent destruction of the high-voltage switch. As the voltage was strictly limited to 3 kV by the power supply, it was deduced that the switch was instead 'welded' in a permanently closed position by a current exceeding 30 A. The capacity of the capacitor bank was reduced again to 0.5 μF in the final build, sacrificing pulse shape for reliability; large current-limiting protection resistors (300 Ω) were also included in the design either side of the replacement switch to reduce the risk of failure. It was noted that the voltage of pulses delivered into sample colonies was significantly lower than that of pulses delivered in air, indicating that part of the stored energy was redirected through the protector resistors instead of the sample with each pulse (**Figure 6.3**). The change did not detectably impact the lysis capability of the apparatus.

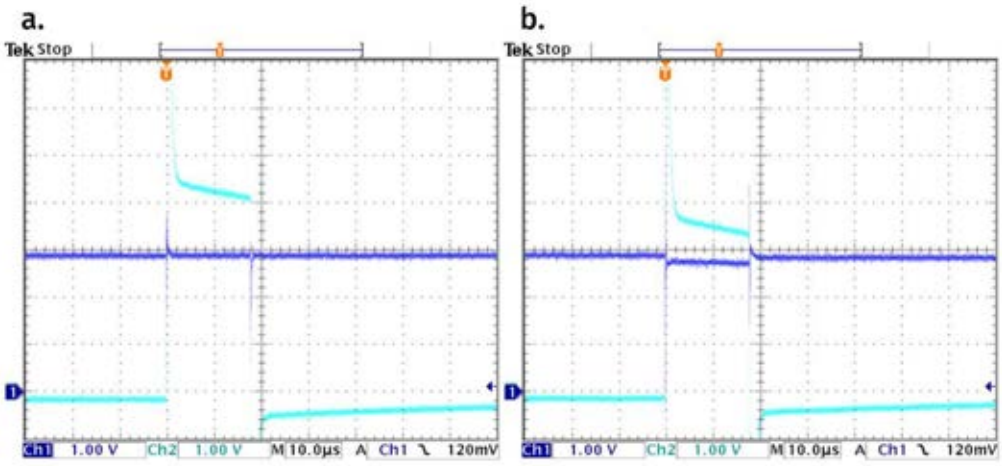


Figure 6.3. Oscilloscope measurement of the pulse shape following the delivery of a 20 µs pulse at 3000 V **a**. in air (load absent) and **b**. into a colony of *S. cerevisiae*, by use of the finalised device. Each major grid division for both channels corresponds to 1000 V on the *x*-axis and 10 µs on the *y*-axis. The dark blue trace (channel 1) measured the charge stored in the capacitors. The cyan trace (channel 2) measured the voltage spike at two probe points accessible through the rear plate of the device, corresponding to the voltage spike induced between the electrodes. The voltage measurement for channel 2 is approximately 1.4 kV higher than the true value due to induction effects within the device (the true value cannot exceed 3 kV).

6.3.1.4 High voltage switching

As described in Chapter 1 section 1.5.3.3, the generation of square wave pulses at the microsecond time scale requires the use of a transistor switch. A Behlke 41-03 single MOSFET switch was used for the prototype. The switch would assume the closed circuit configuration for the duration of the pulse, following which it would open the circuit. As the flow of electrons was simply interrupted following the delivery of the pulse, the capacitors would remain charged and require discharging through a separate safety discharge circuit or by manually shorting the capacitors to ground.

As described above, the 41-03 switch was destroyed during testing. It was subsequently replaced by a 41-03-GSM dual ('push-pull') MOSFET switch. Unlike the 41-03, which simply opens or closes the circuit, the 41-03-GSM switch diverts current through one of the internal transistors at a time. Thus, once a pulse has been fired, the switch automatically directs the

remaining energy of the capacitors to ground, allowing simpler and safer operation of the device.

6.3.1.5 Electrode design

The design of the electrodes had to fulfil several constraints. First, due to the cost and availability of the fast switches and other high-voltage hardware, the output voltage of the device was limited to 3 kV. Thus, in order for the strength of the electric field to reach values capable of rupturing colonies of interest (approximately 15 kV/cm), the distance between the electrodes had to be set to 2 mm or below.

Several geometries were proposed for the probe end. The use of NanoMate conductive pipette tips in conjunction with an electrode surface placed under the agar was initially explored; however, the lack of available specifications for the pipette tips combined with their high resistance at the very end of the tip rendered their use for preliminary experiments impractical. Another proposed solution involved a coaxial setup consisting of an inner needle electrode and an outer ring/pipe electrode. This arrangement, however, did not provide any subsequent adjustment of the inter-electrode distance and was additionally challenging to construct. Thus, the simplest setup, consisting of two planar electrodes, was adopted. Two stainless steel micro-spatulas were split in half and the flat ends were mounted in a plastic socket allowing access to high-voltage wiring, as well as easy removal of the probe assembly for cleaning and sterilisation (**Figure 6.4**). The length of the electrodes was covered in separate isolating sleeves to minimise the risk of arcing in air. The entire probe setup was secured with silicone tubing to an adjustable stand designed for USB microscopes, which provided fine control over the probe height.

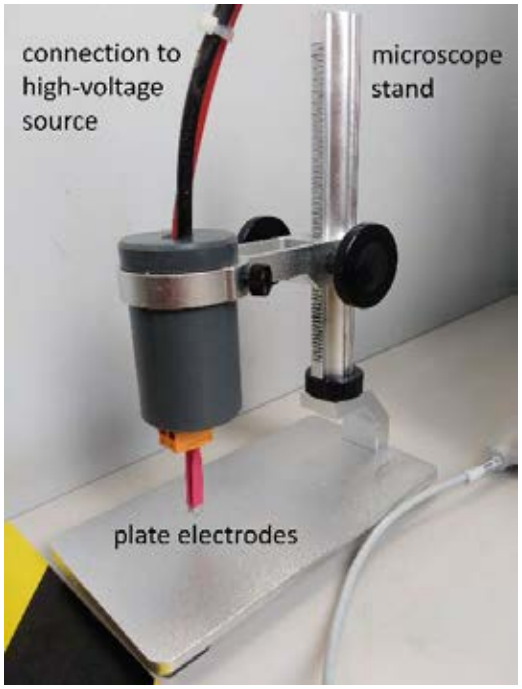


Figure 6.4. Photograph of the finalised electrode assembly.

6.3.1.6 Control software

The device was operated by a simple graphical user interface. Adjustable variables included voltage (up to 3250 V, limited by the maximum operating voltage of the MOSFET switch), pulse width or duration, pulse delay, pulse number and pulse spacing. A repetition function was included by which single pulses or blocks of pulses could be repeated on a longer (seconds) timescale. The function labelled 'pulse number' in the interface was not used in the experiments described this chapter; the number of pulses was instead set by use of the repetition function at the default 1 second interval. The original software used to control the prototype was coded in MATLAB and operated by use of a laptop connected to the device through an optically isolated USB port; this arrangement was later replaced by the Cell Zapper v1.0 beta software (programmed in C# by Andrew Tanner) installed on and operated by a Raspberry Pi-based, dedicated touchscreen interface running the Windows 10 IOT operating system. The final interface is pictured in **Figure 6.5**. The 'Arm' button was implemented as a safety feature to only allow for the preparation of a single pulse (or burst

of pulses) at a time, following which the device would not be capable of generating high voltages until the 'Arm' button was pressed again. The 'Execute' button (executing the code) was used to deliver the pulses.

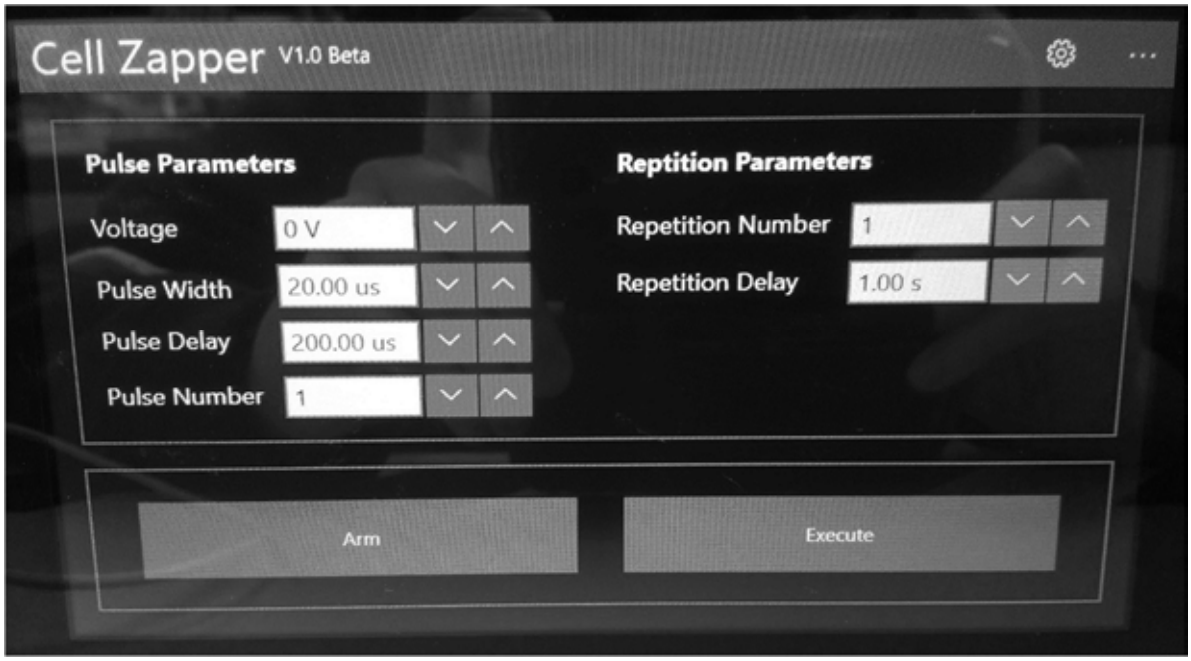


Figure 6.5. Electroporator software interface.

6.3.2 Preliminary results and general observations

According to the literature values,[267] yeast cells (including *Saccharomyces cerevisiae* and *Candida* spp.) were the easiest test subjects for electroporation due to their low resistance to high voltage pulses. A wild-type strain of *S. cerevisiae* was therefore sourced to test the performance of the prototype; as the device was located in the workshop, a BSL1 organism had to be used for testing. The first batch of *S. cerevisiae* colonies used here was incubated at a sub-optimal temperature (37 ℃). All the results described in this section were generated by use of the prototype device (single MOSFET switch).

During initial testing at settings derived from literature (10 pulses, 20 μs each with 1 second spacing, at 3 kV), electrical discharge (arc) in air above the colonies, accompanied by a

characteristic sound, was observed for the first one to four pulses (**Figure 6.6**). The arcs were suspected to result from the initial ionisation of gaseous components emitted by the colony, as they were not observed when electrical pulses at the same parameters were delivered directly to plain agar media. Once the gaseous components above the colony were depleted, the delivery of pulses would proceed without any visual or auditory indication of arcing.

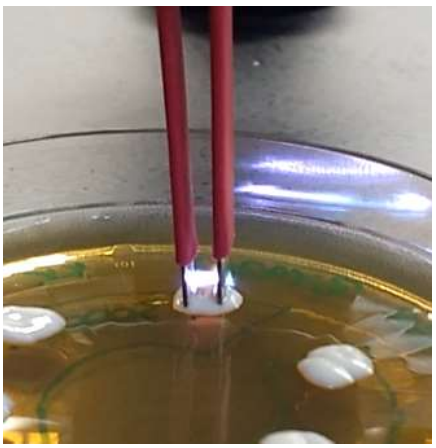


Figure 6.6. Still video capture of an electrical arc occurring above the surface of an electroporated colony of *S. cerevisiae*.

Three colonies were subsequently subjected to a series of 15 pulses, 20 μs each with 1 second spacing, at 3 kV. Three control colonies grown on the same plate were kept intact. The number of pulses was increased to 15 in order to account for the pulses lost to arcing, so that a minimum of 10 correctly executed pulses would be reliably delivered to the colony. During pulse delivery, it was noted that the colony material became visibly liquefied, indicating severe disruption of colony structure (**Figure 6.7**).

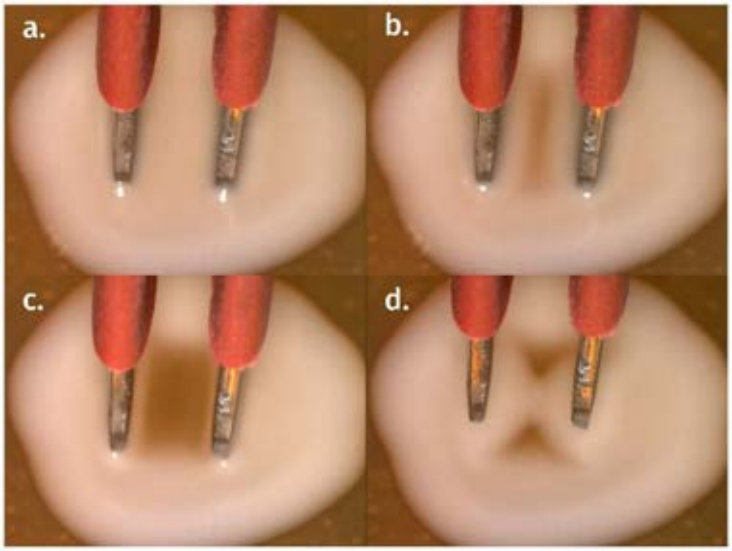


Figure 6.7. Still video capture of the liquefication phenomenon. **a**. Electrodes inserted into the colony prior to the delivery of electric pulses. **b**. Immediately after the delivery of the electric pulses has been completed. **c**. During the withdrawal of the electrodes; the liquefied colony matter is pulled upward and away from the colony centre by surface tension. **d**. Colony matter settling after the contact with electrode tips has been broken.

All three sampling attempts performed on electroporated cells revealed the presence of protein. A representative mass spectrum acquired from an electroporated colony of *S. cerevisiae* is shown in **Figure 6.8;** a mass spectrum acquired from a control colony is provided for comparison. Two protein peaks (*m/z* 774.11 15+, 799.08 11+) were selected for fragmentation and were subsequently identified as 12 kDa heat shock protein HSP12 (MW_{obs} = 11596.48 Da, missing the initiator methionine and containing an acetylation on the N-terminal serine; see **Appendix 3.4.1.1** for fragment assignment) and water deprivation stress response protein SIP18 (MW_{obs} = 8778.76 Da, missing N-terminal methionine, serine acetylated; **Appendix 3.4.1.5**), respectively (**Figure 6.9**). This was consistent with the sub-optimal temperature of incubation for the yeast colonies (37, as opposed to 30 ℃). Crucially, no protein peaks were observed in mass spectra acquired from control colonies, supporting the notion that lysis observed in the electroporated cells was attributable to the electric pulses rather than solvent disruption. Peak assignments for all proteins identified in this chapter are listed in **Appendix 3.4**.

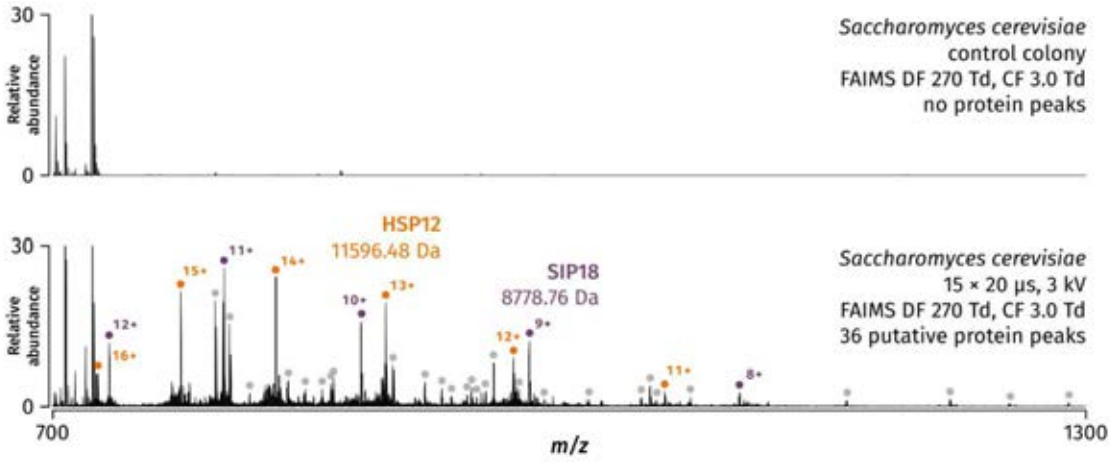


Figure 6.8. Comparison of mass spectra acquired from *S. cerevisiae* using the prototype electroporator. Both mass spectra were magnified to 30% relative abundance due to the presence of highly abundant, singly charged peaks at *m/z* 707.21 and 723.19.

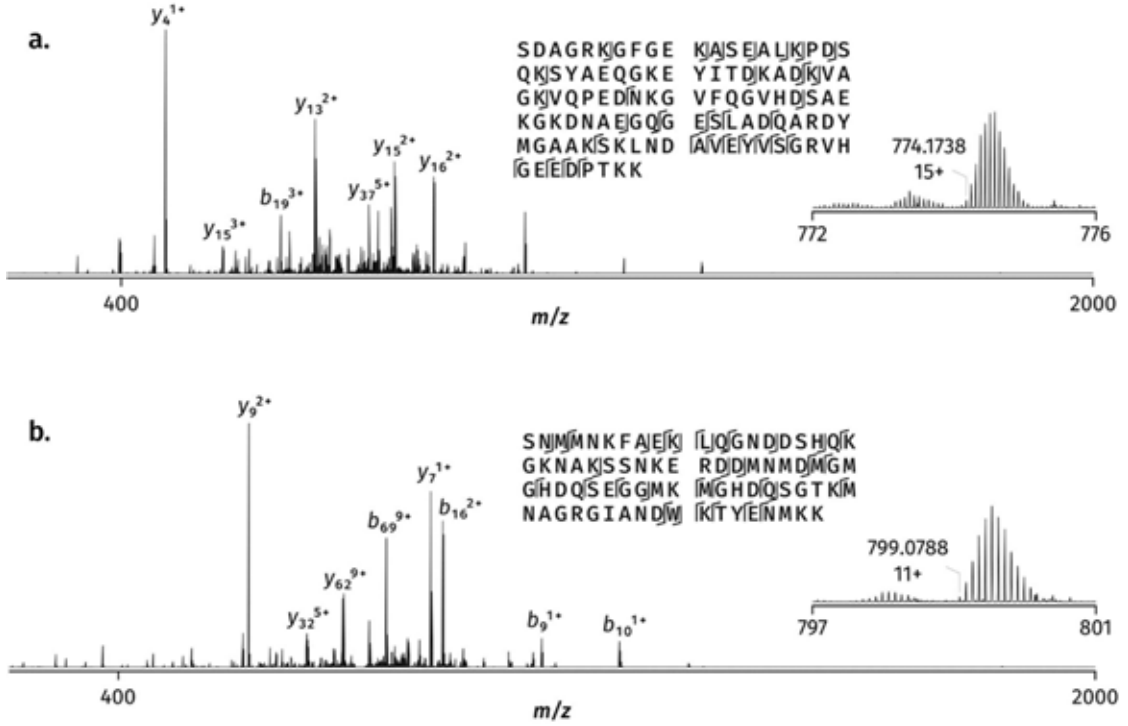


Figure 6.9. MS/MS spectra of the first two proteins identified by LESA tandem mass spectrometry in yeast. **a**. 12 kDa heat shock protein HSP12, isolated at *m/z* 774.11 (15+), MW_{obs} = 11596.48 Da. **b**. Water deprivation stress response protein SIP18, isolated at *m/z* 799.08 (11+), MW_{obs} = 8778.76 Da.

6.3.3 Results generated using the final design

6.3.3.1 Optimisation of parameters

6.3.3.1.1 Variable number of pulses

The optimisation of electroporation parameters for the liberation of protein from colonies of *S. cerevisiae* was carried out as described in section 6.2.3. The purpose of optimisation was to elucidate parameters sufficient for protein extraction without excess energy input, which could otherwise lead to undesirable heating and protein breakdown. Such optimisation was particularly crucial for future work involving the extraction of natively folded proteins, as any excessively harsh treatment could disturb or outright destroy their conformation. Results generated by varying the number of pulses delivered to the colonies are summarised in **Figure 6.10**. Pulse length and voltage were constant (20 μs and 3000 V, respectively). The presence or absence of protein was determined by verifying the presence of two prominent peaks corresponding to HSP12 in the *m/z* 800-900 region (*m/z* 829.78 14+, 893.53 13+), as well as *m/z* 837.65 6+ subsequently identified as a C-terminal fragment of elongation factor 1-α (see section 6.3.3.2.1). Of the three replica plates, one yielded clear protein peaks from the three colonies subjected to 8, 10 and 15 pulses. Protein oxidation was not observed. Protein fragments (e.g. the C-terminal fragment of elongation factor 1-α) were observed, indicating possible decomposition due to electrical treatment; this effect was not further investigated due to time constraints. The remaining two plates yielded exclusively mass spectra dominated by small molecules; no protein peaks in the *m/z* 800-900 region or elsewhere were observed.

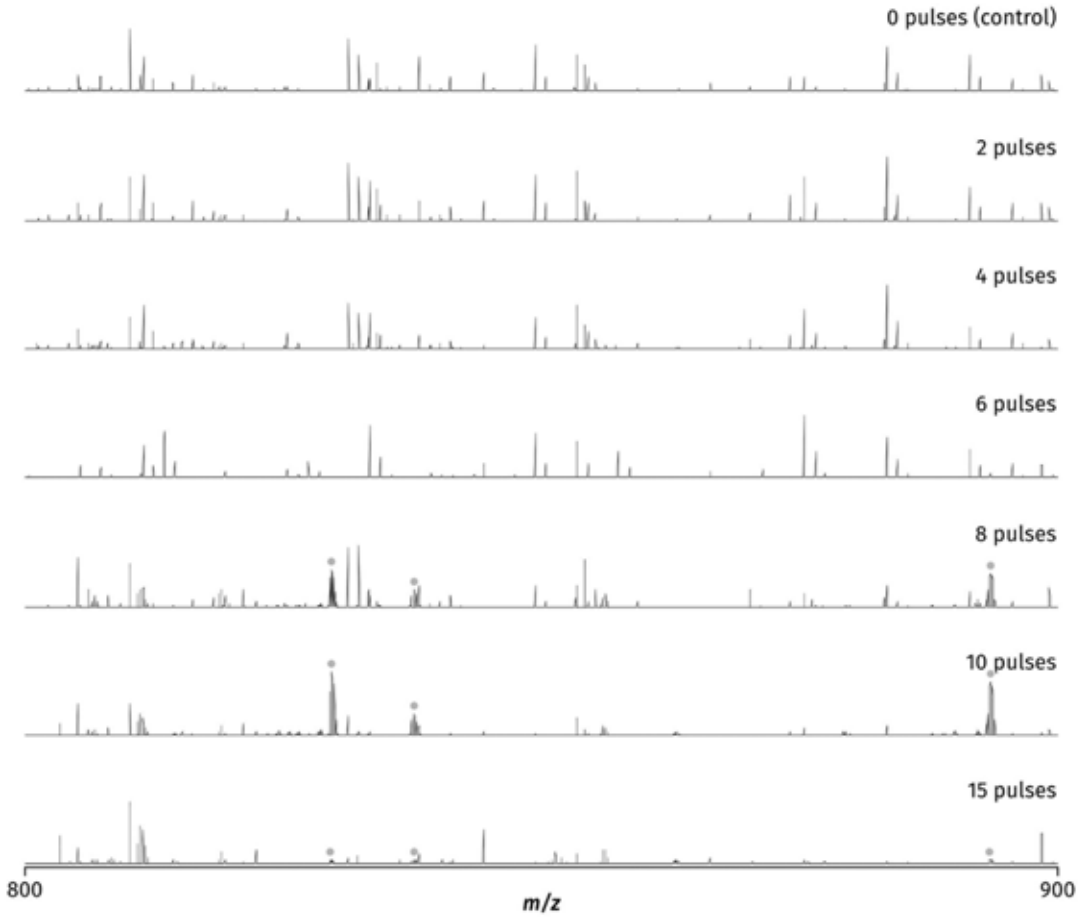


Figure 6.10. Effect of varying the number of pulses delivered to the colony. Monitored protein peaks (*m/z* 829.78 14+, 837.65 6+ and 893.53 13+ from left to right) are marked by grey dots. Pulse duration (20 µs) and voltage (3000 V) were kept constant. All mass spectra were acquired from colonies grown on the same plate.

6.3.3.1.2 Variable length of pulses

Results generated by varying the length of pulses delivered to the colonies are shown in **Figure 6.11**. Pulse number and voltage were constant (10 and 3000 V, respectively). As was the case for the variable pulse number data set, data generated from only one plate contained a detectable protein signal. Protein peaks in the *m/z* 800-900 reporter region, corresponding to the HSP12 protein (*m/z* 829.78 14+, 893.53 13+) were observed from colonies subjected to pulses of 16 µs and 20 µs. All other LESA mass spectra contained peaks corresponding to small molecules only.

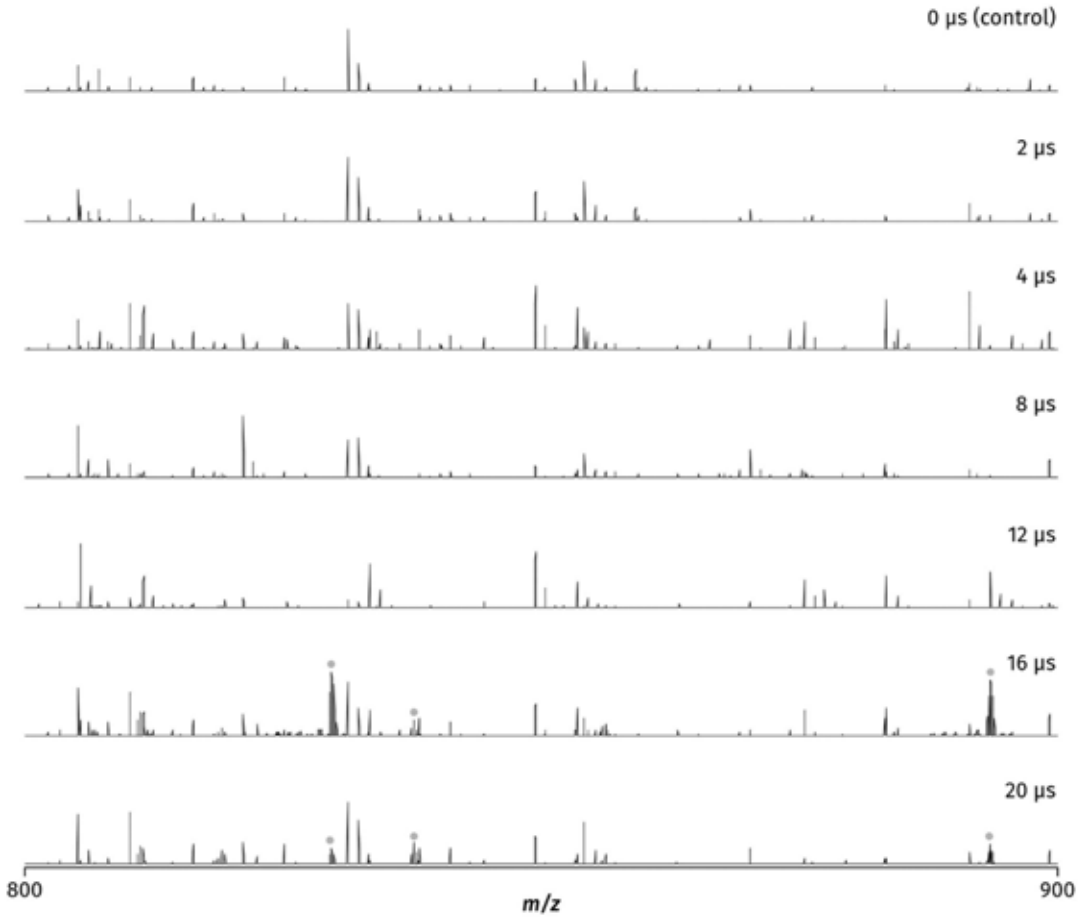


Figure 6.11. Effect of varying the duration of pulses delivered to the colony. Monitored protein peaks (*m/z* 829.78 14+, 837.65 6+ and 893.53 13+ from left to right) are marked by grey dots. Number of pulses (10) and voltage (3000 V) were kept constant. All mass spectra were acquired from colonies grown on the same plate.

6.3.3.1.3 Variable voltage

Results generated by varying the voltage of pulses delivered to the colonies are shown in **Figure 6.12**. Pulse number and length were constant (10 and 20 µs, respectively). All three plates yielded mass spectra containing protein peaks. For two out of three plates, LESA mass spectra containing protein peaks in the *m/z* 800-900 reporter region were acquired from colonies subjected to 2500 and 3000 V pulses. For the third plate, LESA mass spectra containing protein peaks were acquired from colonies subjected to 1000, 2000, 2500 and 3000 V pulses; no protein peaks were detected in the colony subjected to 1500 V pulses.

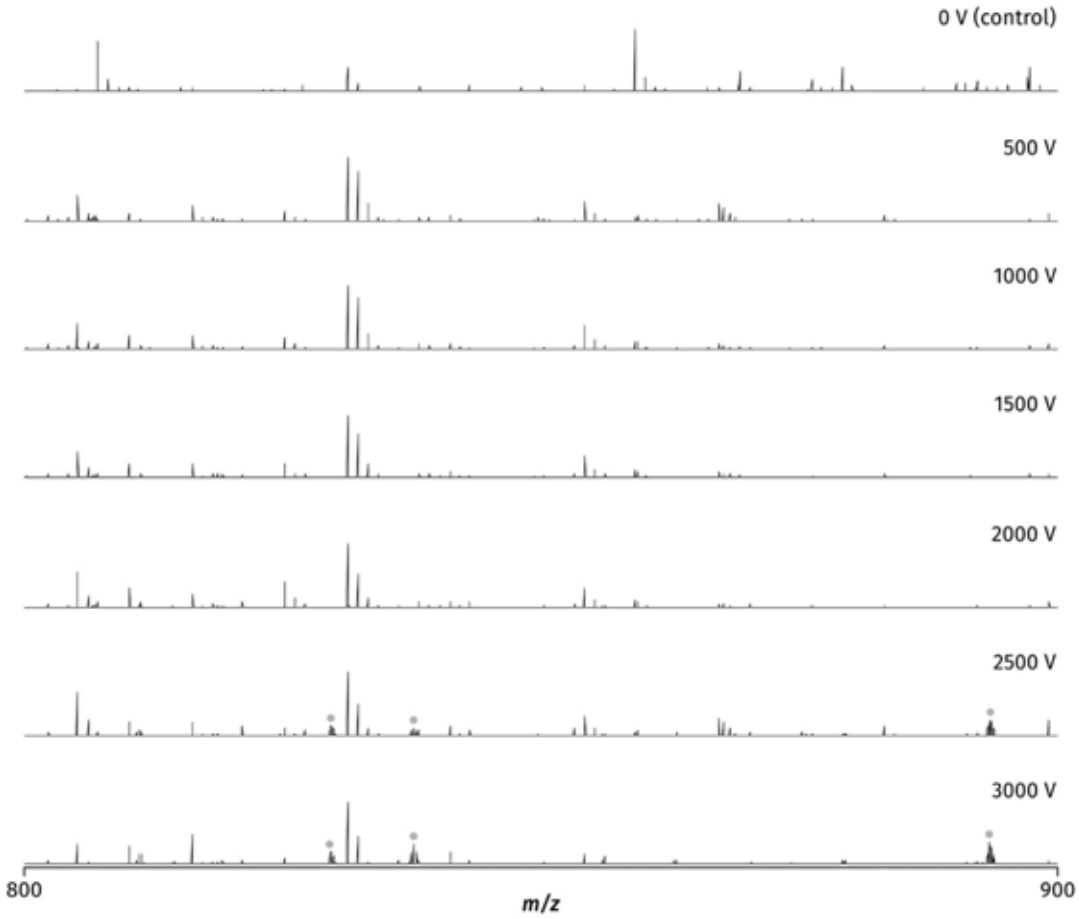


Figure 6.12. Effect of varying the voltage of pulses delivered to the colony. Monitored protein peaks (*m/z* 829.78 14+, 837.65 6+ and 893.53 13+ from left to right) are marked by grey dots. Number of pulses (10) and duration of pulses (20 μs) were kept constant. All mass spectra were acquired from colonies grown on the same plate.

The results indicate significant variability in the current workflow. Both the electroporation step and the LESA sampling are subject to variability associated with manual height adjustment (electrode tips for electroporation, pipette tip for LESA). For electroporation, the selection of the correct height would determine whether the electric pulse is forced to traverse the colony matter, or whether it reaches and travels through the highly conductive agar media instead; as the electrodes were mounted on a microscope stand, fine height adjustment was possible. The interplay of this variability with the known variability of the LESA sampling process itself complicated data analysis in this case; however, in all cases where protein was observed, this occurred at the parameters which resulted in the delivery

of highest energy to the colony. Thus, for the purposes of protein identification by LESA tandem mass spectrometry, 15 pulses, 20 μs each with 1 second spacing, at 3 kV were used for all yeast species to ensure a sufficient degree of lysis.

6.3.3.2 Protein identification

6.3.3.2.1 *S. cerevisiae*

Colonies of *S. cerevisiae* grown at optimum temperature conditions (30 ℃) were subjected to a series of 15 pulses, 20 μs each with 1 second spacing, at 3 kV, as described above. Intact colonies grown on the same plate were used as control samples. FAIMS was not used. Peak assignments for all proteins identified in this section and corresponding supplementary tandem mass spectra are provided in **Appendix 3.4.1**.

In colonies subjected to electric lysis, multiply charged peaks corresponding to protein were observed; 46 peaks attributable to protein were detected in the *m/z* 700-1300 region of the representative mass spectrum shown in **Figure 6.13**. The mass spectra were similar to those recorded before, dominated by the same charge state distribution (*m/z* 725.93 16+, 774.19 15+, 829.34 14+, 893.06 13+, 967.57 12+, 1055.44 11+, 1160.87 10+; MW_{obs} = 11596.64 Da). The peak at *m/z* 893.06 (13+) was fragmented for identification and confirmed as 12 kDa heat shock protein HSP12 (MW_{obs} = 11596.64 Da), missing the initiator methionine and containing an acetylation on the N-terminal serine (**Appendix 3.4.1.2**).[324] The peak at *m/z* 1255.14 (7+; MW_{obs} = 8778.93 Da) was identified as the water deprivation stress response protein SIP18 (**Appendix 3.4.1.6**).[325] Both of these proteins were previously identified in preliminary data acquired using the prototype (see section 6.3.3). Eight further peaks (962.05 13+, 1042.1243 12+, 1004.56 5+, 870.11 10+, 1046.53 7+, 1069.78 5+, 997.54 7+ and 1209.93 12+) were selected for identification.

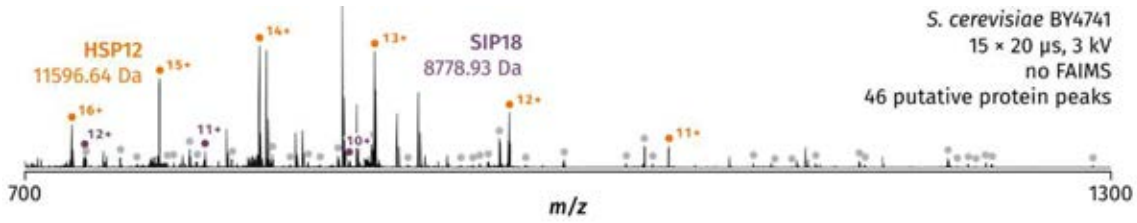


Figure 6.13. Representative full scan mass spectrum of *S. cerevisiae*. Protein peaks are marked with dots.

The protein detected at *m/z* 962.05 (13+) and 1042.12 (12+; MW_{obs} = 12493.40 Da) was identified as zeocin resistance protein 1 (ZEO1), involved in cell wall organisation;[326] the intact mass and the fragmentation pattern were consistent with the cleavage of initiator methionine and the acetylation of the N-terminal serine (fragment assignment provided in **Appendix 3.4.1.9**). The protein detected at *m/z* 1004.56 (5+; MW_{obs} = 5017.78 Da) was identified as a C-terminal fragment of elongation factor 1-α (**Appendix 3.4.1.7**).[327] The protein detected at *m/z* 870.11 (10+; MW_{obs} = 8691.05 Da) was identified as the ribosomal constituent L38 (**Appendix 3.4.1.8**). The protein detected at *m/z* 1046.52 (5+; MW_{obs} = 7318.62 Da) was identified as the mitochondrial protein FMP16 involved in stress response to mild heat shock.[328] Whilst manual sequence analysis listed in UniProt detected a 25 amino acid transit peptide at the N-terminus of the protein (a source for the assertion was not provided), the intact mass and fragmentation pattern recorded here are consistent with a 30-residue transit peptide (the N-terminal residue of the processed protein form is a proline; see **Appendix 3.4.1.10**). A truncated form of a second mitochondrial protein, OM14 (*m/z* 1069.78 5+, MW_{obs} = 5343.88 Da), a membrane protein involved in targeting of proteins to the mitochondrion, was also identified (**Appendix 3.4.1.11**). The protein detected at *m/z* 997.54 (7+; MW_{obs} = 6975.75 Da) was identified as the UPF0495 membrane protein YPR010C-A of uncertain function. The intact mass was consistent with the cleavage of the initial eight amino acids from the N-terminus (**Appendix 3.4.1.12**). The protein detected at *m/z* 1209.93 (7+;

MW_{obs} = 14507.07 Da) could not be identified. In total, eight proteins were successfully identified in *S. cerevisiae*.

Whilst originally suspected to have been induced by sub-optimal incubation temperature (see section 6.3.3), the high abundance of the 12 kDa heat shock protein at optimum growth conditions could be explained by its involvement in a number of other stress response mechanisms, including osmotic and oxidative stress.[324] Nevertheless, an attempt was made to reduce its intensity and reveal additional protein peaks by storing colonies of *S. cerevisiae* at 4 ℃ for three days prior to lysis and sampling. The colonies were subjected to the same electrical lysis regime as colonies sampled immediately following incubation. Whilst the mass spectra were dominated by singly charged species, peaks corresponding to HSP12 remained abundant, consistent with the role of HSP12 in stress response mechanisms other than heat shock (**Figure 6.14**).

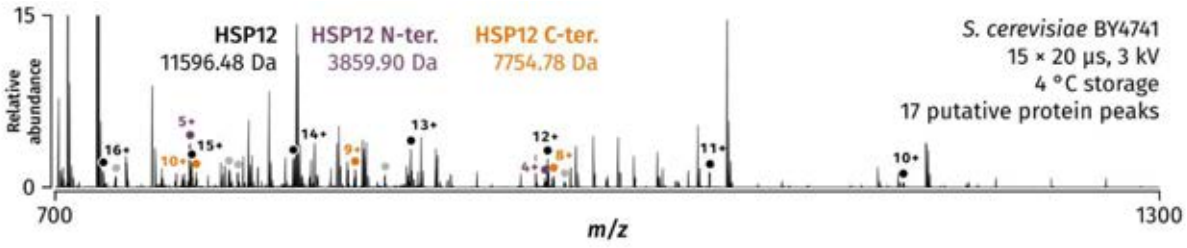


Figure 6.14. Representative mass spectrum of *S. cerevisiae* electroporated and sampled by LESA following three days of storage at 4 ℃.

Two peaks at *m/z* 772.99 5+ and 862.65 9+ were additionally observed; whilst they could also be detected at a lower abundance in colonies electroporated and sampled immediately following incubation, they were not previously selected for fragmentation. The two peaks could not be identified by the absolute mass search in ProSightPC 4.1 alpha (see Chapter 2 section 2.2.5). The biomarker search subsequently identified both peaks as two different fragments of the 12 kDa heat shock protein (**Figure 6.15**). The peak at *m/z* 772.99 5+ (MW_{obs}

= 3859.90 Da) was identified as the N-terminal fragment of the protein containing the N-terminal acetylated serine; its sequence terminated at K36 (**Appendix 3.4.1.3**). The peak at *m/z* 862.65 9+ (MW_{obs} = 7754.78 Da) corresponded to the C-terminal fragment of the protein, residues A37 onwards (**Appendix 3.4.1.4**). Residues K36 and A37 are predicted to be located within an α-helix in the hinge region of the protein; NMR evidence suggests that the protein is intrinsically disordered in solution, but can adopt a four-helical conformation upon binding to lipids, which would be present in LESA yeast samples. The cleavage site was identified by PeptideCutter as a putative site for thermolysin and trypsin; neither protease is, however, endogenously expressed in yeast and thus the mechanism of cleavage remains unknown.

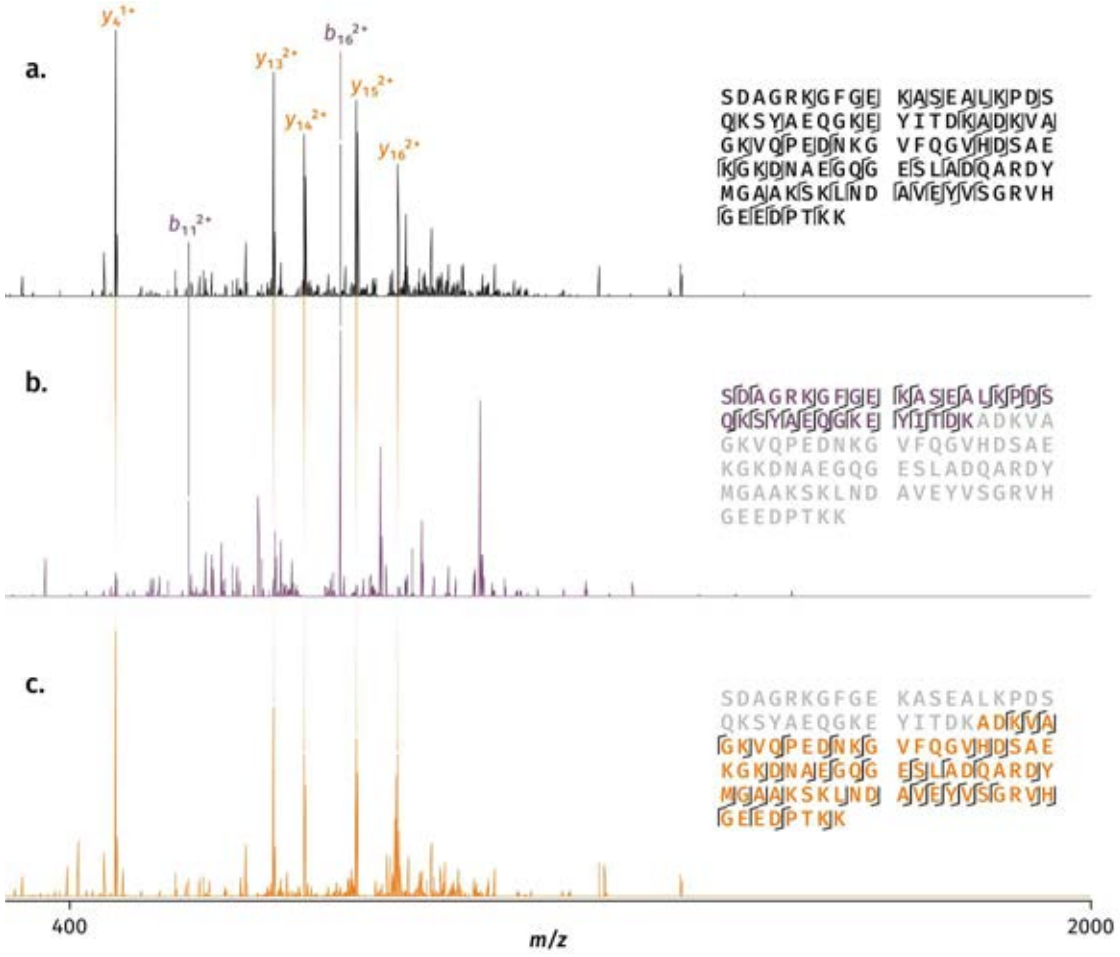


Figure 6.15. MS/MS spectra of HSP12 and its two fragments. **a**. Intact HSP12 isolated at m/z 893.06 (13+), MW_{obs} = 11596.64 Da. **b**. N-terminal fragment of HSP12 isolated at m/z 772.99 (5+), MW_{obs} = 3859.90 Da. **c**. C-terminal fragment of HSP12 isolated at *m/z* 862.65 (9+), MW_{obs} = 7754.78 Da. Representative peaks common to the intact protein and each truncated protein are marked in colour (purple for *b* ions shared with the N-terminal fragment, orange for *y* ions shared with the C-terminal fragment).

6.3.3.2.2 *C. glabrata*

Colonies of *Candida glabrata* were subjected to a series of 15 or 20 pulses, 20 µs each with 1 second spacing, at 3 kV. Intact colonies grown on the same plate were used as controls. Several initial mass spectra were acquired following the delivery of 20 rather than 15 pulses as colonies of *Candida* were expected to require a higher energy to lyse than *S. cerevisiae* (approximately 12 kV/cm in suspension for 0% survival rate, versus 7 kV/cm for *S. cerevisiae*);[267] subsequent sampling of *C. glabrata* following the delivery of 15 pulses yielded

a sufficiently abundant protein signal and thus all subsequent experiments followed the same optimised protocol as described in section 6.3.3.1.

The mass spectrum acquired from the control colony not subjected to electroporation (**Figure 6.16a**), peaks corresponding to a putative protein (*m/z* 1384.61 8+, 1230.88 9+, 1108.28 10+; MW_{obs} = 11068.82 Da) were observed (**Figure 6.16a**, inset). Each charge state was detected as a cluster containing peaks spaced in approximately 18 Da increments (17.9451 or 18.9423 Da), which could not be assigned to any chemical formula; putative sodium and potassium adducts were also observed, creating extensive overlap of isotopic distributions within the cluster. This overlap, combined with the low abundance of peaks and a high contribution of noise, prevented confident identification of monoisotopic peaks within the distribution. No other protein detected in *C. glabrata* exhibited a similar pattern. The 9+ charge state was isolated and fragmented, but the protein could not be identified. No other protein was observed in control colonies sampled immediately following incubation.

A representative mass spectrum acquired from a colony of *C. glabrata* subjected to electroporation, as compared to a control colony grown on the same plate, is shown in **Figure 6.16b**. 38 putative protein peaks were detected. Twelve peaks (*m/z* 1380.0925 8+, 1230.99 9+, 670.38 9+, 739.10 12+, 919.46 15+, 951.20 9+, 924.78 12+, 990.13 7+, 955.38 7+, 870.95 6+, 769.23 10+ and 754.18 8+) were isolated and fragmented for identification.

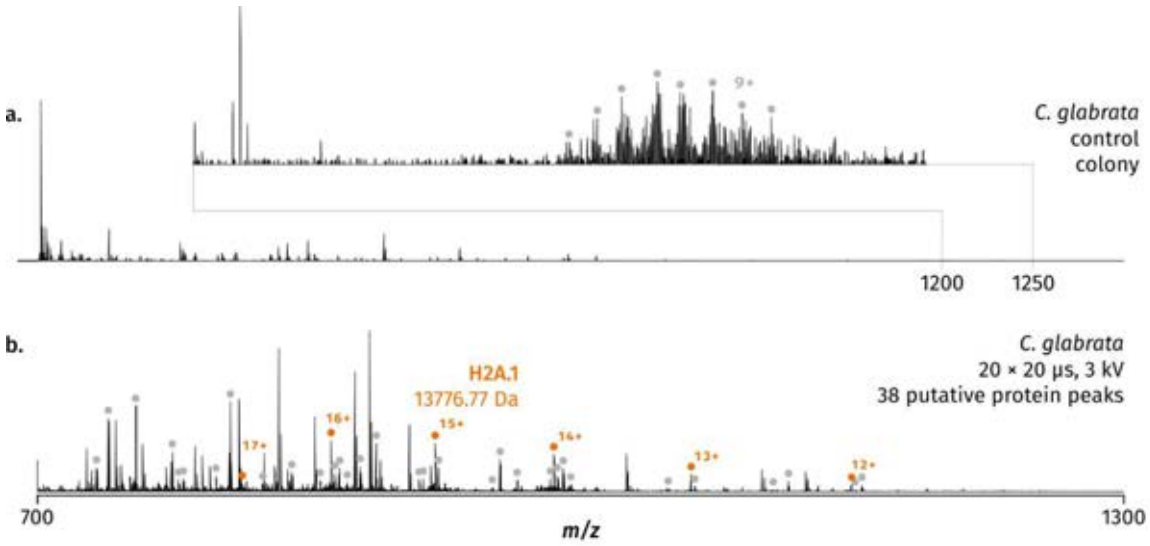


Figure 6.16. Mass spectra acquired from *C. glabrata*. **a**. Control colony. The inset shows the 9+ charge state of the unidentified 11 kDa protein. **b**. Electroporated colony.

Crucially, the peak at *m/z* 919.66 (15+) was identified with high confidence as the histone protein H2A.1 (MW_{obs} = 13776.77 Da, **Figure 6.17a**; fragment assignment given in **Appendix 3.4.2.1**) with its initiator methionine cleaved and the N-terminal serine acetylated, whilst *m/z* 769.23 (10+; MW_{obs} = 7682.26 Da) was identified as a fragment of histone H2A.1 (**Figure 6.17b**), inferred to exist in *C. glabrata* by homology (**Appendix 3.4.2.2**). As histones are localised exclusively to chromatin packaged within the nucleus, the identification of a histone protein in colonies subjected to electric lysis constituted strong evidence that the cell membrane and the nuclear envelope were lysed by the application of the electric field.

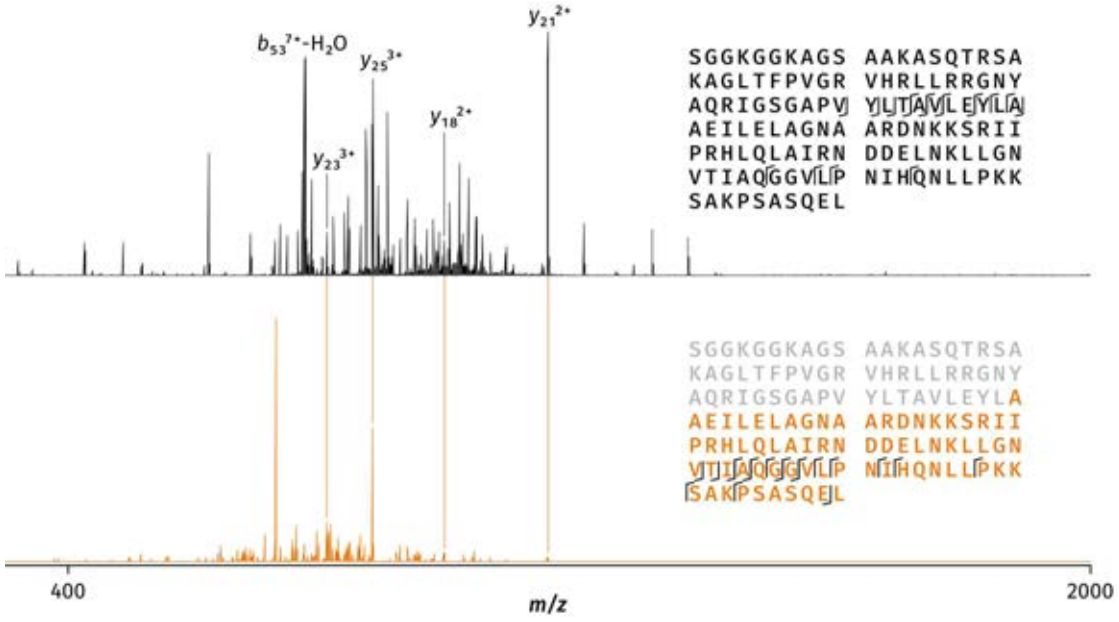


Figure 6.17. LESA tandem mass spectra of histone H2A.1 generated from *C. glabrata*. **a**. LESA tandem mass spectrum of the full length histone H2A.1 generated from *C. glabrata*. **b**. LESA tandem mass spectrum of the C-terminal fragment observed in full scan mass spectra of *C. glabrata*. Representative peaks common to both mass spectra are marked by orange lines.

The protein peak at *m/z* 924.78 (12+; MW_{obs} = 11085.26 Da) was identified as HSP12 (**Appendix 3.4.2.3**), a homolog of the 12 kDa heat shock protein previously identified in *S. cerevisiae*. Two subunits of cytochrome c oxidase (subunit 7A at *m/z* 955.38, 7+, MW_{obs} = 6680.59 Da; putative subunit 7 at *m/z* 990.13, 7+, MW_{obs} = 6923.87 Da) were identified (**Appendix 3.4.2.5** and **3.4.2.6**, respectively). Whilst subunit 7A is a well-understood protein based on its sequence similarity to other eukaryotic homologs, the putative subunit 7 is listed by UniProt as a predicted, uncharacterised protein assigned by its sequence and predicted domain structure to the COX7a family. This protein, or any protein with 90% sequence identity or higher, has not been previously reported in its intact form by any other technique. The protein peak observed at *m/z* 1380.09 (8+; MW_{obs} = 11032.68 Da) was identified as a member of the thioredoxin family (**Figure 6.18**; **Appendix 3.4.2.7**). This enzyme, predicted to be involved in cell redox homeostasis as inferred from electronic annotations, contains a

redox-active disulfide bond separated by three residues. No fragmentation was observed between the two cysteines. The existence of the protein in *C. glabrata* was inferred from homology and its intact form has not been detected by any technique until now. There are no proteins with 50% sequence homology or higher, in *C. glabrata* or any other species, which have been isolated intact and characterised, making this protein uniquely interesting for further characterisation by LESA mass spectrometry. The protein detected at *m/z* 870.95 (6+; MW_{obs} = 5219.67 Da) could not be identified.

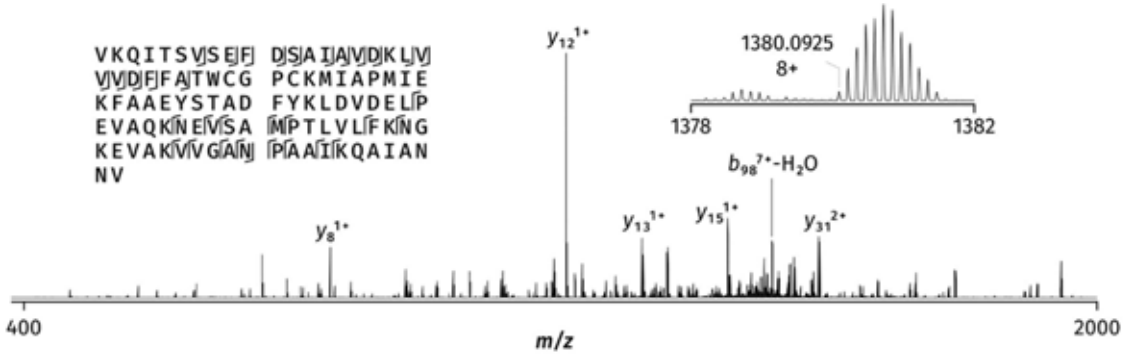


Figure 6.18. LESA tandem mass spectrum of *C. glabrata* thioredoxin isolated at *m/z* 1380.09 (8+), MW_{obs} = 11032.68 Da.

In colonies stored for three days at 4 ℃ prior to electroporation and sampling, 37 protein peaks were detected in the *m/z* 700-1300 region (**Figure 6.19**). The 13+ charge state of the most abundant protein distribution (*m/z* 853.71, MW_{obs} = 11084.22 Da) was isolated and confirmed as HSP12 (**Appendix 3.4.2.3**); six other charge states (*m/z* 740.09 15+, 792.81 14+, 924.77 12+, 1008.84 11+, 1109.63 10+ and 1226.97 9+) were also observed.

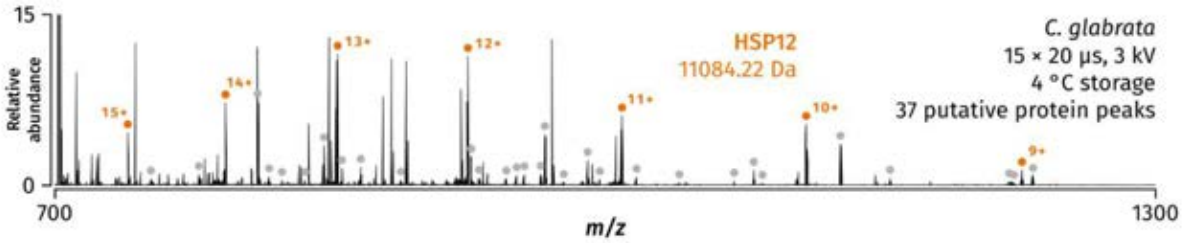


Figure 6.19. Representative mass spectrum of *C. glabrata* electroporated and sampled by LESA following three days of storage at 4 ℃.

6.3.3.2.3 *C. neoformans*

Colonies of *C. neoformans* H99 were subjected to 15 pulses of 20 µs at 3 kV. A representative mass spectrum generated from a colony of *C. neoformans* incubated for 48 hours at 30 ℃ is shown in **Figure 6.20**. Seven peaks (901.65 7+, 692.85 9+, 968.97 7+, 1072.57 5+, 878.94 11+, 741.42 5+ and 860.05 9+) were selected for identification. The protein corresponding to two most abundant peaks (*m/z* 717.21 11+, 788.94 10+; MW_{obs} = 7880.39 Da) was not isolated due to the presence of a second species of lower abundance within the same window for both charge states.

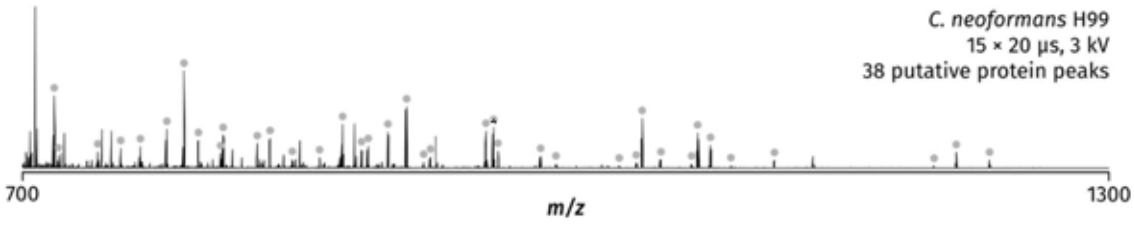


Figure 6.20. Representative LESA mass spectrum of *C. neoformans* following electroporation.

Similarly to *C. glabrata*, evidence of histone proteins was detected in *C. neoformans*. The *m/z* 741.42 peak (5+; MW_{obs} = 3702.06 Da) was identified as a C-terminal fragment of histone H3 (**Figure 6.21**; see **Appendix 3.4.3.1** for fragment assignment). The protein detected at m/z 901.65 (7+; MW_{obs} = 6304.49 Da) was identified as cytochrome c oxidase subunit 7A (**Appendix 3.4.3.2**). Similarly to the putative protein COX7a of *C. glabrata*, neither this

protein nor any closely related (sequence identity ≥90%) homologs have been detected in their intact form by any other technique. Another predicted component of cytochrome c oxidase, COX7c (*m/z* 1072.57, 5+, MW_{obs} = 5357.80 Da; **Appendix 3.4.3.5**), was also detected here for the first time. Two structural constituents of the ribosome, L39 (*m/z* 692.85, 9+; MW_{obs} = 6226.61 Da) and L29 (*m/z* 968.97, 7+; MW_{obs} = 6775.73 Da) were identified (**Appendix 3.4.3.3** and **3.4.3.4**, respectively); their database annotations have been inferred from homology and evidence for intact proteins has not hitherto been reported. The protein detected at *m/z* 878.84 (11+; MW_{obs} = 9657.23 Da) was identified as a truncated form of subunit F of the F-type H-transporting ATPase (**Appendix 3.4.3.6**), likewise never detected before in its intact form by other techniques. The protein detected at *m/z* 860.05 (9+; MW_{obs} = 7731.35 Da) could not be identified.

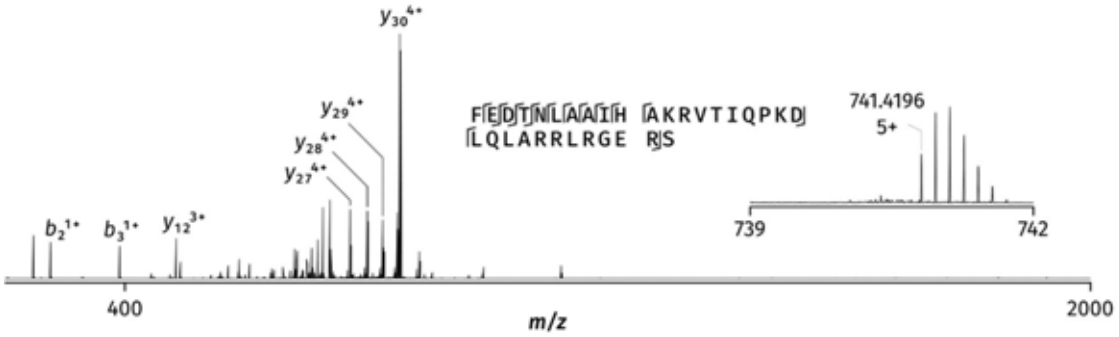


Figure 6.21. Tandem mass spectrum of the C-terminal fragment of the *C. neoformans* histone H3 isolated at *m/z* 741.42 (5+), MW_{obs} = 3702.06 Da.

6.3.4 Viability of cells following electroporation

In a pilot experiment investigating the degree of lysis by application of electricity, colonies of *S. cerevisiae* were subjected to 15 pulses, 20 µs in length, at 3 kV, following which a 1 µl inoculation loop was dipped into the colony matter subjected to the pulse. The loop was withdrawn and the residue was streaked onto a fresh YPD plate and incubated for 24 hours

at 30 ℃. Three control colonies not subjected to electric lysis were streaked in a similar manner as control.

The colony matter subjected to electric lysis contained viable cells, as growth was observed following incubation (**Figure 6.22**). By subjective assessment, the degree of growth did not differ significantly between colony matter subjected to electric lysis and the control colonies. A quantitative analysis of the degree of cell lysis was not performed.

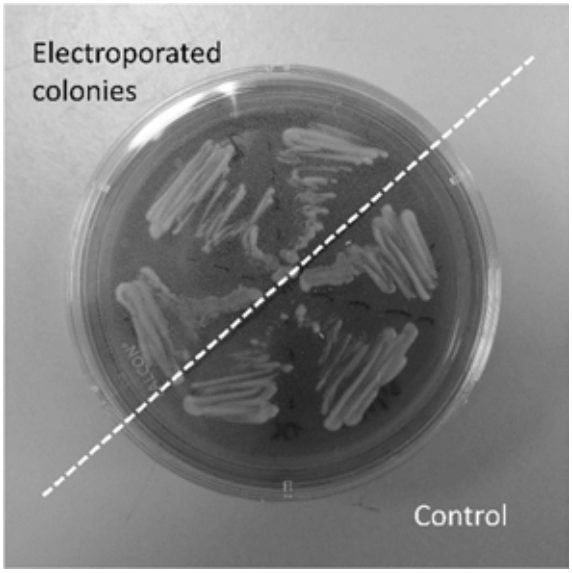


Figure 6.22. Photograph of the plate used to check for cell viability following electroporation. Re-streaked colony matter derived from electroporated colonies is shown to the upper left. Re-streaked colony matter derived from intact control colonies is shown to the lower right.

The survival of the colonies could be explained by two mechanisms. First, the pulse could take the path of least resistance through the colony and thus generate a non-homogeneous electric field between the electrodes. The cells positioned in the path of the electric discharge would be lysed, whilst cells located further from the path would remain intact and therefore viable. Second, it is possible that reversible electroporation may still occur at the parameters used here. The membranes of a subset of electroporated cells may have successfully reformed following pulse delivery and thus the cells remained viable. One proposed experiment to

identify the manner in which viable cells survive electroporation would involve the deposition of a suspension of resistance-conferring plasmid onto the colonies prior to electroporation. The emergence of viable cells resistant to antibiotics would confirm the formation of transient pores in cellular membranes.

6.4 Conclusions

The construction and operation of an electrical apparatus for the lysis of colonies on solid media prior to LESA sampling was demonstrated. Parameters were optimised for the lysis of yeast colonies on the model organism *S. cerevisiae*. Protein extraction and identification, previously impossible by use of LESA alone, was demonstrated for *S. cerevisiae*, *C. glabrata* and *C. neoformans*.

CHAPTER 7

Conclusions and further work

This thesis has discussed the development of liquid extraction surface analysis mass spectrometry for protein analysis directly from microbial colonies grown on solid agar media. A brief discussion of repeatability and reproducibility was included, followed by the optimisation of sampling parameters for the extraction of over 40 proteins from seven known and one unknown species of bacteria, both Gram-positive and Gram-negative. Possible applications for the differentiation of closely related organisms, exemplified by three members of the viridans group of streptococci, were discussed. *De novo* sequencing of a novel protein from an unknown species of *Staphylococcus* on the basis of LESA tandem mass spectra was demonstrated. The coupling of FAIMS to LESA MS was demonstrated, enhancing the capabilities of LESA for protein identification. Finally, analysis of hitherto inaccessible proteins from yeast species was demonstrated by use of a electroporator designed and built in-house to lyse yeast colonies prior to LESA sampling. This chapter contains a detailed discussion of each outcome and provides suggestions for future work as well as possible applications of the methodology in its current state.

Chapter 3 provided a brief summary of the repeatability and reproducibility of the LESA process. The findings derived from the sampling of model surfaces suggested the existence of variability inherent to the LESA apparatus, affecting sampling position in all three dimensions as well as the dispensation of solvent. This was consistent with observations published prior to the start of this project.[116] The reproducibility of sampling on colonies of *E. coli* was also investigated, using a percentage success rate (where success was defined as the maintenance of a stable nanoelectrospray over the entire duration of mass spectrum

acquisition, coupled to the detection of peaks attributable to protein in the resultant mass spectrum) as the primary measure of reproducibility. 25% of sampling attempts resulted in the successful acquisition of protein mass spectra. Issues inherent to the sampling approach were identified. The major, resolvable issue concerned the use of a LESA sampling grid derived from the universal adapter plate supplied with the NanoMate for the selection of locations. The spacing of the grid was large compared to the diameter of the colonies, artificially limiting the availability of sampling locations on the colony. The sampling approach was subsequently modified for the remainder of the project. The sampling grid was no longer used and the targeting of sampling locations was performed manually, rendering the positioning of the pipette tip in all three dimensions fully reliant on the operator's subjective judgement. Any automation of the process for imaging purposes or otherwise would require modifications to the NanoMate, most crucially for the adjustment of sampling height. Automation of this aspect of sampling could be accomplished by the integration of a laser distance measurement attachment or a shear probe, as has already been done for nanoDESI.[102] Similarly, modifications to allow the use of dedicated LESA software for the sampling of colonies, in place of relying on the advanced user interface, would allow for closer grouping of sampling spots on colonies of a lower surface area, thus eliminating the need for manual targeting in the *x, y* plane. Once the significant human factor in sampling is minimised, more representative values for repeatability and reproducibility on microbial colonies may be obtained. Precision of movement on all three axes would benefit from improvement, particularly for imaging purposes, as has been previously suggested; adjustments to the stepper motor responsible for the movements of the mandrel on the *z* axis would be particularly beneficial to the sampling of thin samples with complex topology such

as microbial biofilms, and would likely be required to take full advantage of height adjustment automation.

Further characterisation of the LESA process would be beneficial. The investigation of LESA sampling spots on bacterial colonies on agar by MALDI and DESI MS was proposed and preliminary results were acquired; the conclusions which could be drawn were, however, limited. DESI MS showed more promise of the two techniques and it is likely that further optimisation may yield more informative outcomes. Results generated thus far indicated issues with parts of the dried sample material being ejected from the support surface by the jet of solvent, which led to the generation incomplete images. The ion intensities recorded on agar background were significantly lower than those recorded on the surface of the colony, most likely due to the difference in surface properties between the two areas; plain agar may have distorted or absorbed the jet of solvent, reducing signal intensity. Without normalisation, errors were thus introduced into the measurement of colony-specific compounds versus background. The lack of MS/MS data also reduced confidence in putative peak assignments.

Additional investigation into the three-dimensional extent of solvent spread within colonies sampled in the contact approach, putatively achievable by confocal microscopy, would help define the volume of the colony accessible by each sampling attempt, thus determining the depth of profiling as performed by LESA. The topology of sampling spots could be further investigated, either by confocal microscopy or by scanning electron microscopy. Colonies could also be frozen and sectioned for imaging, an approach previously employed for imaging of fungi.[250]

Chapter 4 demonstrated the application of LESA MS for the extraction and identification of proteins from clinically relevant bacterial strains. A total of 40 proteins were successfully

identified across all samples, which included both Gram-negative and Gram-positive species. Secreted, periplasmic and cytosolic proteins were observed, spanning a range of functions from ribosomal proteins to mitochondrial enzymes, as well as multiple stress response proteins up to 15 kDa. Characteristic protein peaks derived from LESA MS were subsequently used to differentiate between three closely related species of streptococci, a common challenge for MALDI-TOF MS. Whilst such an approach could not, at present, compete with MALDI-TOF MS in terms of speed or depth of identification (genus, species or strain level), one putative advantage of LESA is its suitability for *in situ* analysis, e.g. directly on living patient tissue or contaminated surfaces. To that end, its performance on surfaces other than common nutrient media should be explored.

The project could easily be extended to further clinically relevant species, as well as alternative surfaces for microbial growth. As a result of the work presented here on *S. aureus* and *P. aeruginosa*, a parallel project was devised to focus on the ESKAPE pathogens and alternative growth surfaces, including laboratory-grown skin models and wound models which could be directly probed by non-toxic solvents. The work on secreted peptides, including *de novo* sequencing, could be applied to novel isolates for the characterisation of putative antibiotic peptides. The sampling of zones of inhibition, formed on solid nutrient media where two microbial species interact via secreted compounds, would be particularly amenable to a surface sampling technique; the sensitivity of LESA in particular would further contribute towards the detection of compounds secreted at low concentrations. Additionally, contact sampling lends itself well to the use of volatile solvents, such as chloroform/methanol commonly used for the sampling of lipids. The LESA sampling protocol could thus be adjusted to provide information on the lipid content of bacterial envelopes, particularly valuable for species with unique cell wall compositions such as mycobacteria. Whilst absolute

quantitation by LESA is in its early stages of development, the measurement of relative abundances of lipid classes may be sufficient to provide species identification capability, on its own or in conjunction with protein profiling.

Chapter 5 described an investigation into the use of FAIMS to enhance the detection and identification of proteins by LESA MS from bacterial colonies. Three challenges were initially presented, to which FAIMS was the proposed solution: the interference of high background signal derived from solvent and agar media during the identification of proteins in the low *m/z* range in *E. coli*, the incidence of overlapping protein peaks and the resultant generation of compound tandem mass spectra in *P. aeruginosa*, and the putative suppression of analytes of interest by exceedingly abundant secreted peptides in *S. aureus*. Background signal was successfully eliminated in *E. coli*, granting access to previously undetectable proteins, two of which (L32 and L33) were isolated and identified below *m/z* 800. In *P. aeruginosa*, the resolution of three instances of overlapping protein peaks was demonstrated; in two of these cases, previously inaccessible proteins were isolated and identified. Furthermore, due to the increased detection of medium to low molecular weight proteins in the low *m/z* range, it may be beneficial to revisit the capability for the differentiation of microorganisms based on their LESA mass spectra. MALDI-TOF MS, relying on direct comparisons of fingerprint mass spectra generated from <15 kDa ribosomal and housekeeping proteins, was shown in several studies to be capable of strain-level resolution; thus, it is possible that FAIMS may enhance the capability of LESA for species differentiation, potentially to a similar level as MALDI-TOF MS.

The elimination of the secreted peptide peaks in *S. aureus* was, however, unsuccessful. Two solutions could be proposed to resolve this challenge. First, a cylindrical FAIMS device of higher performance was commercially released by Thermo Scientific shortly prior to the end

of this project. Preliminary results reported by the group indicated that the device could potentially yield a range of proteins an order of magnitude greater than that observed with the device used for this work. Further improvements could therefore be achieved purely through advances in FAIMS technology. Second, a different ion mobility technique could potentially be used to circumvent the large range of ion mobilities in FAIMS fields which characterises δ-haemolysin. Drift tube IMS or travelling wave IMS may afford a higher resolution separation of low molecular weight peptides from larger proteins of interest.

The results generated by top-down LC-MS/MS as an independent comparison to FAIMS-MS/MS were of insufficient quality to provide a fair comparison. Errors in LC-MS/MS protocol optimisation led to poor separation, poor spectral quality and a low number of MS/MS identifications, significantly lower than expected based on literature values. It may, however, be beneficial to develop both a top-down LC-MS/MS approach for LESA extracts of bacterial colonies and a parallel, bottom-up approach in order to determine the classes and masses of proteins which are preferentially extracted by LESA and determine the factors limiting how many proteins are observed in standard LESA mass spectra (extraction efficiency, spectral overlap, ion suppression). Further work using updated FAIMS technology could be performed; ultimately, both high-performance LC (potentially coupled directly to the NanoMate) and next-generation FAIMS could be combined in an automated workflow to provide the most in-depth characterisation of bacterial proteomes. This is not possible using current equipment.

Chapter 6 introduced the home-built electroporator, determined its optimum operating parameters and described its application to three species of yeast. Optimisation was performed on *S. cerevisiae* alone. Whilst the parameters derived through optimisation on a single species enabled the successful liberation and extraction of proteins from all three tested

species, the repeatability of the process during that particular set of experiments was unsatisfactory and thus the measured values for minimum voltage, pulse number and pulse length were not determined to a high degree of confidence. The determination of successful cell lysis was achieved by LESA, which is subject to its own intrinsic variability as determined in Chapter 3, thus complicating the determination of variability intrinsic to the lysis process. A repetition of this work using a different mechanism to gauge the degree of lysis would eliminate the portion of variability attributable to LESA and thus more clearly measure the variability of the lysis process itself. This could be achieved by direct infusion of filtered/centrifuged colony suspensions or by classical cell counting methodologies used in microbiology. The latter approach would also allow the quantitation of cell death within an electroporated colony. Furthermore, an investigation into the formation of transient pores, the main purpose of electroporation performed routinely in suspension, could also be conducted to determine whether the apparatus in its current state is capable of facilitating the transformation of microbial colonies. This could be achieved by the deposition of a suspension of plasmid encoding antibiotic resistance on the colonies prior to electroporation. The re-inoculation of electroporated colonies onto antibiotic-containing nutrient media and subsequent observation of viable growth would indicate uptake of the plasmid and thus successful transformation, which could lead to microbiological applications of this approach outside of mass spectrometry.

An increased understanding of the electric lysis process may also unlock its potential application for native mass spectrometry. The elimination of need for chemical lysis of microbial colonies would enable the use of mild solvents which are, on their own, incapable of extracting proteins through membranes and cell walls. Studies of electroporation in suspension demonstrated that lysis is not due to thermal effects. The local temperature

increase upon application of electric pulses was not expected to exceed 10 ℃ based on theoretical calculations and thus should not lead to protein denaturation. It would, however, be beneficial to experimentally confirm this value for electroporation on solid media; this could be achieved by use of a laser temperature gauge or a high-resolution thermal camera.

In summary, the work described in this thesis contributed additional understanding to the LESA process and its performance as applied to microbial colonies. The range of microbial classes accessible by the technique was expanded to include Gram-positive bacteria and, with the coupling of plate-based electroporation, yeasts. Further work should focus equally on clinically relevant applications and further development of novel workflows, particularly the electroporation-LESA MS workflow which has the potential to unlock the capability for native protein mass spectrometry directly from microbial colonies. This could be applied equally to wild type organisms as well as expression systems producing proteins or protein complexes of interest, which could thus be probed directly from their native environment without any disturbances introduced by purification.

REFERENCES

1. De Hoffmann E, Stroobant V. Mass spectrometry: Principles and applications. 3rd ed: John Wiley & Sons; 2007.
2. Senko MW, McLafferty FW. Mass-spectrometry of macromolecules - has its time now come. Annu Rev Biophys Biomolec Struct. 1994;23:763-85.
3. Nier AO. A mass spectrometer for routine isotope abundance measurements. Rev Sci Instrum. 1940;11(7):212-6.
4. Munson MSB, Field FH. Chemical ionization mass spectrometry. I. General introduction. J Am Chem Soc. 1966;88(12):2621-&.
5. Horning EC, Horning MG, Carroll DI, Dzidic I, Stillwell RN. New picogram detection system based on a mass spectrometer with an external ionization source at atmospheric pressure. Analytical Chemistry. 1973;45(6):936-43.
6. Carroll DI, Dzidic I, Stillwell RN, Haegele KD, Horning EC. Atmospheric pressure ionization mass spectrometry - corona discharge ion source for use in liquid chromatograph-mass spectrometer-computer analytical system. Analytical Chemistry. 1975;47(14):2369-73.
7. Robb DB, Covey TR, Bruins AP. Atmospheric pressure photoionisation: An ionization method for liquid chromatography-mass spectrometry. Analytical Chemistry. 2000;72(15):3653-9.
8. Raffaelli A, Saba A. Atmospheric pressure photoionization mass spectrometry. Mass Spectrom Rev. 2003;22(5):318-31.
9. Benninghoven A. Chemical analysis of inorganic and organic surfaces and thin films by static time-of-flight secondary ion mass-spectrometry (TOF-SIMS). Angew Chem-Int Edit Engl. 1994;33(10):1023-43.
10. Barber M, Bordoli RS, Sedgwick RD, Tyler AN. Fast atom bombardment of solids (FAB) - a new ion-source for mass-spectrometry. J Chem Soc-Chem Commun. 1981(7):325-7.
11. Barber M, Bordoli RS, Elliott GJ, Sedgwick RD, Tyler AN, Green BN. Fast atom bombardment mass-spectrometry of bovine insulin and other large peptides. J Chem Soc-Chem Commun. 1982(16):936-8.
12. Barber M, Bordoli RS, Sedgwick RD, Tyler AN, Whalley ET. Fast atom bombardment mass-spectrometry of bradykinin and related oligopeptides. Biomedical Mass Spectrometry. 1981;8(8):337-42.
13. Barber M, Green BN. The analysis of small proteins in the molecular weight range 10-24 kda by magnetic sector mass spectrometry. Rapid communications in mass spectrometry : RCM. 1987;1(5):80-3.
14. Yang J, Gilmore I. Application of secondary ion mass spectrometry to biomaterials, proteins and cells: a concise review. Mater Sci Technol. 2015;31(2):131-6.
15. Mochiji K, Hashinokuchi M, Moritani K, Toyoda N. Matrix-free detection of intact ions from proteins in argon-cluster secondary ion mass spectrometry. Rapid Communications in Mass Spectrometry. 2009;23(5):648-52.
16. Vestal ML. Studies of ionization mechanisms involved in thermospray LC-MS. International Journal of Mass Spectrometry and Ion Processes. 1983;46(JAN):193-6.
17. Sundqvist B, Macfarlane RD. Cf-252 plasma desorption mass-spectrometry. Mass Spectrom Rev. 1985;4(4):421-60.
18. Sundqvist B, Kamensky I, Hakansson P, Kjellberg J, Salehpour M, Widdiyasekera S, et al. Cf-252 plasma desorption time of flight mass-spectroscopy of proteins. Biomedical Mass Spectrometry. 1984;11(5):242-57.

19. Sundqvist B, Roepstorff P, Fohlman J, Hedin A, Hakansson P, Kamensky I, et al. Molecular-weight determinations of proteins by californium plasma desorption mass-spectrometry. Science. 1984;226(4675):696-8.
20. Jones DS, Krolik ST. Analysis of nonderivatised peptides by thermospray using a magnetic sector mass spectrometer. Rapid communications in mass spectrometry : RCM. 1987;1(4):67-8.
21. Jonsson G, Hedin A, Hakansson P, Sundqvist BU. Competition between protein/protein interactions and protein/substrate interactions studied by plasma desorption mass spectrometry. Rapid communications in mass spectrometry : RCM. 1988;2(8):154-6.
22. Smith RD, Loo JA, Edmonds CG, Barinaga CJ, Udseth HR. New developments in biochemical mass-spectrometry - electrospray ionization. Analytical Chemistry. 1990;62(9):882-99.
23. Karas M, Bachmann D, Bahr U, Hillenkamp F. Matrix-assisted ultraviolet-laser desorption of nonvolatile compounds. International Journal of Mass Spectrometry and Ion Processes. 1987;78:53-68.
24. Yamashita M, Fenn JB. Electrospray ion-source - another variation on the free-jet theme. Journal of Physical Chemistry. 1984;88(20):4451-9.
25. Glish GL, Vachet RW. The basics of mass spectrometry in the twenty-first century. Nat Rev Drug Discov. 2003;2(2):140-50.
26. Kraft P, Alimpiev S, Dratz E, Sunner J. Infrared, surface-assisted laser desorption ionization mass spectrometry on frozen aqueous solutions of proteins and peptides using suspensions of organic solids. Journal of the American Society for Mass Spectrometry. 1998;9(9):912-24.
27. Trimpin S. "Magic" Ionization Mass Spectrometry. Journal of the American Society for Mass Spectrometry. 2016;27(1):4-21.
28. Zenobi R, Knochenmuss R. Ion formation in MALDI mass spectrometry. Mass Spectrom Rev. 1998;17(5):337-66.
29. Knochenmuss R. Ion formation mechanisms in UV-MALDI. Analyst. 2006;131(9):966-86.
30. Chang WC, Huang LCL, Wang YS, Peng WP, Chang HC, Hsu NY, et al. Matrix-assisted laser desorption/ionization (MALDI) mechanism revisited. Analytica Chimica Acta. 2007;582(1):1-9.
31. Lu IC, Lee C, Lee YT, Ni CK. Ionization mechanism of matrix-assisted laser desorption/ionization. In: Cooks RG, Pemberton JE, editors. Annual Review of Analytical Chemistry, Vol 8. Annual Review of Analytical Chemistry. 8. Palo Alto: Annual Reviews; 2015. p. 21-39.
32. Liang SP, Lu IC, Tsai ST, Chen JL, Lee YT, Ni CK. Laser pulse width dependence and ionization mechanism of matrix-assisted laser desorption/ionization. Journal of the American Society for Mass Spectrometry. 2017;28(10):2235-45.
33. Knochenmuss R. The coupled chemical and physical dynamics model of MALDI. In: Bohn PW, Pemberton JE, editors. Annual Review of Analytical Chemistry, Vol 9. Annual Review of Analytical Chemistry. 9. Palo Alto: Annual Reviews; 2016. p. 365-+.
34. Peacock PM, Zhang WJ, Trimpin S. Advances in ionization for mass spectrometry. Analytical Chemistry. 2017;89(1):372-88.
35. Chu KY, Lee S, Tsai MT, Lu IC, Dyakov YA, Lai YH, et al. Thermal proton transfer reactions in ultraviolet matrix-assisted laser desorption/ionization. Journal of the American Society for Mass Spectrometry. 2014;25(3):310-8.

36. Knochenmuss R. Ion yields in the coupled chemical and physical dynamics model of matrix-assisted laser desorption/ionization. Journal of the American Society for Mass Spectrometry. 2015;26(10):1645-8.
37. Bae YJ, Kim MS. A thermal mechanism of ion formation in MALDI. In: Cooks RG, Pemberton JE, editors. Annual Review of Analytical Chemistry, Vol 8. Annual Review of Analytical Chemistry. 8. Palo Alto: Annual Reviews; 2015. p. 41-60.
38. Knochenmuss R. Energetics and kinetics of thermal ionization models of MALDI. Journal of the American Society for Mass Spectrometry. 2014;25(9):1521-7.
39. Gray-Weale A, Ni CK. Comment on: "Energetics and kinetics of thermal ionization models of MALDI" by Richard Knochenmuss. J-Am. Soc. Mass Spectrom. 25, 1521-1527 (2014). Journal of the American Society for Mass Spectrometry. 2015;26(12):2162-6.
40. Knochenmuss R. Reply to the comment on: "Energetics and kinetics of thermal ionization models of MALDI" by Richard Knochenmuss. J-Am. Soc. Mass Spectrom. 25, 1521-1527 (2014). Journal of the American Society for Mass Spectrometry. 2015;26(12):2167-8.
41. Gray-Weale A, Ni CK. Comment to the reply on: "Energetics and kinetics of thermal ionization models of MALDI" by Richard Knochenmuss. J. Am. Soc. Mass spectrom. 25, 1521-1527 (2014). Journal of the American Society for Mass Spectrometry. 2015;26(12):2169-70.
42. Knochenmuss R. Reply to the comment to the reply on: "Energetics and kinetics of thermal ionization models of MALDI" by Richard Knochenmuss, J-Am. Soc. Mass Spectrom. 25, 1521-1527 (2014). Journal of the American Society for Mass Spectrometry. 2015;26(12):2171-.
43. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. Analytical Chemistry. 1988;60(20):2299-301.
44. Pappin DJC, Hojrup P, Bleasby AJ. Rapid identification of proteins by peptide-mass fingerprinting. Curr Biol. 1993;3(6):327-32.
45. Leinweber BD, Tsaprailis G, Monks TJ, Lau SS. Improved MALDI-TOF imaging yields increased protein signals at high molecular mass. Journal of the American Society for Mass Spectrometry. 2009;20(1):89-95.
46. Mange A, Chaurand P, Perrochia H, Roger P, Caprioli RM, Solassol J. Liquid chromatography-tandem and MALDI imaging mass spectrometry analyses of RCL2/CS100-fixed, paraffin-embedded tissues: Proteomics evaluation of an alternate fixative for biomarker discovery. J Proteome Res. 2009;8(12):5619-28.
47. van Remoortere A, van Zeijl RJM, van den Oever N, Franck J, Longuespee R, Wisztorski M, et al. MALDI imaging and profiling MS of higher mass proteins from tissue. Journal of the American Society for Mass Spectrometry. 2010;21(11):1922-9.
48. Madler S, Erba EB, Zenobi R. MALDI-TOF mass spectrometry for studying noncovalent complexes of biomolecules. In: Cai Z, Liu S, editors. Applications of Maldi-Tof Spectroscopy. Topics in Current Chemistry-Series. 331. Berlin: Springer-Verlag Berlin; 2013. p. 1-36.
49. Woods AS, Buchsbaum JC, Worrall TA, Berg JM, Cotter RJ. Matrix-assisted laser desorption/ionization of noncovalently bound compounds. Analytical Chemistry. 1995;67(24):4462-5.
50. Chen F, Gulbakan B, Weidmann S, Fagerer SR, Ibanez AJ, Zenobi R. Applying mass spectrometry to study non-covalent biomolecule complexes. Mass Spectrom Rev. 2016;35(1):48-70.
51. Weidmann S, Barylyuk K, Nespovitaya N, Madler S, Zenobi R. A new, modular mass calibrant for high-mass MALDI-MS. Analytical Chemistry. 2013;85(6):3425-32.

52. Plath LD, Ozdemir A, Aksenov AA, Bier ME. Determination of iron content and dispersity of intact ferritin by superconducting tunnel junction cryodetection mass spectrometry. Analytical Chemistry. 2015;87(17):8985-93.
53. Taylor GI. Disintegration of water drops in electric field. Proc R Soc Lond A. 1964;280:383-97.
54. Dole M, Mack LL, Hines RL. Molecular beams of macroions. J Chem Phys. 1968;49(5):2240-&.
55. Wilm MS, Mann M. Electrospray and taylor-cone theory, doles beam of macromolecules at last. International Journal of Mass Spectrometry. 1994;136(2-3):167-80.
56. Wilm M. Principles of electrospray ionization. Mol Cell Proteomics. 2011;10(7):8.
57. Konermann L, Ahadi E, Rodriguez AD, Vahidi S. Unraveling the mechanism of electrospray ionization. Analytical Chemistry. 2013;85(1):2-9.
58. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass-spectrometry of large biomolecules. Science. 1989;246(4926):64-71.
59. Iribarne JV, Thomson BA. On the evaporation of small ions from charged droplets. J Chem Phys. 1976;64(6):2287-94.
60. Ahadi E, Konermann L. Ejection of solvated ions from electrosprayed methanol/water nanodroplets studied by molecular dynamics simulations. J Am Chem Soc. 2011;133(24):9354-63.
61. Daub CD, Cann NM. How are completely desolvated ions produced in electrospray ionization: Insights from molecular dynamics simulations. Analytical Chemistry. 2011;83(22):8372-6.
62. Ahadi E, Konermann L. Modeling the behavior of coarse-grained polymer chains in charged water droplets: Implications for the mechanism of electrospray ionization. J Phys Chem B. 2012;116(1):104-12.
63. Konermann L, Rodriguez AD, Liu JJ. On the formation of highly charged gaseous ions from unfolded proteins by electrospray ionization. Analytical Chemistry. 2012;84(15):6798-804.
64. Metwally H, Duez Q, Konermann L. Chain ejection model for electrospray ionization of unfolded proteins: Evidence from atomistic simulations and ion mobility spectrometry. Analytical Chemistry. 2018;90(16):10069-77.
65. Hop C, Bakhtiar R. An introduction to electrospray ionization and matrix-assisted laser desorption/ionization mass spectrometry: Essential tools in a modern biotechnology environment. Biospectroscopy. 1997;3(4):259-80.
66. Nadler WM, Waidelich D, Kerner A, Hanke S, Berg R, Trumpp A, et al. MALDI versus ESI: The impact of the ion source on peptide identification. J Proteome Res. 2017;16(3):1207-15.
67. Wilm M, Mann M. Analytical properties of the nanoelectrospray ion source. Analytical Chemistry. 1996;68(1):1-8.
68. Murray KK, Boyd RK, Eberlin MN, Langley GJ, Li L, Naito Y. Definitions of terms relating to mass spectrometry (IUPAC recommendations 2013). Pure Appl Chem. 2013;85(7):1515-609.
69. Laskin J, Lanekoff I. Ambient mass spectrometry imaging using direct liquid extraction techniques. Analytical Chemistry. 2016;88(1):52-73.
70. Huang MZ, Yuan CH, Cheng SC, Cho YT, Shiea J. Ambient ionization mass spectrometry. In: Yeung ES, Zare RN, editors. Annual Review of Analytical Chemistry, Vol 3. Annual Review of Analytical Chemistry. 3. Palo Alto: Annual Reviews; 2010. p. 43-65.

71. Huang MZ, Cheng SC, Cho YT, Shiea J. Ambient ionization mass spectrometry: A tutorial. Analytica Chimica Acta. 2011;702(1):1-15.
72. Ifa DR, Wu CP, Ouyang Z, Cooks RG. Desorption electrospray ionization and other ambient ionization methods: Current progress and preview. Analyst. 2010;135(4):669-81.
73. Badu-Tawiah AK, Eberlin LS, Ouyang Z, Cooks RG. Chemical aspects of the extractive methods of ambient ionization mass spectrometry. In: Johnson MA, Martinez TJ, editors. Annual Review of Physical Chemistry, Vol 64. Annual Review of Physical Chemistry. 64. Palo Alto: Annual Reviews; 2013. p. 481-505.
74. Venter A, Nefliu M, Cooks RG. Ambient desorption ionization mass spectrometry. Trac-Trends Anal Chem. 2008;27(4):284-90.
75. Van Berkel GJ, Pasilis SP, Ovchinnikova O. Established and emerging atmospheric pressure surface sampling/ionization techniques for mass spectrometry. Journal of Mass Spectrometry. 2008;43(9):1161-80.
76. Chen HW, Gamez G, Zenobi R. What can we learn from ambient ionization techniques? Journal of the American Society for Mass Spectrometry. 2009;20(11):1947-63.
77. Harris GA, Galhena AS, Fernandez FM. Ambient sampling/ionization mass spectrometry: Applications and current trends. Analytical Chemistry. 2011;83(12):4508-38.
78. Alberici RM, Simas RC, Sanvido GB, Romao W, Lalli PM, Benassi M, et al. Ambient mass spectrometry: Bringing MS into the "real world". Anal Bioanal Chem. 2010;398(1):265-94.
79. Wu CP, Dill AL, Eberlin LS, Cooks RG, Ifa DR. Mass spectrometry imaging under ambient conditions. Mass Spectrom Rev. 2013;32(3):218-43.
80. Venter AR, Douglass KA, Shelley JT, Hasman G, Honarvar E. Mechanisms of real-time, proximal sample processing during ambient ionization mass spectrometry. Analytical Chemistry. 2014;86(1):233-49.
81. Javanshad R, Venter AR. Ambient ionization mass spectrometry: Real-time, proximal sample processing and ionization. Anal Methods. 2017;9(34):4896-907.
82. Monge ME, Harris GA, Dwivedi P, Fernandez FM. Mass Spectrometry: Recent Advances in Direct Open Air Surface Sampling/Ionization. Chem Rev. 2013;113(4):2269-308.
83. Monge ME, Fernandez FM. An Introduction to Ambient Ionization Mass Spectrometry. In: Domin M, Cody R, editors. Ambient Ionization Mass Spectrometry. New Developments in NMR. Cambridge: Royal Soc Chemistry; 2015. p. 1-22.
84. Takats Z, Wiseman JM, Gologan B, Cooks RG. Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. Science. 2004;306(5695):471-3.
85. Roach PJ, Laskin J, Laskin A. Nanospray desorption electrospray ionization: An ambient method for liquid-extraction surface sampling in mass spectrometry. Analyst. 2010;135(9):2233-6.
86. Van Berkel GJ, Sanchez AD, Quirke JME. Thin-layer chromatography and electrospray mass spectrometry coupled using a surface sampling probe. Analytical Chemistry. 2002;74(24):6216-23.
87. Kertesz V, Ford MJ, Van Berkel GJ. Automation of a surface sampling probe/electrospray mass spectrometry system. Analytical Chemistry. 2005;77(22):7183-9.
88. Kertesz V, Van Berkel GJ. Fully automated liquid extraction-based surface sampling and ionization using a chip-based robotic nanoelectrospray platform. Journal of Mass Spectrometry. 2010;45(3):252-60.
89. Venter A, Sojka PE, Cooks RG. Droplet dynamics and ionization mechanisms in desorption electrospray ionization mass spectrometry. Analytical Chemistry. 2006;78(24):8549-55.

90. Kertesz V, Van Berkel GJ. Improved imaging resolution in desorption electrospray ionization mass spectrometry. Rapid Communications in Mass Spectrometry. 2008;22(17):2639-44.
91. Douglass KA, Venter AR. Protein analysis by desorption electrospray ionization mass spectrometry and related methods. Journal of Mass Spectrometry. 2013;48(5):553-60.
92. Garza KY, Feider CL, Klein DR, Rosenberg JA, Brodbelt JS, Eberlin LS. Desorption electrospray ionization mass spectrometry imaging of proteins directly from biological tissue sections. Anal Chem. 2018;90(13):7785-9.
93. Ambrose S, Housden NG, Gupta K, Fan JY, White P, Yen HY, et al. Native desorption electrospray ionization liberates soluble and membrane protein complexes from surfaces. Angew Chem-Int Edit. 2017;56(46):14463-8.
94. Hsu CC, Chou PT, Zare RN. Imaging of proteins in tissue samples using nanospray desorption electrospray ionization mass spectrometry. Analytical Chemistry. 2015;87(22):11171-5.
95. Lanekoff I, Geydebrekht O, Pinchuk GE, Konopka AE, Laskin J. Spatially resolved analysis of glycolipids and metabolites in living *Synechococcus* sp PCC 7002 using nanospray desorption electrospray ionization. Analyst. 2013;138(7):1971-8.
96. Lanekoff I, Heath BS, Liyu A, Thomas M, Carson JP, Laskin J. Automated platform for high-resolution tissue imaging using nanospray desorption electrospray ionization mass spectrometry. Analytical Chemistry. 2012;84(19):8351-6.
97. Laskin J, Heath BS, Roach PJ, Cazares L, Semmes OJ. Tissue imaging using nanospray desorption electrospray ionization mass spectrometry. Analytical Chemistry. 2012;84(1):141-8.
98. Lanekoff I, Thomas M, Carson JP, Smith JN, Timchalk C, Laskin J. Imaging nicotine in rat brain tissue by use of nanospray desorption electrospray ionization mass spectrometry. Analytical Chemistry. 2013;85(2):882-9.
99. Bergman HM, Lundin E, Andersson M, Lanekoff I. Quantitative mass spectrometry imaging of small-molecule neurotransmitters in rat brain tissue sections using nanospray desorption electrospray ionization. Analyst. 2016;141(12):3686-95.
100. Watrous J, Roach P, Heath B, Alexandrov T, Laskin J, Dorrestein PC. Metabolic profiling directly from the petri dish using nanospray desorption electrospray ionization imaging mass spectrometry. Analytical Chemistry. 2013;85(21):10385-91.
101. Lanekoff I, Burnum-Johnson K, Thomas M, Cha J, Dey SK, Yang PX, et al. Three-dimensional imaging of lipids and metabolites in tissues by nanospray desorption electrospray ionization mass spectrometry. Anal Bioanal Chem. 2015;407(8):2063-71.
102. Nguyen SN, Liyu AV, Chu RK, Anderton CR, Laskin J. Constant-distance mode nanospray desorption electrospray ionization mass spectrometry imaging of biological samples with complex topography. Analytical Chemistry. 2017;89(2):1131-7.
103. Nguyen S, Sontag R, Carson J, Corley R, Ansong C, Laskin J. Towards high-resolution tissue imaging using nanospray desorption electrospray ionization mass spectrometry coupled to shear force microscopy. Journal of the American Society for Mass Spectrometry. 2018;29(2):316-22.
104. ElNaggar MS, Barbier C, Van Berkel GJ. Liquid microjunction surface sampling probe fluid dynamics: Computational and experimental analysis of coaxial intercapillary positioning effects on sample manipulation. Journal of the American Society for Mass Spectrometry. 2011;22(7):1157-66.

105. Baeuml F, Welsch T. Improvement of the long-term stability of polyimide-coated fused-silica capillaries used in capillary electrophoresis and capillary electrochromatography. J Chromatogr A. 2002;961(1):35-44.
106. Griffiths RL, Randall EC, Race AM, Bunch J, Cooper HJ. Raster-mode continuous-flow liquid microjunction mass spectrometry imaging of proteins in thin tissue sections. Analytical Chemistry. 2017;89(11):5684-8.
107. Feider CL, Elizondo N, Eberlin LS. Ambient ionization and FAIMS mass spectrometry for enhanced imaging of multiply charged molecular ions in biological tissues. Anal Chem. 2016;88(23):11533-41.
108. Rao W, Celiz AD, Scurr DJ, Alexander MR, Barrett DA. Ambient DESI and LESA-MS analysis of proteins adsorbed to a biomaterial surface using in-situ surface tryptic digestion. Journal of the American Society for Mass Spectrometry. 2013;24(12):1927-36.
109. Eikel D, Vavrek M, Smith S, Bason C, Yeh S, Korfmacher WA, et al. Liquid extraction surface analysis mass spectrometry (LESA-MS) as a novel profiling tool for drug distribution and metabolism analysis: The terfenadine example. Rapid Communications in Mass Spectrometry. 2011;25(23):3587-96.
110. Xu LX, Wang TT, Geng YY, Wang WY, Li Y, Duan XK, et al. The direct analysis of drug distribution of rotigotine-loaded microspheres from tissue sections by LESA coupled with tandem mass spectrometry. Anal Bioanal Chem. 2017;409(22):5217-23.
111. Swales JG, Tucker JW, Spreadborough MJ, Iverson SL, Clench MR, Webborn PJH, et al. Mapping drug distribution in brain tissue using liquid extraction surface analysis mass spectrometry imaging. Analytical Chemistry. 2015;87(19):10146-52.
112. Swales JG, Strittmatter N, Tucker JW, Clench MR, Webborn PJH, Goodwin RJA. Spatial quantitation of drugs in tissues using liquid extraction surface analysis mass spectrometry imaging. Sci Rep. 2016;6:9.
113. Griffiths RL, Cooper HJ. Direct tissue profiling of protein complexes: Toward native mass spectrometry imaging. Analytical Chemistry. 2016;88(1):606-9.
114. Griffiths RL, Sisley EK, Lopez-Clavijo AF, Simmonds AL, Styles IB, Cooper HJ. Native mass spectrometry imaging of intact proteins and protein complexes in thin tissue sections. International Journal of Mass Spectrometry. 2017.
115. Mikhailov VA, Griffiths RL, Cooper HJ. Liquid extraction surface analysis for native mass spectrometry: Protein complexes and ligand binding. International Journal of Mass Spectrometry. 2016.
116. Randall EC, Race AM, Cooper HJ, Bunch J. MALDI imaging of liquid extraction surface analysis sampled tissue. Analytical Chemistry. 2016;88(17):8433-40.
117. Randall EC, Bunch J, Cooper HJ. Direct analysis of intact proteins from Escherichia coli colonies by liquid extraction surface analysis mass spectrometry. Analytical Chemistry. 2014;86(21):10504-10.
118. Almeida R, Berzina Z, Arnspang EC, Baumgart J, Vogt J, Nitsch R, et al. Quantitative spatial analysis of the mouse brain lipidome by pressurized liquid extraction surface analysis. Analytical Chemistry. 2015;87(3):1749-56.
119. van Hove ERA, Smith DF, Heeren RMA. A concise review of mass spectrometry imaging. J Chromatogr A. 2010;1217(25):3946-54.
120. Spraggins JM, Caprioli R. High-speed MALDI-TOF imaging mass spectrometry: Rapid ion image acquisition and considerations for next generation instrumentation. Journal of the American Society for Mass Spectrometry. 2011;22(6):1022-31.

121. Prentice BM, Chumbley CW, Caprioli RM. High-speed MALDI MS/MS imaging mass spectrometry using continuous raster sampling. Journal of Mass Spectrometry. 2015;50(4):703-10.
122. Gamble LJ, Anderton CR. Secondary Ion Mass Spectrometry Imaging of Tissues, Cells, and Microbial Systems. Microscopy today. 2016;24(2):24-31.
123. Magee CW, Honig RE. Depth profiling by SIMS: depth resolution, dynamic range and sensitivity. Surf Interface Anal. 1982;4(2):35-41.
124. Baryshev SV, Zinovev AV, Tripa CE, Pellin MJ, Peng Q, Elam JW, et al. High-resolution secondary ion mass spectrometry depth profiling of nanolayers. Rapid Communications in Mass Spectrometry. 2012;26(19):2224-30.
125. Dickinson M, Heard PJ, Barker JHA, Lewis AC, Mallard D, Allen GC. Dynamic SIMS analysis of cryo-prepared biological and geological specimens. Appl Surf Sci. 2006;252(19):6793-6.
126. Kuroda K, Fujiwara T, Imai T, Takama R, Saito K, Matsushita Y, et al. The cryo-TOF-SIMS/SEM system for the analysis of the chemical distribution in freeze-fixed Cryptomeria japonica wood. Surf Interface Anal. 2013;45(1):215-9.
127. Caprioli RM, Farmer TB, Gile J. Molecular imaging of biological samples: Localization of peptides and proteins using MALDI-TOF MS. Analytical Chemistry. 1997;69(23):4751-60.
128. Chaurand P, Norris JL, Cornett DS, Mobley JA, Caprioli RM. New developments in profiling and imaging of proteins from tissue sections by MALDI mass spectrometry. J Proteome Res. 2006;5(11):2889-900.
129. Reyzer ML, Caprioli RM. MALDI-MS-based imaging of small molecules and proteins in tissues. Curr Opin Chem Biol. 2007;11(1):29-35.
130. Chatterji B, Pich A. MALDI imaging mass spectrometry and analysis of endogenous peptides. Expert Rev Proteomics. 2013;10(4):381-8.
131. Norris JL, Caprioli RM. Analysis of tissue specimens by matrix-assisted laser desorption/ionization imaging mass spectrometry in biological and clinical research. Chem Rev. 2013;113(4):2309-42.
132. Griffiths RL, Creese AJ, Race AM, Bunch J, Cooper HJ. LESA FAIMS mass spectrometry for the spatial profiling of proteins from tissue. Anal Chem. 2016;88(13):6758-66.
133. Mamyrin BA. Time-of-flight mass spectrometry (concepts, achievements, and prospects). International Journal of Mass Spectrometry. 2001;206(3):251-66.
134. Boesl U. Time-of-flight mass spectrometry: Introduction to the basics. Mass Spectrom Rev. 2017;36(1):86-109.
135. Wiley WC, McLaren IH. Time-of-flight mass spectrometer with improved resolution. Rev Sci Instrum. 1955;26(12):1150-7.
136. Colby SM, King TB, Reilly JP. Improving the resolution of matrix-assisted laser desorption/ionization time-of-flight mass-spectrometry by exploiting the correlation between ion position and velocity. Rapid Communications in Mass Spectrometry. 1994;8(11):865-8.
137. Mamyrin BA, Karataev VI, Shmikk DV, Zagulin VA. Mass-reflectron a new nonmagnetic time-of-flight high-resolution mass-spectrometer. Zhurnal Eksperimentalnoi Teor Fiz. 1973;64(1):82-9.
138. Paul W, Steinwedel H. Ein neues Massenspektrometer ohne Magnetfeld. Z Naturfors Sect A-J Phys Sci. 1953;8(7):448-50.
139. Paul W. Electromagnetic traps for charged and neutral particles. Rev Mod Phys. 1990;62(3):531-40.

140. Douglas DJ, Frank AJ, Mao DM. Linear ion traps in mass spectrometry. Mass Spectrom Rev. 2005;24(1):1-29.
141. Makarov A. Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis. Analytical Chemistry. 2000;72(6):1156-62.
142. Hu QZ, Noll RJ, Li HY, Makarov A, Hardman M, Cooks RG. The orbitrap: A new mass spectrometer. Journal of Mass Spectrometry. 2005;40(4):430-43.
143. Gys T. Micro-channel plates and vacuum detectors. Nucl Instrum Methods Phys Res Sect A-Accel Spectrom Dect Assoc Equip. 2015;787:254-60.
144. Amster IJ. Fourier transform mass spectrometry. Journal of Mass Spectrometry. 1996;31(12):1325-37.
145. Kanu AB, Dwivedi P, Tam M, Matz L, Hill HH. Ion mobility-mass spectrometry. Journal of Mass Spectrometry. 2008;43(1):1-22.
146. Cumeras R, Figueras E, Davis CE, Baumbach JI, Gracia I. Review on ion mobility spectrometry. Part 1: Current instrumentation. Analyst. 2015;140(5):1376-90.
147. Gabelica V, Marklund E. Fundamentals of ion mobility spectrometry. Curr Opin Chem Biol. 2018;42:51-9.
148. Lanucara F, Holman SW, Gray CJ, Eyers CE. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. Nat Chem. 2014;6(4):281-94.
149. Clemmer DE, Hudgins RR, Jarrold MF. Naked protein conformations - cytochrome c in the gas-phase. J Am Chem Soc. 1995;117(40):10141-2.
150. Buryakov IA, Krylov EV, Nazarov EG, Rasulev UK. A new method of separation of multi-atomic ions by mobility at atmospheric-pressure using a high-frequency amplitude-asymmetric strong electric-field. Int J Mass Spectrom Ion Process. 1993;128(3):143-8.
151. Purves RW, Guevremont R, Day S, Pipich CW, Matyjaszczyk MS. Mass spectrometric characterization of a high-field asymmetric waveform ion mobility spectrometer. Rev Sci Instrum. 1998;69(12):4094-105.
152. Barnett DA, Ells B, Guevremont R. Evaluation of carrier gases for use in high-field asymmetric waveform ion mobility spectrometry. Journal of the American Society for Mass Spectrometry. 2000;11(12):1125-33.
153. Shvartsburg AA, Tang KQ, Smith RD. Understanding and designing field asymmetric waveform ion mobility spectrometry separations in gas mixtures. Analytical Chemistry. 2004;76(24):7366-74.
154. Guevremont R, Barnett DA, Purves RW, Viehland LA. Calculation of ion mobilities from electrospray ionization high-field asymmetric waveform ion mobility spectrometry mass spectrometry. J Chem Phys. 2001;114(23):10270-7.
155. Shvartsburg AA, Bryskiewicz T, Purves RW, Tang KQ, Guevremont R, Smith RD. Field asymmetric waveform ion mobility spectrometry studies of proteins: Dipole alignment in ion mobility spectrometry? J Phys Chem B. 2006;110(43):21966-80.
156. Shvartsburg AA, Noskov SY, Purves RW, Smith RD. Pendular proteins in gases and new avenues for characterization of macromolecules by ion mobility spectrometry. Proc Natl Acad Sci U S A. 2009;106(16):6495-500.
157. Shvartsburg AA. Ultrahigh-resolution differential ion mobility separations of conformers for proteins above 10 kDa: Onset of dipole alignment? Analytical Chemistry. 2014;86(21):10608-15.
158. Levin DS, Vouros P, Miller RA, Nazarov EG, Morris JC. Characterization of gas-phase molecular interactions on differential mobility ion behavior utilizing an electrospray

ionization-differential mobility-mass spectrometer system. Analytical Chemistry. 2006;78(1):96-106.
159. Krylov EV, Nazarov EG. Electric field dependence of the ion mobility. International Journal of Mass Spectrometry. 2009;285(3):149-56.
160. Chen YL, Collings BA, Douglas DJ. Collision cross sections of myoglobin and cytochrome c ions with Ne, Ar, and Kr. Journal of the American Society for Mass Spectrometry. 1997;8(7):681-7.
161. Schneider BB, Covey TR, Coy SL, Krylov EV, Nazarov EG. Chemical effects in the separation process of a differential mobility/mass spectrometer system. Analytical Chemistry. 2010;82(5):1867-80.
162. Liu C, Le Blanc JCY, Schneider BB, Shields J, Federico JJ, Zhang H, et al. Assessing Physicochemical Properties of Drug Molecules via Microsolvation Measurements with Differential Mobility Spectrometry. ACS Central Sci. 2017;3(2):101-9.
163. Walker SWC, Anwar A, Psutka JM, Crouse J, Liu C, Le Blanc JCY, et al. Determining molecular properties with differential mobility spectrometry and machine learning. Nat Commun. 2018;9(1):5096.
164. Zhang YY, Fonslow BR, Shan B, Baek MC, Yates JR. Protein analysis by shotgun/bottom-up proteomics. Chem Rev. 2013;113(4):2343-94.
165. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003;422(6928):198-207.
166. Anderson NL, Hofmann JP, Gemmell A, Taylor J. Global approaches to quantitative-analysis of gene-expression patterns observed by use of two-dimensional gel-electrophoresis. Clin Chem. 1984;30(12):2031-6.
167. Hunt DF, Henderson RA, Shabanowitz J, Sakaguchi K, Michel H, Sevilir N, et al. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. Science. 1992;255(5049):1261-3.
168. Brodbelt JS. Ion activation methods for peptides and proteins. Analytical Chemistry. 2016;88(1):30-51.
169. Catherman AD, Skinner OS, Kelleher NL. Top down proteomics: Facts and perspectives. Biochem Biophys Res Commun. 2014;445(4):683-93.
170. Toby TK, Fornelli L, Kelleher NL. Progress in top-down proteomics and the analysis of proteoforms. In: Bohn PW, Pemberton JE, editors. Annual Review of Analytical Chemistry, Vol 9. Annual Review of Analytical Chemistry. 9. Palo Alto: Annual Reviews; 2016. p. 499-519.
171. LeDuc RD, Taylor GK, Kim YB, Januszyk TE, Bynum LH, Sola JV, et al. ProSight PTM: An integrated environment for protein identification and characterization by top-down mass spectrometry. Nucleic Acids Res. 2004;32:W340-W5.
172. Liu X, Sirotkin Y, Shen Y, Anderson G, Tsai YS, Ting YS, et al. Protein identification using top-down spectra. Mol Cell Proteomics. 2012;11(6):M111 008524.
173. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999;20(18):3551-67.
174. Chernushevich IV, Loboda AV, Thomson BA. An introduction to quadrupole-time-of-flight mass spectrometry. Journal of Mass Spectrometry. 2001;36(8):849-65.
175. Makarov A, Denisov E, Lange O. Performance evaluation of a high-field orbitrap mass analyzer. Journal of the American Society for Mass Spectrometry. 2009;20(8):1391-6.
176. Hao ZQ, Zhang T, Xuan Y, Wang HX, Qian J, Lin SH, et al. Intact Antibody Characterization Using Orbitrap Mass Spectrometry. In: Schiel JE, Davis DL, Borisov OV,

editors. State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization, Vol 3: Defining the Next Generation of Analytical and Biophysical Techniques. ACS Symposium Series. 1202. Washington: Amer Chemical Soc; 2015. p. 289-315.
177. Kelleher NL, Lin HY, Valaskovic GA, Aaserud DJ, Fridriksson EK, McLafferty FW. Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. J Am Chem Soc. 1999;121(4):806-12.
178. Katta V, Chait BT. Observation of the heme globin complex in native myoglobin by electrospray-ionization mass-spectrometry. J Am Chem Soc. 1991;113(22):8534-5.
179. Ganem B, Li YT, Henion JD. Detection of noncovalent receptor ligand complexes by mass-spectrometry. J Am Chem Soc. 1991;113(16):6294-6.
180. Leney AC, Heck AJR. Native mass spectrometry: What is in the name? Journal of the American Society for Mass Spectrometry. 2017;28(1):5-13.
181. van de Waterbeemd MV, Fort KL, Boll D, Reinhardt-Szyba M, Routh A, Makarov A, et al. High-fidelity mass analysis unveils heterogeneity in intact ribosomal particles. Nat Methods. 2017;14(3):283-+.
182. Olsen JV, Mann M. Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. Proc Natl Acad Sci U S A. 2004;101(37):13417-22.
183. McLuckey SA. Principles of collisional activation in analytical mass-spectrometry. Journal of the American Society for Mass Spectrometry. 1992;3(6):599-614.
184. Wells JM, McLuckey SA. Collision-induced dissociation (CID) of peptides and proteins. In: Burlingame AL, editor. Biological Mass Spectrometry. Methods in Enzymology. 402. San Diego: Elsevier Academic Press Inc; 2005. p. 148-85.
185. Olsen JV, Macek B, Lange O, Makarov A, Horning S, Mann M. Higher-energy C-trap dissociation for peptide modification analysis. Nat Methods. 2007;4(9):709-12.
186. Sleno L, Volmer DA. Ion activation methods for tandem mass spectrometry. Journal of Mass Spectrometry. 2004;39(10):1091-112.
187. Jones AW, Cooper HJ. Dissociation techniques in mass spectrometry-based proteomics. Analyst. 2011;136(17):3419-29.
188. Roepstorff P, Fohlman J. Proposal for a common nomenclature for sequence ions in mass-spectra of peptides. Biomedical Mass Spectrometry. 1984;11(11):601-.
189. Summerfield SG, Whiting A, Gaskell SJ. Intra-ionic interactions in electrosprayed peptide ions. International Journal of Mass Spectrometry and Ion Processes. 1997;162(1-3):149-61.
190. Wysocki VH, Tsaprailis G, Smith LL, Breci LA. Special feature: Commentary - mobile and localized protons: A framework for understanding peptide dissociation. Journal of Mass Spectrometry. 2000;35(12):1399-406.
191. Lane N. The unseen world: reflections on Leeuwenhoek (1677) 'Concerning little animals'. Philos Trans R Soc B-Biol Sci. 2015;370(1666):10.
192. Gest H. The remarkable vision of Robert Hooke (1635-1703) - first observer of the microbial world. Perspect Biol Med. 2005;48(2):266-72.
193. Silhavy TJ, Kahne D, Walker S. The bacterial cell envelope. Cold Spring Harb Perspect Biol. 2010;2(5):a000414.
194. Sutcliffe IC. A phylum level perspective on bacterial cell envelope architecture. Trends Microbiol. 2010;18(10):464-70.
195. Vollmer W, Blanot D, de Pedro MA. Peptidoglycan structure and architecture. Fems Microbiol Rev. 2008;32(2):149-67.
196. Lesage G, Bussey H. Cell wall assembly in *Saccharomyces cerevisiae*. Microbiol Mol Biol Rev. 2006;70(2):317-+.

197. Orlean P. Architecture and biosynthesis of the *Saccharomyces cerevisiae* cell wall. Genetics. 2012;192(3):775-+.
198. Cassagne C, Cella AL, Suchon P, Normand AC, Ranque S, Piarroux R. Evaluation of four pretreatment procedures for MALDI-TOF MS yeast identification in the routine clinical laboratory. Med Mycol. 2013;51(4):371-7.
199. Costerton JW, Irvin RT, Cheng KJ. The bacterial glycocalyx in nature and disease. Annu Rev Microbiol. 1981;35:299-324.
200. Costerton JW, Lewandowski Z, Caldwell DE, Korber DR, Lappinscott HM. Microbial biofilms. Annu Rev Microbiol. 1995;49:711-45.
201. Davies DG, Chakrabarty AM, Geesey GG. Exopolysaccharide production in biofilms - substratum activation of alginate gene-expression by *Pseudomonas aeruginosa*. Appl Environ Microbiol. 1993;59(4):1181-6.
202. Mah TFC, O'Toole GA. Mechanisms of biofilm resistance to antimicrobial agents. Trends Microbiol. 2001;9(1):34-9.
203. Lewis K. Riddle of biofilm resistance. Antimicrob Agents Chemother. 2001;45(4):999-1007.
204. Davies D. Understanding biofilm resistance to antibacterial agents. Nat Rev Drug Discov. 2003;2(2):114-22.
205. Hoiby N, Bjarnsholt T, Givskov M, Molin S, Ciofu O. Antibiotic resistance of bacterial biofilms. Int J Antimicrob Agents. 2010;35(4):322-32.
206. Reiner E. Identification of bacterial strains by pyrolysis - gas-liquid chromatography. Nature. 1965;206(4990):1272-&.
207. Reiner E. Studies on differentiation of microorganisms by pyrolysis-gas-liquid chromatography. Journal of Gas Chromatography. 1967;5(2):65-&.
208. Oyama VI, Carle GC. Pyrolysis gas chromatography application to life detection and chemotaxonomy. Journal of Gas Chromatography. 1967;5(3):151-&.
209. Reiner E, Ewing WH. Chemotaxonomic studies of some Gram negative bacteria by means of pyrolysis-gas-liquid chromatography. Nature. 1968;217(5124):191-&.
210. Reiner E, Kubica GP. Predictive value of pyrolysis-gas-liquid chromatography in differentiation of mycobacteria. Am Rev Respir Dis. 1969;99(1):42-&.
211. Reiner E, Beam RE, Kubica GP. Pyrolysis-gas-liquid chromatography studies for classification of mycobacteria. Am Rev Respir Dis. 1969;99(5):750-&.
212. Reiner E, Hicks JJ, Beam RE, David HL. Recent studies on mycobacterial differentiation by means of pyrolysis-gas-liquid chromatography. Am Rev Respir Dis. 1971;104(5):656-&.
213. Reiner E, Ball MM, Hicks JJ, Martin WJ. Rapid characterization of salmonella organisms by means of pyrolysis gas-liquid chromatography. Analytical Chemistry. 1972;44(6):1058-&.
214. Meuzelaar HL, Intveld RA. Technique for curie-point pyrolysis-gas chromatography of complex biological samples. J Chromatogr Sci. 1972;10(4):213-+.
215. Haddadin JM, Stirland RM, Preston NW, Collard P. Identification of *Vibrio cholerae* by pyrolysis gas-liquid chromatography. Applied Microbiology. 1973;25(1):40-3.
216. Schulten HR, Beckey HD, Meuzelaar HL, Boerboom AJ. High-resolution field ionization mass-spectrometry of bacterial pyrolysis products. Analytical Chemistry. 1973;45(1):191-5.
217. Simmonds PG. Whole microorganisms studied by pyrolysis-gas chromatography-mass spectrometry - significance for extraterrestrial life detection experiments. Applied Microbiology. 1970;20(4):567-&.

218. Meuzelaa.Hl, Kistemak.Pg. Technique for fast and reproducible fingerprinting of bacteria by pyrolysis mass-spectrometry. Analytical Chemistry. 1973;45(3):587-90.
219. Anhalt JP, Fenselau C. Identification of bacteria using mass-spectrometry. Analytical Chemistry. 1975;47(2):219-25.
220. Heller DN, Cotter RJ, Fenselau C. Profiling of bacteria by fast-atom-bombardment mass-spectrometry. Analytical Chemistry. 1987;59(23):2806-9.
221. Gutteridge CS, Puckey DJ. Discrimination of some Gram-negative bacteria by direct probe mass-spectrometry. J Gen Microbiol. 1982;128(APR):721-30.
222. Shaw BG, Puckey DJ, Macfie HJH, Bolt SJ. Classification of some lactic-acid bacteria from vacuum-packed meats by direct probe mass-spectrometry. J Appl Bacteriol. 1985;59(2):157-65.
223. Claydon MA, Davey SN, EdwardsJones V, Gordon DB. The rapid identification of intact microorganisms using mass spectrometry. Nature Biotechnology. 1996;14(11):1584-6.
224. Holland RD, Wilkes JG, Rafii F, Sutherland JB, Persons CC, Voorhees KJ, et al. Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry. 1996;10(10):1227-32.
225. Demirev PA, Ho YP, Ryzhov V, Fenselau C. Microorganism identification by mass spectrometry and protein database searches. Analytical Chemistry. 1999;71(14):2732-8.
226. Fenselau C, Demirev PA. Characterization of intact microorganisms by MALDI mass spectrometry. Mass Spectrom Rev. 2001;20(4):157-71.
227. Krishnamurthy T, Ross PL, Rajamani U. Detection of pathogenic and non-pathogenic bacteria by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry. 1996;10(8):883-8.
228. Krishnamurthy T, Ross PL. Rapid identification of bacteria by direct matrix-assisted laser desorption/ionization mass spectrometric analysis of whole cells. Rapid Communications in Mass Spectrometry. 1996;10(15):1992-6.
229. Egli A, Tschudin-Sutter S, Oberle M, Goldenberger D, Frei R, Widmer AF. Matrix-assisted laser desorption/ionization time of flight mass-spectrometry (MALDI-TOF MS) based typing of extended-spectrum beta-lactamase producing *E. coli* - a novel tool for real-time outbreak investigation. Plos One. 2015;10(4):6.
230. Valentine N, Wunschel S, Wunschel D, Petersen C, Wahl K. Effect of culture conditions on microorganism identification by matrix-assisted laser desorption ionization mass spectrometry. Appl Environ Microbiol. 2005;71(1):58-64.
231. Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, et al. Ongoing revolution in bacteriology: Routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clinical Infectious Diseases. 2009;49(4):543-51.
232. Dubois D, Leyssene D, Chacornac JP, Kostrzewa M, Schmit PO, Talon R, et al. Identification of a variety of *Staphylococcus* species by matrix-assisted laser desorption ionization-time of flight mass spectrometry. Journal of Clinical Microbiology. 2010;48(3):941-5.
233. Bizzini A, Durussel C, Bille J, Greub G, Prod'hom G. Performance of matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of bacterial strains routinely isolated in a clinical microbiology laboratory. Journal of Clinical Microbiology. 2010;48(5):1549-54.

234. Prod'hom G, Bizzini A, Durussel C, Bille J, Greub G. Matrix-assisted laser desorption ionization-time of flight mass spectrometry for direct bacterial identification from positive blood culture pellets. Journal of Clinical Microbiology. 2010;48(4):1481-3.
235. Cherkaoui A, Hibbs J, Emonet S, Tangomo M, Girard M, Francois P, et al. Comparison of two matrix-assisted laser desorption ionization-time of flight mass spectrometry methods with conventional phenotypic identification for routine identification of bacteria to the species level. Journal of Clinical Microbiology. 2010;48(4):1169-75.
236. Saffert RT, Cunningham SA, Ihde SM, Jobe KEM, Mandrekar J, Patel R. Comparison of Bruker Biotyper matrix-assisted laser desorption ionization-time of flight mass spectrometer to BD Phoenix automated microbiology system for identification of Gram-negative bacilli. Journal of Clinical Microbiology. 2011;49(3):887-92.
237. Santos T, Capelo JL, Santos HM, Oliveira I, Marinho C, Goncalves A, et al. Use of MALDI-TOF mass spectrometry fingerprinting to characterize *Enterococcus* spp. and *Escherichia coli* isolates. Journal of Proteomics. 2015;127:321-31.
238. Singhal N, Kumar M, Kanaujia PK, Virdi JS. MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. Frontiers in Microbiology. 2015;6:16.
239. Angeletti S, Dicuonzo G, Avola A, Crea F, Dedej E, Vailati F, et al. Viridans group streptococci clinical isolates: MALDI-TOF mass spectrometry versus gene sequence-based identification. Plos One. 2015;10(3):12.
240. Girard V, Mailler S, Welker M, Arsac M, Celliere B, Cotte-Pattat PJ, et al. Identification of *Mycobacterium* spp. and *Nocardia* spp. from solid and liquid cultures by matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS). Diagn Microbiol Infect Dis. 2016;86(3):277-83.
241. Wieser A, Schubert S. MALDI-TOF MS entering the microbiological diagnostic laboratory - from fast identification to resistance testing. Trac-Trends Anal Chem. 2016;84:80-7.
242. Angeletti S. Matrix assisted laser desorption time of flight mass spectrometry (MALDI-TOF MS) in clinical microbiology. J Microbiol Methods. 2017;138:20-9.
243. Levesque S, Dufresne PJ, Soualhine H, Domingo MC, Bekal S, Lefebvre B, et al. A side by side comparison of Bruker Biotyper and VITEK MS: Utility of MALDI-TOF MS technology for microorganism identification in a public health reference laboratory. Plos One. 2015;10(12):21.
244. TeKippe EM, Burnham CAD. Evaluation of the Bruker Biotyper and VITEK MS MALDI-TOF MS systems for the identification of unusual and/or difficult-to-identify microorganisms isolated from clinical specimens. European Journal of Clinical Microbiology & Infectious Diseases. 2014;33(12):2163-71.
245. Anderton CR, Chu RK, Tolic N, Creissen A, Pasa-Tolic L. Utilizing a robotic sprayer for high lateral and mass resolution MALDI FT-ICR msi of microbial cultures. Journal of the American Society for Mass Spectrometry. 2016;27(3):556-9.
246. Hoffmann T, Dorrestein PC. Homogeneous matrix deposition on dried agar for MALDI imaging mass spectrometry of microbial cultures. Journal of the American Society for Mass Spectrometry. 2015;26(11):1959-62.
247. Debois D, Ongena M, Cawoy H, De Pauw E. MALDI-FTICR MS imaging as a powerful tool to identify paenibacillus antibiotics involved in the inhibition of plant pathogens. Journal of the American Society for Mass Spectrometry. 2013;24(8):1202-13.
248. Yang JY, Phelan VV, Simkovsky R, Watrous JD, Trial RM, Fleming TC, et al. Primer on agar-based microbial imaging mass spectrometry. J Bacteriol. 2012;194(22):6023-8.

249. Moree WJ, Phelan VV, Wu CH, Bandeira N, Cornett DS, Duggan BM, et al. Interkingdom metabolic transformations captured by microbial imaging mass spectrometry. Proc Natl Acad Sci U S A. 2012;109(34):13811-6.
250. Figueroa M, Jarmusch AK, Raja HA, El-Elimat T, Kavanaugh JS, Horswill AR, et al. Polyhydroxyanthraquinones as quorum sensing inhibitors from the guttates of *Penicillium restrictum* and their analysis by desorption electrospray ionization mass spectrometry. J Nat Prod. 2014;77(6):1351-8.
251. Tata A, Perez C, Campos ML, Bayfield MA, Eberlin MN, Ifa DR. Imprint desorption electrospray ionization mass spectrometry imaging for monitoring secondary metabolites production during antagonistic interaction of fungi. Analytical Chemistry. 2015;87(24):12298-304.
252. Watrous J, Hendricks N, Meehan M, Dorrestein PC. Capturing bacterial metabolic exchange using thin film desorption electrospray ionization-imaging mass spectrometry. Analytical Chemistry. 2010;82(5):1598-600.
253. Meetani MA, Shin YS, Zhang SF, Mayer R, Basile F. Desorption electrospray ionization mass spectrometry of intact bacteria. Journal of Mass Spectrometry. 2007;42(9):1186-93.
254. Song YS, Talaty N, Tao WA, Pan ZZ, Cooks RG. Rapid ambient mass spectrometric profiling of intact, untreated bacteria using desorption electrospray ionization. Chem Commun. 2007(1):61-3.
255. Jackson AU, Werner SR, Talaty N, Song Y, Campbell K, Cooks RG, et al. Targeted metabolomic analysis of *Escherichia coli* by desorption electrospray ionization and extractive electrospray ionization mass spectrometry. Anal Biochem. 2008;375(2):272-81.
256. Song Y, Talaty N, Datsenko K, Wanner BL, Cooks RG. In vivo recognition of *Bacillus subtilis* by desorption electrospray ionization mass spectrometry (DESI-MS). Analyst. 2009;134(5):838-41.
257. Sica VP, Raja HA, El-Elimat T, Oberlies NH. Mass spectrometry imaging of secondary metabolites directly on fungal cultures. RSC Adv. 2014;4(108):63221-7.
258. Angolini CFF, Vendramini PH, Araujo FDS, Araujo WL, Augusti R, Eberlin MN, et al. Direct protocol for ambient mass spectrometry imaging on agar culture. Analytical Chemistry. 2015;87(13):6925-30.
259. Watrous J, Roach P, Alexandrov T, Heath BS, Yang JY, Kersten RD, et al. Mass spectral molecular networking of living microbial colonies. Proc Natl Acad Sci U S A. 2012;109(26):E1743-E52.
260. Rath CM, Alexandrov T, Higginbottom SK, Song J, Milla ME, Fischbach MA, et al. Molecular analysis of model gut microbiotas by imaging mass spectrometry and nanodesorption electrospray ionization reveals dietary metabolite transformations. Analytical Chemistry. 2012;84(21):9259-67.
261. Rath CM, Yang JY, Alexandrov T, Dorrestein PC. Data-independent microbial metabolomics with ambient ionization mass spectrometry. Journal of the American Society for Mass Spectrometry. 2013;24(8):1167-76.
262. Traxler MF, Watrous JD, Alexandrov T, Dorrestein PC, Kolter R. Interspecies interactions stimulate diversification of the streptomyces coelicolor secreted metabolome. mBio. 2013;4(4):12.
263. Hsu CC, ElNaggar MS, Peng Y, Fang JS, Sanchez LM, Mascuch SJ, et al. Real-time metabolomics on living microorganisms using ambient electrospray ionization flow-probe. Analytical Chemistry. 2013;85(15):7014-8.

264. Sica VP, Raja HA, El-Elimat T, Kertesz V, Van Berkel GJ, Pearce CJ, et al. Dereplicating and spatial mapping of secondary metabolites from fungal cultures in situ. J Nat Prod. 2015;78(8):1926-36.
265. Sica VP, Rees ER, Tchegnon E, Bardsley RH, Raja HA, Oberlies NH. Spatial and temporal profiling of griseofulvin production in Xylaria cubensis using mass spectrometry mapping. Frontiers in Microbiology. 2016;7:14.
266. Sarsby J, Griffiths RL, Race AM, Bunch J, Randall EC, Creese AJ, et al. Liquid extraction surface analysis mass spectrometry coupled with field asymmetric waveform ion mobility spectrometry for analysis of intact proteins from biological substrates. Analytical Chemistry. 2015;87(13):6794-800.
267. Sale AJH, Hamilton WA. Effects of high electric fields on microorganisms. I. Killing of bacteria and yeasts. Biochimica Et Biophysica Acta. 1967;148(3):781-&.
268. Hamilton WA, Sale AJH. Effects of high electric fields on microorganisms. 2. Mechanism of action of lethal effect. Biochimica Et Biophysica Acta. 1967;148(3):789-&.
269. Sale AJH, Hamilton WA. Effects of high electric fields on micro-organisms. 3. Lysis of erythrocytes and protoplasts. Biochimica Et Biophysica Acta. 1968;163(1):37-&.
270. Haberl S, Miklavcic D, Sersa G, Frey W, Rubinsky B. Cell membrane electroporation-part 2: The applications. IEEE Electr Insul Mag. 2013;29(1):29-37.
271. Jiang CL, Davalos RV, Bischof JC. A review of basic to clinical studies of irreversible electroporation therapy. IEEE Trans Biomed Eng. 2015;62(1):4-20.
272. Hsiao CY, Huang KW. Irreversible electroporation: A novel ultrasound-guided modality for non-thermal tumor ablation. J Med Ultrasound. 2017;25(4):195-200.
273. Vroomen L, Petre EN, Cornelis FH, Solomon SB, Srimathveeravalli G. Irreversible electroporation and thermal ablation of tumors in the liver, lung, kidney and bone: What are the differences? Diagn Interv Imaging. 2017;98(9):609-17.
274. Lyu TC, Wang XF, Su ZL, Shangguan JJ, Sun C, Figini M, et al. Irreversible electroporation in primary and metastatic hepatic malignancies a review. Medicine (Baltimore). 2017;96(17):7.
275. Tasu JP, Vesselle G, Herpe G, Richer JP, Boucecbi S, Velasco S, et al. Irreversible electroporation for locally advanced pancreatic cancer where do we stand in 2017? Pancreas. 2017;46(3):283-7.
276. Tian G, Zhao QY, Chen F, Jiang TN, Wang WL. Ablation of hepatic malignant tumors with irreversible electroporation: A systematic review and meta-analysis of outcomes. Oncotarget. 2017;8(4):5853-60.
277. Ansari D, Kristoffersson S, Andersson R, Bergenfeldt M. The role of irreversible electroporation (IRE) for locally advanced pancreatic cancer: A systematic review of safety and efficacy. Scand J Gastroenterol. 2017;52(11):1165-71.
278. Sugrue A, Maor E, Ivorra A, Vaidya V, Witt C, Kapa S, et al. Irreversible electroporation for the treatment of cardiac arrhythmias. Expert Rev Cardiovasc Ther. 2018;16(5):349-60.
279. Bakonyi M, Berko S, Eros G, Varju G, Dehelean CA, Budai-Szucs M, et al. A review of electroporation-based antitumor skin therapies and investigation of betulinic acid-loaded ointment. Anti-Cancer Agents Med Chem. 2018;18(5):693-701.
280. Kraynyak KA, Bodles-Brakhop A, Bagarazzi M. Tapping the potential of DNA delivery with electroporation for cancer immunotherapy. In: Savelyeva N, Ottensmeier C, editors. Cancer Vaccines. Current Topics in Microbiology and Immunology. 405. Cham: Springer International Publishing Ag; 2017. p. 55-78.

281. Kotnik T, Kramar P, Pucihar G, Miklavcic D, Tarek M. Cell membrane electroporation-part 1: The phenomenon. IEEE Electr Insul Mag. 2012;28(5):14-23.
282. Tieleman DP, Leontiadou H, Mark AE, Marrink SJ. Simulation of pore formation in lipid bilayers by mechanical stress and electric fields. J Am Chem Soc. 2003;125(21):6382-3.
283. Tarek M. Membrane electroporation: A molecular dynamics simulation. Biophys J. 2005;88(6):4045-53.
284. Tieleman DP, Berendsen HJC, Sansom MSP. Voltage-dependent insertion of alamethicin at phospholipid/water and octane/water interfaces. Biophys J. 2001;80(1):331-46.
285. Melikov KC, Frolov VA, Shcherbakov A, Samsonov AV, Chizmadzhev YA, Chernomordik LV. Voltage-induced nonconductive pre-pores and metastable single pores in unmodified planar lipid bilayer. Biophys J. 2001;80(4):1829-36.
286. Koronkiewicz S, Kalinowski S, Bryl K. Programmable chronopotentiometry as a tool for the study of electroporation and resealing of pores in bilayer lipid membranes. Biochim Biophys Acta-Biomembr. 2002;1561(2):222-9.
287. Deipolyi AR, Golberg A, Yarmush ML, Arellano RS, Oklu R. Irreversible electroporation: Evolution of a laboratory technique in interventional oncology. Diagn Interv Radiol. 2014;20(2):147-54.
288. Rebersek M, Miklavcic D, Bertacchini C, Sack M. Cell membrane electroporation-part 3: The equipment. IEEE Electr Insul Mag. 2014;30(3):8-18.
289. Horowitz P, Hill W. The art of electronics. 3rd ed: Cambridge University Press; 2015.
290. Ulaby FT, Ravaioli U, Michielssen E. Fundamentals of applied electromagnetics. 7th ed: Pearson; 2014.
291. Lander CW. Power electronics. 3rd ed: McGraw-Hill Education; 1993.
292. Ho J, Jow TR, Boggs S. Historical introduction to capacitor technology. IEEE Electr Insul Mag. 2010;26(1):20-5.
293. Gurdak E, Green FM, Rakowska PD, Seah MP, Salter TL, Gilmore IS. VAMAS interlaboratory study for desorption electrospray ionization mass spectrometry (DESI MS) intensity repeatability and constancy. Analytical Chemistry. 2014;86(19):9603-11.
294. Gilmore IS, Seah MP. Static SIMS - surface-charge stabilization of insulators for highly repeatable spectra when using a quadrupole mass-spectrometer. Surf Interface Anal. 1995;23(4):191-203.
295. Cheng KD, Chui HX, Domish L, Hernandez D, Wang GH. Recent development of mass spectrometry and proteomics applications in identification and typing of bacteria. Proteomics Clinical Applications. 2016;10(4):346-57.
296. Zhou ML, Yang QW, Kudinha T, Zhang L, Xiao M, Kong FR, et al. Using matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) complemented with selected 16S rRNA and GyrB genes sequencing to practically identify clinical important viridans group streptococci (VGS). Frontiers in Microbiology. 2016;7.
297. Williams TL, Monday SR, Feng PC, Musser SM. Identifying new pcr targets for pathogenic bacteria using top-down LC/MS protein discovery. J Biomol Tech. 2005;16(2):134-42.
298. Fidel PL, Vazquez JA, Sobel JD. *Candida glabrata*: Review of epidemiology, pathogenesis, and clinical disease with comparison to c-albicans. Clin Microbiol Rev. 1999;12(1):80-+.
299. Csank C, Haynes K. *Candida glabrata* displays pseudohyphal growth. FEMS Microbiol Lett. 2000;189(1):115-20.
300. Lachke SA, Joly S, Daniels K, Soll DR. Phenotypic switching and filamentation in *Candida glabrata*. Microbiology-(UK). 2002;148:2661-74.

301. Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, et al. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature. 2000;406(6799):959-64.
302. Yeats C, Bateman A. The BON domain: A putative membrane-binding domain. Trends BiochemSci. 2003;28(7):352-5.
303. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The Proteomics Protocols Handbook. Totowa, NJ: Humana Press; 2005. p. 571-607.
304. Peschel A, Otto M. Phenol-soluble modulins and staphylococcal infection. Nature Reviews Microbiology. 2013;11(10):667-73.
305. Raghunathan G, Seetharamulu P, Brooks BR, Guy HR. Models of delta-hemolysin membrane channels and crystal-structures. Proteins. 1990;8(3):213-25.
306. Wang R, Braughton KR, Kretschmer D, Bach THL, Queck SY, Li M, et al. Identification of novel cytolytic peptides as key virulence determinants for community-associated MRSA. Nat Med. 2007;13(12):1510-4.
307. Holden MTG, Feil EJ, Lindsay JA, Peacock SJ, Day NPJ, Enright MC, et al. Complete genomes of two clinical *Staphylococcus aureus* strains: Evidence for the rapid evolution of virulence and drug resistance. Proc Natl Acad Sci U S A. 2004;101(26):9786-91.
308. Gillaspy AF, Worrell V, Orvis J, Roe BA, Dyer DW, Iandolo JJ. The *Staphylococcus aureus* NCTC 8325 genome. In: Fischetti V, Novick R, Ferretti J, Portnoy D, Rood J, editors. Gram positive pathogens. 2 ed. Washington D.C.: ASM Press; 2006. p. 381-412.
309. Babu M, Arnold R, Bundalovic-Torma C, Gagarinova A, Wong KS, Kumar A, et al. Quantitative genome-wide genetic interaction screens reveal global epistatic relationships of protein complexes in *Escherichia coli*. PLoS Genet. 2014;10(2):15.
310. Polissi A, De Laurentis W, Zangrossi S, Briani F, Longhi V, Pesole G, et al. Changes in Escherichia coli transcriptome during acclimatization at low temperature. Res Microbiol. 2003;154(8):573-80.
311. Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, et al. The complete genome sequence of *Escherichia coli* K-12. Science. 1997;277(5331):1453-&.
312. Oogai Y, Matsuo M, Hashimoto M, Kato F, Sugai M, Komatsuzawa H. Expression of virulence factors by *Staphylococcus aureus* grown in serum. Appl Environ Microbiol. 2011;77(22):8097-105.
313. Wittmannliebold B, Pannenbecker R. Primary structure of protein L33 from large subunit of *Escherichia coli* ribosome. FEBS Lett. 1976;68(1):115-8.
314. Link AJ, Robison K, Church GM. Comparing the predicted and observed properties of proteins encoded in the genome of *Escherichia coli* k-12. Electrophoresis. 1997;18(8):1259-313.
315. Agafonov DE, Spirin AS. The ribosome-associated inhibitor a reduces translation errors. Biochem Biophys Res Commun. 2004;320(2):354-8.
316. Vila-Sanjurjo A, Schuwirth BS, Hau CW, Cate JHD. Structural basis for the control of translation initiation during stress. Nat Struct Mol Biol. 2004;11(11):1054-9.
317. Pasloske BL, Finlay BB, Paranchych W. Cloning and sequencing of the *Pseudomonas aeruginosa* PAK pilin gene. FEBS Lett. 1985;183(2):413-6.
318. Islam MS, Aryasomayajula A, Selvaganapathy PR. A Review on Macroscale and Microscale Cell Lysis Methods. Micromachines. 2017;8(3):27.
319. Van der Paal J, Neyts EC, Verlackt CCW, Bogaerts A. Effect of lipid peroxidation on membrane permeability of cancer and normal cells subjected to oxidative stress. Chem Sci. 2016;7(1):489-98.

320. Yusupov M, Bogaerts A, Huygh S, Snoeckx R, van Duin ACT, Neyts EC. Plasma-Induced Destruction of Bacterial Cell Wall Components: A Reactive Molecular Dynamics Simulation. J Phys Chem C. 2013;117(11):5993-8.
321. Wang M, Holmes B, Cheng XQ, Zhu W, Keidar M, Zhang LG. Cold Atmospheric Plasma for Selectively Ablating Metastatic Breast Cancer Cells. Plos One. 2013;8(9):11.
322. Dezest M, Chavatte L, Bourdens M, Quinton D, Camus M, Garrigues L, et al. Mechanistic insights into the impact of Cold Atmospheric Pressure Plasma on human epithelial cell lines. Sci Rep. 2017;7:17.
323. Turrini E, Laurita R, Stancampiano A, Catanzaro E, Calcabrini C, Maffei F, et al. Cold Atmospheric Plasma Induces Apoptosis and Oxidative Stress Pathway Regulation in T-Lymphoblastoid Leukemia Cells. Oxidative Med Cell Longev. 2017:13.
324. Welker S, Rudolph B, Frenzel E, Hagn F, Liebisch G, Schmitz G, et al. Hsp12 is an intrinsically unstructured stress protein that folds upon membrane association and modulates membrane function. Mol Cell. 2010;39(4):507-20.
325. Dang NX, Hincha DK. Identification of two hydrophilins that contribute to the desiccation and freezing tolerance of yeast (*Saccharomyces cerevisiae*) cells. Cryobiology. 2011;62(3):188-93.
326. Green R, Lesage G, Sdicu AM, Menard P, Bussey H. A synthetic analysis of the *Saccharomyces cerevisiae* stress sensor Mid2p, and identification of a Mid2p-interacting protein, ZEO1p, that modulates the PKC1-MPK1 cell integrity pathway. Microbiology-(UK). 2003;149:2487-99.
327. Sandbaken MG, Culbertson MR. Mutations in elongation-factor EF-1-alpha affect the frequency of frameshifting and amino-acid misincorporation in *Saccharomyces cerevisiae*. Genetics. 1988;120(4):923-34.
328. Sakaki K, Tashiro K, Kuhara S, Mihara K. Response of genes associated with mitochondrial function to mild heat stress in yeast *Saccharomyces cerevisiae*. J Biochem. 2003;134(3):373-84.

APPENDIX 1

Publications

K.I. Kocurek, L. Stones, J. Bunch, R.C. May and H.J. Cooper. Top-down LESA mass spectrometry protein analysis of Gram-positive and Gram-negative bacteria. *JASMS* 2017, 28: 2066–2077.

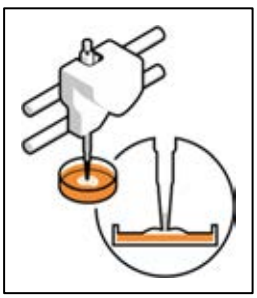
J. Am. Soc. Mass Spectrom. (2017)
DOI: 10.1007/s13361-017-1718-8

RESEARCH ARTICLE

Top-Down LESA Mass Spectrometry Protein Analysis of Gram-Positive and Gram-Negative Bacteria

Klaudia I. Kocurek,[1] Leanne Stones,[1,2] Josephine Bunch,[3,4] Robin C. May,[1,2] Helen J. Cooper[1]

[1]School of Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK
[2]Institute of Microbiology and Infection, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK
[3]National Physical Laboratory, Hampton Road, Teddington, TW11 0LW, UK
[4]School of Pharmacy, University of Nottingham, University Park, Nottingham, NG7 2RD, UK



Abstract. We have previously shown that liquid extraction surface analysis (LESA) mass spectrometry (MS) is a technique suitable for the top-down analysis of proteins directly from intact colonies of the Gram-negative bacterium *Escherichia coli* K-12. Here we extend the application of LESA MS to Gram-negative *Pseudomonas aeruginosa* PS1054 and Gram-positive *Staphylococcus aureus* MSSA476, as well as two strains of *E. coli* (K-12 and BL21 mCherry) and an unknown species of *Staphylococcus*. Moreover, we demonstrate the discrimination between three species of Gram-positive *Streptococcus* (*Streptococcus pneumoniae* D39, and the viridans group *Streptococcus oralis* ATCC 35037 and *Streptococcus gordonii* ATCC35105), a recognized challenge for matrix-assisted laser desorption ionization time-of-flight MS. A range of the proteins detected were selected for top-down LESA MS/MS. Thirty-nine proteins were identified by top-down LESA MS/MS, including 16 proteins that have not previously been observed by any other technique. The potential of LESA MS for classification and characterization of novel species is illustrated by the de novo sequencing of a new protein from the unknown species of *Staphylococcus*.

Keywords: Top-down protein analysis, LESA, Ambient mass spectrometry, Gram-positive, Gram-negative

Received: 5 January 2017/Revised: 15 May 2017/Accepted: 16 May 2017

Introduction

The application of mass spectrometry (MS) for the identification and characterization of bacterial species has a history of more than 40 years [1]. It has emerged as a mature and powerful technique for the study of microorganisms, shedding light on the chemical composition of their cells and colonies, from metabolites through lipids to entire proteomes.

Detection of characteristic compounds in bacterial samples allows their identification down to a genus or, in rarer instances, species and strain level [2]. Matrix-assisted laser desorption ionization (MALDI) time-of-flight (TOF) MS, the most commonly used technique [3], has been in development for that purpose since the late 1990s [4, 5] (recently reviewed in [6]). It relies primarily on spectral fingerprinting—matching mass spectra acquired from unknown samples against databases filled with mass spectra of previously identified microorganisms. Thanks to its long development time and resulting robustness, it is currently being rolled out worldwide as a clinical and diagnostic tool with the potential to eventually replace classic microbiological identification methods [7, 8], although discrimination between certain strains and species (e.g., *Escherichia coli* and *Shigella* or the viridans group streptococci) remains a challenge for the current protocols [6].

The MALDI-TOF approach does, however, have drawbacks that limit its use for the characterization of proteins from microbial colonies and biofilms. The most significant obstacle lies in sample preparation for analysis in a vacuum; whereas the routine preparation of bacterial smears for the purposes of clinical identification is trivial, the study of intact biofilms directly on growth media has proven challenging. Even though multiple groups have developed protocols for MALDI imaging analysis of bacterial colonies on agar [9–12], there is no consistency in the sample preparation methods across the literature,

Electronic supplementary material The online version of this article (doi:10.1007/s13361-017-1718-8) contains supplementary material, which is available to authorized users.

Correspondence to: Helen J. Cooper; *e-mail:* h.j.cooper@bham.ac.uk

Published online: 05 July 2017

even where identical equipment has been used. The problems reported include insufficient matrix saturation, uneven coverage, and flaking of the sample under a vacuum, the latter issue being by far the most serious as it may occasionally lead to instrument damage [12].

The challenge of sample preparation can be eliminated by use of ambient ionization techniques, although frequently at the cost of the capability for protein detection. Desorption electrospray ionization (DESI) has been applied to the analysis of bacterial suspensions dried onto solid substrates [13, 14], and colonies grown on filter membranes [15]; however, direct measurements on agar could not be easily obtained without drying of the sample because of the soft and conductive nature of the nutrient medium [16]. Rapid evaporative ionization MS has been extensively tested for the characterization of lipid profiles directly from cultured bacterial colonies, and as a high-throughput identification platform [17–19]. Laser ablation electrospray ionization has been used to directly probe biofilms on filter paper, providing information largely limited to lipids and metabolites [16, 20]. Similarly, paper spray MS of bacterial colony smears, generating mass spectra primarily consisting of phospholipid peaks, was shown to provide enough information for bacterial identification at the species level [21]. NanoDESI and the liquid microjunction surface sampling probe, both making use of liquid microjunction extraction, have been shown to extract small (up to 4.5 kDa), secreted peptides as well [22, 23].

One related liquid microjunction method coupled with high-resolution MS, liquid extraction surface analysis (LESA), has consistently proven to be suitable for protein extraction. To date, it has been used to characterize a wide array of samples, ranging from dried blood spots [24–27] through tissue sections [28, 29] to living bacterial colonies [29, 30]. A straightforward, solvent-based technique, suitable for manual application although usually performed by use of the dedicated TriVersa NanoMate automatic pipette system [31], LESA relies on the deposition of a droplet of solvent on the surface to be sampled. The droplet is held in contact with the surface, forming a liquid microjunction through which analytes diffuse into the solvent. The droplet is then withdrawn into the electroconductive pipette tip and introduced into the mass spectrometer by nanoelectrospray ionization. The distinct advantage of LESA is the ability to simultaneously extract multiple types of analytes from a single location with little to no sample preparation. This versatility means that a single instrument could potentially be capable of analyzing a vast range of clinically relevant sample types for different biomarkers. The extraction of multiple classes of molecules may in some cases produce overly complex mass spectra or lead to extensive ion suppression [32]; to circumvent this issue, separation methods such as nano liquid chromatography (LC)–MS [28] or high-field asymmetric waveform ion mobility separation (FAIMS) [26, 29] may be used before MS analysis.

We have previously reported top-down LESA MS identification of proteins extracted directly from colonies of *E. coli* K-12 grown on agar plates [29, 30]. Seven proteins were identified: three DNA-binding proteins and four stress response proteins, one of which required FAIMS separation of the sample to be detected. Crucially, the protein signal could not be observed unless a modified protocol, bringing the tip of the sampling pipette in contact with the colony surface ("contact" LESA), was used. It has been suggested that the inability to detect proteins when standard LESA is used arises from the presence of the extracellular matrix around and within the colony that requires physical disruption to allow the extraction of proteins from the bacterial cells.

Here we present the application of this approach to two clinical isolate species, the Gram-negative *Pseudomonas aeruginosa* PS1054 and the Gram-positive *Staphylococcus aureus* MSSA476, as well as three representatives of a group of closely related species of Gram-positive streptococci, known to cause nonidentification and misidentification issues in clinically deployed MALDI-TOF MS systems [33, 34], and an unknown laboratory strain of *Staphylococcus*. Two strains of *E. coli*, K-12 and BL21 labeled with mCherry, were also analyzed. The results show the broad applicability of LESA MS to the analysis of proteins in bacterial colonies. Importantly, differentiation of the streptococci was straightforward by LESA MS. Top-down LESA MS is demonstrated by the identification of nearly 40 proteins, more than 40% of which have not been observed previously by any other biochemical or MS technique. De novo identification of a further protein from an unknown *Staphylococcus* species is also demonstrated, illustrating the potential of LESA MS for classification and characterization of novel species.

Experimental

Materials

E. coli K-12, *E. coli* BL21 mCherry, *P. aeruginosa* PS1054, *S. aureus* MSSA476, and an unidentified *Staphylococcus* species were cultured on solid Lysogeny broth agar (LBA) medium in 6-cm-diameter petri dishes. (The size of the plates is limited by the size of the sampling tray that forms part of the Advion TriVersa NanoMate system.) *Streptococcus pneumoniae* D39, *Streptococcus oralis* ATCC 35037, and *Streptococcus gordonii* ATCC 35105 were cultured on blood agar either in open air or under semianaerobic conditions by use of a candle jar. All strains were incubated at 37 °C for 24–48 h and stored at room temperature or 4 °C if necessary; deliberate deviations from the established protocols are noted in the text.

The solvent system used for liquid extraction consisted of acetonitrile (J.T.Baker, Deventer, Netherlands), water (J.T.Baker, Netherlands), and formic acid (Sigma-Aldrich, Gillingham, UK); the relative proportions were either 39.5:59.5:1 or 40:60:1 for Gram-negative species, and either 40:60:1 (for high nanoelectrospray stability, allowing acquisition of tandem mass spectra) or 50:45:5 (for rapid, reliable generation of full mass spectra) for Gram-positive species as noted.

Liquid Extraction Surface Analysis

Petri dishes containing colonies of interest were placed in the sample tray of a TriVersa NanoMate robot (Advion, Ithaca, NY, USA), adjacent to a quarter of a 96-well microtiter plate holding the solvent system. The position of the colonies was indicated with use of the advanced user interface (AUI) of the program ChipSoft controlling the movements of the robotic pipette. Solvent (3 µl) was aspirated from a well of the microtiter plate. The pipette tip was moved above the colony to the position specified by the advanced user interface and brought into contact with the colony surface by the lowering of it to a height of approximately -10 mm; the exact distance would be adjusted for each individual agar plate and reassessed periodically to account for the gradual changes in the height of the solid agar medium. On contact, 2 µl of the solvent was dispensed onto the surface of the colony encompassed by the pipette tip. Contact was maintained for a minimum of 3 s, following which the solution was reaspirated into the pipette tip.

LESA MS was undertaken in an Advisory Committee on Dangerous Pathogens Containment Level 2 laboratory. No viable bacterial cells are present in the LESA droplet following sampling. To verify that, solvent droplets following LESA extraction from live bacterial colonies were spotted onto agar plates and incubated overnight. No growth was observed (see Supplementary File 1).

Mass Spectrometry

All MS experiments were performed with an Orbitrap Elite instrument (Thermo Fisher Scientific, Bremen, Germany) at a resolution of 120,000 at *m*/*z* 400. Samples were introduced into the instrument by use of the TriVersa NanoMate integrated, chip-based nanoelectrospray system at a pressure of 0.3 psi and a voltage of 1.75 kV. Full-scan mass spectra in the *m*/*z* range from 600 to 2000 were acquired for a minimum of 5 min. Each scan comprised ten coadded microscans. Precursor ions were selected for fragmentation with an isolation window of 2, 3, or 5 *m*/*z* as appropriate. Collision-induced dissociation was performed in the ion trap with use of helium gas at a normalized collision energy of 35%, and the fragments were detected in the Orbitrap. MS/MS spectra were recorded for 5 min. Each MS/MS scan comprised 30 coadded microscans. Automatic gain control targets were 1×10^6 charges for full-scan mass spectra and 5×10^5 charges for MS/MS spectra.

Protein Identification

Top-down protein identification was performed with ProSight 3.0 software (Thermo Fisher Scientific, Bremen, Germany). MS/MS spectra were deconvoluted by THRASH at a signal-to-noise ratio of 3. These were then matched against custom databases constructed for each species from full proteome data available via UniProt.

Each database was constructed as a standard top-down database, taking into account the cleavage of initial methionines and N-terminal acetylation but ignoring formylation. SNPs and all available posttranslational modifications were considered, with up to 13 features per sequence and a maximum mass of 70 kDa. For both strains of *E. coli*, the complete, annotated reference proteome of the K-12 strain was used (UniProt ID UP000000625, 4306 protein entries). For *S. aureus* MSSA476, two databases were generated: one based on the reference proteome of a representative strain (NCTC 8325, UniProt ID UP000008816, 2889 protein entries), making use of the more complete, reviewed annotation, and one specific to the MSSA476 strain (UniProt ID UP000002201, 2598 protein entries). For *P. aeruginosa* PS1054 the reference proteome (UniProt ID UP000002438, 5563 protein entries) was used exclusively. *S. pneumoniae* D39 was searched against the reference proteome of a closely related, nonpathogenic strain (UniProt ID UP000000586, 2030 protein entries). The *S. oralis* database was based on the only available reference proteome for this species (UniProt ID UP000005621, 2022 protein entries). Two *S. gordonii* databases were used, one based on the annotated chromosome (UniProt ID UP000001131, 2050 protein entries) and one based on the whole genome sequence (UniProt ID UP000069207, 2095 protein entries). A multispecies database was also constructed based on the pan-proteomes of *S. pneumoniae* ATCC BAA-255/R6 (UniProt ID UP000000586) and *S. gordonii* Challis (UniProt ID UP000001131; a total of 61,367 protein entries). Tandem mass spectra of the unknown species of *Staphylococcus* were searched against the reference *Staphylococcus epidermidis* proteome (UniProt ID UP000000531, 2492 protein entries) and a concatenated multispecies database comprising all available nonredundant proteomes in the *Staphylococcus* genus (145,289 protein entries), referred to as the "all *Staphylococcus*" database in the text.

For each database, a broad absolute mass search was specified as a starting point—proteins were considered within 2 kDa of the measured intact mass, with both Δm and disulfide modes active and all available posttranslational modifications taken into account; fragment tolerance was set to ±15 ppm. Putative hits, as well as any posttranslational modifications, were verified by narrowing down the search criteria and optimizing any remaining matches in ProSight's Sequence Gazer, followed by manual peak assignment with an acceptance threshold of ±5 ppm.

The protein sequence generated de novo from the unknown *Staphylococcus* species was searched by standard protein BLAST against all nonredundant protein sequences belonging to the *Staphylococcus* genus (taxonomic ID 1279), with use of the protein–protein BLAST algorithm with the default BLOSUM62 scoring matrix.

Results and Discussion

LESA Sampling of Bacteria

The successful generation of protein mass spectra following LESA sampling of bacteria appears to be dependent on the

physical properties of the colonies, in particular the tendency of colony material to adhere to the pipette tip, as well as colony size and surface variability. *E. coli* K-12, the model bacterium used to first develop the protocol and chosen to measure sampling reproducibility, was also by far the most challenging from which to successfully generate protein mass spectra. Initial experiments comprising a total of 125 sampling attempts across colonies grown and stored under a range of conditions yielded 62 mass spectra, of which 31 contained a clear protein signal, an overall success rate of 25%. In the later refrigeration experiments (see below), the success rate was 49% (30 protein mass spectra from 61 sampling attempts). The failure to acquire mass spectra following sampling was universally due to the nanoelectrospray nozzles becoming blocked with aspirated colony matter. A careful choice of sampling height (touching rather than piercing the colony) alleviated the issue although it is currently limited by the minimum vertical step size of 0.2 mm on the TriVersa NanoMate robot.

Pseudomonas colonies were less sensitive to the choice of sampling height, reliably generating protein-rich mass spectra. For example, in the refrigeration experiments described later, 45 sampling attempts resulted in 37 protein mass spectra (82% success rate). The Gram-positive staphylococci were similarly amenable to sampling, although rather than cytosolic proteins, they yielded mainly secreted proteins. Initial experiments with the staphylococci using the same solvent system as for the Gram-negative species (i.e., 40:60:1 acetonitrile–water–formic acid) had a success rate of approximately 50%. Optimization of the solvent system (to 50:45:5 acetonitrile–water–formic acid) increased the sampling success rates, possibly due to increased extraction as a result of cell lysis and/or increased ionization efficiency. With this solvent system, the sampling success rate in the refrigeration experiments was 97% (28 protein mass spectra obtained from 29 sampling attempts). The same solvent was applied to the sampling of Gram-positive streptococci.

General Overview of the Mass Spectra

Figure 1 shows representative mass spectra of seven bacterial colonies corresponding to seven strains: Gram-negative *E. coli* K-12, *E. coli* BL21 mCherry, and *P. aeruginosa* PS1054, and Gram-positive *S. aureus* MSSA476, *S. pneumoniae* D39, *S. oralis* ATCC 35037, and *S. gordonii* ATCC 35105. All colonies were sampled immediately following incubation at 37 °C, with use of either the 40:60:1 acetonitrile-based solvent system for Gram-negative bacteria or the 50:45:5 variant for Gram-positive bacteria.

P. aeruginosa gave rise to highly reproducible mass spectra, particularly easy to acquire and rich in protein peaks. Prominent, consistent protein peaks corresponding to five proteins appeared in every acquired mass spectrum. The total number of protein peaks observed in a representative mass spectrum generated following use of optimum incubation conditions was 66, corresponding to 20 unique masses, excluding suspected adducts and modifications. Fifteen of these were selected for MS/MS identification (see later). The total number of protein peaks in a representative spectrum of *E. coli* was 44, corresponding to 14 unique masses; alteration of the growth conditions, such as subjecting the colonies to cold or room temperature storage, increased the number of unique masses to at least 19 by inducing the expression of multiple stress response proteins (see later). The set of predictable protein peaks in *E. coli* corresponded to the DNA-binding proteins HU-α and HU-β in up to four charge states. Whereas the number of observed proteins is markedly lower than the 150+ proteins reported by top-down LC–MS of *E. coli* K-12 extracts [35], the acquisition of data by LESA MS is much faster and more straightforward as it involves no sample preparation nor lengthy separation before analysis. Addition of a separation step, such as FAIMS [29] or LC [27], into the LESA MS workflow is possible, and would likely expand the range of detected proteins.

Crucially, direct comparison of LESA mass spectra acquired from the two Gram-negative species revealed considerable promise for the technique to be used as an in situ microbial identification tool. Although the signal at the lower end of the *m*/*z* range was similar for both species, because of the presence of peaks derived from the Lysogeny broth agar (LBA) growth medium as well as common lipids and metabolites, unambiguous differentiation between the mass spectra was still possible as the range of protein peaks found in each species was unique.

The mass spectra of *E. coli* K-12 and *E. coli* BL21 mCherry were highly similar. Peaks corresponding to the characteristic DNA-binding proteins HU-α, HU-β, and H-NS as well as the putative periplasmic protein YahO were seen in mass spectra of both strains; numerous low-mass species were also common to both strains and occurred at similar abundances. Whereas the BL21 strain should be differentiated by the presence of the 28.8-kDa mCherry fluorescent protein, no peaks that could be confidently assigned to this protein were observed.

The mass spectra of *S. aureus* were dominated by peaks corresponding to secreted peptides, the most abundant of which was the well-characterized toxin δ-hemolysin. A representative mass spectrum acquired with 5% formic acid, as seen in Figure 1, contained 26 peaks corresponding to 11 unique protein and peptide masses.

Differentiation of Streptococci

As an additional challenge, colonies of three species of *Streptococcus*—*S. pneumoniae* D39, *S. oralis* ATCC 35037, and *S. gordonii* ATCC 35105—were subjected to LESA MS characterization. Streptococci are encapsulated Gram-positive bacteria, seen in nature both as pathogens and as harmless commensals. *S. pneumoniae* and the viridans group streptococci (including *S. oralis* and *S. gordonii*, which constitute a major part of dental plaque) are of particular clinical interest, as the current method for their differentiation, which is required for the identification of a correct treatment, relies primarily on an optochin test and similar microbiological techniques. Distinguishing these two groups of streptococci by MALDI-TOF MS presents a challenge because of the high similarity of their fingerprint mass spectra; the performance of any

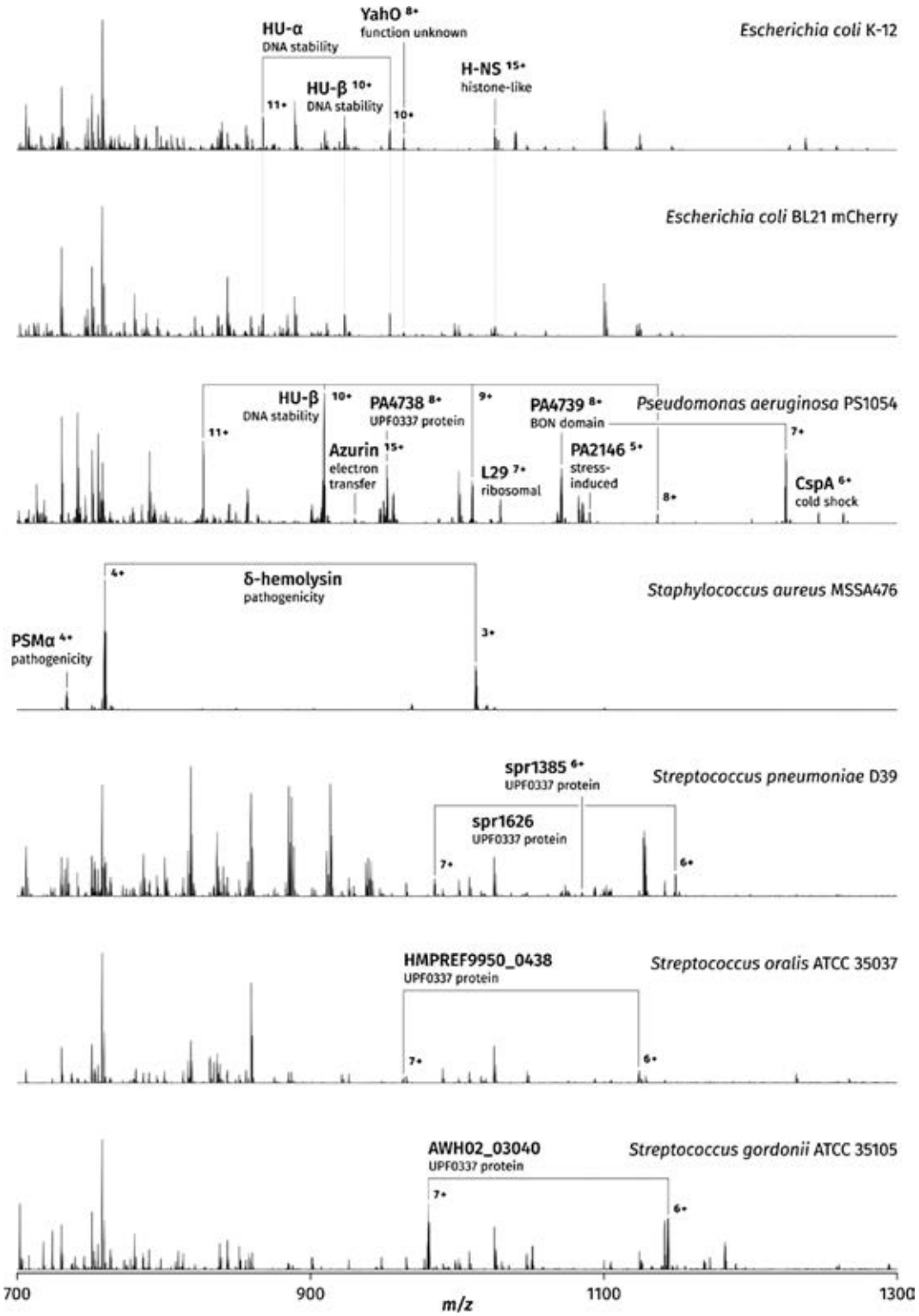


Figure 1. Liquid extraction surface analysis (LESA) mass spectra of seven bacterial strains: *Escherichia coli* K-12, *Escherichia coli* BL21 mCherry, *Pseudomonas aeruginosa* PS1054, *Staphylococcus aureus* MSSA476, *Streptococcus pneumoniae* D39, *Streptococcus oralis* ATCC 35037, and *Streptococcus gordonii* ATCC 35105. All mass spectra were acquired immediately following incubation of the colonies at 37 °C, with use of the 40:60:1 acetonitrile–water–formic acid solvent system for Gram-negative bacteria and the 50:45:5 variant for Gram-positive bacteria. The streptococci were incubated under semianaerobic conditions for optimum growth; the remaining strains were grown in open air

individual identification platform in this task is largely dependent on the spectral databases used [3, 33, 34]. An unambiguous identification approach based on defined, individual peaks may therefore prove superior.

The thick capsule surrounding streptococcal cells was expected to hinder lysis and therefore reduce the efficiency of protein extraction. Predictably, the protein signal generated from the colonies was much weaker than in the case of all other investigated strains; 1% and 5% formic acid solvents were both used, with no marked improvement observed at the higher formic acid concentration. Nevertheless, eight peaks corresponding to five protein masses were detected in *S. pneumoniae*, and nine peaks corresponding to six protein masses were detected in *S. oralis*; *S. gordonii* gave rise to richer mass spectra, with 26 peaks and 15 protein masses detected. Only one protein (observed mass 7984 Da) was common to all three species; all other protein masses observed were unique to their species. That is, differentiation of these species on the basis of their LESA mass spectra was straightforward. Four proteins were selected for MS/MS identification; none of these have thus far been directly observed as intact proteins (for details see the following section). Fragmentation data were searched against a combined protein database comprising multiple strains of *S. pneumoniae*, *S. oralis*, *Streptococcus mitis*, *Streptococcus pseudopneumoniae*, *S. gordonii*, and other closely related species. All four proteins were confidently assigned (expectation value less than 1×10^{-70}) to the correct species.

Top-Down Protein Identification

Table 1 lists the proteins identified in all bacterial strains investigated. Fragment assignments are given in Supplementary File 2. Sample growth and storage conditions were deliberately varied, allowing the observation of a wider range of proteins, particularly ones involved in stress response; the conditions under which the tandem mass spectrum of any given protein was acquired are listed in Table 1 (although these were not necessarily the only conditions under which the protein was observed). Notably, numerous proteins reported here have not been directly observed in their intact form by any other technique. Novel protein identifications are shown in bold type in Table 1.

P. aeruginosa was arguably the most amenable to LESA MS analysis among the species chosen, owing to the fast growth of colonies, the ease of sampling, and the wide range of detected proteins. Several of those, such as the DNA-binding protein HU-β or the ribosomal constituent L29, were homologs of proteins detected in *E. coli*. PA2146, a KGG stress response motif protein predicted from the full genome sequence [36], is noteworthy as a relative of *E. coli* YciG and YmdF, both newly identified here. Most notably, however, the UPF0337 family of stress response proteins, represented in *P. aeruginosa* by PA4738, was observed both in *E. coli* (YjbJ) and in *S. aureus* (SAOUHSC_00845) as well as in all three streptococci.

In addition to PA2146, *P. aeruginosa* yielded multiple proteins whose existence is predicted on the basis of genome data, in several instances bolstered by the presence of homologous sequences in other species, but which have not previously been detected as proteins. A prime example of such a protein was peptidylprolyl isomerase PpiC2, a small predicted enzyme related to *E. coli* parvulin [36]; PpiC2 was identified here in its intact form, missing the N-terminal methionine. PA4739, another predicted protein detected by LESA MS, contains a putative BON domain, which suggests it is localized to either the periplasm or the outer membrane, and is likely involved in stress response [37]. Automatic gene annotation by UniProt (accession number Q9HV60) suggests it contains a 25 amino acid signal peptide; however, on the basis of the intact mass of the observed species as well as its fragmentation pattern, we can infer that the signal peptide is actually cleaved at position 32. The tandem mass spectrum of PA0039, similarly detected with a missing signal peptide, showed robust evidence of a hitherto unannotated disulfide bridge between the cysteines at positions 4 and 42. No fragmentation was observed between these two positions, and the observed intact mass was 2.04 Da lower than predicted [36], matching the intact mass of the disulfide-containing protein form to within 0.1 ppm. PA5178, the second predicted BON domain protein observed in *P. aeruginosa*, was particularly curious; the C-terminal fragment, cleaved off at the boundary between the two predicted domains of the protein, was observed in two colonies subjected to an incubation–cold storage–incubation cycle (see later), though not the third, adjacent colony grown on the same plate and subjected to the same conditions. It was also detected and identified by MS/MS in a colony stored for 3 months at 4 °C, which yielded extremely poor protein mass spectra consistent with extensive protein breakdown. A predicted thermolysin cleavage site was identified by PeptideCutter [38] at the site of observed fragmentation; it is therefore possible that the cleavage is conducted by an as yet unidentified protease and may be functionally relevant. Finally, two examples of ribosomal proteins hitherto inferred from homology [36], L35 (with its N-terminal methionine cleaved) and L36, were also seen.

Sixteen proteins from *P. aeruginosa* were fragmented, and 15 of these were successfully identified. One protein, observed in charge states 6+ and 7+ at *m*/*z* 1237 and *m*/*z* 1060 respectively (measured mass 7419.73 Da), could not be identified by the current method. The only sequence tag generated by ProSight 3.0 from multiple acquired tandem mass spectra consisted of five amino acids, QTAVQ—a much shorter signature than expected of a protein of this size. This observation may suggest a highly modified protein, or one with an unusual structure, and implies that the nonspecificity of the proteome database to this particular strain is not the reason for the lack of identification; however, the repeated detection of this protein in multiple mass spectra suggests it may be linked to cold stress response—it was observed only in colonies that had been stored at 4 °C or room temperature (conditions suboptimal

Table 1. Proteins Identified by Liquid Extraction Surface Analysis Mass Spectrometry in Sampled Colonies

m/z	Charge	Molecular weight observed	Mass difference (ppm)	ID	UniProt accession no.	Cov (%)	Conditions	Modifications
Escherichia coli BL21								
959.1288	+6	5748.73	-0.8	YmdF	P56614	66	Incubation: 24 h, 37 °C	-Met
978.9706	+6	5867.78	-2.0	YciG	P21361	36	Storage: 4 days, room temperature	-Met
1101.2759	+7	7701.88	-1.5	YahO	P75694	91		-signal peptide
1189.5904	+7	8320.08	-1.3	UPF0337 protein YjbJ[a]	P68206	80		
1254.2614	+7	8772.78	-2.2	YdfK	P76154	31		
1494.9477	+7	10,457.58	+1.02 Da	**YbgS**	P0AAV6	45		-signal peptide; putative disulfide 71–76 and deamidation
Escherichia coli K-12								
923.0049	+10	9219.98	-2.0	HU-β[a]	P0ACF4	43	Incubation: 24 h, 37 °C Storage: 2 days, room temperature	
953.9263	+10	9529.93	0.2	HU-α[a]	P0ACF0	40	Incubation: 24 h, 37 °C Storage: 15 days, 4 °C	
963.5065	+16	15,399.99	+1.02 Da	H-NS[a]	P0ACF8	30	Incubation: 24h, 37 °C	-Met
1019.5584	+5	5092.76	-1.9	SRA	P68191	11	Storage: 2 days, room temperature	
1039.4342	+7	7268.99	0.8	L29	P0A7M6	41	Incubation: 24 h, 37 °C	
1088.8935	+6	6527.32	-1.1	BhsA[a]	P0AB40	35	Storage: 15 days, 4 °C	-signal peptide
1094.2255	+8	8745.75	-0.6	YnaE	P76073	62		
1212.1269	+6	7266.72	-0.4	CspC	P0A9Y6	37	Incubation: 24 h, 37 °C Storage: 2 days, room temperature	-Met
1356.6814	+3	4067.02	0.3	CydX	P56100	46	Incubation: 24 h, 37 °C Storage: 18 days, 4 °C	fMet
Pseudomonas aeruginosa PS1054								
672.2674	+9	6041.34	-1.8	L33	Q9HTN9	31	Incubation and storage: 4 days, room temperature	-Met
723.8214	+10	7228.14	-2.0	**L35**	Q9I0A1	41		
739.5860	+6	4431.47	-1.2	**L36**	Q9HWF6	45		-Met
759.9734	+11	8348.63	-1.4	S21	Q9I5V8	29		
826.5565	+11	9081.04	-0.3	HU-β	P05384	51	Incubation: 48 h, 37 °C	
951.8583	+8	7606.81	0.0	UPF0337 protein PA4738	Q9HV61	58	Sampled fresh	
956.3376	+6	5731.98	0.1	**PA0039**	Q9I793	24	Incubation: 24 h, 37 °C Sampled fresh	-signal peptide, 4–42 disulfide
958.5127	+16	15,320.09	3.5	**PA5178**	Q9HU11	27	Incubation: 48 h, 37 °C Sampled fresh	-Met
983.5923	+10	9825.85	-1.6	**Peptidylprolyl isomerase**	Q9HWK5	45	Incubation: 24 h, 37 °C	-Met
990.3747	+8	7914.94	-1.4	L31	Q9HUD0	31	Storage: 4 days, room temperature	
996.4895	+14	13,936.75	-0.8	Azurin	P00282	14		-signal peptide, disulfide
1029.1389	+7	7196.92	-0.6	L29	Q9HWE3	29	Incubation: 48 h, 37 °C	
1090.1381	+5	5445.65	0.4	**PA2146**	Q9I1W9	67	Sampled fresh	-Met
1223.5075	+7	8557.50	-0.7	**PA4739**	Q9HV60	61		-signal peptide (1–32)
1245.9733	+6	7469.80	-0.2	**CspA**	P95459	31		-Met
Staphylococcus aureus MSSA476								
733.7894	+3	2198.35	0.6	Phenol-soluble modulin α_4	P0C826	100	Incubation: 24 h, 37 °C Sampled fresh	fMet
759.6682	+4	3034.64	0.8	δ-Hemolysin	Q9I5V8	100	Incubation: 48 h, 37 °C Sampled fresh	fMet
921.1190	+6	5520.67	2.0		Q2FXV2	96	Incubation: 24 h, 37 °C	

Table 1 (continued)

m/z	Charge	Molecular weight observed	Mass difference (ppm)	ID	UniProt accession no.	Cov (%)	Conditions	Modifications
				Uncharacterized protein SAOUHSC_01729			Sampled fresh	
1136.8509	+4	4521.39	2.0	**Uncharacterized protein SAOUHSC_01135**	Q2FZA4	82		fMet, sodium adduct
1148.2490	+6	6883.45	1.7	**UPF0337 protein SAOUHSC_00845**	Q2FZY9	89		
Streptococcus pneumoniae D39								
930.0688	+7	6503.43	0.3	**UPF0337 protein spr1385**	Q8CYJ5	85	Incubation: 24 h, 37 °C	-Met
984.6548	+7	6884.53	0.4	**UPF0337 protein spr1626**	Q8CYD7	76	Sampled fresh	
Streptococcus oralis								
963.0736	+7	6734.46	-2.0	**UPF0337 protein HMPREF9950_0438**	F9Q496	82	Incubation: 24 h, 37 °C Sampled fresh	-Met
Streptococcus gordonii								
980.2427	+7	6854.65	-0.6	**UPF0337 protein AWH02_03040**	A0A0F5MKW3	71	Incubation: 48 h, 37 °C, semianaerobic Sampled fresh	-Met

Protein names in bold signify proteins that have been directly observed for the first time.
Met formylmethionine, *Met* methionine
[a]Proteins identified in the previous study on *Escherichia coli* K-12 using liquid extraction surface analysis mass spectrometry

for *P. aeruginosa*, which thrives at 37 °C). Other software packages, such as MSAlign [39], may facilitate the identification of uncharacterized, highly modified proteins.

All proteins selected for fragmentation from *S. aureus* were successfully identified. Five proteins were identified, three of which belong to the phenol-soluble modulin family of small secreted proteins found in virulent staphylococci [40]. δ-Hemolysin, a 26 amino acid peptide that oligomerizes into pore-like structures [41] and subsequently lyses erythrocytes, is one of two well-characterized examples on the list. The other example, phenol-soluble modulin α_4, is implicated in the lysis of neutrophils [42]; in this particular strain, its existence has so far been inferred only from homology, although in the methicillin-resistant *S. aureus* strain USA300 an identical gene product has been directly observed and studied in detail. A related, predicted protein, SAOUHSC_01135 [43], was also identified; all three retain N-terminal formylmethionine characteristic of this family [40]. SAOUHSC_00845 [44] has been assigned by sequence homology to the previously mentioned UPF0337 family of stress response proteins, detected in all three species. Finally, the predicted protein SAOUHSC_01729 [43], identified here with no modifications, is completely uncharacterized; identical gene products have been found by a standard protein BLAST search in numerous strains of *S. aureus*, as well as one strain of *Staphylococcus haemolyticus*.

Sixteen proteins from *E. coli* (BL21 and K-12) were selected for fragmentation, with 15 successfully identified; the one failure, a 30-kDa protein observed in a single mass spectrum, was due to the poor quality of the MS/MS spectrum. All proteins previously detected in *E. coli* K-12 by direct LESA MS were also detected here. Five of these were selected for MS/MS identification (see Table 1). In addition to these, one previously unseen protein and nine proteins identified by other techniques (five by peptide mass fingerprinting by MALDI-TOF MS, as part of large-scale interaction studies [45, 46], and four by gel electrophoresis [47] and top-down MALDI MS [48, 49]) were identified. Notably, UniProt does not accept the interaction studies mentioned above as sufficient experimental evidence for the existence of the five proteins (YmdF, YciG, YahO, YdfK, and YnaE) that are clearly identified here. YbgS, a protein inferred from homology and shown to display genetic interactions with several partners [50], was detected with its signal peptide cleaved. MS/MS data suggested the presence of a disulfide bond as well as possible deamidation, although insufficient information was available to determine the exact position of the modifications.

The improvements in software for top-down protein analysis play a major part in the method. YdfK, a hypothetical cold shock protein [51], could not be identified in a previous study using the search modes and databases provided by ProSight 2.0 [30]. Highly confident hits with the correct intact mass for this protein were obtained here with use of a custom *E. coli* K-12 proteome database in ProSight 3.0. In addition, YahO was initially identified only by the biomarker search, which takes into account subsequences of the archived proteins as well as their intact forms; the absolute mass search proved, however, sufficient

when an updated (2017) database was used. YahO was identified following cleavage of a 21 amino acid sequence peptide from the intact protein, previously inferred from protein sequence [52] and experimentally confirmed here by LESA MS.

Of the proteins observed in streptococci, four were selected for identification: two in *S. pneumoniae* and one each in *S. oralis* and *S. gordonii*. All four were found to be members of the ubiquitous but poorly studied UPF0337 stress response family of proteins; none have thus far been observed in their intact form by other techniques. Their markedly different masses make them suitable as identification markers.

De Novo Sequencing of an Unknown Protein from an Unknown Staphylococcus Species

Thus far, all experiments were conducted on well-characterized species, for which full genome sequences are readily available. The usefulness of LESA MS in bacterial work is, however, not limited to the identification of unknown proteins; with the use of de novo protein sequencing, it can potentially be extended to the characterization of novel microbial species. This capability was demonstrated on the original *Staphylococcus* sample selected for top-down protein analysis and originally thought to be an in-house strain of *S. epidermidis*. In overview, the full mass spectra of this species were dominated by small secreted proteins, similarly to *S. aureus* (Figure 2a). More than 26 peaks were observed, corresponding to at least eight unique protein and peptide masses. On acquisition of tandem mass spectra of several highly abundant protein peaks, however, it was discovered that none of these precisely matched any known *S. epidermidis* proteins. A subsequent search against an all *Staphylococcus* protein database comprising more than 100 strains similarly returned no matches.

To shed light on the true identity of the bacterium, a 4-kDa peptide observed at *m/z* ~1124 was analyzed in detail (Figure 2b). Matching of the deconvoluted tandem mass spectrum directly against custom ProSight 3.0 databases derived from a reference *S. epidermidis* proteome placed the unknown protein in the phenol-soluble modulin β family. The strongest hit returned by the software (see Supplementary File 3 for a complete list of candidate sequences and their associated scores) was then used as a guide for de novo sequencing. With this approach, an almost complete sequence was elucidated. Two gaps (AGN and TSI; Figure 2b) were identified in which the nature but not the sequence of the amino acids could be deduced; these were arranged to best match the most closely related homologs found by a standard protein BLAST search. The final sequence is 75% identical to the *S. epidermidis* guide sequence identified by ProSight, 68% identical to the best match returned in the all *Staphylococcus* protein database search (*S. aureus* NCTC 8325, uncharacterized protein SAOUHSC_1135), and 80% identical to a *Staphylococcus capitis* homolog. In the same vein, several partial sequences of the other isolated proteins were acquired, although none were sufficiently complete to allow a meaningful BLAST search. Nevertheless, with some improvements to the software to avoid the most time-consuming step of manual de novo sequencing, as well as to make use of incomplete protein sequences, the ability of this technique to rapidly classify and characterize novel species could be greatly expanded.

The Influence of Refrigeration

The capability of LESA MS (and any other technique relying on the measurement of protein signal) to identify bacteria has a potential weakness: any variability in the range of the protein peaks observed may introduce uncertainties into the identification, unless this variability is measured and taken into account. The most likely source of such variability in a laboratory or clinical setting is a change in the temperature to which the sample colonies are exposed as a result of transport or storage, prompting changes in protein expression.

On the other hand, should LESA MS detect differences in protein expression under varying conditions or following specific stimuli, it could potentially be used as a tool to shed light on the function of unknown proteins; more generally, it may provide valuable, real-time information on the interactions of bacterial colonies and biofilms and their responses to cues in their environment.

Accordingly, it has previously been demonstrated that the storage of *E. coli* K-12 colonies at 4 °C before analysis has a marked effect on the protein signal in the acquired mass

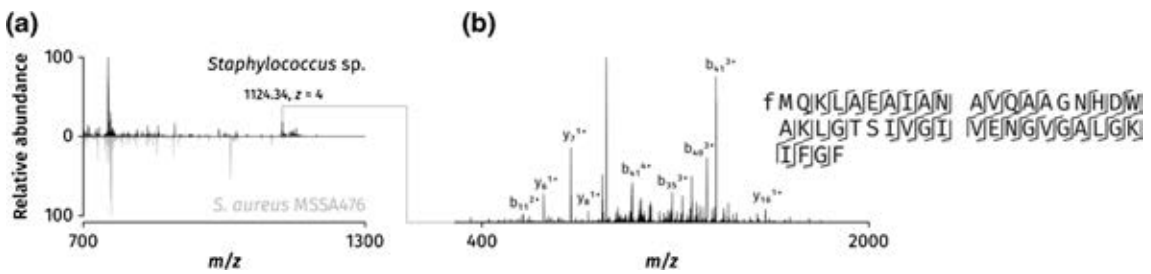


Figure 2. Characterization of the unknown species of *Staphylococcus*. (**a**) Comparison of full LESA mass spectra of *Staphylococcus* sp. and *S. aureus*. (**b**) Collision-induced dissociation mass spectrum of the ion centered at *m/z* 1124.34 (4+), identified as a putative phenol-soluble modulin β protein on the basis of de novo sequencing

spectra, as compared with mass spectra of colonies sampled immediately after incubation [30].

Figure 3 shows the results of a systematic investigation into this effect, as observed in the three bacterial species tested, *P. aeruginosa*, *E. coli*, and *S. aureus*. For each species, nine colonies grown on a single Lysogeny broth agar (LBA) plate were subjected to sequential changes in temperature: 37 °C for 24 h (standard incubation protocol), 4 °C for 4 days, and finally 37 °C for 24 h once again; the second incubation step was included to investigate whether any putative changes introduced into the mass spectra by cold storage might be reversible. Following each period, three of the colonies were sampled by LESA MS. Each set of three mass spectra, corresponding to the three periods of incubation or storage, were normalized. Experiments were performed in triplicate, with three identical plates per species, to ensure reproducibility.

Previously, we reported the appearance of a multiple stress response protein, BhsA, in mass spectra of *E. coli* colonies stored at 4 °C before sampling. In the controlled study presented here, as expected from these observations, the BhsA signal was absent in colonies sampled fresh after incubation but prominent peaks corresponding to the 7+ and 6+ charge states of the protein were detected in all colonies sampled following cold storage. Crucially, the peaks disappeared following the second incubation step, confirming that the detection of BhsA in *E. coli* colonies is dependent on the exposure of the colonies to low temperatures immediately before sampling, consistent with its expected expression during cold stress. Two additional peaks corresponding to YnaE and YdfK, the two cold shock proteins known from their transcripts, were seen alongside BhsA following refrigeration—however, unlike BhsA, these proteins remained detectable even after the second incubation. This suggests that refrigeration may, in some cases, lead to persistent changes in the sample that can be detected by LESA MS.

The mass spectra of *P. aeruginosa* displayed behavior similar to that of *E. coli*. Most of the peaks corresponding to proteins remained consistent across all conditions, with additional peaks corresponding to an unknown protein of mass 7419 Da observed in the mass spectra collected immediately following cold storage. The protein, mentioned briefly in the previous section, could not be identified with the current method because of the unusually short sequence tag generated from its tandem mass spectra. Its appearance following refrigeration points to its potential role in cold shock response; we propose that in a similar approach, putative function could be assigned to other unknown proteins following the deliberate exposure of the colonies to particular environmental conditions before sampling. These could include changes in temperature, pH, or exposure to chemical signals and drugs, including antibiotics, which could potentially prompt a resistance response detectable by LESA MS.

The changes observed in the mass spectra of *S. aureus* were different in nature from those seen in the Gram-negative species. No peaks corresponding to putative cold shock proteins were observed. The most abundant species for all conditions remained δ-hemolysin; however, once the set of mass spectra were normalized, it became apparent that the abundance of δ-hemolysin peaks increased over the course of the experiment. The most intense signal was invariably observed following the second incubation. This observation is consistent with the literature suggesting that the highest expression of δ-

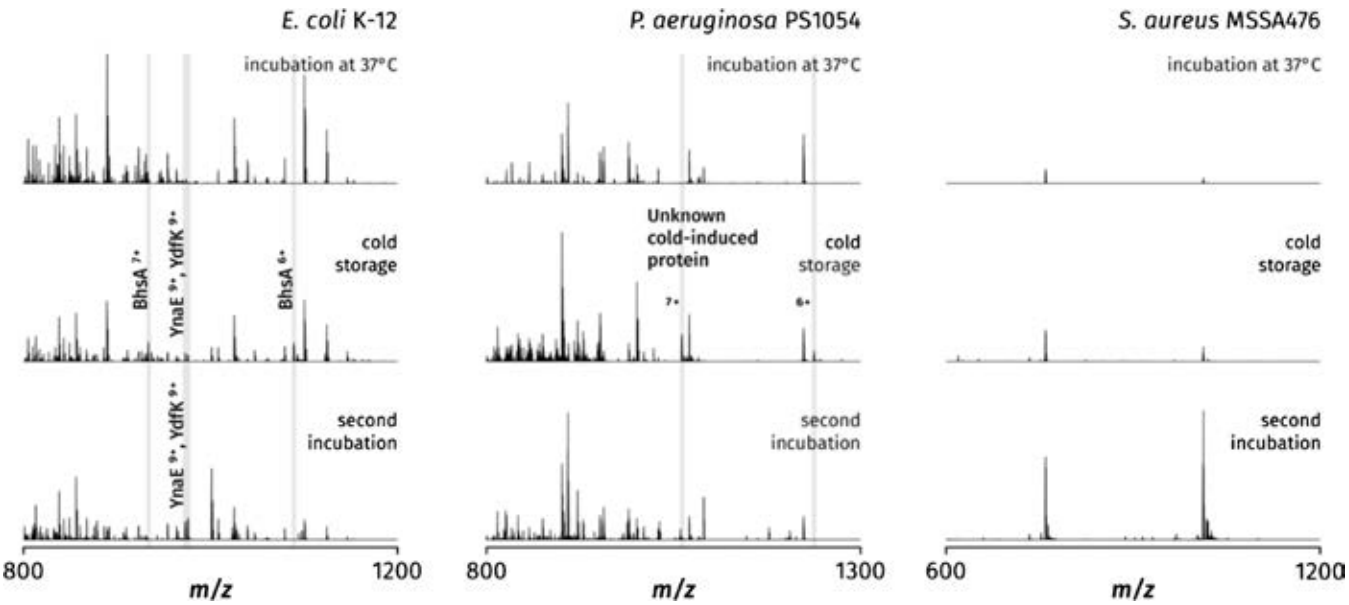


Figure 3. Investigation into the effects of refrigeration on LESA mass spectra of three bacterial species. Each set of three mass spectra corresponds to three colonies grown and sampled from the same agar plate; for each species, the first sampled set of replicates is shown. The solvent system used was 40:60:1 acetonitrile–water–formic acid for *E. coli* and *P. aeruginosa*, and 50:45:5 acetonitrile–water–formic acid for *S. aureus*

hemolysin and related, secreted proteins occurs during stationary phase [53].

Conclusions

We have demonstrated that LESA MS is a powerful tool for the analysis of intact proteins from bacterial colonies. The approach can be applied to Gram-positive bacteria as well as Gram-negative bacteria. The LESA mass spectra were reproducible within species and allowed differentiation of *E. coli*, *P. aeruginosa*, *S. aureus*, *S. pneumoniae*, *S. oralis*, and *S. gordonii*, but it was not possible to differentiate between the two strains of *E. coli* analyzed: K-12 and BL21 mCherry. The effect of refrigeration on the three species was determined. Forty proteins were identified by top-down MS/MS, of which 16 had previously only been predicted on the basis of genome sequencing and had not been observed as proteins and one was entirely novel. Of the total proteins selected for fragmentation and searched against proteome databases with use of ProSight, only two of 41 were not identified, an identification rate of 95%. We have also demonstrated that LESA MS is a suitable tool for the identification of proteins from unknown species.

Acknowledgements

H.J.C. is an EPSRC Established Career Fellow (EP/L023490/1). K.I.K. is in receipt of an EPSRC studentship in collaboration with the National Physical Laboratory. RCM and LS were supported by project MitoFun, funded by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement No. 614562 and by a Wolfson Research Merit Award from the Royal Society (to RCM). The Advion TriVersa NanoMate and Thermo Fisher Orbitrap Elite mass spectrometer used in this research were funded through Birmingham Science City Translational Medicine, Experimental Medicine Network of Excellence Project with support from Advantage West Midlands. Supplementary data supporting this research are openly available from the University of Birmingham data archive at http://findit.bham.ac.uk/. *P. aeruginosa* and *S. aureus* were provided by Mark Webber at the Institute of Microbiology and Infection at the University of Birmingham. *S. pneumoniae* was provided by Tim Mitchell and Andrea Mitchell at the Institute of Microbiology and Infection, and *S. oralis* and *S. gordonii* were provided by Rachel Sammons at the University of Birmingham School of Dentistry.

Open Access

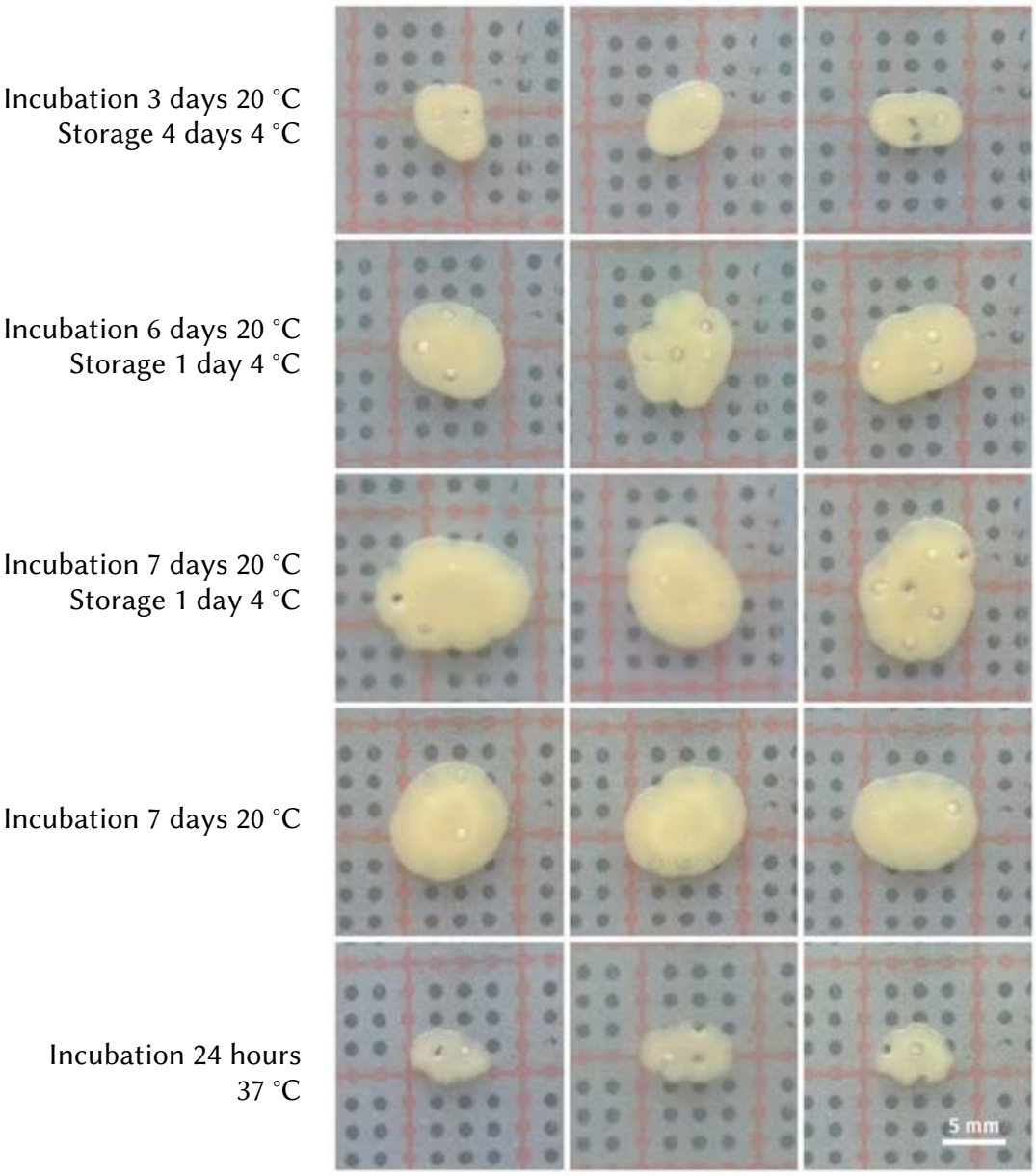
References

1. Anhalt, J.P., Fenselau, C.: Identification of bacteria using mass spectrometry. Anal. Chem. **47**, 219–225 (1975)
2. Krasny, L., Hynek, R., Hochel, I.: Identification of bacteria using mass spectrometry techniques. Int. J. Mass Spectrom. **353**, 67–79 (2013)
3. Cheng, K.D., Chui, H.X., Domish, L., Hernandez, D., Wang, G.H.: Recent development of mass spectrometry and proteomics applications in identification and typing of bacteria. Proteomics Clin. Appl. **10**, 346–357 (2016)
4. Holland, R.D., Wilkes, J.G., Rafii, F., Sutherland, J.B., Persons, C.C., Voorhees, K.J., Lay, J.O.: Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. **10**, 1227–1232 (1996)
5. Claydon, M.A., Davey, S.N., EdwardsJones, V., Gordon, D.B.: The rapid identification of intact microorganisms using mass spectrometry. Nat. Biotechnol. **14**, 1584–1586 (1996)
6. Singhal, N., Kumar, M., Kanaujia, P.K., Virdi, J.S.: MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Front. Microbiol. **6**, 16 (2015)
7. Bizzini, A., Durussel, C., Bille, J., Greub, G., Prod'hom, G.: Performance of matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of bacterial strains routinely isolated in a clinical microbiology laboratory. J. Clin. Microbiol. **48**, 1549–1554 (2010)
8. Seng, P., Drancourt, M., Gouriet, F., La Scola, B., Fournier, P.E., Rolain, J.M., Raoult, D.: Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin. Infect. Dis. **49**, 543–551 (2009)
9. Anderton, C.R., Chu, R.K., Tolic, N., Creissen, A., Pasa-Tolic, L.: Utilizing a robotic sprayer for high lateral and mass resolution MALDI FT-ICR MSI of microbial cultures. J. Am. Soc. Mass Spectrom. **27**, 556–559 (2016)
10. Hoffmann, T., Dorrestein, P.C.: Homogeneous matrix deposition on dried agar for MALDI imaging mass spectrometry of microbial cultures. J. Am. Soc. Mass Spectrom. **26**, 1959–1962 (2015)
11. Debois, D., Ongena, M., Cawoy, H., De Pauw, E.: MALDI-FTICR MS imaging as a powerful tool to identify Paenibacillus antibiotics involved in the inhibition of plant pathogens. J. Am. Soc. Mass Spectrom. **24**, 1202–1213 (2013)
12. Yang, J.Y., Phelan, V.V., Simkovsky, R., Watrous, J.D., Trial, R.M., Fleming, T.C., Wenter, R., Moore, B.S., Golden, S.S., Pogliano, K., Dorrestein, P.C.: Primer on agar-based microbial imaging mass spectrometry. J. Bacteriol. **194**, 6023–6028 (2012)
13. Takats, Z., Wiseman, J.M., Gologan, B., Cooks, R.G.: Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. Science **306**, 471–473 (2004)
14. Song, Y.S., Talaty, N., Tao, W.A., Pan, Z.Z., Cooks, R.G.: Rapid ambient mass spectrometric profiling of intact, untreated bacteria using desorption electrospray ionization. Chem. Commun., 61–63 (2007)
15. Song, Y., Talaty, N., Datsenko, K., Wanner, B.L., Cooks, R.G.: In vivo recognition of Bacillus subtilis by desorption electrospray ionization mass spectrometry (DESI-MS). Analyst **134**, 838–841 (2009)
16. Angolini, C.F.F., Vendramini, P.H., Araujo, F.D.S., Araujo, W.L., Augusti, R., Eberlin, M.N., de Oliveira, L.G.: Direct protocol for ambient mass spectrometry imaging on agar culture. Anal. Chem. **87**, 6925–6930 (2015)
17. Bolt, F., Cameron, S.J.S., Karancsi, T., Simon, D., Schaffer, R., Rickards, T., Hardiman, K., Burke, A., Bodai, Z., Perdones-Montero, A., Rebec, M., Balog, J., Takats, Z.: Automated high-throughput identification and characterization of clinically important bacteria and fungi using rapid evaporative ionization mass spectrometry. Anal. Chem. **88**, 9419–9426 (2016)
18. Strittmatter, N., Jones, E.A., Veselkov, K.A., Rebec, M., Bundy, J.G., Takats, Z.: Analysis of intact bacteria using rapid evaporative ionisation mass spectrometry. Chem. Commun. **49**, 6188–6190 (2013)
19. Strittmatter, N., Rebec, M., Jones, E.A., Golf, O., Abdolrasouli, A., Balog, J., Behrends, V., Veselkov, K.A., Takats, Z.: Characterization and identification of clinically relevant microorganisms using rapid evaporative ionization mass spectrometry. Anal. Chem. **86**, 6555–6562 (2014)

20. Dean, S.N., Walsh, C., Goodman, H., van Hoek, M.L.: Analysis of mixed biofilm (Staphylococcus aureus and Pseudomonas aeruginosa) by laser ablation electrospray ionization mass spectrometry. Biofouling **31**, 151–161 (2015)
21. Hamid, A.M., Jarmusch, A.K., Pirro, V., Pincus, D.H., Clay, B.G., Gervasi, G., Cooks, R.G.: Rapid discrimination of bacteria by paper spray mass spectrometry. Anal. Chem. **86**, 7500–7507 (2014)
22. Watrous, J., Roach, P., Heath, B., Alexandrov, T., Laskin, J., Dorrestein, P.C.: Metabolic profiling directly from the Petri dish using nanospray desorption electrospray ionization imaging mass spectrometry. Anal. Chem. **85**, 10385–10391 (2013)
23. Hsu, C.C., ElNaggar, M.S., Peng, Y., Fang, J.S., Sanchez, L.M., Mascuch, S.J., Moller, K.A., Alazzeh, E.K., Pikula, J., Quinn, R.A., Zeng, Y., Wolfe, B.E., Dutton, R.J., Gerwick, L., Zhang, L.X., Liu, X.T., Mansson, M., Dorrestein, P.C.: Real-time metabolomics on living microorganisms using ambient electrospray ionization flow-probe. Anal. Chem. **85**, 7014–7018 (2013)
24. Edwards, R.L., Creese, A.J., Baumert, M., Griffiths, P., Bunch, J., Cooper, H.J.: Hemoglobin variant analysis via direct surface sampling of dried blood spots coupled with high-resolution mass spectrometry. Anal. Chem. **83**, 2265–2270 (2011)
25. Edwards, R.L., Griffiths, P., Bunch, J., Cooper, H.J.: Top-down proteomics and direct surface sampling of neonatal dried blood spots: diagnosis of unknown hemoglobin variants. J. Am. Soc. Mass Spectrom. **23**, 1921–1930 (2012)
26. Griffiths, R.L., Dexter, A., Creese, A.J., Cooper, H.J.: Liquid extraction surface analysis field asymmetric waveform ion mobility spectrometry mass spectrometry for the analysis of dried blood spots. Analyst **14**, 6879–6885 (2015)
27. Martin, N.J., Bunch, J., Cooper, H.J.: Dried blood spot proteomics: surface extraction of endogenous proteins coupled with automated sample preparation and mass spectrometry analysis. J. Am. Soc. Mass Spectrom. **24**, 1242–1249 (2013)
28. Quanico, J., Franck, J., Dauly, C., Strupat, K., Dupuy, J., Day, R., Salzet, M., Fournier, I., Wisztorski, M.: Development of liquid microjunction extraction strategy for improving protein identification from tissue sections. J. Proteome **79**, 200–218 (2013)
29. Sarsby, J., Griffiths, R.L., Race, A.M., Bunch, J., Randall, E.C., Creese, A.J., Cooper, H.J.: Liquid extraction surface analysis mass spectrometry coupled with field asymmetric waveform ion mobility spectrometry for analysis of intact proteins from biological substrates. Anal. Chem. **87**, 6794–6800 (2015)
30. Randall, E.C., Bunch, J., Cooper, H.J.: Direct analysis of intact proteins from escherichia coli colonies by liquid extraction surface analysis mass spectrometry. Anal. Chem. **86**, 10504–10510 (2014)
31. Kertesz, V., Van Berkel, G.J.: Fully automated liquid extraction-based surface sampling and ionization using a chip-based robotic nanoelectrospray platform. J. Mass Spectrom. **45**, 252–260 (2010)
32. Tomlinson, L., Fuchser, J., Futterer, A., Baumert, M., Hassall, D.G., West, A., Marshall, P.S.: Using a single, high mass resolution mass spectrometry platform to investigate ion suppression effects observed during tissue imaging. Rapid Commun. Mass Spectrom. **28**, 995–1003 (2014)
33. Angeletti, S., Dicuonzo, G., Avola, A., Crea, F., Dedej, E., Vailati, F., Farina, C., De Florio, L.: Viridans group streptococci clinical isolates: MALDI-TOF mass spectrometry versus gene sequence-based identification. PLoS One. **10** (2015)
34. Zhou, M.L., Yang, Q.W., Kudinha, T., Zhang, L., Xiao, M., Kong, F.R., Zhao, Y.P., Xu, Y.C.: Using matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) complemented with selected 16S rRNA and gyrB genes sequencing to practically identify clinical important viridans group streptococci (VGS). Front. Microbiol. **7** (2016)
35. Williams, T.L., Monday, S.R., Feng, P.C.H., Musser, S.M.: Identifying new PCR targets for pathogenic bacteria using top-down LC/MS protein discovery. J. Biomol. Tech. **16**, 134–142 (2005)
36. Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K.S., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E.W., Lory, S., Olson, M.V.: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. Nature **406**, 959–964 (2000)
37. Yeats, C., Bateman, A.: The BON domain: a putative membrane-binding domain. Trends Biochem. Sci. **28**, 352–355 (2003)
38. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S.E., Wilkins, M.R., Appel, R.D., Bairoch, A.: Protein identification and analysis tools on the ExPASy Server. In: Walker JM (ed.) Humana, Totowa The Proteomics Protocols Handbook. 571–607 (2005)
39. Liu, X., Sirotkin, Y., Shen, Y., Anderson, G., Tsai, Y.S., Ting, Y.S., Goodlett, D.R., Smith, R.D., Bafna, V., Pevzner, P.A.: Protein identification using top-down spectra. Mol. Cell. Proteomics **11**, M111.008524 (2012)
40. Peschel, A., Otto, M.: Phenol-soluble modulins and staphylococcal infection. Nat. Rev. Microbiol. **11**, 667–673 (2013)
41. Raghunathan, G., Seetharamulu, P., Brooks, B.R., Guy, H.R.: Models of δ-hemolysin membrane channels and crystal-structures. Proteins **8**, 213–225 (1990)
42. Wang, R., Braughton, K.R., Kretschmer, D., Bach, T.H.L., Queck, S.Y., Li, M., Kennedy, A.D., Dorward, D.W., Klebanoff, S.J., Peschel, A., Deleo, F.R., Otto, M.: Identification of novel cytolytic peptides as key virulence determinants for community-associated MRSA. Nat. Med. **13**, 1510–1514 (2007)
43. Holden, M.T.G., Feil, E.J., Lindsay, J.A., Peacock, S.J., Day, N.P.J., Enright, M.C., Foster, T.J., Moore, C.E., Hurst, L., Atkin, R., Barron, A., Bason, N., Bentley, S.D., Chillingworth, C., Chillingworth, T., Churcher, C., Clark, L., Corton, C., Cronin, A., Doggett, J., Dowd, L., Feltwell, T., Hance, Z., Harris, B., Hauser, H., Holroyd, S., Jagels, K., James, K.D., Lennard, N., Line, A., Mayes, R., Moule, S., Mungall, K., Ormond, D., Quail, M.A., Rabbinowitsch, E., Rutherford, K., Sanders, M., Sharp, S., Simmonds, M., Stevens, K., Whitehead, S., Barrell, B.G., Spratt, B.G., Parkhill, J.: Complete genomes of two clinical Staphylococcus aureus strains: evidence for the rapid evolution of virulence and drug resistance. Proc. Natl. Acad. Sci. U. S. A. **101**, 9786–9791 (2004)
44. Gillaspy, A.F., Worrell, V., Orvis, J., Roe, B.A., Dyer, D.W., Iandolo, J.J.: The Staphylococcus aureus NCTC 8325 genome. In: Fischetti V, Novick R, Ferretti J, Portnoy D, Rood J (eds.) ASM Press, Washington Gram-Positive Pathogens, Second Edition. 381–412 (2006)
45. Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K., Takita, C., Saito, R., Ara, T., Nakahigashi, K., Huang, H.C., Hirai, A., Tsuzuki, K., Nakamura, S., Altaf-Ul-Amin, M., Oshima, T., Baba, T., Yamamoto, N., Kawamura, T., Ioka-Nakamichi, T., Kitagawa, M., Tomita, M., Kanaya, S., Wada, C., Mori, H.: Large-scale identification of protein-protein interaction of Escherichia coli K-12. Genome Res. **16**, 686–691 (2006)
46. Butland, G., Peregrin-Alvarez, J.M., Li, J., Yang, W.H., Yang, X.C., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J., Emili, A.: Interaction network containing conserved and essential protein complexes in Escherichia coli. Nature **433**, 531–537 (2005)
47. Wada, A.: Analysis of Escherichia coli ribosomal proteins by an improved two dimensional gel electrophoresis. I. Detection of four new proteins. J. Biochem. **100**, 1583–1594 (1986)
48. Arnold, R.J., Reilly, J.P.: Observation of Escherichia coli ribosomal proteins and their posttranslational modifications by mass spectrometry. Anal. Biochem. **269**, 105–112 (1999)
49. Reid, G.E., Shang, H., Hogan, J.M., Lee, G.U., McLuckey, S.A.: Gas-phase concentration, purification, and identification of whole proteins from complex mixtures. J. Am. Chem. Soc. **124**, 7353–7362 (2002)
50. Babu, M., Arnold, R., Bundalovic-Torma, C., Gagarinova, A., Wong, K.S., Kumar, A., Stewart, G., Samanfar, B., Aoki, H., Wagih, O., Vlasblom, J., Phanse, S., Lad, K., Yu, A.Y.H., Graham, C., Jin, K., Brown, E., Golshani, A., Kim, P., Moreno-Hagelsieb, G., Greenblatt, J., Houry, W.A., Parkinson, J., Emili, A.: Quantitative genome-wide genetic interaction screens reveal global epistatic relationships of protein complexes in Escherichia coli. PLoS Genet. **10**, 15 (2014)
51. Polissi, A., De Laurentis, W., Zangrossi, S., Briani, F., Longhi, V., Pesole, G., Deho, G.: Changes in Escherichia coli transcriptome during acclimatization at low temperature. Res. Microbiol. **154**, 573–580 (2003)
52. Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., ColladoVides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., Shao, Y.: The complete genome sequence of Escherichia coli K-12. Science **277**, 1453–1462 (1997)
53. Oogai, Y., Matsuo, M., Hashimoto, M., Kato, F., Sugai, M., Komatsuzawa, H.: Expression of virulence factors by Staphylococcus aureus grown in serum. Appl. Environ. Microbiol. **77**, 8097–8105 (2011)

APPENDIX 2

Photographs of bacterial colonies



Photographs of bacterial colonies used for the determination of reproducibility (Chapter 3, section 3.3.3). All images were acquired at the same scale (shown in the bottom right panel).

APPENDIX 3

Tandem mass spectra and fragment assignments

3.1 Notes

Section 3.2 lists tandem mass spectra and fragment assignments referred to in Chapter 4. Fragment assignment was performed manually, based on raw data.

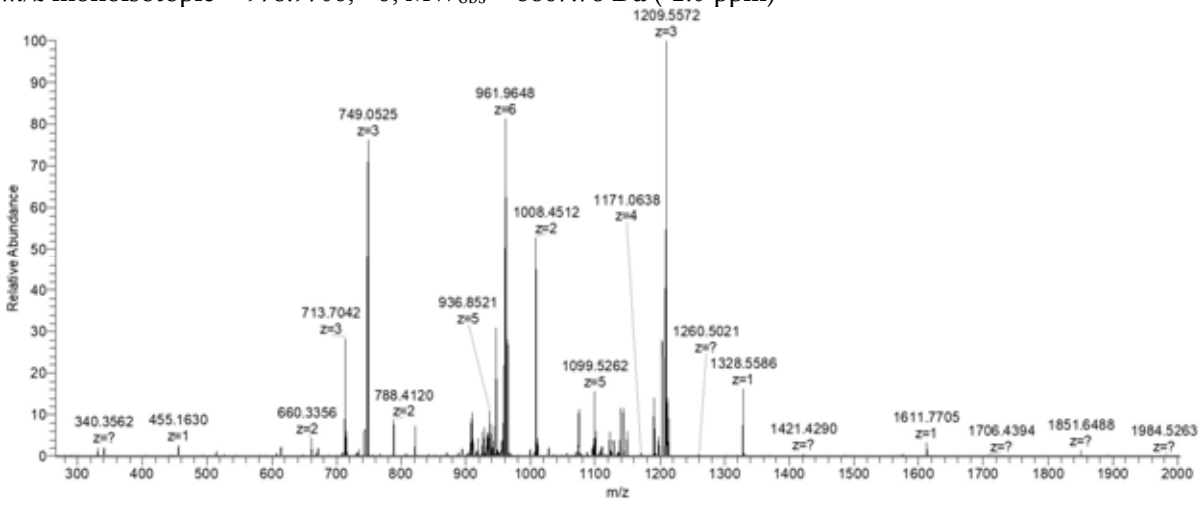
Section 3.3 lists tandem mass spectra and fragment assignments referred to in Chapter 5. Section 3.4 lists tandem mass spectra referred to in Chapter 6. For both sections, fragment assignment was performed automatically by use of ProSightPC, based on deconvoluted data.

3.2 Chapter 4

3.2.1 *E. coli*

3.2.1.1 YciG

UniProt P21361, YciG (con-10)
Seq: AEHRGGSGNFAEDREKASDAGRKGGQHSGGNFKNDPQRASEAGKKGGQQSGGNKSGKS
m/z monoisotopic = 978.9706, +6; MW_{obs} = 5867.78 Da (-2.0 ppm)



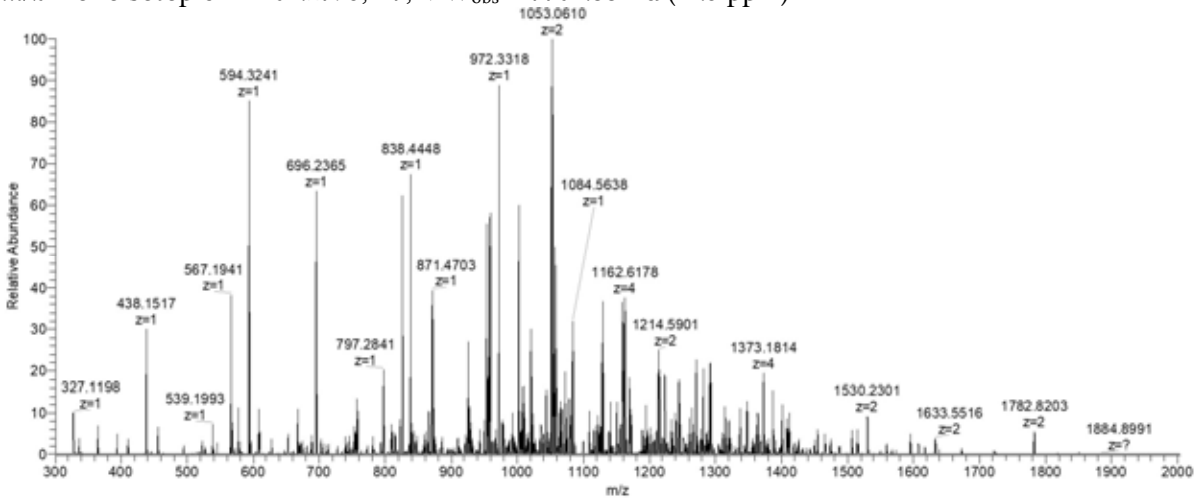
m/z meas	z	m/z calc	Δppm	assignment
378.1980	1	378.1983	-0.79	y4 +1

494.2461	1	494.2470	-1.82	b4 +1
660.3359	2	660.3367	-1.21	y14 +2
742.7148	3	742.7160	-1.62	y23-H2O +3
748.7185	3	748.7195	-1.34	y23 +3
767.0226	3	767.0235	-1.17	b22 +3
788.4124	2	788.4135	-1.40	y17 +2
807.3551	2	807.3562	-1.36	b15 +2
821.4101	1	821.4112	-1.34	y9 +1
849.9021	4	849.9036	-1.76	b33 +4
905.6503	5	905.6495	0.88	y45-NH3 +5
907.1698	4	907.1710	-1.32	b35 +4
909.0537	5	909.0548	-1.21	y45 +5
910.8322	5	910.8338	-1.76	b44 +5
930.6093	6	930.6127	-3.65	b55 +6
932.0588	5	932.0602	-1.50	y46 +5
933.0500	5	933.0475	2.68	b45-NH3 +5
936.4520	5	936.4528	-0.85	b45 +5
940.1190	6	940.1163	2.87	b56 +6
942.6218	6	942.6242	-2.55	y56-H2O +6
945.6244	6	945.6260	-1.69	y56 +6
958.4624	6	958.4637	-1.36	b57-H2O +6
961.4637	6	961.4655	-1.87	b57 +6
964.4816	4	964.4830	-1.45	y39 +4
967.1298	6	967.1331	-3.41	y57 +6
998.9451	2	998.9464	-1.30	b19-H2O +2
1007.9502	2	1007.9517	-1.49	b19 +2
1010.4839	5	1010.4848	-0.89	b49 +5
1027.8903	5	1027.8913	-0.97	b50 +5
1073.5095	5	1073.5084	1.02	b53+5
1074.2509	4	1074.2521	-1.12	b41 +4
1095.7237	5	1095.7221	1.46	b54-NH3 +5
1099.1255	5	1099.1274	-1.73	b54 +5
1113.1303	5	1113.1285	1.62	b55-NH3 +5
1122.5742	2	1122.5756	-1.25	y23 +2
1124.5339	5	1124.5328	0.98	b56-NH3 +5
1127.9358	5	1127.9381	-2.04	b56 +5
1138.2890	4	1138.2905	-1.32	b44 +4
1170.3108	4	1170.3142	-2.91	b45 +4
1191.5680	1	1191.5713	-2.77	y13 +1
1203.2202	3	1203.2221	-1.58	b35-H2O +3
1209.2235	3	1209.2256	-1.74	b35 +3
1213.5330	1	1213.5345	-1.24	b12 +1
1319.6648	1	1319.6662	-1.06	y14 +1

1328.5587	1	1328.5614	-2.03	b13 +1

3.2.1.2 YahO

UniProt P75694, YahO
Seq:
AELMTKAEFEKVESQYEKIGDISTSNEMSTADAKEDLIKKADEKGADVLVLTSGQTDNKIHGTANIY
KKK
m/z monoisotopic = 1101.2795, +7; MW_{obs} = 7701.88 Da (-1.5 ppm)



m/z meas	z	m/z calc	Δppm	assignment
546.2586	1	546.2592	1.10	b5 +1
566.3656	1	566.3661	0.88	y4 +1
580.3322	2	580.3327	0.86	y10 +2
636.8735	2	636.8748	2.04	y11 +2
656.3432	1	656.3436	0.61	b6-H2O +1
674.3535	1	674.3542	1.04	b6 +1
680.3516	2	680.3525	1.32	b12-H2O +2
689.3572	2	689.3578	0.87	b12 +2
700.9210	2	700.9223	1.85	y12 +2
702.0420	3	702.0430	1.42	y19 +3
727.3799	1	727.3807	1.10	b7-H2O +1
733.7334	3	733.7342	1.09	y20-H2O +3
739.7368	3	739.7377	1.22	y20 +3
745.3904	1	745.3913	1.21	b7 +1
746.3761	3	746.3768	0.94	b19-H2O +3
748.9377	2	748.9384	0.93	y13-H2O +2
752.3795	3	752.3803	1.06	b19 +3
753.8783	2	753.8791	1.06	b13 +2
757.9429	2	757.9437	1.06	y13 +2

771.3864	3	771.3875	1.43	b20 +3
772.7596	3	772.7605	1.16	y21 +3
788.3895	2	788.3898	0.38	b14-H2O +2
793.4924	1	793.4931	0.88	y6 +1
809.7287	3	809.7298	1.36	b21 +3
815.4564	2	815.4572	0.98	y14 +2
836.4548	4	836.4559	1.32	y31 +4
838.0919	3	838.0928	1.07	a22 +3
847.4233	3	847.4245	1.42	b22 +3
852.4175	2	852.4191	1.88	b15-H2O +2
856.9749	2	856.9758	1.05	y15-H2O +2
861.4233	2	861.4244	1.28	b15 +2
864.5291	1	864.5302	1.27	y7 +1
865.9800	2	865.9810	1.15	y15 +2
868.4786	4	868.4796	1.15	y32 +4
870.4311	3	870.4316	0.57	b23-H2O +3
874.4332	1	874.4339	0.80	b8 +1
876.4342	3	876.4352	1.14	b23 +3
903.6866	5	903.6878	1.33	y42 +5
921.0040	2	921.0050	1.09	y16-H2O +2
924.5051	3	924.5065	1.51	y26 +3
925.0203	4	925.0216	1.41	y34 +4
930.0092	2	930.0103	1.18	y16 +2
933.1237	3	933.1249	1.29	b25-H2O +3
939.1274	3	939.1284	1.06	b25 +3
942.9548	2	942.9560	1.27	b16 +2
949.5151	2	949.5158	0.74	y17-H2O +2
958.5198	2	958.5211	1.36	y17 +2
965.5767	1	965.5778	1.14	y8 +1
967.2037	3	967.2048	1.14	y27 +3
971.1381	3	971.1392	1.13	b26-H2O +3
977.1415	3	977.1427	1.23	b26 +3
978.5109	5	978.5130	2.15	y45 +5
986.0374	4	986.0390	1.62	y36 +4
993.0305	2	993.0318	1.31	y18-H2O +2
995.9170	5	995.9194	2.41	y46 +5
1002.0357	2	1002.0371	1.40	y18 +2
1004.2139	3	1004.2155	1.59	y28-H2O +3
1007.4760	2	1007.4773	1.29	b17 +2
1010.2175	3	1010.2190	1.48	y28 +3
1010.8242	3	1010.8253	1.09	a27 +3
1014.1520	3	1014.1534	1.38	b27-H2O +3
1018.0613	4	1018.0628	1.47	y37 +4

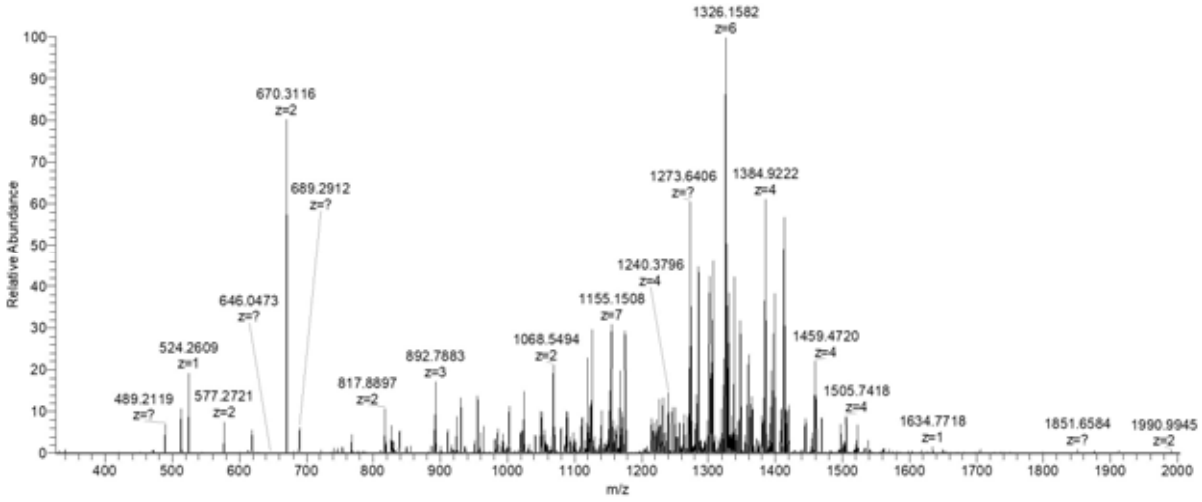
1020.1556	3	1020.1569	1.27	b27 +3
1022.5980	1	1022.5993	1.27	y9 +1
1031.5672	4	1031.5654	-1.74	y38-NH3 +4
1033.5328	5	1033.5354	2.52	y48 +5
1035.8201	4	1035.8220	1.83	y38 +4
1035.9027	5	1035.9033	0.58	b47 +5
1043.5552	2	1043.5556	0.38	y19-H2O +2
1048.5607	3	1048.5614	0.67	y29 +3
1052.5596	2	1052.5609	1.24	y19 +2
1056.1514	5	1056.1522	0.76	y49 +5
1057.8308	3	1057.8335	2.55	b28-H2O +3
1063.8354	3	1063.8371	1.60	b28 +3
1064.5775	4	1064.5788	1.22	y39 +4
1066.2355	3	1066.2369	1.31	y30-H2O +3
1068.0296	6	1068.0343	4.40	a59 +6
1071.5237	2	1071.5248	1.03	b18 +2
1072.6903	7	1072.6948	4.20	y68 +7
1074.9299	5	1074.9284	-1.40	b49-NH3 +5
1078.3331	5	1078.3338	0.65	b49 +5
1080.4040	7	1080.4054	1.30	b69 +7
1082.3361	4	1082.3380	1.76	y40 +4
1109.1013	2	1109.1029	1.44	y20 +2
1114.9366	3	1114.9387	1.88	y31 +3
1119.0602	2	1119.0616	1.25	b19-H2O +2
1120.7614	5	1120.7642	2.50	b51 +5
1123.0541	4	1123.0555	1.25	b40 +4
1124.8539	4	1124.8553	1.24	y42-H2O +4
1128.0651	2	1128.0668	1.51	b19 +2
1129.3556	4	1129.3580	2.13	y42 +4
1137.5697	5	1137.5685	-1.05	b52-NH3 +5
1138.7958	5	1138.7977	1.67	y53 +5
1140.8121	4	1140.8147	2.28	b41 +4
1140.9722	5	1140.9738	1.40	b52 +5
1150.5433	1	1150.5449	1.39	b10 +1
1154.9780	5	1154.9749	-2.68	b53-NH3 +5
1156.5756	2	1156.5776	1.73	b20 +2
1157.6116	4	1157.6155	3.37	y43-H2O +4
1157.8639	4	1157.8615	-2.07	y43-NH3 +4
1158.6371	2	1158.6371	0.00	y21 +2
1162.1159	4	1162.1181	1.89	y43 +4
1164.6043	5	1164.6062	1.63	y54 +5
1165.0707	4	1165.0688	-1.63	b42-H2O +4
1166.3809	5	1166.3792	-1.46	b54-NH3 +5

1169.5688	4	1169.5715	2.31	b42 +4
1188.5503	3	1188.5516	1.09	b32 +3
1193.7860	6	1193.7806	-4.52	y65 +6
1194.3770	4	1194.3787	1.42	y44 +4
1197.2186	5	1197.2189	0.25	y55 +5
1201.8308	4	1201.8321	1.08	b43 +4
1205.0863	2	1205.0858	-0.41	b21-H2O +2
1214.0886	2	1214.0910	1.98	b21 +2
1214.3896	4	1214.3868	-2.31	y45-H2O +4
1215.1772	2	1215.1792	1.65	y22 +2
1222.8879	4	1222.8895	1.31	y45 +4
1233.0249	3	1233.0264	1.22	y34 +3
1233.8532	4	1233.8559	2.19	b44 +4
1238.6105	5	1238.6111	0.48	b57 +5
1238.9520	6	1238.9559	3.15	b68 +6
1240.2365	5	1240.2370	0.40	y57 +5
1240.3922	4	1240.3908	-1.13	y46-NH3 +4
1244.6451	4	1244.6475	1.93	y46 +4
1251.3100	6	1251.3093	-0.56	y68 +6
1255.7060	2	1255.7081	1.67	y23-H2O +2
1260.3080	6	1260.3050	-2.38	b69 +6
1264.7114	2	1264.7134	1.58	y23 +2
1269.9064	4	1269.9094	2.36	y47 +4
1270.6293	2	1270.6331	2.99	b22 +2
1278.6350	1	1278.6399	3.83	b11 +1
1285.8609	5	1285.8592	-1.32	y59 +5
1287.4109	4	1287.4108	-0.08	y48-NH3 +4
1291.6646	4	1291.6674	2.17	y48 +4
1297.9429	3	1297.9432	0.23	b35 +3
1305.1425	2	1305.1438	1.00	b23-H2O +2
1306.2539	5	1306.2502	-2.83	b60-NH3 +5
1308.7111	3	1308.7074	-2.83	y36-NH3 +3
1309.6527	5	1309.6555	2.14	b60 +5
1314.1466	2	1314.1491	1.90	b23 +2
1314.3810	3	1314.3829	1.45	y36 +3
1319.3946	4	1319.3944	-0.15	b48 +4
1319.9366	4	1319.9384	1.36	y49 +4
1322.2249	2	1322.2269	1.51	y24 +2
1336.2828	3	1336.2855	2.02	b36 +3
1343.4100	4	1343.4087	-0.97	b49-NH3 +4
1347.6611	4	1347.6654	3.19	b49 +4
1355.6648	2	1355.6677	2.14	b24-H2O +2
1357.7458	2	1357.7454	-0.29	y25 +2

1362.9485	4	1362.9505	1.47	y51 +4
1364.6702	2	1364.6729	1.98	b24 +2
1368.1771	4	1368.1758	-0.95	b50-NH3 +4
1372.4303	4	1372.4325	1.60	b50 +4
1373.9778	3	1373.9802	1.75	b37 +3
1377.2482	2	1377.2509	1.96	y26-H2O +2
1377.7069	1	1377.7083	1.02	b12 +1
1380.7564	3	1380.7603	2.82	y38 +3
1386.2536	2	1386.2561	1.80	y26 +2
1391.2181	4	1391.2215	2.44	y52 +4
1396.4443	4	1396.4469	1.86	b51-NH3 +4
1399.1819	2	1399.1837	1.29	b25-H2O +2
1400.6984	4	1400.7035	3.64	b51 +4
1406.0019	3	1405.9994	-1.78	b38-NH3 +3
1408.1865	2	1408.1889	1.70	b25 +2
1411.6723	3	1411.6749	1.84	b38 +3
1419.1008	3	1419.1026	1.27	y39 +3
1421.7086	4	1421.7088	0.14	b52-NH3 +4
1442.7756	3	1442.7816	4.16	y40 +3
1450.3016	2	1450.3036	1.38	y27 +2
1454.3717	3	1454.3732	1.03	b39 +3
1465.2088	2	1465.2104	1.09	b26 +2
1476.4608	3	1476.4642	2.30	y41 +3
1505.4723	3	1505.4749	1.73	y42 +3
1506.7484	1	1506.7509	1.66	b13 +1
1514.8229	2	1514.8249	1.32	y28 +2
1514.8772	1	1514.8802	1.98	y13 +1
1521.2244	2	1521.2184	-3.94	b27-NH3 +2
1529.7283	2	1529.7317	2.22	b27 +2
1559.0895	3	1559.0929	2.18	b42 +3
1572.3364	2	1572.3384	1.27	y29 +2
1595.2500	2	1595.2519	1.19	b28 +2
1607.8542	2	1607.8570	1.74	y30 +2
1721.8397	1	1721.8415	1.05	b15 +1
1782.3206	2	1782.3238	1.80	b32 +2

3.2.1.3 YjbJ

UniProt P68206, YjbJ
Seq: MNKDEAGGNWKQFKGKVKEQWGKLTDDDMTIIEGKRDQLVGKIQERYGYQKDQAEKEVVD
m/z monoisotopic = 1189.5904, +7; MW_{obs} = 8320.08 (-1.3 ppm)



m/z meas	z	m/z calc	Δppm	assignment
357.1592	1	357.1591	0.28	b3-NH3 +1
374.1856	1	374.1857	-0.27	b3 +1
471.2018	1	471.2020	-0.42	b4-H2O +1
512.7512	2	512.7516	-0.78	y7 +2
524.2613	1	524.2616	-0.57	y3 +1
568.2672	2	568.2676	-0.70	y8-H2O +2
577.2724	2	577.2729	-0.87	y8 +2
600.2443	1	600.2446	-0.50	b5-H2O +1
618.2547	1	618.2552	-0.81	b5 +1
637.3109	3	637.3112	-0.47	y14 +3
661.7986	2	661.7993	-1.06	y9-NH3 +2
670.3120	2	670.3125	-0.75	y9 +2
680.3087	2	680.3091	-0.59	b12 +2
689.2917	1	689.2923	-0.87	b6 +1
704.0040	3	704.0045	-0.71	y16 +3
727.8254	2	727.8260	-0.82	y10 +2
740.6864	3	740.6871	-0.95	y17-H2O +3
746.3130	1	746.3138	-1.07	b7 +1
746.6899	3	746.6907	-1.07	y17 +3
753.8427	2	753.8434	-0.93	b13 +2
767.3464	1	767.3471	-0.91	y5 +1
777.3595	2	777.3602	-0.90	y11 +2
785.0322	3	785.0330	-1.02	y18 +3
787.8766	4	787.8772	-0.76	y24 +4
803.3342	1	803.3352	-1.24	b8 +1
809.3767	2	809.3776	-1.11	b14-NH3 +2
817.8900	2	817.8908	-0.98	b14 +2
825.0744	3	825.0742	0.24	b21-NH3 +3

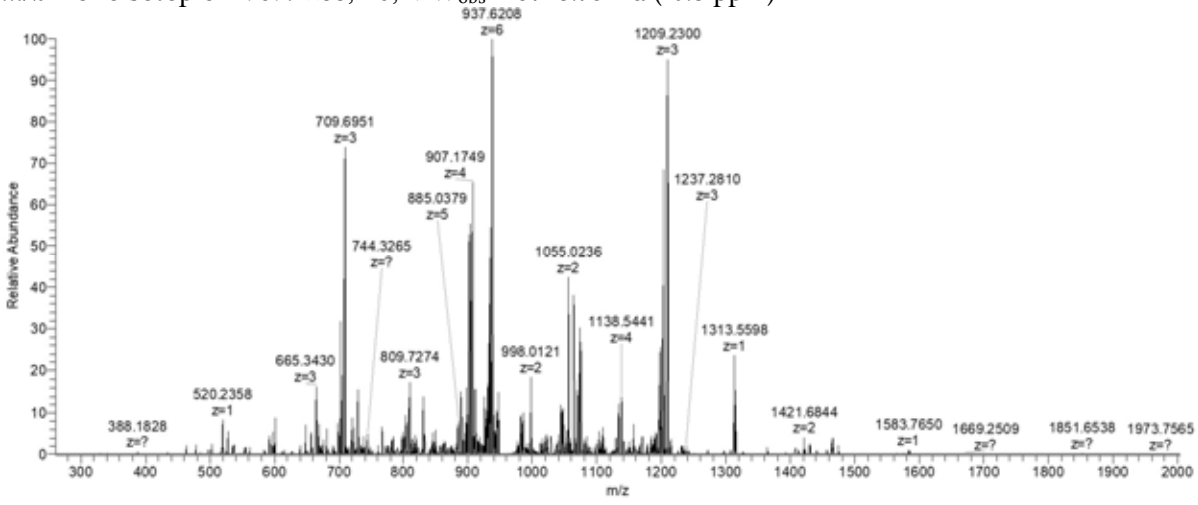
826.8936	2	826.8944	-0.97	y12 +2
830.7490	3	830.7497	-0.84	b21 +3
846.4004	2	846.4016	-1.42	b15 +2
849.7560	3	849.7569	-1.06	b22 +3
867.4174	4	867.4184	-1.15	b30 +4
880.4230	4	880.4236	-0.68	y27 +4
882.4090	2	882.4104	-1.59	y13-H2O +2
886.4505	3	886.4517	-1.35	b23-H2O +3
891.4148	2	891.4157	-1.01	y13 +2
892.4541	3	892.4552	-1.23	b23 +3
900.8533	5	900.8547	-1.55	y36 +5
910.4480	2	910.4490	-1.10	b16 +2
917.3770	1	917.3782	-1.31	b9 +1
924.1452	3	924.1464	-1.30	b24-H2O +3
930.1488	3	930.1499	-1.18	b24 +3
950.9768	2	950.9780	-1.26	b17-H2O +2
955.4621	2	955.4632	-1.15	y14 +2
957.8277	3	957.8289	-1.25	b25-H2O +3
959.9819	2	959.9833	-1.46	b17 +2
963.8323	3	963.8325	-0.21	b25 +3
971.8958	5	971.8968	-1.03	y39 +5
972.8180	6	972.8181	-0.10	y48 +6
974.4948	7	974.4959	-1.13	y56 +7
992.1047	5	992.1064	-1.71	y40 +5
995.6745	6	995.6765	-2.01	b51 +6
998.1319	3	998.1335	-1.60	y23 +3
1002.1734	3	1002.1748	-1.40	b26 +3
1008.2520	8	1008.2551	-3.07	y67-NH3 +8
1010.3811	8	1010.3834	-2.28	y67 +8
1014.8481	6	1014.8476	0.49	b52 +6
1017.7581	8	1017.7589	-0.79	b68+H2O +8
1018.3129	5	1018.3145	-1.57	y41 +5
1019.9827	2	1019.9845	-1.76	y15 +2
1022.3368	6	1022.3367	0.10	y50-NH3 +6
1024.0303	2	1024.0307	-0.39	b18 +2
1024.6348	8	1024.6388	-3.90	y68 +8
1050.1661	3	1050.1672	-1.05	y24 +3
1055.5015	2	1055.5031	-1.52	y16 +2
1060.9216	5	1060.9199	1.60	y43-NH3 +5
1064.3234	5	1064.3252	-1.69	y43 +5
1068.0295	6	1068.0307	-1.12	y52 +6
1078.8579	3	1078.8594	-1.39	b28 +3
1083.1174	7	1083.1195	-1.94	y62 +7

1087.3292	5	1087.3306	-1.29	y44 +5
1088.5531	2	1088.5520	1.01	b19 +2
1088.8340	7	1088.8331	0.83	y63-NH3 +7
1090.8837	6	1090.8865	-2.57	b56 +6
1091.2639	7	1091.2655	-1.47	y63 +7
1098.9796	7	1098.9813	-1.55	y64-NH3 +7
1101.4115	7	1101.4136	-1.91	y64 +7
1103.4559	1	1103.4575	-1.45	b10 +1
1104.1363	5	1104.1349	1.27	y45-NH3 +5
1105.8890	6	1105.8912	-1.99	y54 +6
1107.5362	5	1107.5402	-3.61	y45 +5
1111.0196	2	1111.0191	0.45	y17-NH3 +2
1112.5573	6	1112.5571	0.18	y55-NH3 +6
1114.8410	7	1114.8413	-0.27	b66 +7
1117.2736	7	1117.2753	-1.52	y65-H2O +7
1117.4174	7	1117.4159	1.34	y65-NH3 +7
1119.5305	2	1119.5323	-1.61	y17 +2
1119.8461	7	1119.8483	-1.96	y65 +7
1125.8152	4	1125.8165	-1.15	y36 +4
1126.7537	5	1126.7517	1.78	y46-NH3 +5
1128.9034	6	1128.9050	-1.42	b58 +6
1130.1571	5	1130.1570	0.09	y46 +5
1133.8480	7	1133.8483	-0.26	y66-NH3 +7
1136.2787	7	1136.2807	-1.76	y66 +7
1154.5768	7	1154.5800	-2.77	y67 +7
1155.7749	5	1155.7760	-0.95	y47 +5
1158.0782	4	1158.0772	0.86	y37 +4
1161.2529	6	1161.2554	-2.15	y57 +6
1164.5833	6	1164.5876	-3.69	b60 +6
1167.1780	5	1167.1803	-1.97	y48 +5
1170.8675	7	1170.8718	-3.67	y68 +7
1173.5608	3	1173.5623	-1.28	y27 +3
1177.0443	2	1177.0458	-1.27	y18 +2
1179.5948	6	1179.5967	-1.61	y58-H2O +6
1182.0926	4	1182.0916	0.85	y38-NH3 +4
1182.5936	6	1182.5985	-4.14	y58 +6
1186.3463	4	1186.3482	-1.60	y38 +4
1191.2095	5	1191.2050	3.78	b51-NH3 +5
1193.9145	3	1193.9168	-1.93	b31 +3
1194.6075	5	1194.6103	-2.34	b51 +5
1204.3936	5	1204.3961	-2.08	y49 +5
1214.2110	5	1214.2104	0.49	b52-NH3 +5
1217.1063	6	1217.1079	-1.31	b62 +6

1226.6043	5	1226.6025	1.47	y50-NH3 +5
1231.5505	1	1231.5524	-1.54	b11 +1
1234.2231	5	1234.2231	0.00	a53-NH3 +5
1235.2646	3	1235.2678	-2.59	y29 +3
1239.8790	4	1239.8811	-1.69	y40 +4
1241.0914	2	1241.0933	-1.53	y19 +2
1245.6180	2	1245.6210	-2.41	b21 +2
1252.4108	5	1252.4111	-0.24	y51-NH3 +5
1268.2880	3	1268.2906	-2.05	y30 +3
1272.6390	4	1272.6413	-1.81	y41 +4
1274.6245	3	1274.6257	-0.94	b33 +3
1283.2421	5	1283.2434	-1.01	b55 +5
1299.8601	5	1299.8581	1.54	a56-NH3 +5
1301.3952	4	1301.3980	-2.15	y42 +4
1322.5892	1	1322.5913	-1.59	y9-NH3 +1
1326.9969	3	1326.9995	-1.96	a35 +3
1330.1527	4	1330.1547	-1.50	y43 +4
1339.6159	1	1339.6178	-1.42	y9 +1
1358.9084	4	1358.9115	-2.28	y44 +4
1384.1705	4	1384.1734	-2.10	y45 +4
1412.4410	4	1412.4444	-2.41	y46 +4
1506.6773	1	1506.6794	-1.39	b13 +1
1634.7709	1	1634.7710	-0.06	y12-H2O +1

3.2.1.4 YmdF

UniProt P56614, YmdF (con-10)
Seq: MANHRGGSGNFAEDRERASEAGKKGGQHSGGNFKNDPQRASEAGKKGGKSSHGKSDN
m/z monoisotopic = 959.1288, +6; MW_{obs} = 5748.73 Da (-0.8 ppm)



m/z meas	z	m/z calc	Δppm	assignment
446.1879	1	446.1882	0.67	y4-NH3 +1
462.2207	1	462.2208	0.22	b4-NH3 +1
463.2146	1	463.2147	0.22	y4 +1
479.2472	1	479.2473	0.21	b4 +1
502.2256	1	502.2256	0.00	y5-H2O +1
503.2095	1	503.2096	0.20	y5-NH3 +1
520.2358	1	520.2362	0.77	y5 +1
528.2469	2	528.2469	0.00	y11-H2O +2
528.7388	2	528.7389	0.19	y11-NH3 +2
535.2495	2	535.2497	0.37	b11 +2
537.2519	2	537.2522	0.56	y11 +2
579.6032	3	579.6035	0.52	b16-H2O +3
585.6067	3	585.6070	0.51	b16 +3
592.2941	2	592.2944	0.51	y12-H2O +2
592.7859	2	592.7864	0.84	y12-NH3 +2
601.2992	2	601.2996	0.67	y12 +2
638.2964	3	638.2968	0.63	b18 +3
640.2682	1	640.2685	0.47	y6-NH3 +1
657.2946	1	657.2951	0.76	y6 +1
671.9790	3	671.9793	0.45	a19 +3
675.3069	3	675.3074	0.74	b19-H2O +3
681.3103	3	681.3110	1.03	b19 +3
685.3449	2	685.3446	-0.44	y14-NH3 +2
697.5840	4	697.5846	0.86	b27-H2O +4
702.0867	4	702.0873	0.85	b27 +4
703.3574	3	703.3579	0.71	y21-H2O +3
709.3609	3	709.3614	0.70	y21 +3
719.3419	4	719.3426	0.97	b28-H2O +4
720.8620	2	720.8631	1.53	y15-NH3 +2
723.8446	4	723.8453	0.97	b28 +4
723.9969	3	723.9971	0.28	b21 +3
727.3000	1	727.3006	0.82	y7-NH3 +1
729.3758	2	729.3764	0.82	y15 +2
733.5975	4	733.5980	0.68	b29-H2O +4
735.3342	2	735.3351	1.22	b14 +2
738.1001	4	738.1006	0.68	b29 +4
744.3265	1	744.3271	0.81	y7 +1
747.8525	4	747.8534	1.20	b30-H2O +4
752.3559	4	752.3560	0.13	b30 +4
760.6913	3	760.6919	0.79	b22-H2O +3
766.6947	3	766.6955	1.04	b22 +3
771.3779	3	771.3754	-3.24	y30 +4

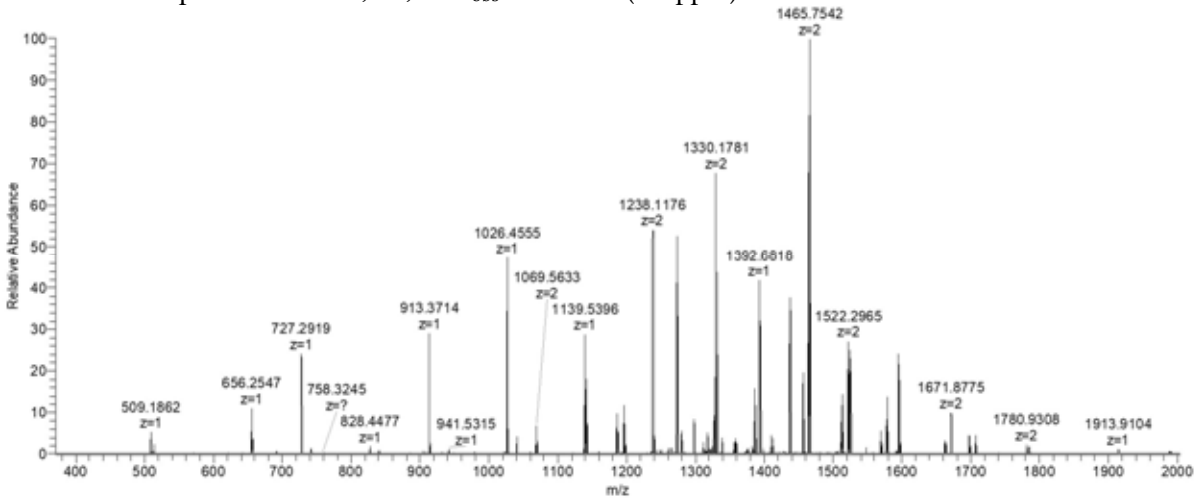
776.3637	4	776.3641	0.52	b31-H2O +4
780.8660	4	780.8667	0.90	b31 +4
785.8580	2	785.8589	1.15	a15 +2
790.8505	2	790.8511	0.76	b15-H2O +2
799.8556	2	799.8564	1.00	b15 +2
803.3895	3	803.3903	1.00	b23-H2O +3
809.3930	3	809.3938	0.99	b23 +3
813.1303	4	813.1312	1.11	b32-H2O +4
814.3320	1	814.3326	0.74	y8-NH3 +1
817.6330	4	817.6338	0.98	b32 +4
822.3961	3	822.3974	1.58	b24-H2O +3
827.3989	4	827.3981	-0.97	y33-H2O +4
831.3584	1	831.3591	0.84	y8 +1
845.1541	4	845.1549	0.95	b33-H2O +4
847.4081	3	847.4081	0.00	b25 +3
849.6567	4	849.6576	1.06	b33 +4
851.3860	1	851.3867	0.82	b9 +1
855.8029	5	855.8042	1.52	b41-H2O +5
859.4052	5	859.4063	1.28	b41 +5
859.6681	4	859.6679	-0.23	y34-NH3 +4
863.9235	4	863.9245	1.16	y34 +4
868.9005	2	868.9016	1.27	b16-H2O +2
869.5880	6	869.5896	1.84	b51-H2O +6
870.0104	5	870.0116	1.38	b42-H2O +5
872.5910	6	872.5914	0.46	b51 +6
873.6136	5	873.6138	0.23	b42 +5
873.6652	4	873.6657	0.57	b34-H2O +4
877.4383	3	877.4392	1.03	y25 +3
878.1685	4	878.1683	-0.23	b34 +4
884.0900	3	884.0908	0.90	b26-H2O +3
884.8380	5	884.8378	-0.23	y43-NH3 +5
888.2422	5	888.2431	1.01	y43 +5
890.0932	3	890.0943	1.24	b26 +3
900.4421	6	900.4424	0.33	b53-H2O +6
902.4212	4	902.4224	1.33	b35-H2O +4
903.4432	6	903.4441	1.00	b53 +6
906.9240	4	906.9250	1.10	b35 +4
907.0339	5	907.0349	1.10	b44-H2O +5
910.6358	5	910.6370	1.32	b44 +5
914.9461	6	914.9477	1.75	b54-H2O +6
917.9483	6	917.9495	1.31	b54 +6
932.6526	5	932.6539	1.39	b45-H2O +5
934.1176	6	934.1189	1.39	b55-H2O +6

936.2550	5	936.2560	1.07	b45 +5
937.1192	6	937.1206	1.49	b55 +6
976.8028	3	976.8030	0.20	y29-NH3 +3
977.2754	5	977.2728	-2.66	y47-NH3 +5
981.0799	5	981.0815	1.63	b48-H2O +5
982.4781	3	982.4785	0.41	y29 +3
984.6829	5	984.6836	0.71	b48 +5
1002.0903	5	1002.0900	-0.30	b49 +5
1012.4563	2	1012.4575	1.19	b19-H2O +2
1015.8923	5	1015.8943	1.97	b50-H2O +5
1019.4957	5	1019.4964	0.69	b50 +5
1021.4616	2	1021.4628	1.17	b19 +2
1022.1597	3	1022.1612	1.47	y30-H2O +3
1028.1634	3	1028.1648	1.36	y30 +3
1040.8189	3	1040.8199	0.96	b31 +3
1043.3047	5	1043.3061	1.34	b51-H2O +5
1046.9068	5	1046.9082	1.34	b51 +5
1054.5322	2	1054.5332	0.95	y21-H2O +2
1056.4688	1	1056.4705	1.61	y11-NH3 +1
1058.3121	5	1058.3125	0.38	b52 +5
1063.5371	2	1063.5385	1.32	y21 +2
1069.5020	4	1069.5035	1.40	b41-H2O +4
1073.4959	1	1073.4970	1.02	y11 +1
1074.0046	4	1074.0061	1.40	b41 +4
1080.3270	5	1080.3294	2.22	b53-H2O +5
1083.9297	5	1083.9315	1.66	b53 +5
1087.2606	4	1087.2627	1.93	b42-H2O +4
1091.7635	4	1091.7654	1.74	b42 +4
1101.5173	4	1101.5181	0.73	b43-H2O +4
1103.1896	3	1103.1898	0.18	y33-NH3 +3
1106.0201	4	1106.0207	0.54	b43 +4
1108.8633	3	1108.8653	1.80	y33 +3
1112.0455	2	1112.0467	1.08	y22-H2O +2
1138.0428	4	1138.0445	1.49	b44 +4
1140.5320	2	1140.5343	2.02	b22-H2O +2
1145.8883	3	1145.8881	-0.17	y34-NH3 +3
1149.5378	2	1149.5396	1.57	b22 +2
1151.5601	3	1151.5636	3.04	y34 +3
1164.5512	3	1164.5518	0.52	b34-H2O +3
1165.8140	4	1165.8116	-2.06	b45-NH3 +4
1169.5594	2	1169.5602	0.68	y23-NH3 +2
1178.0719	2	1178.0734	1.27	y23 +2
1184.5642	1	1184.5654	1.01	y12-NH3 +1

1201.5904	1	1201.5920	1.33	y12 +1
1202.8922	3	1202.8941	1.58	b35-H2O +3
1208.8958	3	1208.8976	1.49	b35 +3
1213.5850	2	1213.5870	1.65	b23 +2
1230.9449	3	1230.9446	-0.24	y37-H2O +3
1233.6059	2	1233.6076	1.38	y24-NH3 +2
1236.9464	3	1236.9481	1.37	y37 +3
1242.1194	2	1242.1209	1.21	y24 +2
1296.5332	1	1296.5352	1.54	b13-NH3 +1
1306.6543	2	1306.6498	-3.44	y25-H2O +2
1313.5598	1	1313.5618	1.52	b13 +1
1363.6684	2	1363.6713	2.13	y26-H2O +2
1369.6817	1	1369.6819	0.15	y14-NH3 +1
1392.6726	2	1392.6740	1.01	y27-NH3 +2
1421.1829	2	1421.1848	1.34	y28-NH3 +2
1429.6954	2	1429.6980	1.82	y28 +2
1440.7167	1	1440.7190	1.60	y15-NH3 +1
1457.7432	1	1457.7455	1.58	y15 +1
1464.6999	2	1464.7008	0.61	y29-NH3 +2
1473.2109	2	1473.2141	2.17	y29 +2
1533.2264	2	1533.2302	2.48	y30-NH3 +2

3.2.1.5 CydX

UniProt P56100, CydX
Seq: fMWYFAWILGTLLACSFGVITALALEHVESGKAGQEDI
m/z monoisotopic = 1356.6814, +3; MW_{obs} = 4067.02 (0.3 ppm)

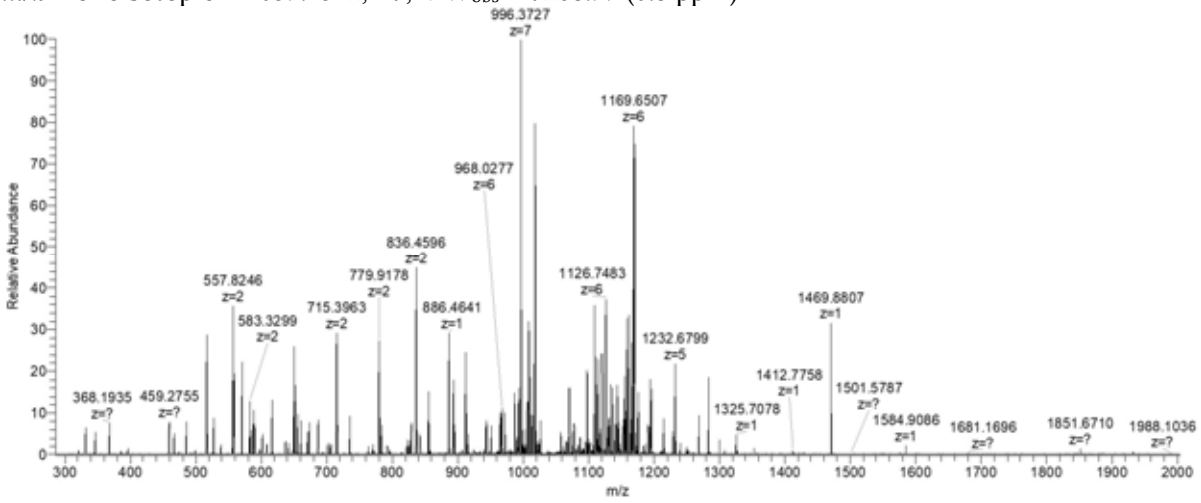


m/z meas	z	m/z calc	Δppm	assignment
509.1862	1	509.1853	-1.77	b3 +1

656.2547	1	656.2537	-1.52	b4 +1
727.2919	1	727.2908	-1.51	b5 +1
817.4056	1	817.4050	-0.73	y8 +1
904.4385	1	904.4371	-1.55	y9 +1
913.3714	1	913.3701	-1.42	b6 +1
1026.4555	1	1026.4542	-1.27	b7 +1
1040.5511	2	1040.5497	-1.35	y20 +2
1040.5511	2	1040.5497	-1.35	y20 +2
1069.0618	2	1069.0604	-1.31	y21 +2
1069.0618	2	1069.0604	-1.31	y21 +2
1139.5396	1	1139.5383	-1.14	b8 +1
1142.5949	2	1142.5946	-0.26	y22 +2
1186.1118	2	1186.1106	-1.01	y23 +2
1196.5608	1	1196.5597	-0.92	b9 +1
1237.6161	2	1237.6152	-0.73	y24 +2
1273.1346	2	1273.1338	-0.63	y25 +2
1279.5983	1	1279.5968	-1.17	b10-H2O +1
1297.6087	1	1297.6074	-1.00	b10 +1
1329.6765	2	1329.6758	-0.53	y26 +2
1386.2185	2	1386.2178	-0.50	y27 +2
1392.6818	1	1392.6809	-0.65	b11-H2O +1
1410.6927	1	1410.6915	-0.85	b11 +1
1436.7424	2	1436.7417	-0.49	y28 +2
1456.2475	2	1456.2471	-0.27	y29-H2O +2
1465.2530	2	1465.2524	-0.41	y29 +2
1512.7894	2	1512.7892	-0.13	y30-H2O +2
1521.7948	2	1521.7944	-0.26	y30 +2
1523.7742	1	1523.7755	0.85	b12 +1
1569.3314	2	1569.3312	-0.13	y31-H2O +2
1578.3360	2	1578.3365	0.32	y31 +2
1594.8131	1	1594.8127	-0.25	b13 +1
1662.3709	2	1662.3709	0.00	y32-H2O +2
1671.3758	2	1671.3761	0.18	y32 +2
1697.8218	1	1697.8218	0.00	b14 +1
1706.8946	2	1706.8947	0.06	y33 +2
1780.4275	2	1780.4289	0.79	y34 +2
1784.8538	1	1784.8539	0.06	b15 +1
1913.9104	1	1913.9117	0.68	b16-H2O +1
1988.9446	1	1988.9437	-0.45	b17 +1

3.2.1.6 L29

UniProt P0A7M6, L29
Seq: MKAKELREKSVEELNTELLNLLREQFNLRMQAASGQLQQSHLLKQVRRDVARVKTLLNEKAGA
m/z monoisotopic = 1039.4342, +7; MW_{obs} = 7268.99 (0.8 ppm)

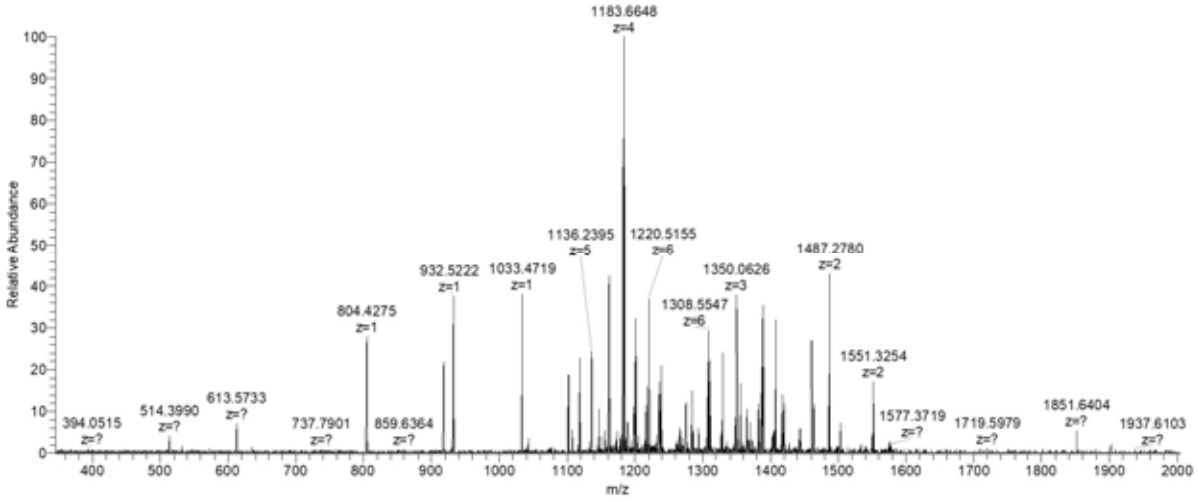


m/z meas	z	m/z calc	Δppm	assignment
331.1805	1	331.1798	-2.11	b3 +1
346.2091	1	346.2085	-1.73	y4 +1
459.2755	1	459.2748	-1.52	b4 +1
475.2520	1	475.2511	-1.89	y5 +1
557.8246	2	557.8237	-1.61	b9 +2
570.3078	1	570.3068	-1.75	b5-H2O +1
588.3181	1	588.3174	-1.19	b5 +1
601.3404	2	601.3397	-1.16	b10 +2
636.8775	2	636.8765	-1.57	a11 +2
641.8693	2	641.8686	-1.09	b11-H2O +2
650.8748	2	650.8739	-1.38	b11 +2
701.4020	1	701.4015	-0.71	b6 +1
715.3963	2	715.3952	-1.54	b12 +2
735.4444	2	735.4436	-1.09	y14 +2
770.9126	2	770.9112	-1.82	b13-H2O +2
779.9178	2	779.9165	-1.67	b13 +2
792.9578	2	792.9570	-1.01	y15 +2
815.4631	1	815.4621	-1.23	y8 +1
822.4623	2	822.4611	-1.46	a14 +2
836.4596	2	836.4585	-1.32	b14 +2
857.5037	1	857.5026	-1.28	b7 +1
893.4812	2	893.4800	-1.34	b15 +2
916.5111	1	916.5098	-1.42	y9 +1
944.0050	2	944.0038	-1.27	b16 +2

967.6949	6	967.6924	-2.58	b49 +6
986.5464	1	986.5452	-1.22	b8 +1
1005.9884	7	1005.9859	-2.49	b60-NH3 +7
1008.4198	7	1008.4182	-1.59	b60 +7
1016.1368	7	1016.1340	-2.76	b61-NH3 +7
1018.5681	7	1018.5664	-1.67	b61 +7
1026.7134	7	1026.7123	-1.07	b62 +7
1077.6050	5	1077.6053	0.28	y47 +5
1097.8165	5	1097.8148	-1.55	y48 +5
1111.6237	6	1111.6189	-4.32	b57-NH3 +6
1120.6248	5	1120.6234	-1.25	y49 +5
1133.4648	6	1133.4638	-0.88	b58 +6
1136.1293	6	1136.1265	-2.46	y59 +6
1143.2357	5	1143.2402	3.94	y50 +5
1154.6412	6	1154.6379	-2.86	y60-NH3 +6
1157.4765	6	1157.4757	-0.69	y60 +6
1165.6474	5	1165.6434	-3.43	y51-NH3 +5
1166.4818	6	1166.4775	-3.69	y61-NH3 +6
1169.0534	5	1169.0487	-4.02	y51 +5
1169.3167	6	1169.3152	-1.28	y61 +6
1176.3203	6	1176.3201	-0.17	b60 +6
1194.8618	5	1194.8572	-3.85	y52 +5
1228.4735	5	1228.4752	1.38	y54-H2O +5
1232.0786	5	1232.0773	-1.06	y54 +5
1308.2876	2	1308.2826	-3.82	y23 +2
1469.8807	1	1469.8798	-0.61	y14 +1
1584.9086	1	1584.9068	-1.14	y15 +1

3.2.1.7 YdfK

UniProt P76154, YdfK
Seq:
MKSKDTLKWFPAQLPEVRIILGDAVVEVAKQGRPINTRTLLDYIEGNIKKKSWLDNKELLQTAISVLK
DNQNLNGKM
m/z monoisotopic = 1254.2614, +7; MW_{obs} = 8772.78 Da (-2.2 ppm)

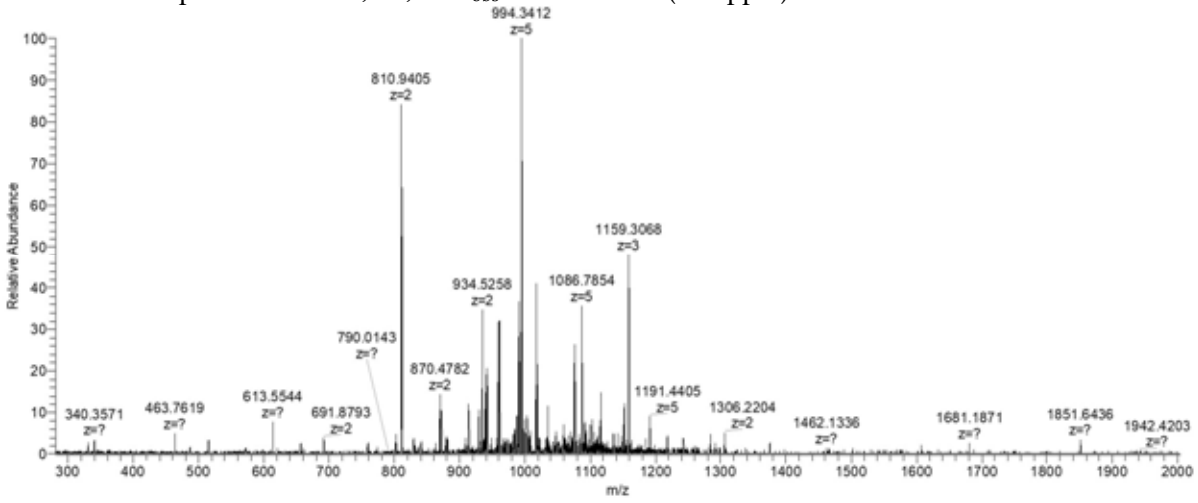


m/z meas	z	m/z calc	Δppm	assignment
804.4275	1	804.4284	-1.12	b7 +1
918.4452	1	918.4462	-1.09	y8 +1
932.5222	1	932.5234	-1.29	b8 +1
1033.4719	1	1033.4731	-1.16	y9 +1
1100.5917	1	1100.5921	-0.36	b9-H2O +1
1115.0962	2	1115.0990	-2.51	y20 +2
1118.6014	1	1118.6027	-1.16	b9 +1
1161.2561	5	1161.2582	-1.81	b51 +5
1161.5668	1	1161.5681	-1.12	y10 +1
1182.9134	4	1182.9151	-1.44	b42 +4
1186.3253	6	1186.3248	0.42	b62-NH3 +6
1198.1639	6	1198.1643	-0.33	b63-NH3 +6
1200.9991	6	1201.0021	-2.50	b63 +6
1219.8487	6	1219.8494	-0.57	b64 +6
1236.1659	2	1236.1680	-1.70	y22 +2
1238.4967	5	1238.4973	-0.48	b54 +5
1265.6682	1	1265.6711	-2.29	b10 +1
1266.7117	6	1266.7118	-0.08	b67-H2O +6
1274.6501	1	1274.6521	-1.57	y11 +1
1293.6793	2	1293.6814	-1.62	y23 +2
1307.8907	6	1307.8876	2.37	y69 +6
1329.2377	6	1329.2410	-2.48	b70 +6
1349.3933	3	1349.3963	-2.22	y35 +3
1369.7574	2	1369.7567	0.51	b24 +2
1373.7165	1	1373.7206	-2.98	y12 +1
1385.4313	6	1385.4368	-3.97	b73-H2O +6
1387.2716	4	1387.2734	-1.30	b49 +4
1406.5798	5	1406.5841	-3.06	b61 +5

1417.1166	6	1417.1133	2.33	y75-NH3 +6
1460.7511	1	1460.7526	-1.03	y13 +1
1486.7768	2	1486.7791	-1.55	y26 +2
1550.8236	2	1550.8266	-1.93	y27 +2

3.2.1.8 SRA

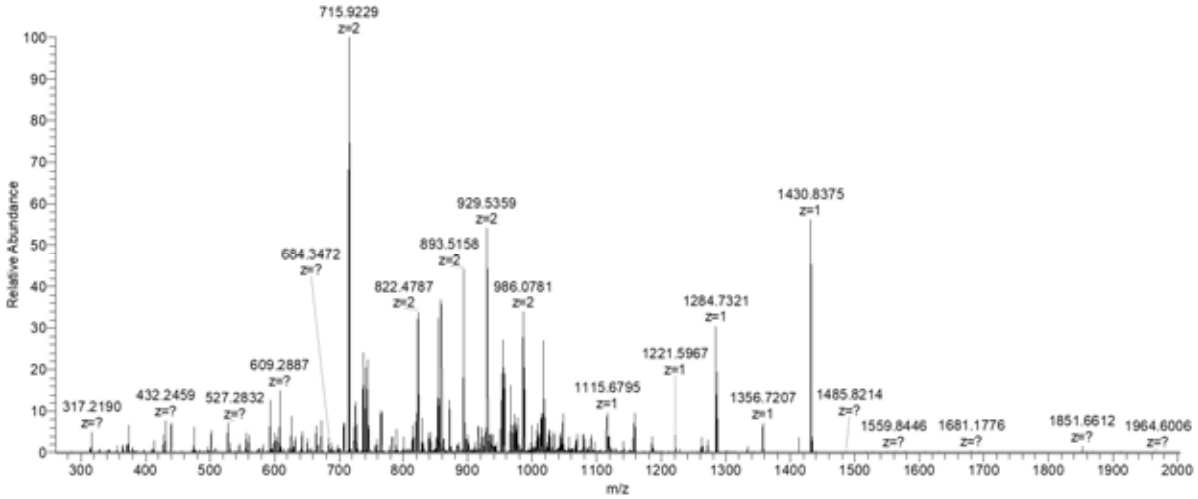
UniProt P68191, SRA
Seq: MKSNRQARHILGLDHKISNQRKIVTEGDKSSVVNNPTGRKRPAEK
m/z monoisotopic = 1019.5584, +5; MW_{obs} = 5092.76 Da (-1.9 ppm)



m/z meas	**z**	**m/z calc**	**Δppm**	**assignment**
801.9347	2	801.9359	1.50	b14-H2O +2
810.9405	2	810.9412	0.86	b14 +2
828.7866	3	828.7874	0.97	b21 +3
870.4782	2	870.4792	1.15	y16 +2
934.5258	2	934.5267	0.96	y17 +2
990.3381	5	990.3392	1.11	b44 +5
993.9413	5	993.9431	1.81	b44+H2O +5
1152.6353	3	1152.6363	0.87	y31-H2O +3
1158.6381	3	1158.6398	1.47	y31 +3
1305.7180	2	1305.7197	1.30	y24 +2

3.2.1.9 HU-α

UniProt P0ACF0, HU-α
Seq: MNKTQLIDVIAEKAELSKTQAKAALESTLAAITESLKEGDAVQLVGFGTFKVNHRAERTG
RNPQTGKEIKIAAANVPAFVSGKALKDAVK
m/z monoisotopic = 953.9263, +10; MW_{obs} = 9529.93 Da (0.2 ppm)

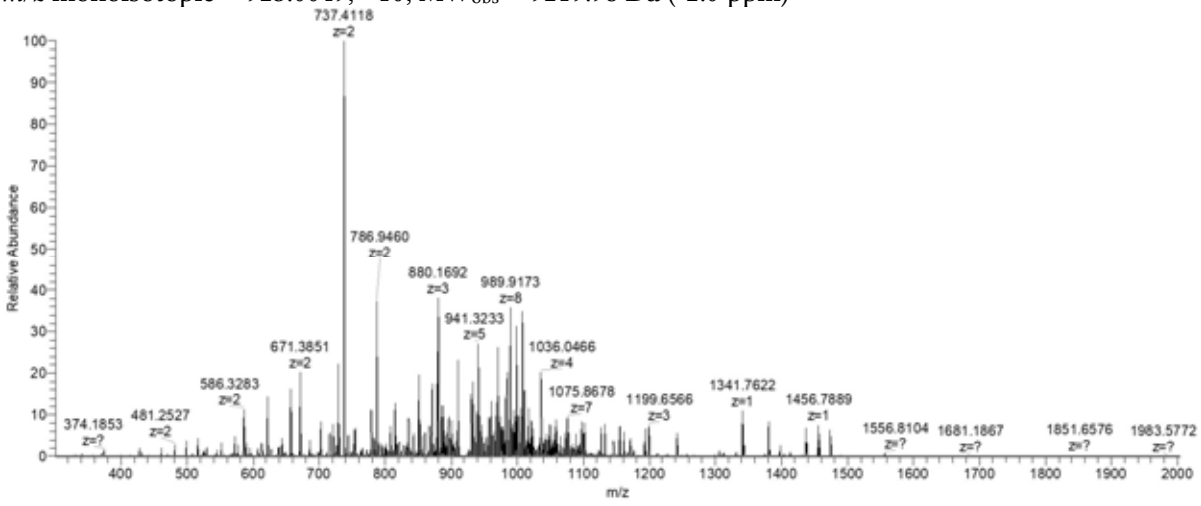


m/z meas	z	m/z calc	Δppm	assignment
317.2190	1	317.2183	-2.21	y3 +1
374.1864	1	374.1857	-1.87	b3 +1
414.2354	1	414.2347	-1.69	y4-H2O +1
432.2459	1	432.2453	-1.39	y4 +1
475.2340	1	475.2333	-1.47	b4 +1
508.8094	2	508.8086	-1.57	y10 +2
560.3409	1	560.3402	-1.25	y5 +1
603.2926	1	603.2919	-1.16	b5 +1
631.8777	2	631.8770	-1.11	y12 +2
673.4252	1	673.4243	-1.34	y6 +1
715.9229	2	715.9219	-1.40	y14 +2
716.3771	1	716.3760	-1.54	b6 +1
744.4625	1	744.4614	-1.48	y7 +1
765.4575	2	765.4561	-1.83	y15 +2
814.4539	3	814.4526	-1.60	b22 +3
822.4787	2	822.4776	-1.34	y16 +2
829.4613	1	829.4600	-1.57	b7 +1
838.1326	3	838.1316	-1.19	b23 +3
857.9973	2	857.9962	-1.28	y17 +2
861.8117	3	861.8107	-1.16	b24 +3
872.5580	1	872.5564	-1.83	y8 +1
893.5158	2	893.5147	-1.23	y18 +2
915.9140	5	915.9135	-0.55	y43 +5
929.0346	2	929.0333	-1.40	y19 +2
929.5790	1	929.5778	-1.29	y9 +1
944.4886	1	944.4870	-1.69	b8 +1
976.5457	5	976.5451	-0.61	y46 +5
985.5765	2	985.5753	-1.22	y20 +2

999.1624	5	999.1619	-0.50	y47 +5
1007.0573	2	1007.0561	-1.19	b18 +2
1016.6110	1	1016.6099	-1.08	y10 +1
1018.3433	9	1018.3418	-1.47	y87 +9
1025.7296	6	1025.7289	-0.68	y58 +6
1043.5560	1	1043.5554	-0.57	b9 +1
1066.5887	3	1066.5865	-2.06	b30 +3
1090.2670	3	1090.2656	-1.28	b31 +3
1115.6795	1	1115.6783	-1.08	y11 +1
1156.6406	1	1156.6395	-0.95	b10 +1
1262.7488	1	1262.7467	-1.66	y12 +1
1333.7850	1	1333.7838	-0.90	y13 +1
1412.8275	1	1412.8260	-1.06	y14-H2O +1
1430.8375	1	1430.8366	-0.63	y14 +1

3.2.1.10 HU-β

UniProt P0ACF4, HU-β
Seq:
MNKSQLIDKIAAGADISKAAAGRALDAIIASVTESLKEGDDVALVGFGTFAVKERAARTGRNPQTGK
EITIAAAKVPSFRAGKALKDAVN
m/z monoisotopic = 923.0049, +10; MW_{obs} = 9219.98 Da (-2.0 ppm)



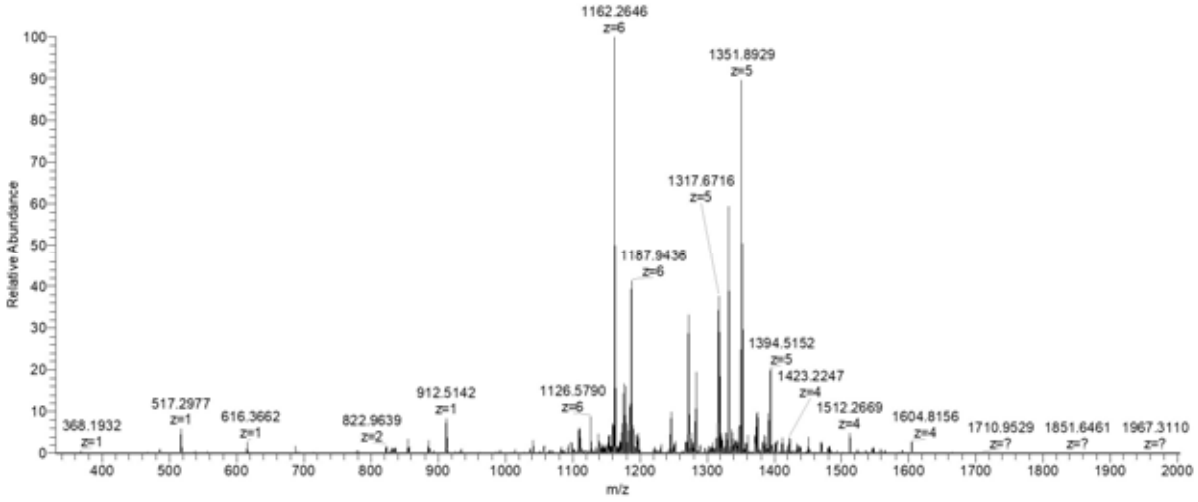
m/z meas	z	m/z calc	Δppm	assignment
524.9662	3	524.9667	-0.95	y15 +3
529.7864	2	529.7868	-0.76	b9 +2
572.3309	2	572.3314	-0.87	a10 +2
586.3283	2	586.3288	-0.85	b10 +2
621.8470	2	621.8474	-0.64	b11 +2
657.3656	2	657.3659	-0.46	b12 +2

659.3717	1	659.3723	-0.91	y6 +1
685.8762	2	685.8767	-0.73	b13 +2
702.3599	1	702.3603	-0.57	b6 +1
721.3952	2	721.3952	0.00	b14 +2
728.4068	2	728.4070	-0.27	y14-H2O +2
737.4118	2	737.4123	-0.68	y14 +2
753.6282	5	753.6289	-0.93	y36 +5
771.0889	3	771.0899	-1.30	b23 +3
778.4322	2	778.4332	-1.28	y15-NH3 +2
778.9083	2	778.9087	-0.51	b15 +2
786.9460	2	786.9465	-0.64	y15 +2
800.7808	6	800.7818	-1.25	y46 +6
815.4437	1	815.4444	-0.86	b7 +1
819.6283	6	819.6291	-0.98	y47 +6
835.4501	2	835.4507	-0.72	b16 +2
845.1417	6	845.1423	-0.71	y49-NH3 +6
850.9933	2	850.9939	-0.71	y16 +2
857.1503	6	857.1512	-1.05	y50 +6
870.8053	3	870.8059	-0.69	b26 +3
877.9983	2	877.9992	-1.03	y17-NH3 +2
886.5121	2	886.5125	-0.45	y17 +2
894.4843	3	894.4850	-0.78	b27 +3
900.1039	5	900.1050	-1.22	y43 +5
909.1246	8	909.1252	-0.66	y71 +8
930.4708	1	930.4713	-0.54	b8 +1
932.1789	3	932.1796	-0.75	b28 +3
937.5182	5	937.5177	0.53	y45-NH3 +5
938.6812	6	938.6822	-1.07	y54 +6
940.9219	5	940.9230	-1.17	y45 +5
944.8952	8	944.8957	-0.53	y74 +8
949.0362	2	949.0363	-0.11	y19-NH3 +2
951.3961	8	951.3941	2.10	a75-NH3 +8
960.7357	5	960.7367	-1.04	y46 +5
967.2781	8	967.2770	1.14	b76-NH3 +8
969.4042	8	969.4054	-1.24	b76 +8
980.0364	8	980.0379	-1.53	y77-H2O +8
983.9888	9	983.9869	1.93	y87 +9
987.1643	8	987.1656	-1.32	y78-H2O +8
993.5529	3	993.5534	-0.50	b30 +3
996.1694	8	996.1683	1.10	y79-NH3 +8
998.2947	8	998.2966	-1.90	y79 +8
1004.9287	8	1004.9249	3.78	y80-H2O +8
1007.1747	8	1007.1762	-1.49	y80 +8

1010.3824	6	1010.3833	-0.89	y58 +6
1014.0502	2	1014.0513	-1.08	b20 +2
1021.3124	8	1021.3117	0.69	y81 +8
1022.5630	3	1022.5640	-0.98	b31 +3
1035.5453	4	1035.5468	-1.45	b41 +4
1040.3817	5	1040.3800	1.63	y50 +5
1049.5817	3	1049.5833	-1.52	b32-H2O +3
1053.2389	6	1053.2395	-0.57	y61 +6
1058.5659	1	1058.5663	-0.38	b9 +1
1065.1430	7	1065.1479	-4.60	b73 +7
1069.0816	6	1069.0851	-3.27	y62-H2O +6
1075.3012	7	1075.2960	4.84	b74 +7
1078.0722	4	1078.0732	-0.93	b43 +4
1089.2684	3	1089.2694	-0.92	b33 +3
1093.5920	7	1093.5953	-3.02	b75 +7
1100.6007	5	1100.5982	2.27	y53 +5
1102.7699	6	1102.7737	-3.45	y64 +6
1122.8102	5	1122.8119	-1.51	y54-NH3 +5
1126.2150	5	1126.2172	-1.95	y54 +5
1132.2826	3	1132.2836	-0.88	b34 +3
1155.2897	3	1155.2908	-0.95	b35-H2O +3
1161.2933	3	1161.2943	-0.86	b35 +3
1198.9884	3	1198.9890	-0.50	b36 +3
1241.6838	3	1241.6873	-2.82	b37 +3
1305.7044	2	1305.7052	-0.61	b26 +2
1380.3921	3	1380.3933	-0.87	b41 +3
1397.7649	2	1397.7658	-0.64	b28 +2
1437.0944	3	1437.0951	-0.49	b43 +3
1456.7907	1	1456.7889	1.24	y14-NH3 +1
1473.8158	1	1473.8172	-0.95	y14 +1

3.2.1.11 CspC

UniProt P0A9Y6, CspC
Seq:
AKIKGQVKWFNESKGFGFITPADGSKDVFVHFSAIQGNGFKTLAEGQNVEFEIQDGQKGPAAVNVT
AI
m/z monoisotopic = 1212.1269, +6; MW_{obs} = 7266.72 Da (-0.4 ppm)

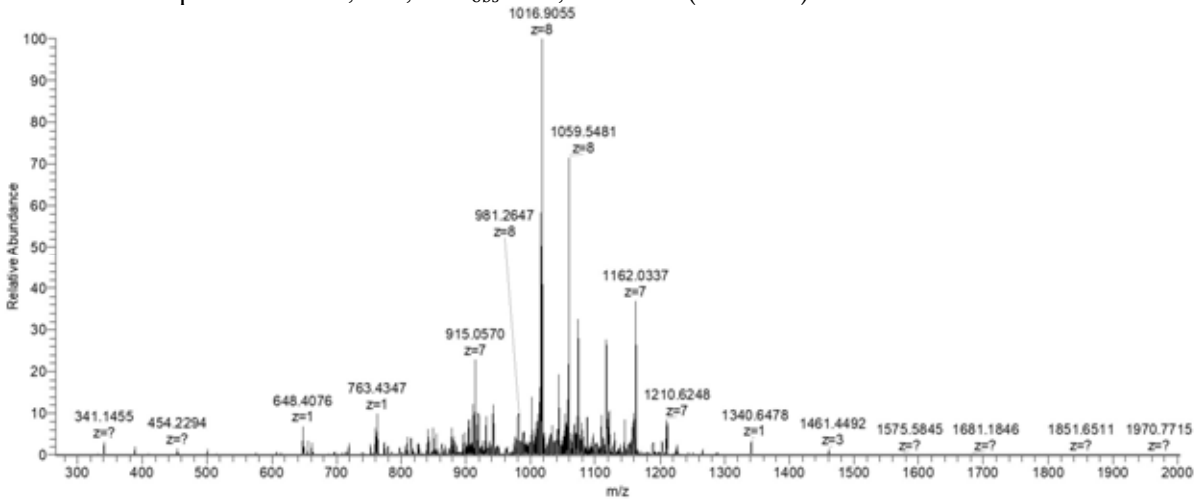


m/z meas	z	m/z calc	Δppm	assignment
517.2977	1	517.2980	-0.58	y5 +1
616.3662	1	616.3665	-0.49	y6 +1
687.4034	1	687.4036	-0.29	y7 +1
734.8910	2	734.8914	-0.54	y15 +2
791.4328	2	791.4334	-0.76	y16 +2
822.9639	2	822.9647	-0.97	b14 +2
855.4923	1	855.4934	-1.29	y9 +1
912.5142	1	912.5149	-0.77	y10 +1
925.0087	2	925.0096	-0.97	b16 +2
929.4885	2	929.4889	-0.43	y18 +2
960.2580	4	960.2583	-0.31	b35 +4
992.2714	4	992.2730	-1.61	b36 +4
994.0092	2	994.0102	-1.01	y19 +2
1027.0537	2	1027.0545	-0.78	b18 +2
1040.6089	1	1040.6099	-0.96	y11 +1
1057.1488	5	1057.1503	-1.42	b49 +5
1069.7118	6	1069.7130	-1.12	b59 +6
1082.9572	5	1082.9588	-1.48	b50 +5
1083.5956	2	1083.5966	-0.92	b19 +2
1097.7268	6	1097.7280	-1.09	b61 +6
1109.5662	6	1109.5675	-1.17	b62 +6
1112.3716	5	1112.3725	-0.81	b51 +5
1126.0772	6	1126.0789	-1.51	b63 +6
1138.1791	5	1138.1810	-1.67	b52 +5
1154.0929	6	1154.0939	-0.87	a65-NH3 +6
1161.5966	6	1161.5975	-0.77	b65 +6
1187.2749	6	1187.2765	-1.35	b67-H2O +6
1243.0251	5	1243.0256	-0.40	b57-NH3 +5

1246.4286	5	1246.4309	-1.85	b57 +5
1272.0480	5	1272.0499	-1.49	b58 +5
1275.6769	2	1275.6788	-1.49	b23 +2
1283.4515	5	1283.4542	-2.10	b59 +5
1302.8623	5	1302.8647	-1.84	b60 +5
1313.6680	5	1313.6669	0.84	b61-NH3 +5
1317.0702	5	1317.0722	-1.52	b61 +5
1321.1846	4	1321.1860	-1.06	b49 +4
1327.8748	5	1327.8743	0.38	b62-NH3 +5
1331.2774	5	1331.2796	-1.65	b62 +5
1342.0873	5	1342.0890	-1.27	a63-NH3 +5
1347.4943	5	1347.4912	2.30	b63-H2O +5
1351.0907	5	1351.0933	-1.92	b63 +5
1353.4442	4	1353.4466	-1.77	b50 +4
1393.7134	5	1393.7155	-1.51	b65 +5
1450.7429	4	1450.7454	-1.72	b53 +4
1468.7732	1	1468.7754	-1.50	y15 +1
1481.1194	3	1481.1130	4.32	a41 +3
1482.7551	4	1482.7601	-3.37	b54 +4
1546.8124	3	1546.8147	-1.49	a43-NH3 +3
1558.0338	4	1558.0290	3.08	y59 +4
1604.0591	4	1604.0659	-4.24	b59 +4

3.2.1.12 H-NS

UniProt P0ACF8, H-NS
Seq:
SEALKILNNIRTLRAQARECTLETLEEMLEKLEVVVNERREEESAAAAEVEERTRKLQQYREMLIADGI
DPNELLNSLAAVKSGTKAKRAQRPAKYSYVDENGETKTWTGQGRTPAVIKKAMDEQGKSLDDFLIK
Q
m/z monoisotopic = 963.5065, +16; MW_{obs} = 15,399.99 Da (+1.02 Da)

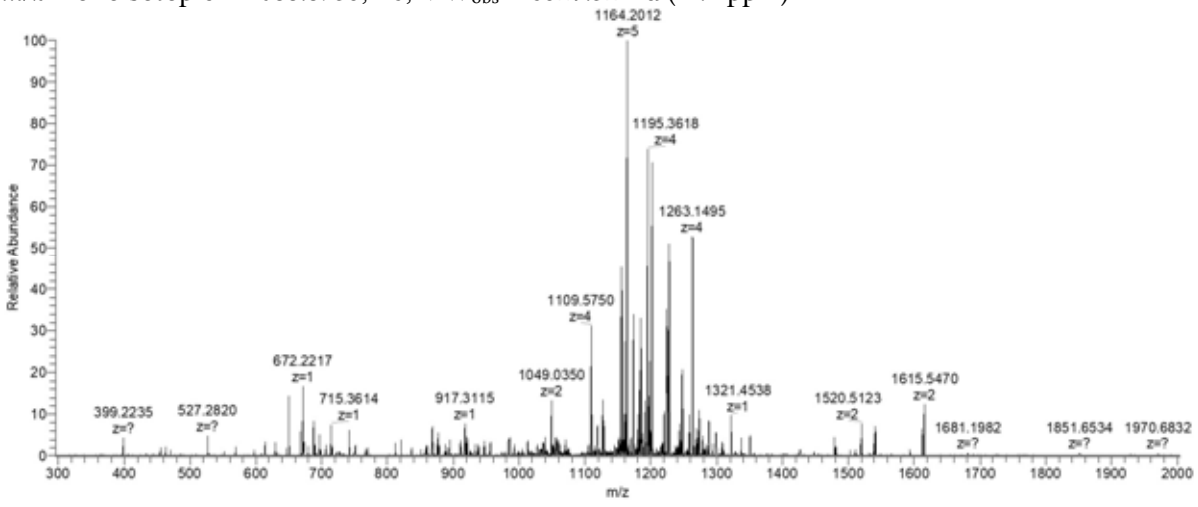


m/z meas	z	m/z calc	Δppm	assignment
388.2553	1	388.2554	0.26	y3 +1
501.3391	1	501.3395	0.80	y4 +1
648.4076	1	648.4079	0.46	y5 +1
656.0028	3	656.0032	0.61	y17 +3
687.8664	2	687.8668	0.58	y12-NH3 +2
696.3796	2	696.3801	0.72	y12 +2
718.9853	5	718.9858	0.70	y32 +5
760.9009	2	760.9014	0.66	y13 +2
763.4347	1	763.4349	0.26	y6 +1
774.9162	4	774.9172	1.29	b27 +4
779.0065	5	779.0072	0.90	y35 +5
796.4221	4	796.4220	-0.13	a28-NH3 +4
800.6755	4	800.6786	3.87	a28 +4
804.8145	5	804.8157	1.49	y36 +5
809.4305	9	809.4306	0.12	y66 +9
815.9325	4	815.9329	0.49	y29 +4
818.4137	2	818.4149	1.47	y14 +2
839.6948	8	839.6953	0.60	y61 +8
841.1942	4	841.1948	0.71	y30 +4
853.8269	8	853.8308	4.57	y62 +8
861.8790	7	861.8797	0.81	y54 +7
867.9621	8	867.9663	4.84	y63 +8
873.2169	4	873.2185	1.83	y31 +4
875.4803	3	875.4815	1.37	b23 +3
878.4612	1	878.4618	0.68	y7 +1
880.1753	7	880.1790	4.20	y55 +7
883.9342	2	883.9351	1.02	y15 +2
894.3308	7	894.3316	0.89	y56 +7
903.5818	9	903.5833	1.66	b70 +9
910.4826	8	910.4836	1.10	y66 +8
914.6273	7	914.6280	0.77	y58 +7
919.4530	2	919.4537	0.76	y16 +2
930.7830	7	930.7828	-0.21	y59 +7
941.3756	9	941.3764	0.85	b73 +9
943.2142	7	943.2160	1.91	y60 +7
975.6608	7	975.6627	1.95	y62 +7
980.7627	8	980.7637	1.02	b67 +8
983.5002	2	983.5011	0.92	y17 +2
986.8424	9	986.8467	4.36	b77-H2O +9
987.8916	8	987.8914	-0.20	b68 +8
988.8475	9	988.8479	0.40	b77 +9
1014.0310	8	1014.0342	3.16	y74 +8

1016.4043	8	1016.4053	0.98	b70 +8
1030.2037	6	1030.2013	-2.33	b54-H2O +6
1032.8862	3	1032.8872	0.97	b27 +3
1042.7902	8	1042.7923	2.01	b72 +8
1056.7959	8	1056.7943	-1.51	b73-NH3 +8
1058.9210	8	1058.9226	1.51	b73 +8
1066.8966	6	1066.8981	1.41	y58 +6
1073.0570	8	1073.0581	1.03	b74 +8
1078.5774	1	1078.5779	0.46	y8 +1
1085.0599	8	1085.0653	4.98	b75-NH3 +8
1095.8386	4	1095.8371	-1.37	b38 +4
1110.0785	8	1110.0767	-1.62	b77-H2O +8
1112.3268	8	1112.3280	1.08	b77 +8
1116.7807	5	1116.7817	0.90	b49 +5
1120.7280	7	1120.7290	0.89	b67 +7
1126.4420	7	1126.4425	0.44	b68-NH3 +7
1128.8732	7	1128.8749	1.51	b68 +7
1145.0291	7	1145.0297	0.52	b69 +7
1159.0321	7	1159.0298	-1.98	b70-NH3 +7
1161.4612	7	1161.4622	0.86	b70 +7
1162.4025	5	1162.4039	1.20	b51 +5
1188.2148	5	1188.2125	-1.94	b52 +5
1210.0488	7	1210.0533	3.72	b73 +7

3.2.1.13 BhsA

UniProt P0AB40, BhsA
Seq: AVEVQSTPEGQQKVGTISANAGTNLGSLEEQLAQKADEMGAKSFRITSVTGPNTLHGTAVIYK
m/z monoisotopic = 1088.8935, +6; MW_{obs} = 6527.32 Da (-1.1 ppm)

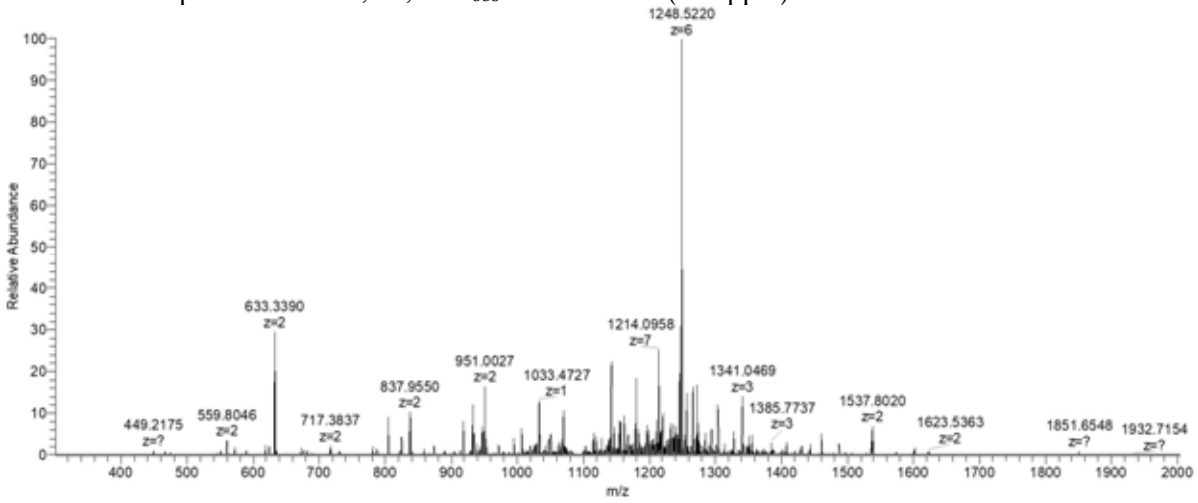


m/z meas	z	m/z calc	Δppm	assignment
399.2235	1	399.2238	-0.75	b4 +1
527.2820	1	527.2824	-0.76	b5 +1
569.2926	1	569.2930	-0.70	a6-NH3 +1
614.3139	1	614.3144	-0.81	b6 +1
697.3513	1	697.3515	-0.29	b7-H2O +1
715.3614	1	715.3621	-0.98	b7 +1
727.3852	2	727.3859	-0.96	a14 +2
741.3829	2	741.3834	-0.67	b14 +2
751.4342	1	751.4349	-0.93	y7 +1
769.8938	2	769.8941	-0.39	b15 +2
811.4118	2	811.4127	-1.11	b16-H2O +2
820.4173	2	820.4179	-0.73	b16 +2
867.9538	2	867.9547	-1.04	b17-H2O +2
876.9590	2	876.9600	-1.14	b17 +2
888.4933	1	888.4938	-0.56	y8 +1
911.4696	2	911.4707	-1.21	b18-H2O +2
920.4753	2	920.4760	-0.76	b18 +2
926.8239	3	926.8251	-1.29	y26 +3
946.9887	2	946.9893	-0.63	b19-H2O +2
955.9930	2	955.9945	-1.57	b19 +2
1004.0101	2	1004.0107	-0.60	b20-H2O +2
1012.7749	4	1012.7757	-0.79	y38 +4
1036.1978	6	1036.1989	-1.06	y60-NH3 +6
1039.0358	6	1039.0366	-0.77	y60 +6
1041.0463	4	1041.0467	-0.38	y39 +4
1048.5334	2	1048.5346	-1.14	b21 +2
1055.7532	5	1055.7544	-1.14	y51 +5
1057.5424	6	1057.5420	0.38	y61-H2O +6
1104.8201	4	1104.8181	1.81	y42-NH3 +4
1109.0736	4	1109.0747	-0.99	y42 +4
1114.9790	5	1114.9768	1.97	y54-NH3 +5
1118.3802	5	1118.3821	-1.70	y54 +5
1126.8326	4	1126.8340	-1.24	y43 +4
1151.0909	4	1151.0881	2.43	y44-NH3 +4
1155.3430	4	1155.3447	-1.47	y44 +4
1159.9966	5	1159.9990	-2.07	y56-H2O +5
1163.5992	5	1163.6012	-1.72	y56 +5
1168.8445	4	1168.8473	-2.40	y45-NH3 +4
1173.1020	4	1173.1040	-1.70	y45 +4
1180.4068	5	1180.4054	1.19	y57-NH3 +5
1183.8088	5	1183.8107	-1.60	y57 +5
1194.8602	4	1194.8620	-1.51	y46 +4

1197.6112	5	1197.6150	-3.17	y58-H2O +5
1201.2159	5	1201.2171	-1.00	y58 +5
1223.1315	4	1223.1330	-1.23	y47 +4
1223.2242	5	1223.2267	-2.04	y59-H2O +5
1226.8262	5	1226.8288	-2.12	y59 +5
1243.2386	5	1243.2372	1.13	y60-NH3 +5
1246.6411	5	1246.6425	-1.12	y60 +5
1248.3930	4	1248.3949	-1.52	y48 +4
1258.3949	4	1258.3937	0.95	y49-NH3 +4
1262.6476	4	1262.6503	-2.14	y49 +4
1269.0476	5	1269.0457	1.50	y61-NH3 +5
1282.9126	4	1282.9148	-1.71	y50-H2O +4
1287.4163	4	1287.4174	-0.85	y50 +4
1350.0317	3	1350.0318	-0.07	y38 +3
1478.4249	3	1478.4305	-3.79	y42 +3
1481.7593	1	1481.7594	-0.07	b14 +1
1592.8103	3	1592.8136	-2.07	y46 +3

3.2.1.14 YnaE

UniProt P76073, YnaE
Seq:
MKSKDTLKWFPAQLPEVRIILGDAVVEVAKQGRPINTRTLLDYIEGNIKKTSWLDNKELLQTAISVLK
DNQNLNGKM
m/z monoisotopic = 1094.2255, +8; MW_{obs} = 8745.75 Da (-0.6 ppm)



m/z meas	z	m/z calc	Δppm	assignment
449.2173	1	449.2177	-0.89	y4 +1
475.2693	1	475.2697	-0.84	b4 +1
562.3013	1	562.3017	-0.71	y5 +1

572.2856	1	572.2861	-0.87	b5-H2O +1
590.2961	1	590.2967	-1.02	b5 +1
659.3171	1	659.3181	-1.52	y6-NH3 +1
676.3443	1	676.3447	-0.59	y6 +1
691.3438	1	691.3443	-0.72	b6 +1
730.8792	2	730.8799	-0.96	y13 +2
786.4174	1	786.4178	-0.51	b7-H2O +1
804.4280	1	804.4284	-0.50	b7 +1
837.9548	2	837.9554	-0.72	b14 +2
901.4192	1	901.4196	-0.44	y8-NH3 +1
914.5115	1	914.5128	-1.42	b8-H2O +1
918.4458	1	918.4462	-0.44	y8 +1
932.5229	1	932.5234	-0.54	b8 +1
951.0023	2	951.0031	-0.84	b16 +2
994.0345	2	994.0357	-1.21	y18 +2
1016.4459	1	1016.4466	-0.69	y9-NH3 +1
1022.5669	3	1022.5667	0.20	b27 +3
1033.4726	1	1033.4731	-0.48	y9 +1
1050.5770	2	1050.5777	-0.67	y19 +2
1061.7879	5	1061.7844	3.30	b47 +5
1062.2283	3	1062.2272	1.04	y28-NH3 +3
1067.9016	3	1067.9027	-1.03	y28 +3
1072.3542	4	1072.3544	-0.19	b38 +4
1078.5866	2	1078.5879	-1.21	b18 +2
1100.5914	1	1100.5921	-0.64	b9-H2O +1
1110.0191	5	1110.0202	-0.99	b49 +5
1115.0985	2	1115.0990	-0.45	y20 +2
1127.6180	6	1127.6230	-4.43	b59 +6
1132.2373	5	1132.2339	3.00	b50-NH3 +5
1135.6381	5	1135.6392	-0.97	b50 +5
1138.9633	6	1138.9654	-1.84	y61-H2O +6
1141.9651	6	1141.9672	-1.84	y61 +6
1144.5404	1	1144.5415	-0.96	y10-NH3 +1
1146.4691	6	1146.4703	-1.05	b60 +6
1152.2437	5	1152.2466	-2.52	b51-H2O +5
1155.8479	5	1155.8487	-0.69	b51 +5
1161.5674	1	1161.5681	-0.60	y10 +1
1164.9804	6	1164.9757	4.03	b61-NH3 +6
1167.8086	6	1167.8134	-4.11	b61 +6
1169.6495	5	1169.6530	-2.99	b52-H2O +5
1173.2539	5	1173.2551	-1.02	b52 +5
1177.1514	6	1177.1511	0.25	a62-NH3 +5
1178.4105	4	1178.4125	-1.70	b42-H2O +4

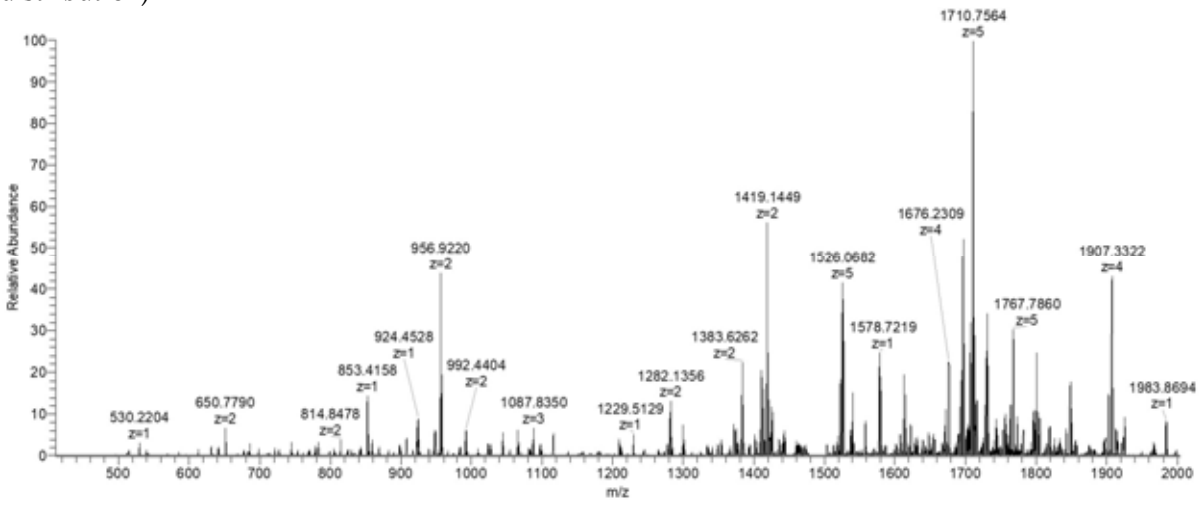
1178.6604	4	1178.6585	1.61	b42-NH3 +4
1181.6535	6	1181.6529	0.51	b62-H2O +6
1182.9140	4	1182.9151	-0.93	b42 +4
1188.9916	6	1188.9906	0.84	a63-NH3 +6
1193.4905	6	1193.4925	-1.68	b63-H2O +6
1196.4929	6	1196.4942	-1.09	b63 +6
1207.0666	5	1207.0657	0.75	b53-NH3 +5
1210.4693	5	1210.4710	-1.40	b53 +5
1212.5054	6	1212.5038	1.32	b64-NH3 +6
1213.3774	7	1213.3809	-2.88	y75 +7
1215.3399	6	1215.3416	-1.40	b64 +6
1227.6522	2	1227.6547	-2.04	y22-NH3 +2
1229.8428	6	1229.8469	-3.33	b65 +6
1233.0858	5	1233.0878	-1.62	b54 +5
1236.1663	2	1236.1680	-1.38	y22 +2
1256.0922	5	1256.0932	-0.80	b55 +5
1257.6255	1	1257.6256	-0.08	y11-NH3 +1
1262.3675	6	1262.3679	-0.32	b67-NH3 +6
1265.2005	6	1265.2056	-4.03	b67 +6
1265.6699	1	1265.6711	-0.95	b10 +1
1269.5328	6	1269.5288	3.15	y68-NH3 +5
1272.3724	6	1272.3665	4.64	y68 +6
1274.6512	1	1274.6521	-0.71	y11 +1
1278.9046	5	1278.9018	2.19	b56 +5
1279.9592	4	1279.9560	2.50	b45-NH3 +4
1284.2117	4	1284.2126	-0.70	b45 +4
1285.1685	2	1285.1682	0.23	y23-NH3 +2
1293.6798	2	1293.6814	-1.24	y23 +2
1302.8879	6	1302.8882	-0.23	b69-NH3 +6
1305.7219	6	1305.7260	-3.14	b69 +6
1324.7294	6	1324.7331	-2.79	b70 +6
1326.9771	4	1326.9787	-1.21	b47 +4
1330.3292	5	1330.3293	-0.08	b58 +5
1334.2366	2	1334.2382	-1.20	b23 +2
1334.7047	3	1334.7050	-0.22	y35-NH3 +3
1340.3785	3	1340.3805	-1.49	y35 +3
1343.5731	6	1343.5763	-2.38	y71 +6
1346.0784	6	1346.0762	1.63	b71 +6
1348.2476	4	1348.2510	-2.52	a48 +4
1350.2212	2	1350.2235	-1.70	y24 +2
1352.9451	5	1352.9461	-0.74	b59 +5
1355.2475	4	1355.2497	-1.62	b48 +4
1357.5786	6	1357.5798	-0.88	y72-NH3 +6

1360.4142	6	1360.4176	-2.50	y72 +6
1362.0826	6	1362.0816	0.73	b72-H2O +6
1365.0843	6	1365.0834	0.66	b72 +6
1369.7559	2	1369.7567	-0.58	b24 +2
1371.7559	3	1371.7609	-3.64	b37-NH3 +3
1372.1618	5	1372.1576	3.06	b60-NH3 +5
1373.7195	1	1373.7206	-0.80	y12 +1
1375.5611	5	1375.5629	-1.31	b60 +5
1379.5910	6	1379.5887	1.67	y73 +6
1381.0908	6	1381.0930	-1.59	b73-NH3 +6
1383.9290	6	1383.9307	-1.23	b73 +6
1387.2707	4	1387.2734	-1.95	b49 +4
1400.1040	6	1400.1001	2.79	b74-NH3 +6
1401.1734	5	1401.1747	-0.93	b61 +5
1402.9342	6	1402.9379	-2.64	b74 +6
1409.6044	6	1409.6037	0.50	b75-NH3 +6
1412.4388	6	1412.4414	-1.84	b75 +6
1417.7779	5	1417.7821	-2.96	b62-H2O +5
1419.2918	2	1419.2909	0.63	b25 +2
1419.2918	4	1419.2972	-3.80	b50 +4
1421.3832	5	1421.3842	-0.70	b62 +5
1434.7510	2	1434.7498	0.84	y25-NH3 +2
1435.5899	5	1435.5916	-1.18	b63 +5
1443.2611	2	1443.2631	-1.39	y25 +2
1443.7247	1	1443.7260	-0.90	y13-NH3 +1
1454.8049	5	1454.8031	1.24	b64-NH3 +5
1458.2053	5	1458.2084	-2.13	b64 +5
1460.7512	1	1460.7526	-0.96	y13 +1
1472.2119	5	1472.2095	1.63	b65-NH3 +5
1475.6145	5	1475.6148	-0.20	b65 +5
1478.2657	2	1478.2659	-0.14	y26-NH3 +2
1486.7778	2	1486.7791	-0.87	y26 +2
1491.8219	5	1491.8264	-3.02	b66-H2O +5
1493.6247	5	1493.6226	1.41	y67-H2O +5
1495.4260	5	1495.4285	-1.67	b66 +5
1497.2259	5	1497.2247	0.80	y67 +5
1518.0445	5	1518.0453	-0.53	b67 +5
1526.6361	5	1526.6384	-1.51	y68 +5
1528.7924	2	1528.7897	1.77	y27-NH3 +2
1533.3453	2	1533.3465	-0.78	b27 +2
1537.3004	2	1537.3030	-1.69	y27 +2
1556.8102	1	1556.8101	0.06	y14-NH3 +1
1560.4552	5	1560.4490	3.97	y69-NH3 +5

1569.8664	4	1569.8647	1.08	b55 +4
1573.8352	1	1573.8367	-0.95	y14 +1
1576.8838	3	1576.8844	-0.38	b42 +3
1592.8384	2	1592.8372	0.75	y28-NH3 +2
1601.3485	2	1601.3505	-1.25	y28 +2
1644.8734	1	1644.8738	-0.24	y15 +1
1674.9026	1	1674.9036	-0.60	b14 +1
1745.9207	1	1745.9214	-0.40	y16 +1
1900.9977	1	1900.9990	-0.68	b16 +1

3.2.1.15 YbgS

UniProt P0AAV6, YbgS
Seq:
ADSGAQTNNGQANAAADAGQVAPDARENVAPNNVDNNGVNTGSGGTMLHSDGSSMNNDGMT
KDEEHKNTMCKDGRCPDINKKVQTGDGINNDVDTKTDGTTQ
m/z monoisotopic = 1494.9477, +7; MW_{obs} = 10,457.58 Da (+1.02 Da)
(* in the last column denotes peaks which do not correspond to the monoisotopic peaks within the distribution)



m/z meas	z	m/z calc	Δppm	assignment	not mono?
512.2100	1	512.2100	0.00	b6-H2O +1	
530.2204	1	530.2205	0.19	b6 +1	
613.2577	1	613.2575	-0.33	b7-H2O +1	
622.2675	1	622.2679	0.64	y6 +1	
631.2682	1	631.2682	0.00	b7 +1	
727.3004	1	727.3006	0.27	b8-H2O +1	
745.3109	1	745.3111	0.27	b8 +1	
833.3997	1	833.3999	0.24	y8-H2O +1	
841.3434	1	841.3435	0.12	b9-H2O +1	
851.4102	1	851.4105	0.35	y8 +1	

859.3536	1	859.3541	0.58	b9 +1	
916.3751	1	916.3755	0.44	b10 +1	
966.4371	1	966.4374	0.31	y9 +1	
1026.4230	1	1026.4235	0.49	b11-H2O +1	
1044.4333	1	1044.4341	0.77	b11 +1	
1047.4935	1	1047.4953	1.72	y10-H2O +1	
1065.5051	1	1065.5059	0.75	y10 +1	
1097.4599	1	1097.4606	0.64	b12-H2O +1	
1115.4701	1	1115.4712	0.99	b12 +1	
1180.5319	1	1180.5328	0.76	y11 +1	
1210.3037	4	1210.3076	3.22	y44 +4	*
1211.5031	1	1211.5036	0.41	b13-H2O +1	
1212.0493	2	1212.0502	0.74	b26 +2	
1212.4861	1	1212.4876	1.24	b13-NH3 +1	
1229.5129	1	1229.5141	0.98	b13 +1	
1272.6295	2	1272.6304	0.71	y24-H2O +2	
1276.5701	2	1276.5715	1.10	b27 +2	
1281.6344	2	1281.6357	1.01	y24 +2	
1294.5739	1	1294.5757	1.39	y12 +1	
1297.9704	5	1297.9726	1.69	y61 +5	*
1300.5497	1	1300.5512	1.15	b14 +1	
1318.1797	5	1318.1821	1.82	y62 +5	*
1333.5907	2	1333.5929	1.65	b28 +2	
1340.9869	5	1340.9907	2.83	y63 +5	*
1353.5759	1	1353.5778	1.40	b15-H2O +1	
1371.5866	1	1371.5884	1.31	b15 +1	
1372.2047	5	1372.2087	2.92	y65 +5	*
1374.1200	2	1374.1218	1.31	b29-H2O +2	
1378.2797	6	1378.2833	2.61	y78 +6	*
1383.1252	2	1383.1271	1.37	b29 +2	
1394.6110	5	1394.6167	4.09	y79-NH3 +6	
1410.1351	2	1410.1324	-1.91	b30-NH3 +2	
1413.6256	6	1413.6299	3.04	y80 +6	*
1418.6443	2	1418.6457	0.99	b30 +2	
1425.4655	6	1425.4694	2.74	y81 +6	*
1435.9928	3	1435.9971	2.99	y39 +3	*
1440.8270	5	1440.8312	2.91	y68 +5	*
1441.9763	6	1441.9808	3.12	y82 +6	*
1442.6205	1	1442.6255	3.47	b16 +1	
1460.3173	6	1460.3222	3.36	y83-H2O +6	*
1469.9872	6	1469.9897	1.70	y84-NH3 +6	
1471.3913	4	1471.3967	3.67	y54 +4	*
1472.8221	6	1472.8275	3.67	y84 +6	

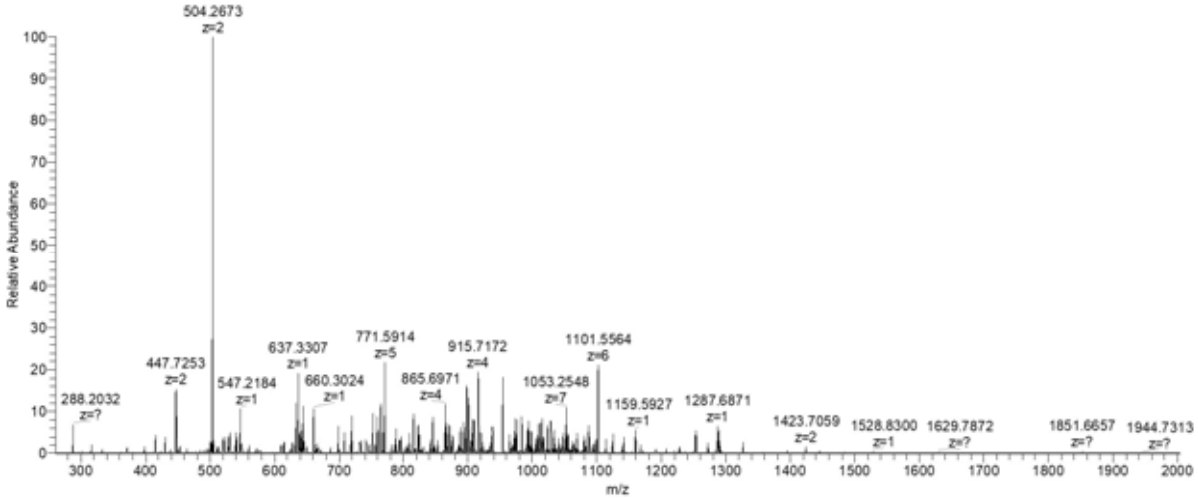
1525.6672	5	1525.6726	3.54	y72 +5	*
1536.2727	5	1536.2779	3.38	y73-H2O +5	*
1539.8760	5	1539.8801	2.66	y73 +5	*
1557.6498	1	1557.6524	1.67	b17 +1	
1578.7219	1	1578.7242	1.46	y15 +1	
1586.1954	4	1586.2005	3.22	y59 +4	*
1601.0503	6	1601.0501	-0.12	y93 +6	
1607.3967	3	1607.4041	4.60	y44-H2O +3	*
1613.4015	3	1613.4077	3.84	y44 +3	*
1617.7059	4	1617.7113	3.34	y61-H2O +4	*
1628.6860	1	1628.6895	2.15	b18 +1	
1630.7442	2	1630.7492	3.07	b34 +2	
1642.9686	4	1642.9732	2.80	y62-H2O +4	*
1647.4701	4	1647.4758	3.46	y62 +4	*
1650.1333	5	1650.1364	1.88	y78-H2O +5	
1653.0671	6	1653.0679	0.48	y96-NH3 +6	
1653.7321	5	1653.7385	3.87	y78 +5	
1655.9013	6	1655.9057	2.66	y96 +6	
1671.4771	4	1671.4839	4.07	y63-H2O +4	*
1675.9805	4	1675.9865	3.58	y63 +4	*
1685.7063	1	1685.7110	2.79	b19 +1	
1688.2605	2	1688.2627	1.30	b35 +2	
1692.5466	5	1692.5523	3.37	y80-H2O +5	*
1696.1487	5	1696.1544	3.36	y80 +5	*
1706.7529	5	1706.7597	3.98	y81-H2O +5	*
1710.3559	5	1710.3619	3.51	y81 +5	*
1715.0028	4	1715.0090	3.62	y65 +4	*
1726.5660	5	1726.5734	4.29	y82-H2O +5	
1730.1696	5	1730.1755	3.41	y82 +5	*
1739.0116	4	1739.0171	3.16	y66-H2O +4	*
1743.5140	4	1743.5197	3.27	y66 +4	
1752.3759	5	1752.3819	3.42	y83-NH3 +5	*
1755.7814	5	1755.7873	3.36	y83 +5	*
1763.5828	5	1763.5894	3.74	y84-H2O +5	*
1767.1853	5	1767.1915	3.51	y84 +5	*
1772.0246	4	1772.0305	3.33	y67 +4	*
1777.9891	5	1777.9937	2.59	y85-NH3 +5	*
1781.3918	5	1781.3990	4.04	y85 +5	*
1796.2773	4	1796.2846	4.06	y68-H2O +4	*
1800.7812	4	1800.7872	3.33	y68 +4	*
1804.3978	5	1804.4044	3.66	y86 +5	*
1813.7659	1	1813.7696	2.04	b20 +1	
1815.2032	5	1815.2065	1.82	y87-NH3 +5	

1818.6036	5	1818.6118	4.51	y87 +5	*
1821.0437	4	1821.0517	4.39	y69-H2O +4	
1825.5482	4	1825.5543	3.34	y69 +4	*
1829.4103	5	1829.4139	1.97	y88-NH3 +5	
1832.8132	5	1832.8192	3.27	y88 +5	*
1836.4658	3	1836.4658	0.00	b57 +3	
1842.4760	3	1842.4837	4.18	y51-H2O +3	*
1848.4792	3	1848.4873	4.38	y51 +3	*
1854.0571	4	1854.0651	4.31	y70 +4	*
1874.8026	3	1874.8081	2.93	b58 +3	
1882.5664	4	1882.5758	4.99	y71 +4	*
1894.8231	1	1894.8274	2.27	b21-H2O +1	
1902.3281	4	1902.3363	4.31	y72-H2O +4	*
1906.8312	4	1906.8390	4.09	y72 +4	*
1912.8340	1	1912.8380	2.09	b21 +1	
1920.0867	4	1920.0956	4.64	y73-H2O +4	*
1920.3377	4	1920.3416	2.03	y73-NH3 +4	*
1924.5907	4	1924.5983	3.95	y73 +4	*
1949.3570	4	1949.3654	4.31	y74 +4	*
1961.5187	3	1961.5266	4.03	y54 +3	*
1965.8617	1	1965.8645	1.42	b22-H2O +1	
1983.8694	1	1983.8751	2.87	b22 +1	

3.2.2 *P. aeruginosa*

3.2.2.1 PA4738

UniProt Q9HV61, PA4738
Seq:
MNSDVIKGKWKQLTGKIKERWGDLTDDDLQAADGHAEYLVGKLQERYGWSKERAEQEVRDFSDR
L
m/z monoisotopic = 951.8583, +8; MW_{obs} = 7606.81 Da (0.0 ppm)

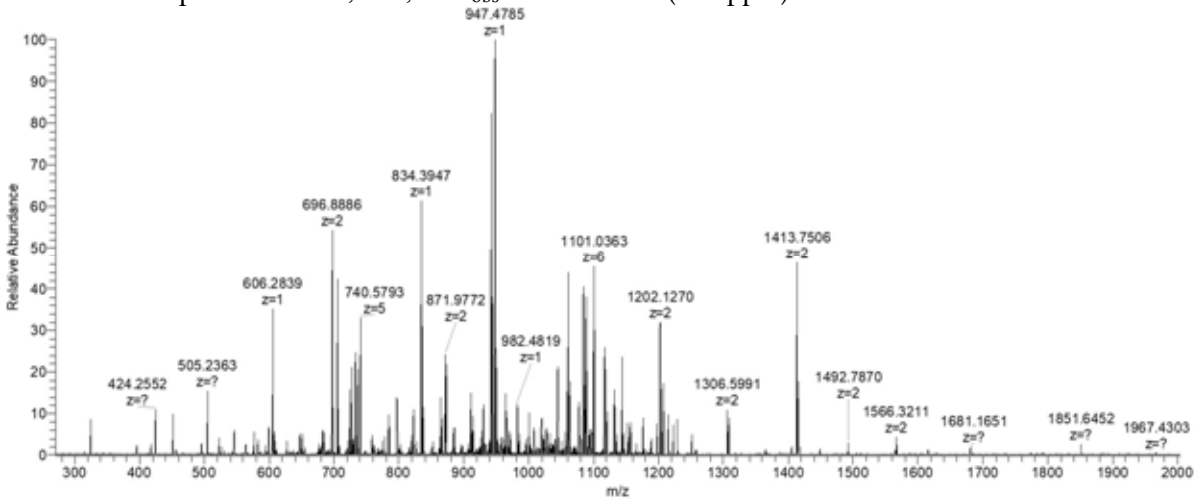


m/z meas	z	m/z calc	Δppm	assignment
288.2032	1	288.2030	0.69	y2 +1
333.1230	1	333.1227	0.90	b3 +1
430.1393	1	430.1391	0.46	b4-H2O +1
431.1234	1	431.1231	0.70	b4-NH3 +1
448.1499	1	448.1497	0.45	b4 +1
504.2673	2	504.2671	0.40	y8 +2
519.2234	1	519.2232	0.39	a5 +1
529.2077	1	529.2075	0.38	b5-H2O +1
530.1918	1	530.1915	0.57	b5-NH3 +1
540.9414	3	540.9411	0.55	y13 +3
547.2184	1	547.2181	0.55	b5 +1
632.8178	2	632.8177	0.16	y10 +2
637.3307	1	637.3304	0.47	y5 +1
642.2918	1	642.2916	0.31	b6-H2O +1
643.2760	1	643.2756	0.62	b6-NH3 +1
644.3477	2	644.3475	0.31	b11 +2
660.3024	1	660.3021	0.45	b6 +1
699.3716	2	699.3715	0.14	b12-H2O +2
708.3771	2	708.3768	0.42	b12 +2
732.8577	2	732.8575	0.27	y12 +2
734.3471	1	734.3468	0.41	y6-H2O +1
752.3575	1	752.3573	0.27	y6 +1
755.9137	2	755.9136	0.13	b13-H2O +2
759.7858	5	759.7858	0.00	y31 +5
764.9191	2	764.9189	0.26	b13 +2
771.1903	5	771.1901	0.26	y32 +5
781.1107	3	781.1106	0.13	b20 +3
788.3974	1	788.3971	0.38	b7 +1

794.1957	5	794.1955	0.25	y33 +5
796.1587	4	796.1589	-0.25	y26 +4
806.4376	2	806.4374	0.25	b14-H2O +2
808.4034	5	808.4029	0.62	y34 +5
810.9083	2	810.9081	0.25	y13 +2
815.4428	4	815.4427	0.12	b14 +2
822.6103	5	822.6101	0.24	y35 +5
824.4295	4	824.4299	-0.49	y27 +4
827.4080	1	827.4080	0.00	b8-H2O +1
843.9543	2	843.9534	1.07	b15 +2
848.2225	5	848.2220	0.59	y36 +5
865.1962	4	865.1957	0.58	y28 +4
870.8385	5	870.8388	-0.34	y37 +5
875.7053	4	875.7055	-0.23	b30 +4
897.4564	4	897.4564	0.00	y29 +4
900.4861	3	900.4865	-0.44	b23 +3
908.0011	2	908.0009	0.22	b16 +2
915.2152	4	915.2157	-0.55	y30 +4
921.2235	8	921.2229	0.65	y63 +8
935.4695	8	935.4715	-2.14	b64 +8
937.7227	8	937.7228	-0.11	b64+H2O +8
963.7359	4	963.7358	0.10	y32 +4
973.5139	1	973.5135	0.41	b9 +1
982.6816	5	982.6813	0.31	y42 +5
992.4920	4	992.4925	-0.50	y33 +4
993.5119	7	993.5092	2.72	y59 +7
997.3290	6	997.3301	-1.10	y51 +6
1009.6641	7	1009.6640	0.10	y60 +7
1014.1731	6	1014.1714	1.68	y52 +6
1023.8153	7	1023.8167	-1.37	y61 +7
1028.0109	4	1028.0111	-0.19	y35 +4
1033.0172	6	1033.0188	-1.55	y53 +6
1050.2516	7	1050.2499	1.62	y63-NH3 +7
1051.5243	6	1051.5241	0.19	y54-NH3 +6
1054.3612	6	1054.3619	-0.66	y54 +6
1086.8907	3	1086.8907	0.00	b28 +3
1093.1115	2	1093.1117	-0.18	b19 +2
1101.0557	6	1101.0554	0.27	b57 +6
1141.5827	1	1141.5823	0.35	b10-H2O +1
1159.5927	1	1159.5928	-0.09	b10 +1
1252.9720	3	1252.9720	0.00	b33 +3
1271.9789	3	1271.9791	-0.16	b34 +3
1528.8300	1	1528.8304	-0.26	b13 +1

3.2.2.2 Peptidylprolyl isomerase

UniProt Q9HWK5, Peptidylprolyl isomerase
Seq:
ARATARHILVSSEAKCNELKTAIEGGADFAEVAREHSSCPSGRDGGNLGSFGPGQMVREFDQVVFSA
PLNVVQGPVKTQFGYHLLEVTSRQD
m/z monoisotopic = 983.5923, +10; MW_{obs} = 9825.85 Da (-1.6 ppm)



m/z meas	z	m/z calc	Δppm	assignment
418.2043	1	418.2045	-0.48	y3 +1
495.8013	2	495.8014	-0.20	b9 +2
505.2363	1	505.2365	-0.40	y4 +1
545.3352	2	545.3356	-0.73	b10 +2
599.3144	2	599.3148	-0.67	y10 +2
606.2839	1	606.2842	-0.49	y5 +1
609.8429	4	609.8431	-0.33	a23 +4
627.3680	1	627.3685	-0.80	b6 +1
646.6706	3	646.6707	-0.15	b18 +3
649.1020	4	649.1024	-0.62	b24 +4
678.6892	3	678.6898	-0.88	b19-NH3 +3
682.8911	2	682.8915	-0.59	a13 +2
684.3656	3	684.3653	0.44	b19 +3
688.3729	4	688.3737	-1.16	a27 +4
696.8886	2	696.8889	-0.43	b13 +2
705.3522	1	705.3526	-0.57	y6 +1
707.1289	4	707.1291	-0.28	y25 +4
718.4088	2	718.4100	-1.67	a14 +2
724.1288	4	724.1292	-0.55	b28 +4
727.0631	3	727.0637	-0.83	b20 +3

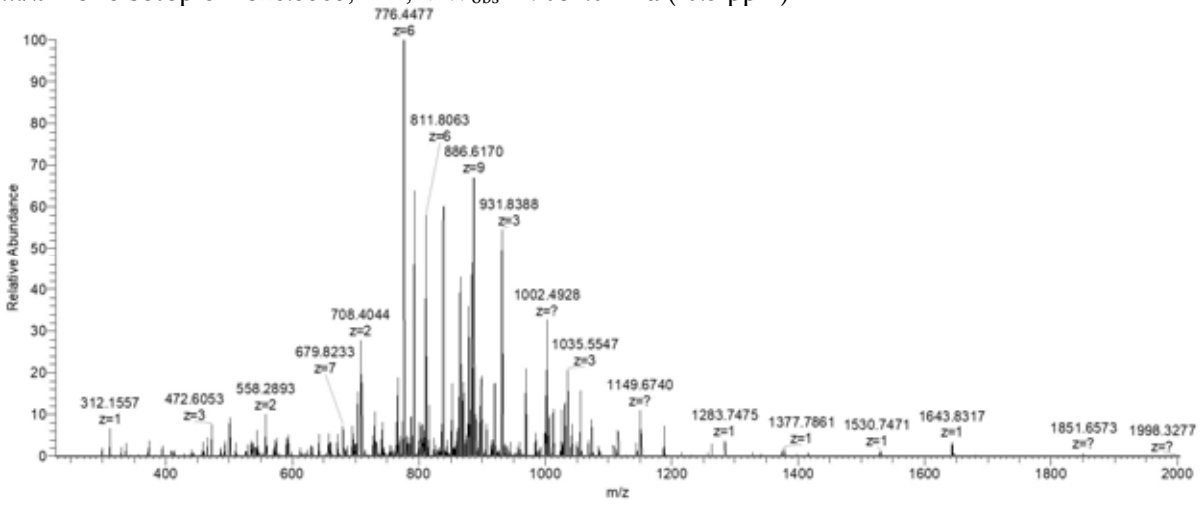
732.4069	2	732.4075	-0.82	b14 +2
736.5767	5	736.5790	-3.12	b35-NH3 +5
767.3939	5	767.3960	-2.74	b36 +5
784.4250	3	784.4253	-0.38	b22 +3
796.4548	2	796.4550	-0.25	b15 +2
801.4198	3	801.4203	-0.62	y21 +3
816.3843	1	816.3846	-0.37	y7-H2O +1
819.2088	5	819.2086	0.24	b39-H2O +5
822.1192	3	822.1199	-0.85	b23 +3
822.8092	5	822.8107	-1.82	b39 +5
834.3947	1	834.3952	-0.60	y7 +1
865.1335	3	865.1341	-0.69	b24 +3
911.0129	7	911.0138	-0.99	b61 +7
929.3051	7	929.3079	-3.01	b62 +7
936.8278	3	936.8275	0.32	y25-NH3 +3
942.5021	3	942.5030	-0.95	y25 +3
947.4785	1	947.4793	-0.84	y8 +1
1018.8988	5	1018.8969	1.86	b50-NH3 +5
1044.2143	3	1044.2155	-1.15	y28 +3
1059.5516	2	1059.5526	-0.94	y18 +2
1060.5622	1	1060.5633	-1.04	y9 +1
1077.2378	3	1077.2383	-0.46	y29 +3
1084.0238	6	1084.0246	-0.74	b62 +6
1088.0627	2	1088.0633	-0.55	y19 +2
1100.5353	6	1100.5360	-0.64	b63 +6
1117.2134	6	1117.2176	-3.76	y62-H2O +6
1143.0877	2	1143.0873	0.35	y20-H2O +2
1143.5793	2	1143.5793	0.00	y20-NH3 +2
1152.0914	2	1152.0926	-1.04	y20 +2
1201.6254	2	1201.6268	-1.17	y21 +2
1206.5866	6	1206.5881	-1.24	y68 +6
1251.1614	2	1251.1610	0.32	y22 +2
1308.1825	2	1308.1825	0.00	y23 +2
1364.7241	2	1364.7245	-0.29	y24 +2
1413.2494	2	1413.2509	-1.06	y25 +2
1448.7694	2	1448.7694	0.00	y26 +2
1492.2841	2	1492.2854	-0.87	y27 +2
1565.8203	2	1565.8196	0.45	y28 +2

3.2.2.3 HU-β

UniProt P05384, HU-β

Seq:
MNKSELIDAIAASADIPKAVAGRALDAVIESVTGALKAGDSVVLVGFGTFAVKERAARTGRNPQTGK
PIKIAAAKIPGFKAGKALKDAVN

m/z monoisotopic = 826.5565, +11; MW_{obs} = 9081.04 Da (-0.3 ppm)



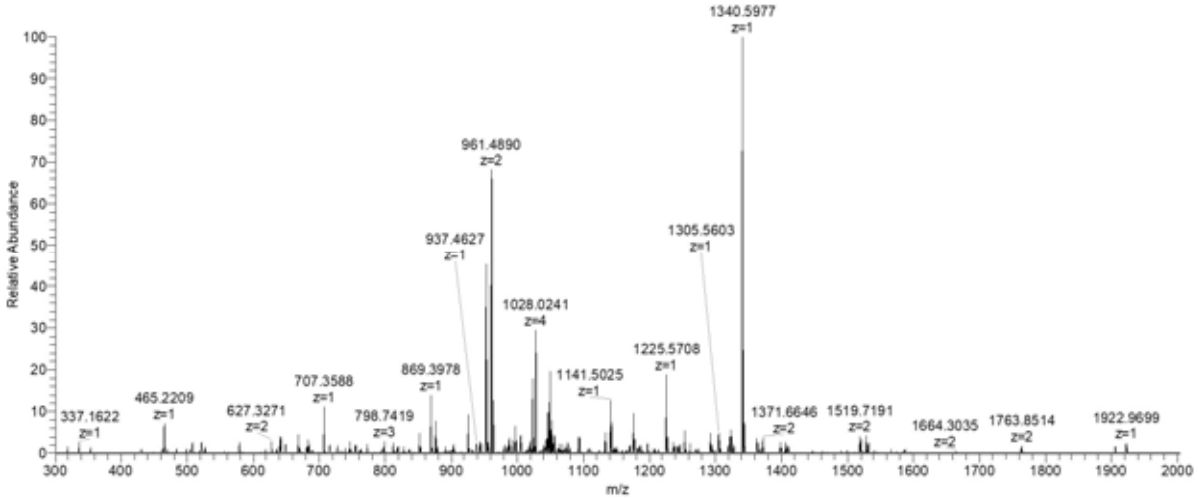
m/z meas	z	m/z calc	Δppm	assignment
374.1860	1	374.1857	0.80	b3 +1
394.7207	2	394.7204	0.76	a7 +2
408.7182	2	408.7178	0.98	b7 +2
461.2182	1	461.2177	1.08	b4 +1
466.2316	2	466.2313	0.64	b8 +2
472.6053	3	472.6050	0.63	y14 +3
487.7527	2	487.7524	0.62	a9 +2
492.7450	2	492.7446	0.81	b9-H2O +2
493.7855	2	493.7851	0.81	y10 +2
501.7503	2	501.7499	0.80	b9 +2
510.3001	3	510.2997	0.78	y15 +3
529.2621	1	529.2617	0.76	y5-NH3 +1
544.2949	2	544.2944	0.92	a10 +2
546.2886	2	546.2882	0.73	y5 +1
549.2870	2	549.2866	0.73	b10-H2O +2
558.2893	2	558.2919	-4.66	b10 +2
579.8133	2	579.8130	0.52	a11 +2
584.8059	2	584.8052	1.20	b11-H2O +2
590.2606	1	590.2603	0.51	b5 +1
593.8108	2	593.8105	0.51	b11 +2
600.3556	3	600.3561	-0.83	y18 +3
620.0354	6	620.0347	1.13	y36 +6
624.0355	3	624.0351	0.64	y19 +3
629.3293	2	629.3290	0.48	b12 +2
659.3727	1	659.3723	0.61	y6 +1

661.7302	3	661.7298	0.60	y20 +3
679.5363	7	679.5363	0.00	y46 +7
695.6913	7	695.6912	0.14	y47 +7
698.8799	4	698.8798	0.14	b28 +4
703.3448	1	703.3443	0.71	b6 +1
708.4044	2	708.4039	0.71	y14 +2
709.8443	7	709.8438	0.70	y48 +7
727.1509	4	727.1508	0.14	b29 +4
730.4096	1	730.4094	0.27	y7 +2
733.6731	8	733.6732	-0.14	y58 +8
742.0983	6	742.0981	0.27	y43 +6
743.8406	5	743.8402	0.54	y36 +5
764.9465	2	764.9459	0.78	y15 +2
766.6101	6	766.6095	0.78	y44 +6
773.2764	6	773.2754	1.29	y45-NH3 +6
776.1134	6	776.1131	0.39	y45 +6
781.1699	4	781.1695	0.51	b31 +6
787.2018	8	787.2016	0.25	y62 +8
792.6247	6	792.6245	0.25	y46 +6
801.4341	4	801.4340	0.12	b32-H2O +4
811.4722	6	811.4719	0.37	y47 +6
815.7594	7	815.7584	1.23	y56 +7
816.4288	1	816.4284	0.49	b7 +1
838.3401	7	838.3397	0.48	y58 +7
844.6596	5	844.6587	1.07	b43 +5
852.4931	7	852.4924	0.82	y59 +7
859.0000	6	858.9989	1.28	y50 +6
860.7147	9	860.7149	-0.23	y77 +9
862.3519	7	862.3526	-0.81	y60-H2O +7
864.9258	7	864.9255	0.35	y60 +7
870.3851	9	870.3852	-0.11	y78 +9
874.8030	3	874.8022	0.91	b26 +3
876.3875	9	876.3863	1.37	y79-NH3 +9
878.2775	9	878.2782	-0.80	y79 +9
883.3602	7	883.3602	0.00	y61 +7
884.1709	9	884.1700	1.02	y80-H2O +9
886.1713	9	886.1712	0.11	y80 +9
887.6746	6	887.6747	-0.11	y52 +6
896.0088	10	896.0077	1.23	y89 +10
898.4816	3	898.4812	0.45	b27 +3
898.7362	9	898.7361	0.11	y81 +9
899.5150	7	899.5167	-1.89	y62 +7
906.6279	9	906.6291	-1.32	y82 +9

915.5262	1	915.5258	0.44	y9 +1
919.4099	9	919.4099	0.00	y83 +9
931.5044	3	931.5040	0.43	b28 +3
935.5496	2	935.5491	0.53	y19 +2
941.2690	4	941.2672	1.91	b38 +4
945.1546	8	945.1532	1.48	b75 +8
955.5209	4	955.5226	-1.78	b39 +4
969.1989	3	969.1987	0.21	b29 +3
984.2796	4	984.2793	0.30	b40 +4
986.5636	1	986.5629	0.71	y10 +1
992.0919	2	992.0911	0.81	y20 +2
1001.5343	4	1001.5347	-0.40	b41-H2O +4
1002.4928	1	1002.4925	0.30	b9 +1
1006.0372	4	1006.0373	-0.10	b41 +4
1012.2134	3	1012.2129	0.49	b30 +3
1023.8069	4	1023.8057	1.17	a42 +4
1026.3014	4	1026.3018	-0.39	b42-H2O +4
1030.8044	4	1030.8044	0.00	b42 +4
1035.2204	3	1035.2201	0.29	b31-H2O +3
1041.2238	3	1041.2236	0.19	b31 +3
1051.0679	4	1051.0689	-0.95	b43-H2O +4
1055.5716	4	1055.5715	0.09	b43 +4
1068.2429	3	1068.2429	0.00	b32-H2O +3
1074.2468	3	1074.2464	0.37	b32 +3
1083.8440	4	1083.8425	1.38	b44 +4
1107.9300	3	1107.9289	0.99	b33 +3
1115.5770	1	1115.5765	0.45	b10 +1
1150.6156	3	1150.6151	0.43	b35 +3
1186.6146	1	1186.6136	0.84	b11 +1
1188.3099	3	1188.3098	0.08	b36 +3
1257.6511	1	1257.6507	0.32	b12 +1
1415.8017	1	1415.8005	0.85	y14 +1
1530.7471	1	1530.7468	0.20	b15 +1
1643.8317	1	1643.8309	0.49	b16 +1

3.2.2.4 PA2146

UniProt Q9I1W9, PA2146
Seq: AQHQGGKGNFAEDPKRASEAGKKGGQASGGNFKNDPQRASEAGKKGGQRSHGGN
m/z monoisotopic = 1090.1381, +5; MW_{obs} = 5445.65 Da (0.4 ppm)



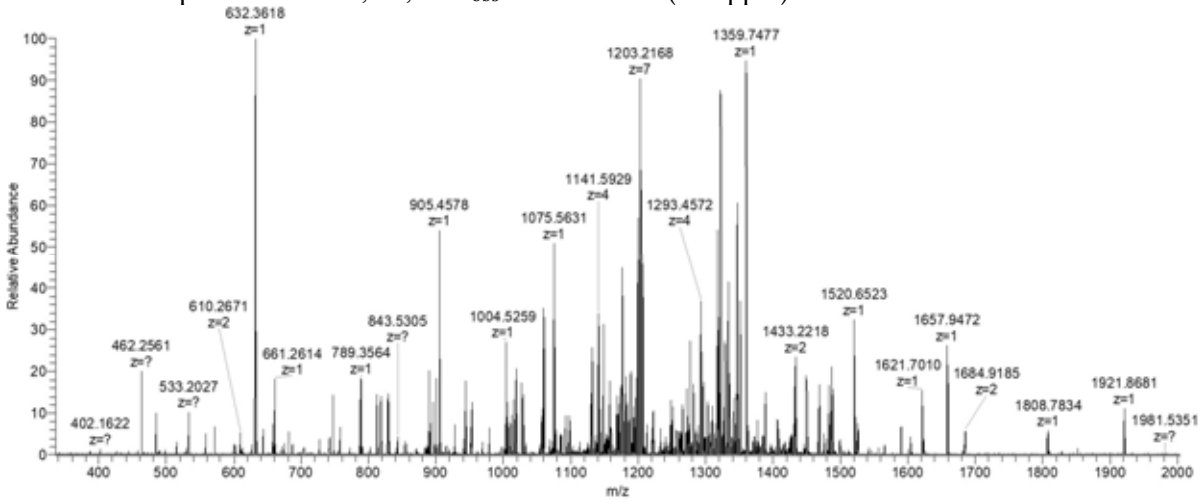
m/z meas	z	m/z calc	Δppm	assignment
337.1622	1	337.1619	0.89	b3 +1
384.1633	1	384.1626	1.82	y4 +1
465.2209	1	465.2205	0.86	b4 +1
471.1955	1	471.1946	1.91	y5 +1
499.2507	2	499.2497	2.00	y10 +2
522.2425	1	522.2419	1.15	b5 +1
579.2638	1	579.2634	0.69	b6 +1
618.8138	2	618.8132	0.97	y13-NH3 +2
627.2965	1	627.2957	1.28	y6 +1
627.3271	2	627.3265	0.96	y13 +2
635.6534	3	635.6526	1.26	y19-NH3 +3
641.3285	3	641.3282	0.47	y19 +3
707.3588	1	707.3583	0.71	b7 +1
755.3549	1	755.3543	0.79	y7 +1
755.7095	3	755.7088	0.93	b22 +3
764.3803	1	764.3798	0.65	b8 +1
792.4044	3	792.4036	1.01	b23-H2O +3
795.3500	1	795.3492	1.01	y8-NH3 +1
798.4077	3	798.4071	0.75	b23 +3
812.3764	1	812.3758	0.74	y8 +1
817.4152	3	817.4143	1.10	b24 +3
836.4218	3	836.4215	0.36	b25 +3
852.3712	1	852.3707	0.59	y9-NH3 +1
861.4282	2	861.4270	1.39	b16 +2
869.3978	1	869.3973	0.58	y9 +1
879.1083	3	879.1076	0.80	b26 +3
896.9465	2	896.9455	1.11	b17 +2
902.7871	3	902.7867	0.44	b27 +3

931.7977	3	931.7974	0.32	b28 +3
952.9755	2	952.9753	0.21	y19-NH3 +2
961.4890	2	961.4886	0.42	y19 +2
980.4662	1	980.4657	0.51	y10-NH3 +1
997.4927	1	997.4922	0.50	y10 +1
1004.9831	2	1004.9828	0.30	b19 +2
1018.8248	3	1018.8247	0.10	y31 +3
1023.2669	4	1023.2660	0.88	y41-NH3 +4
1027.5228	4	1027.5226	0.19	y41 +4
1046.7149	5	1046.7164	-1.43	y52-H2O +5
1050.3186	5	1050.3185	0.10	y52 +5
1052.0238	4	1052.0227	1.05	y42-NH3 +4
1056.2795	4	1056.2794	0.09	y42 +4
1061.5230	3	1061.5231	-0.09	y32 +3
1063.7274	5	1063.7270	0.38	b53 +5
1067.3298	5	1067.3291	0.66	b53+H2O +5
1075.9250	5	1075.9303	-4.93	y53 +5
1133.0598	2	1133.0596	0.18	b22 +2
1145.3219	4	1145.3228	-0.79	b45 +4
1146.9071	3	1146.9076	-0.44	y35 +3
1169.9007	3	1169.9002	0.43	b35-H2O +3
1175.9037	3	1175.9037	0.00	b35 +3
1180.5483	1	1180.5494	-0.93	a12-NH3 +1
1181.5786	4	1181.5759	2.29	y47-NH3 +4
1182.6094	1	1182.6086	0.68	y12 +1
1185.8317	4	1185.8325	-0.67	y47 +4
1197.1081	2	1197.1071	0.84	b23 +2
1208.5444	1	1208.5443	0.08	b12-NH3 +1
1225.5708	1	1225.5709	-0.08	b12 +1
1236.6190	1	1236.6192	-0.16	y13-NH3 +1
1242.1112	4	1242.1103	0.72	y50-NH3 +4
1246.3670	4	1246.3670	0.00	y50 +4
1253.6457	1	1253.6458	-0.08	y13 +1
1270.6230	4	1270.6267	-2.91	y24 +2
1274.1258	4	1274.1250	0.63	y51-NH3 +4
1278.3822	4	1278.3816	0.47	y51 +4
1290.6244	2	1290.6242	0.15	y25-NH3 +2
1299.1385	2	1299.1374	0.85	y25 +2
1319.1345	2	1319.1349	-0.30	y26-NH3 +2
1322.5869	1	1322.5872	-0.23	b13-H2O +1
1327.6481	2	1327.6482	-0.08	y26 +2
1340.5977	1	1340.5978	-0.07	b13 +1
1362.6508	2	1362.6509	-0.07	y27-NH3 +2

1371.1635	2	1371.1642	-0.51	y27 +2
1398.1690	2	1398.1695	-0.36	y28-NH3 +2
1406.6818	2	1406.6827	-0.64	y28 +2
1462.2000	2	1462.1987	0.89	y29-NH3 +2
1469.7207	1	1469.7204	0.20	y15 +1
1470.7150	2	1470.7120	2.04	y29 +2
1490.7083	2	1490.7095	-0.80	y30-NH3 +2
1499.2230	2	1499.2228	0.13	y30 +2
1519.2198	2	1519.2202	-0.26	y31-NH3 +2
1527.7324	2	1527.7335	-0.72	y31 +2
1540.7564	1	1540.7575	-0.71	y16 +1
1565.7447	1	1565.7455	-0.51	b15 +1
1583.2677	2	1583.2678	-0.06	y32-NH3 +2
1591.7790	2	1591.7810	-1.26	y32 +2
1763.3514	2	1763.3520	-0.34	b35 +2
1904.9426	1	1904.9434	-0.42	y19-NH3 +1
1921.9671	1	1921.9699	-1.46	y19 +1

3.2.2.5 PA4739

UniProt Q9HV60, PA4739
Seq:
MKFSKLAIAAATATALSFSMANAFAAQPISLAANDTMQKTEEAVSDTWITSKVKSSLIANKNVSGVD
IKVETNKGVVSLSGNVKSDAERDLAIETAKGIKGVKAVSADGLKSVE
m/z monoisotopic = 1223.5075, +7; MW_{obs} = 8557.50 Da (-0.7 ppm)



m/z meas	z	m/z calc	Δppm	assignment
402.1622	1	402.1619	0.75	b4 +1
444.2454	1	444.2453	0.23	y4-H2O +1
462.2561	1	462.2558	0.65	y4 +1

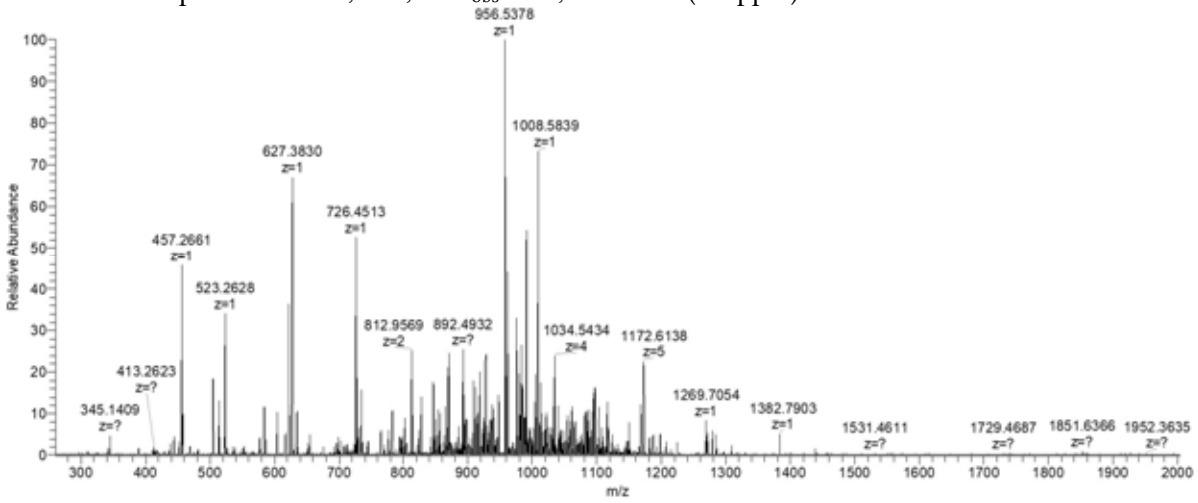
515.1923	1	515.1919	0.78	b5-H2O +1
533.2027	1	533.2024	0.56	b5 +1
610.2671	2	610.2666	0.82	b11 +2
614.3510	1	614.3508	0.33	y6-H2O +1
632.3618	1	632.3614	0.63	y6 +1
643.2505	1	643.2504	0.16	b6-H2O +1
644.2345	1	644.2345	0.00	b6-NH3 +1
661.2614	1	661.2610	0.60	b6 +1
747.3887	1	747.3883	0.54	y7 +1
772.3293	1	772.3294	-0.13	b7-NH3 +1
789.3564	1	789.3560	0.51	b7 +1
802.3491	2	802.3489	0.25	b15-H2O +2
811.3546	2	811.3541	0.62	b15 +2
818.4254	1	818.4257	-0.37	y8 +1
829.4781	2	829.4778	0.36	y17 +2
885.8394	3	885.8395	-0.11	y26 +3
890.4043	1	890.4036	0.79	b8 +1
895.3888	2	895.3885	0.34	b16-H2O +2
905.4578	1	905.4575	0.33	y9 +1
952.5331	3	952.5327	0.42	y28 +3
979.5678	2	979.5677	0.10	y20 +2
1004.5259	1	1004.5259	0.00	y10 +1
1019.4469	1	1019.4462	0.69	b9 +1
1059.8748	3	1059.8747	0.09	b29 +3
1075.5631	1	1075.5630	0.09	y11 +1
1091.8851	3	1091.8855	-0.37	b30-H2O +3
1097.8887	3	1097.8890	-0.27	b30 +3
1102.8057	5	1102.8073	-1.45	y54 +5
1130.9121	3	1130.9118	0.27	b31 +3
1136.5906	4	1136.5892	1.23	b42-H2O +4
1141.0890	4	1141.0918	-2.45	b42 +4
1143.1220	4	1143.1238	-1.57	y45 +4
1148.4890	1	1148.4888	0.17	b10 +1
1152.6295	3	1152.6283	1.04	y34 +3
1157.6129	7	1157.6157	-2.42	b78 +7
1162.4407	5	1162.4402	0.43	y57 +5
1168.5576	2	1168.5574	0.17	b21 +2
1173.3413	7	1173.3421	-0.68	b79-H2O +7
1175.9127	7	1175.9150	-1.96	b79 +7
1178.2001	7	1178.2036	-2.97	y79-NH3 +7
1180.1188	4	1180.1143	3.81	b44 +4
1185.9178	7	1185.9158	1.69	b80-NH3 +7
1188.3493	7	1188.3481	1.01	b80 +7

1192.6926	2	1192.6916	0.84	y24 +2
1199.9240	7	1199.9278	-3.17	b81-H2O +7
1202.4986	7	1202.5008	-1.83	b81 +7
1232.6045	2	1232.6049	-0.32	b22 +2
1248.1645	6	1248.1659	-1.12	b71 +6
1250.2955	3	1250.2948	0.56	b35 +3
1260.0060	6	1260.0055	0.40	b72 +6
1261.6987	5	1261.7003	-1.27	y62-H2O +5
1265.2987	5	1265.3025	-3.00	y62 +5
1271.1947	4	1271.1978	-2.44	y50 +4
1276.5167	5	1276.5169	-0.16	b73 +6
1281.3667	3	1281.3672	-0.39	y38 +3
1291.0191	6	1291.0222	-2.40	b74 +6
1292.9551	4	1292.9558	-0.54	y51 +4
1295.8669	6	1295.8671	-0.15	y75 +6
1302.7267	1	1302.7264	0.23	y13 +1
1302.8628	6	1302.8617	0.84	b75 +6
1319.1977	6	1319.1951	1.97	b76-NH3 +6
1322.0325	6	1322.0329	-0.30	b76 +6
1325.1341	5	1325.1321	1.51	y65-NH3 +5
1327.7256	3	1327.7216	3.01	y40-NH3 +3
1328.5380	5	1328.5374	0.45	y65 +5
1333.3965	3	1333.3971	-0.45	y40 +3
1341.9820	4	1341.9770	3.73	y53-NH3 +4
1346.2334	4	1346.2336	-0.15	y53 +4
1350.3858	6	1350.3838	1.48	b78 +6
1351.1538	5	1351.1542	-0.30	y66 +5
1359.7477	1	1359.7478	-0.07	y14 +1
1376.1778	2	1376.1789	-0.80	b25 +2
1378.2562	4	1378.2574	-0.87	y54 +4
1405.6275	1	1405.6264	0.78	b13 +1
1406.7712	4	1406.7681	2.20	y55 +4
1408.5738	5	1408.5796	-4.12	y68 +5
1432.7209	2	1432.7209	0.00	b26 +2
1448.9964	5	1448.9914	3.45	y70 +5
1468.2392	2	1468.2395	-0.20	b27 +2
1487.8424	1	1487.8428	-0.27	y15 +1
1520.6523	1	1520.6533	-0.66	b14 +1
1525.2602	2	1525.2610	-0.52	b28 +2
1589.3077	2	1589.3084	-0.44	b29 +2
1600.9286	1	1600.9268	1.12	y16 +1
1603.6906	1	1603.6904	0.12	b15-H2O +1
1621.7010	1	1621.7010	0.00	b15 +1

1657.9472	1	1657.9483	-0.66	y17 +1
1807.7816	1	1807.7803	0.72	b16 +1
1920.8653	1	1920.8644	0.47	b17 +1

3.2.2.6 PA5178

UniProt Q9HU11, PA5178
Seq:
MGIFAFVKEAGEKLWDTLTGHEAQAAESLKEHVAKVGLGNPNIQVSVEGDKVIASGEVASQEEKEK
ILLALGNVAGVSSVEDKITVAQAAPEARFVTVKKGDTLSAIAKAEYGNANAYMKIFEANKPMLSHPD
KIYPGQVLRIPE
m/z monoisotopic = 958.5127, +16; MW_{obs} = 15,320.09 Da (3.5 ppm)

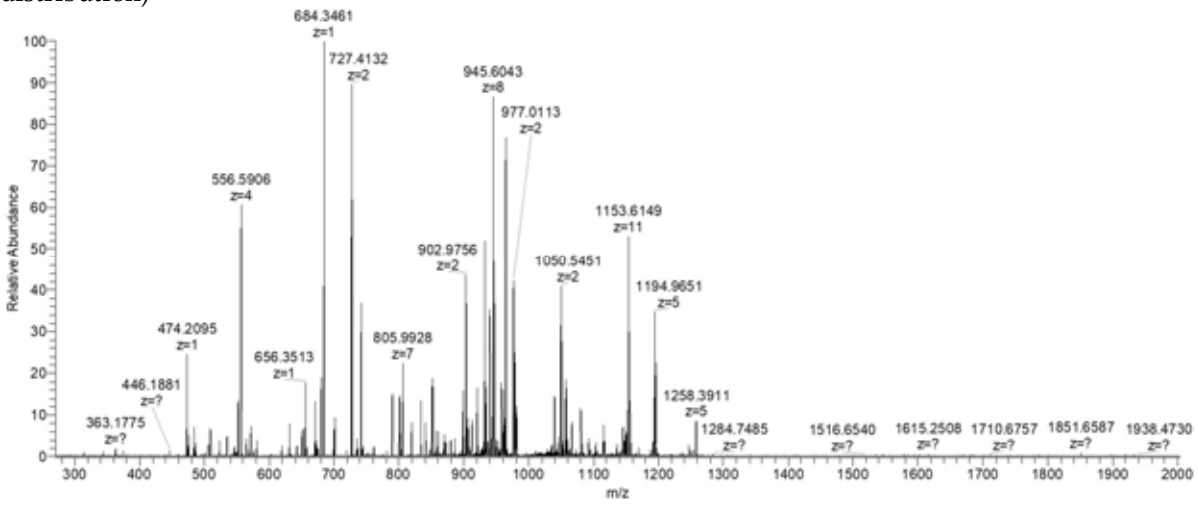


m/z meas	**z**	**m/z calc**	**Δppm**	**assignment**
389.2186	1	389.2183	0.77	b4 +1
418.7535	2	418.7531	0.96	y7-H2O +2
504.7959	2	504.7955	0.79	y9 +2
514.2989	1	514.2984	0.97	y4 +1
536.2874	1	536.2867	1.31	b5 +1
586.3277	2	586.3271	1.02	y10 +2
627.3830	1	627.3824	0.96	y5 +1
635.3557	1	635.3552	0.79	b6 +1
695.8906	2	695.8901	0.72	b13 +2
706.9169	2	706.9166	0.42	y12 +2
724.8793	6	724.8796	-0.41	b41 +6
726.4513	1	726.4509	0.55	y6 +1
730.7303	3	730.7295	1.09	y19 +3
743.7278	6	743.7269	1.21	b42 +6
777.3986	3	777.3987	-0.13	b21 +3
812.9569	2	812.9565	0.49	y14 +2
824.0303	5	824.0296	0.85	b39-NH3 +5

827.4349	5	827.4349	0.00	b39 +5
842.4503	7	842.4511	-0.95	y53 +7
846.4440	2	846.4432	0.95	b15 +2
854.5098	1	854.5094	0.47	y7 +1
869.6546	5	869.6541	0.57	b41 +5
881.4859	2	881.4860	-0.11	y15 +2
892.2710	5	892.2709	0.11	b42 +5
892.4932	1	892.4927	0.56	b8 +1
904.1057	8	904.1046	1.22	y66-NH3 +8
911.5315	1	911.5309	0.66	y8 +1
912.7350	8	912.7348	0.22	b69 +8
913.3399	7	913.3408	-0.99	y58 +7
917.8827	5	917.8826	0.11	b43 +5
924.6193	8	924.6190	0.32	b70-H2O +8
926.8691	8	926.8703	-1.29	b70 +8
927.9946	8	927.9909	3.99	y68 +8
931.8711	8	931.8697	1.50	b71-NH3 +8
934.0007	8	933.9980	2.89	b71 +8
940.3751	8	940.3745	0.64	y69 +8
947.5021	8	947.5021	0.00	y70 +8
963.5295	1	963.5298	-0.31	b9 +1
983.0235	8	983.0207	2.85	y73 +8
990.1496	8	990.1484	1.21	y74 +8
1004.2859	8	1004.2839	1.99	y75 +8
1008.5839	1	1008.5837	0.20	y9 +1
1020.3759	6	1020.3743	1.57	y55 +6
1023.5400	2	1023.5384	1.56	b19-H2O +2
1027.0233	4	1027.0223	0.97	y36 +4
1032.5472	2	1032.5437	3.39	b19 +2
1034.0430	4	1034.0418	1.16	b39 +4
1060.4215	7	1060.4171	4.15	y68 +7
1077.2352	6	1077.2359	-0.65	y59 +6
1080.5692	5	1080.5672	1.85	b51 +5
1083.5747	7	1083.5742	0.46	b72 +7
1086.8194	4	1086.8157	3.40	b41 +4
1095.5923	2	1095.5906	1.55	y19 +2
1115.0862	4	1115.0868	-0.54	b42 +4
1149.5944	1	1149.5939	0.43	b11 +1
1172.0115	5	1172.0107	0.68	b56 +5
1284.7325	1	1284.7310	1.17	y11 +1

3.2.2.7 Azurin

UniProt P00282, Azurin
Seq:
AECSVDIQGNDQMQFNTNAITVDKSCKQFTVNLSHPGNLPKNVMGHNWVLSTAADMQGVVTDG
MASGLDKDYLKPDDSRVIAHTKLIGSGEKDSVTFDVSKLKEGEQYMFFCTFPGHSALMKGTLTLK
m/z monoisotopic = 996.4895, +14; MW_{obs} = 13,936.75 Da (-0.8 ppm)
(* in the last column denotes peaks which do not correspond to the monoisotopic peaks within the distribution)

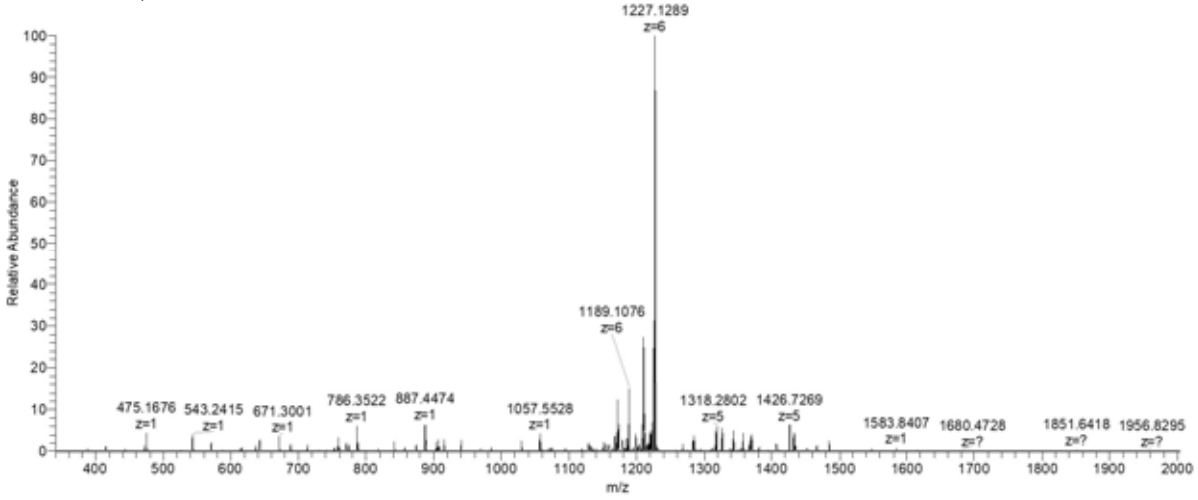


m/z meas	z	m/z calc	Δppm	assignment	not mono?
485.2779	3	485.2780	-0.21	y14 +3	
602.3197	3	602.3198	-0.17	y17 +3	
632.3974	1	632.3978	-0.63	y6 +1	
651.3421	3	651.3426	-0.77	y18 +3	
700.3649	3	700.3654	-0.71	y19 +3	
718.4077	1	718.4081	-0.56	y14-H2O +2	
727.4132	2	727.4134	-0.27	y14 +2	
744.0451	3	744.0456	-0.67	y20 +3	
800.9473	2	800.9476	-0.37	y15 +2	
805.5626	7	805.5634	-0.99	y51 +7	
851.4710	2	851.4715	-0.59	y16 +2	
869.8859	9	869.8842	1.95	y72-NH3 +9	
871.7737	9	871.7761	-2.75	y72 +9	
902.9756	2	902.9760	-0.44	y17 +2	
932.7195	8	932.7201	-0.64	y68 +8	
939.6558	6	939.6561	-0.32	y51 +6	
942.9721	8	942.9753	-3.39	y69-NH3 +8	
945.1027	8	945.1036	-0.95	y69 +8	
957.4892	8	957.4872	2.09	y70 +8	
962.3655	8	962.3635	2.08	y71-H2O +8	
964.6138	8	964.6148	-1.04	y71 +8	

976.5096	2	976.5103	-0.72	y18 +2	
978.4946	8	978.4939	0.72	y72-NH3 +8	
980.6201	8	980.6222	-2.14	y72 +8	
1045.4894	6	1045.4928	-3.25	b58-H2O +6	
1048.4925	6	1048.4945	-1.91	b58 +6	*
1050.0437	2	1050.0445	-0.76	y19 +2	
1115.5642	2	1115.5647	-0.45	y20 +2	
1194.7647	5	1194.7679	-2.68	b55 +5	*
1246.5853	5	1246.5877	-1.93	b57 +5	*
1257.9891	5	1257.9920	-2.31	b58 +5	*

3.2.2.8 CspA

UniProt P95459, CspA
Seq:
SNRQNGTVKWFNDAKGFGFITPESGNDLFVHFRSIQGTGFKSLQEGQKVSFVVVNGQKGLQADEVQVV
m/z monoisotopic = 1245.9733, +6; MW_{obs} = 7469.80 Da (-0.2 ppm)
(* in the last column denotes peaks which do not correspond to the monoisotopic peaks within the distribution)

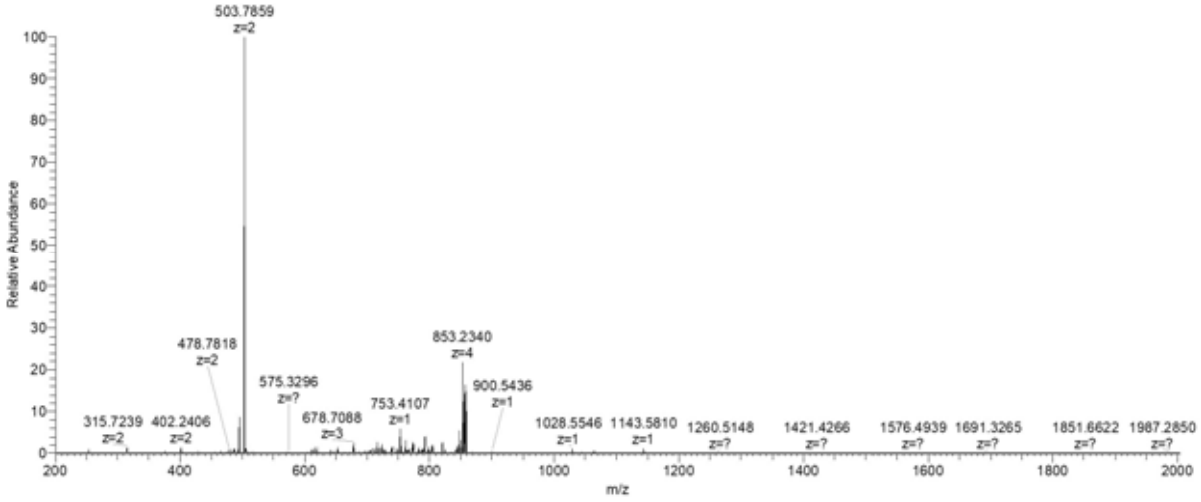


m/z meas	z	m/z calc	Δppm	assignment	not mono?
688.3517	1	688.3512	0.73	y6 +1	
759.3890	1	759.3883	0.92	y7 +1	
774.3716	2	774.3711	0.65	b13 +2	
790.3978	3	790.3977	0.13	b21 +3	
857.4225	1	857.4224	0.12	b8 +1	
873.9377	2	873.9372	0.57	b15 +2	
887.4474	1	887.4469	0.56	y8 +1	
985.5179	1	985.5174	0.51	b9 +1	
1057.5528	1	1057.5524	0.38	y10 +1	

1088.6065	6	1088.6058	0.64	b65 +6	
1094.0815	4	1094.0815	0.00	y40-NH3 +4	
1119.5730	6	1119.5767	-3.30	b61 +6	
1134.5694	2	1134.5691	0.26	b20 +2	
1150.5861	6	1150.5873	-1.04	b63 +6	
1158.5873	5	1158.5868	0.43	b52 +5	
1167.8101	5	1167.8117	-1.37	y54-H2O +5	*
1169.2600	6	1169.2567	2.82	b64-NH3 +6	
1171.4122	5	1171.4138	-1.37	y54 +5	*
1172.0935	6	1172.0944	-0.77	b64 +6	
1178.4000	5	1178.4005	-0.42	b53 +5	
1181.0987	6	1181.1023	-3.05	a65-NH3 +6	
1185.0936	2	1185.0929	0.59	b21 +2	
1185.7710	6	1185.7681	2.45	b65-NH3 +6	
1198.2151	5	1198.2141	0.83	b54 +5	
1207.1139	6	1207.1112	2.24	b66-NH3 +6	
1209.9494	6	1209.9489	0.41	b66 +6	
1223.6252	6	1223.6226	2.12	b67-NH3 +6	
1226.4603	6	1226.4603	0.00	b67 +6	
1283.6527	5	1283.6577	-3.90	b58 +5	
1317.6777	5	1317.6788	-0.83	b60 +5	
1343.2906	5	1343.2905	0.07	b61 +5	
1357.4932	5	1357.4980	-3.54	b62 +5	
1370.7283	1	1370.7274	0.66	y13 +1	
1466.7605	1	1466.7598	0.48	y14-H2O +1	
1484.7714	1	1484.7703	0.74	y14 +1	
1547.7368	1	1547.7350	1.16	b13 +1	
1583.8407	1	1583.8388	1.20	y15 +1	

3.2.2.9 L36

UniProt Q9HWF6, L36
Seq: MKVRASVKKLCRNCKIIRRDGIVRVICSAEPRHKQRQG
m/z monoisotopic = 739.5860, +6; MW_{obs} = 4431.47 Da (-1.2 ppm)

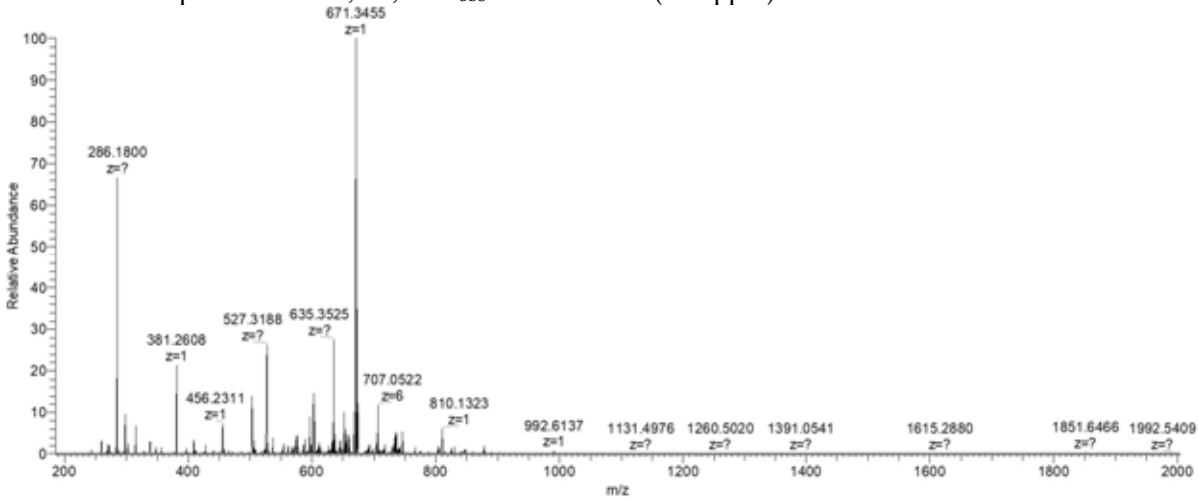


m/z meas	z	m/z calc	Δppm	assignment
360.1987	1	360.1990	-0.83	y3 +1
377.2091	2	377.2094	-0.80	y6 +2
455.2599	2	455.2599	0.00	y7 +2
471.2308	1	471.2310	-0.42	y4-NH3 +1
488.2573	1	488.2576	-0.61	y4 +1
503.7859	2	503.7863	-0.79	y8 +2
515.3118	1	515.3122	-0.78	b4 +1
577.0653	4	577.0659	-1.04	y20 +4
599.3258	1	599.3260	-0.33	y5-NH3 +1
616.3521	1	616.3525	-0.65	y5 +1
647.3414	2	647.3422	-1.24	y11 +2
678.7088	3	678.7094	-0.88	y18 +3
698.8460	2	698.8468	-1.14	y12 +2
700.9203	2	700.9207	-0.57	b12 +2
702.4155	4	702.4162	-1.00	b24-H2O +4
704.4097	3	704.4128	-4.40	b18-NH3 +3
706.9182	4	706.9188	-0.85	b24 +4
710.0876	3	710.0883	-0.99	b18 +3
717.0512	3	717.0518	-0.84	y19 +3
753.4107	1	753.4114	-0.93	y6 +1
755.3882	2	755.3888	-0.79	y13 +2
762.1213	3	762.1220	-0.92	b19 +3
800.4634	3	800.4643	-1.12	b20 +3
804.9220	2	804.9230	-1.24	y14 +2
825.2256	4	825.2265	-1.09	b29 +4
852.9834	4	852.9845	-1.29	b30-H2O +4
857.4861	4	857.4872	-1.28	b30 +4
900.5436	1	900.5448	-1.33	b8 +1

986.5775	2	986.5783	-0.81	b17 +2
1064.6279	2	1064.6288	-0.85	b18 +2

3.2.2.10 L33

UniProt Q9HTN9, L33
Seq: MRELIRLVSSAGTGHFYTTDKNKRTKPEKIEIKKYDPVVRQHVIYKEAKIK
m/z monoisotopic = 672.2674, +9; MW_{obs} = 6041.34 Da (-1.8 ppm)

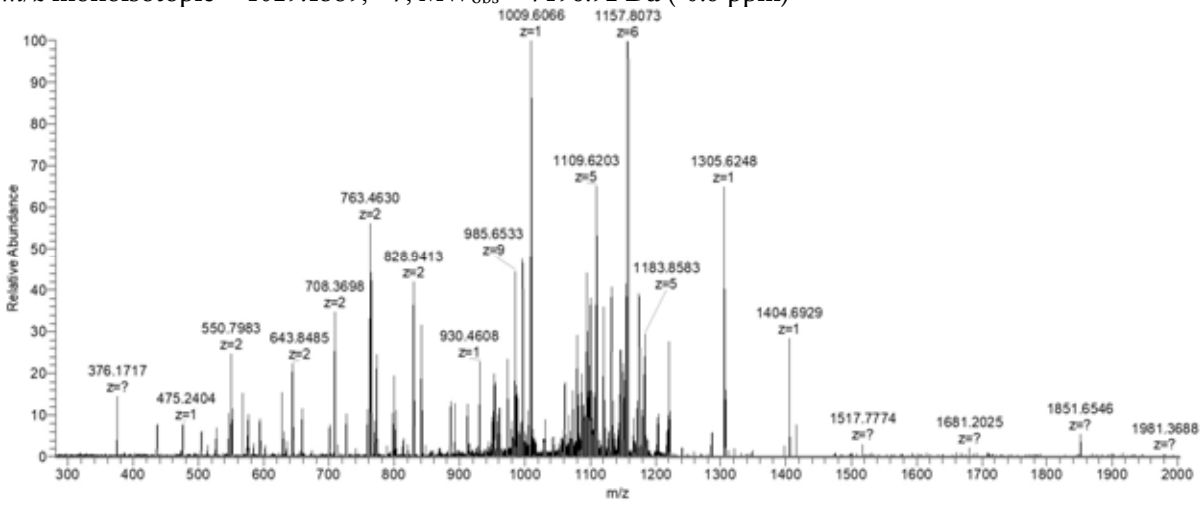


m/z meas	z	m/z calc	Δppm	assignment
400.2336	2	400.2340	-1.00	b6 +2
431.2630	2	431.2633	-0.70	y7-H2O +2
440.2683	2	440.2686	-0.68	y7 +2
536.9591	3	536.9595	-0.74	b15 +3
570.3607	5	570.3610	-0.53	y5-H2O +5
585.9818	3	585.9823	-0.85	b16 +3
588.3710	1	588.3715	-0.85	y5 +1
597.3643	3	597.3648	-0.84	y15-H2O +3
603.3678	3	603.3684	-0.99	y15 +3
610.1158	4	610.1165	-1.15	y20-H2O +4
632.3782	6	632.3790	-1.27	y31-H2O +6
635.3796	6	635.3808	-1.89	y31 +6
696.0645	3	696.0651	-0.86	y17 +3
706.7176	6	706.7187	-1.56	b36 +6
716.4664	1	716.4665	-0.14	y6 +1
718.3872	5	718.3881	-1.25	b31 +5
736.4053	2	736.4061	-1.09	b14 +2
746.0434	3	746.0442	-1.07	b20 +3
748.6535	4	748.6541	-0.80	b26 +4
766.6228	5	766.6239	-1.43	b33 +5

804.9347	2	804.9356	-1.12	b15 +2
826.1196	6	826.1201	-0.61	b42 +6
847.8599	5	847.8610	-1.30	b36 +5
861.5185	1	861.5193	-0.93	y7-H2O +1
879.5292	1	879.5298	-0.68	y7 +1
904.5481	2	904.5489	-0.88	y15 +2
992.6137	1	992.6139	-0.20	y8 +1

3.2.2.11 L29

UniProt Q9HWE3, L29
Seq:
MKANELREKSVEQLNEQLLGLLRDQFNLRMQKATGQLGQSHLLSQVKRDIARVKTVLNQQAGK
m/z monoisotopic = 1029.1389, +7; MW_{obs} = 7196.92 Da (-0.6 ppm)

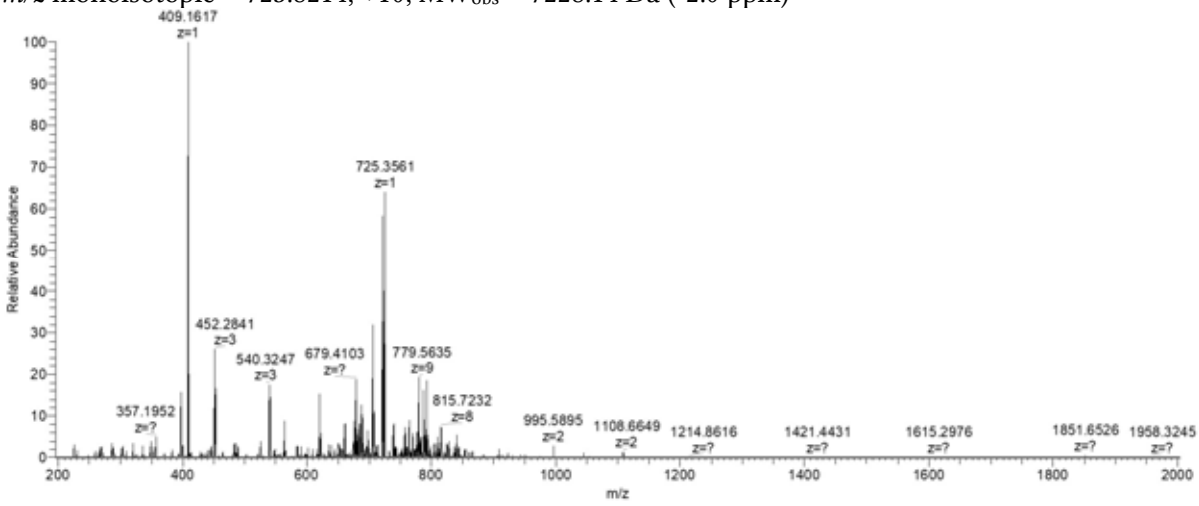


m/z meas	z	m/z calc	Δppm	assignment
514.2629	1	514.2620	1.75	y5-NH3 +1
550.7983	2	550.7977	1.09	b9 +2
574.2655	1	574.2654	0.17	b5 +1
594.3142	2	594.3137	0.84	b10 +2
643.8485	2	643.8479	0.93	b11 +2
645.3325	1	645.3315	1.55	y6 +1
708.3698	2	708.3692	0.85	b12 +2
758.4160	2	758.4155	0.66	y7 +1
763.4630	2	763.4623	0.92	y14 +2
772.3992	2	772.3985	0.91	b13 +2
814.9441	2	814.9431	1.23	a14 +2
819.9357	2	819.9352	0.61	b14-H2O +2
828.9413	2	828.9405	0.97	b14 +2
885.9628	2	885.9620	0.90	b15 +2

950.4844	2	950.4833	1.16	b16 +2
958.5329	1	958.5316	1.36	y9 +1
958.8848	3	958.8843	0.52	y26 +3
972.4940	1	972.4931	0.93	b8 +1
1086.4116	5	1086.4105	1.01	y48 +5
1100.5885	1	1100.5881	0.36	b9 +1
1105.8157	5	1105.8138	1.72	y49-NH3 +5
1109.2198	5	1109.2191	0.63	y49 +5
1131.8358	5	1131.8359	-0.09	y50 +5
1145.4604	6	1145.4661	-4.98	y60 +6
1154.8045	6	1154.8008	3.20	b60 +6
1157.3067	6	1157.3056	0.95	y61 +6
1183.2570	5	1183.2561	0.76	y52 +5
1185.6967	1	1185.6950	1.43	y11 +1
1220.4787	5	1220.4762	2.05	y54 +5
1286.6898	1	1286.6885	1.01	b11 +1
1415.7329	1	1415.7311	1.27	b12 +1

3.2.2.12 L35

UniProt Q9I0A1, L35
Seq: PKMKTKSGAAKRFKKTAGGLKHKHAFKSHILTKMTTKRKRQLRGTSMLNKSDVARVERSLRLR
m/z monoisotopic = 723.8214, +10; MW_{obs} = 7228.14 Da (-2.0 ppm)



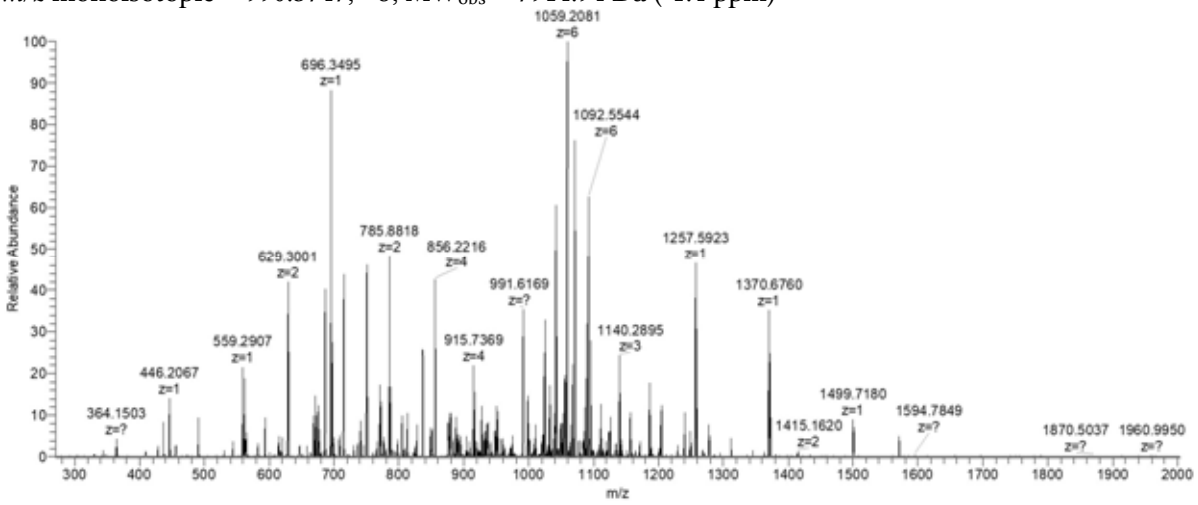
m/z meas	z	m/z calc	Δppm	assignment
226.1550	1	226.1550	0.00	b2 +1
357.1952	1	357.1955	-0.84	b3 +1
400.7638	2	400.7643	-1.25	y6 +2
446.2805	3	446.2810	-1.12	y11-H2O +3
452.2841	3	452.2845	-0.88	y11 +3

485.2896	1	485.2905	-1.85	b4 +1
554.5857	4	554.5860	-0.54	b21 +4
557.3390	3	557.3397	-1.26	b15-H2O +3
557.3875	1	557.3882	-1.26	y4 +1
563.3426	3	563.3433	-1.24	b15 +3
584.3474	4	584.3480	-1.03	b22-H2O +4
586.3373	1	586.3381	-1.36	b5 +1
597.0253	3	597.0258	-0.84	b16 +3
616.3712	4	616.3718	-0.97	b23-H2O +4
620.7044	3	620.7049	-0.81	b17 +3
620.8736	4	620.8744	-1.29	b23 +4
638.3702	5	638.3711	-1.41	b29 +5
644.4197	1	644.4202	-0.78	y5 +1
650.6358	4	650.6365	-1.08	b24-H2O +4
655.3884	5	655.3889	-0.76	a30 +5
655.7155	6	655.7161	-0.92	y33 +6
657.3853	5	657.3858	-0.76	b30-H2O +5
660.9869	5	660.9879	-1.51	b30 +5
672.8979	4	672.8984	-0.74	b25 +4
674.5627	6	674.5634	-1.04	y34 +6
679.3896	7	679.3912	-2.36	y40 +7
683.6038	5	683.6047	-1.32	b31 +5
711.9112	6	711.9119	-0.98	y36 +6
739.1112	6	739.1122	-1.35	b21 +3
743.8302	5	743.8315	-1.75	y31 +5
760.4388	5	760.4389	-0.13	y32-H2O +5
764.0398	5	764.0410	-1.57	y32 +5
766.7755	6	766.7742	1.70	y39-NH3 +6
768.0151	7	768.0163	-1.56	y46 +7
769.6104	6	769.6120	-2.08	y39 +6
777.1159	9	777.1181	-2.83	y61-H2O +9
779.1169	9	779.1193	-3.08	y61 +9
780.4638	2	780.4638	0.00	b14 +2
783.0549	5	783.0557	-1.02	y33-H2O +5
786.6564	5	786.6578	-1.78	y33 +5
789.4534	6	789.4534	0.00	y40-H2O +6
792.4531	6	792.4551	-2.52	y40 +6
793.3485	9	793.3520	-4.41	y62 +9
797.7115	4	797.7120	-0.63	b29 +4
804.4670	8	804.4695	-3.11	y56 +8
809.2737	5	809.2746	-1.11	y34 +5
813.2209	8	813.2202	0.86	y57-NH3 +8
813.8044	6	813.8043	0.12	y41 +6

815.3473	8	815.3485	-1.47	y57 +8
825.9818	4	825.9830	-1.45	b30 +4
827.4960	3	827.4968	-0.97	b23 +3
836.6460	6	836.6474	-1.67	y42 +6
837.6252	7	837.6235	2.03	b52-H2O +7
840.1948	7	840.1964	-1.90	b52 +7
843.9906	8	843.9913	-0.83	y59 +8
844.5107	2	844.5113	-0.71	b15 +2
854.0915	5	854.0928	-1.52	y36 +5
879.7096	5	879.7118	-2.50	y37 +5
909.1243	5	909.1255	-1.32	y38 +5
950.7419	5	950.7447	-2.95	y40 +5

3.2.2.13 L31

UniProt Q9HUD0, L31
Seq:
MKADIHPTYEAIEATCSCGNVIKTRSTLCKPIHLDVCSECHPFYTGKQKVLDTGGRIDRFKQRFGVFG
ATK
m/z monoisotopic = 990.3747, +8; MW_{obs} = 7914.94 Da (-1.4 ppm)



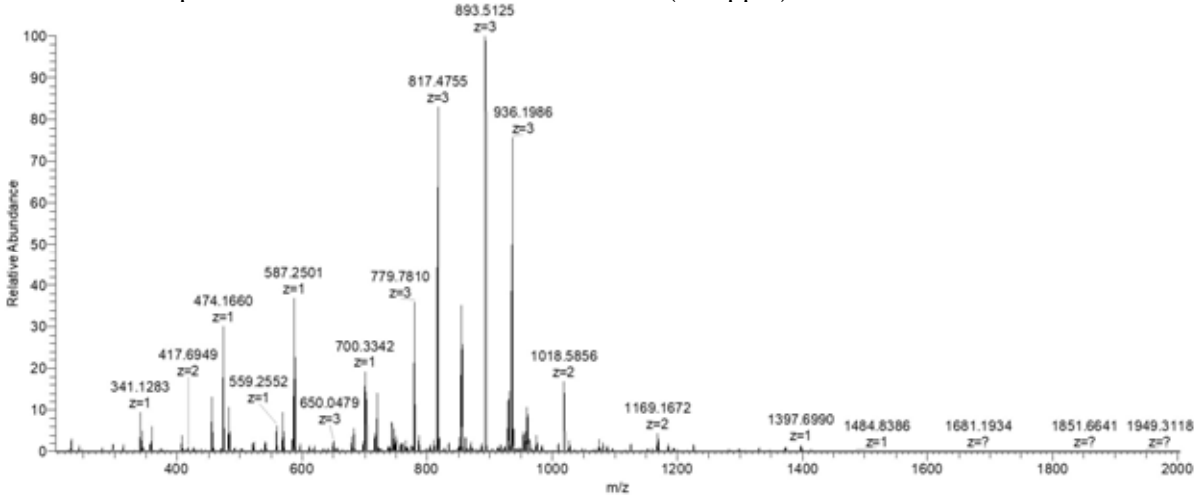
m/z meas	z	m/z calc	Δppm	assignment
331.1799	1	331.1798	0.30	b3 +1
446.2067	1	446.2068	-0.22	b4 +1
531.2957	1	531.2959	-0.38	a5 +1
559.2907	1	559.2908	-0.18	b5 +1
593.7814	2	593.7817	-0.51	b10 +2
615.3022	2	615.3028	-0.98	a11 +2
620.2945	2	620.2950	-0.81	b11-H2O +2
629.3001	2	629.3003	-0.32	b11 +2

651.3282	2	651.3283	-0.15	a6-NH3 +2
663.3314	2	663.3316	-0.30	a12-NH3 +2
668.3546	1	668.3548	-0.30	a6 +1
671.8447	2	671.8448	-0.15	a12 +2
678.3390	1	678.3392	-0.29	b6-H2O +6
685.8420	2	685.8423	-0.44	b12 +2
696.3495	1	696.3498	-0.43	b6 +2
708.3917	3	708.3922	-0.71	y19-H2O +3
714.3954	3	714.3957	-0.42	y19 +3
736.3657	2	736.3661	-0.54	a13 +2
741.3575	2	741.3583	-1.08	b13-H2O +2
750.3633	2	750.3636	-0.40	b13 +2
771.4383	2	771.4386	-0.39	y13 +2
776.8764	2	776.8769	-0.64	b14-H2O +2
785.8818	2	785.8821	-0.38	b14 +2
822.4089	2	822.4085	0.49	a15 +2
827.4001	2	827.4007	-0.73	b15-H2O +2
836.4056	2	836.4060	-0.48	b15 +2
850.7160	4	850.7164	-0.47	y30-H2O +4
855.2186	4	855.2190	-0.47	y30 +4
878.9048	2	878.9053	-0.57	b16-H2O +2
887.9098	2	887.9106	-0.90	b16 +2
889.0544	5	889.0547	-0.34	y39 +5
894.4497	1	894.4502	-0.56	b8 +1
903.9497	4	903.9503	-0.66	b33 +4
915.2352	4	915.2360	-0.87	y32 +4
931.4260	4	931.4266	-0.64	b17 +2
932.2198	4	932.2213	-1.61	b34 +4
933.4723	7	933.4744	-2.25	y59-H2O +7
952.2019	7	952.2022	-0.32	y60 +7
1010.0269	6	1010.0259	0.99	y54 +6
1021.6959	6	1021.6935	2.35	y55-NH3 +6
1024.5292	6	1024.5312	-1.95	y55 +6
1029.9490	7	1029.9475	1.46	y65-NH3 +7
1032.3797	7	1032.3798	-0.10	y65 +7
1038.6968	6	1038.6977	-0.87	y56-H2O +6
1041.6985	6	1041.6994	-0.86	y56 +6
1048.5332	4	1048.5308	2.29	y37 +4
1055.5414	6	1055.5390	2.27	y57-H2O +6
1058.5394	6	1058.5407	-1.23	y57 +6
1065.5148	4	1065.5161	-1.22	b39 +4
1067.5441	6	1067.5425	1.50	y58-NH3 +6
1070.3789	6	1070.3802	-1.21	y58 +6

1076.7998	4	1076.8018	-1.86	y38 +4
1084.5503	7	1084.5470	3.04	y68 +7
1088.8845	6	1088.8856	-1.01	y59-H2O +6
1091.8864	6	1091.8873	-0.82	y59 +6
1094.6941	7	1094.6951	-0.91	y69 +7
1107.8970	6	1107.8969	0.09	y60-NH3 +6
1110.7343	6	1110.7347	-0.36	y60 +6
1118.5353	4	1118.5344	0.80	a41 +4
1125.5308	4	1125.5331	-2.04	b41 +4
1133.9522	3	1133.9527	-0.44	y30-H2O +3
1137.2032	5	1137.2012	1.76	y50 +5
1139.9556	3	1139.9563	-0.61	y30 +3
1155.9628	5	1155.9643	-1.30	b52 +5
1159.2424	3	1159.2450	-2.24	b32 +3
1163.5994	4	1163.6008	-1.20	y41 +4
1186.5552	1	1186.5561	-0.76	b10 +1
1204.2723	6	1204.2753	-2.49	y65 +6
1229.2403	5	1229.2360	3.50	y55 +5
1239.5841	1	1239.5827	1.13	b11-H2O +1
1257.5923	1	1257.5932	-0.72	b11 +1
1370.6760	1	1370.6773	-0.95	b12 +1
1499.7180	1	1499.7199	-1.27	b13 +1
1570.7562	1	1570.7570	-0.51	b14 +1

3.2.2.14 S21

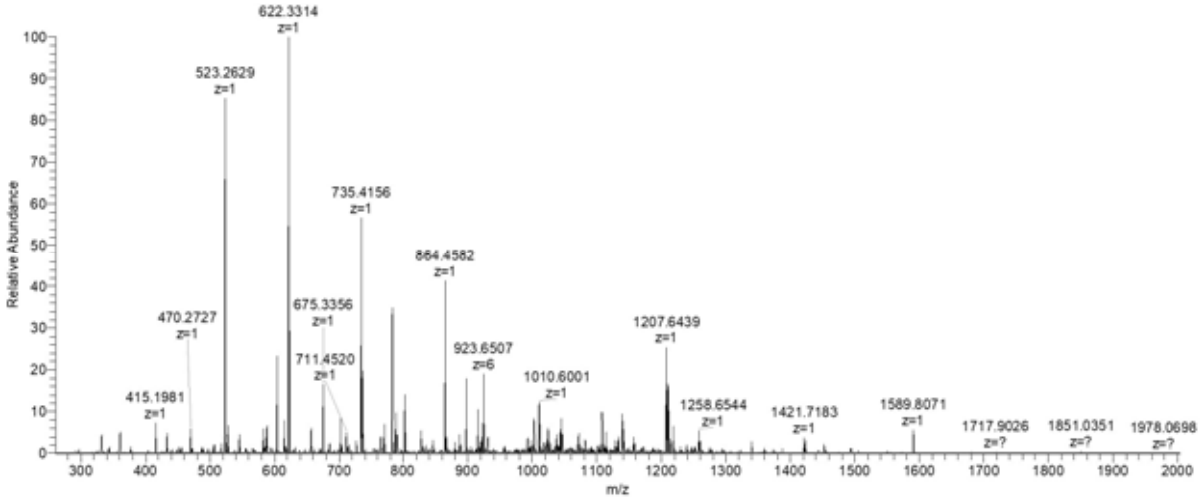
UniProt Q9I5V8, S21
Seq:
PAVKVKENEPFDVALRRFKRSCEKAGVLAEVRSREFYEKPTAERKRKAAAAVKRHAKKVQREQRRR
ERLY
m/z monoisotopic = 759.9734, +11; MW_{obs} = 8348.63 Da (-1.4 ppm)



m/z meas	z	m/z calc	Δppm	assignment
268.1656	1	268.1656	0.00	b3 +1
367.7313	2	367.7316	-0.82	b7-H2O +2
396.2602	1	396.2605	-0.76	b4 +1
446.7589	2	446.7592	-0.67	y6 +2
451.2660	1	451.2663	-0.66	y3 +1
495.3286	1	495.3289	-0.61	b5 +1
498.2793	2	498.2796	-0.60	b9 +2
524.8095	2	524.8098	-0.57	y7 +2
562.2979	1	562.2984	-0.89	y4-H2O +1
580.3084	1	580.3089	-0.86	y4 +1
623.4234	1	623.4239	-0.80	b6 +1
644.3544	2	644.3551	-1.09	y9-H2O +2
647.3910	5	647.3921	-1.70	y26 +5
653.3904	6	653.3907	-0.46	y32 +6
677.8530	2	677.8537	-1.03	b12 +2
678.6114	5	678.6123	-1.33	y27 +5
712.5520	7	712.5530	-1.40	y40 +7
718.4130	3	718.4140	-1.39	y16 +3
736.4086	1	736.4100	-1.90	y5 +1
778.2222	9	778.2235	-1.67	y58 +9
817.2606	10	817.2611	-0.61	y68-H2O +10
818.1279	9	818.1288	-1.10	y61 +9
819.0605	10	819.0622	-2.08	y68 +10
826.1657	10	826.1659	-0.24	y69 +10
842.9607	4	842.9620	-1.54	b30 +4
977.5418	1	977.5415	0.31	b9-H2O +1
995.5511	1	995.5520	-0.90	b9 +1
1354.6990	1	1354.7001	-0.81	b12 +1

3.2.2.15 PA0039

UniProt Q9I793, PA0039
Seq: AKPCEELKAEIDAKIKANGVPAYTLEIVDKGSVTDKKVVGTCDGGTKEIVYQRG
m/z monoisotopic = 956.3376, +6; MW_{obs} = 5731.98 Da (0.1 ppm)



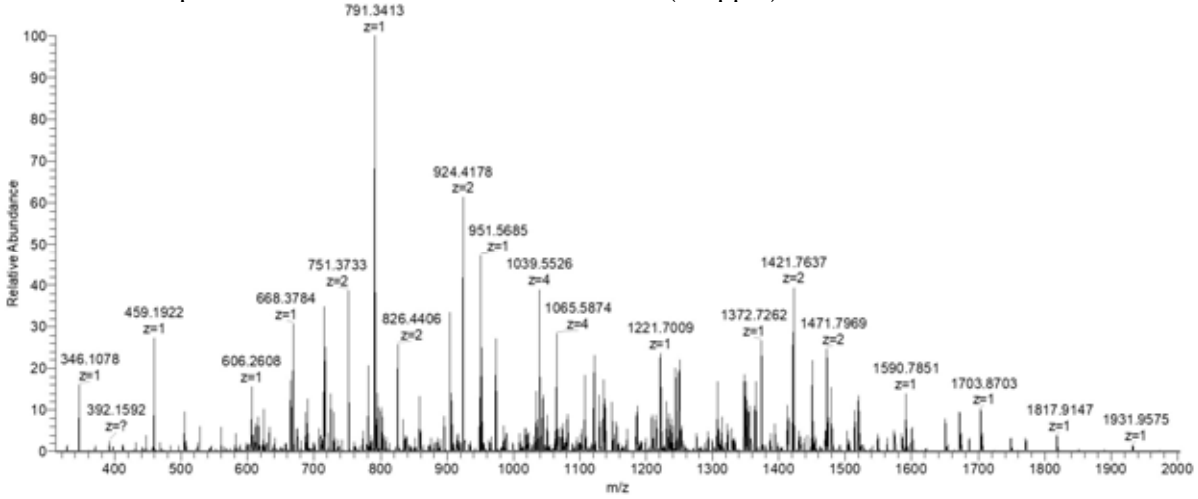
m/z meas	z	m/z calc	Δppm	assignment
360.1995	1	360.1990	1.39	y3 +1
523.2629	1	523.2623	1.15	y4 +1
604.3258	2	604.3251	1.16	y11 +2
622.3314	1	622.3307	1.12	y5 +1
735.4156	1	735.4148	1.09	y6 +1
864.4582	1	864.4574	0.93	y7 +1
906.9725	6	906.9732	-0.77	y51 +6
920.3142	6	920.3109	3.59	y52-NH3 +6
923.1492	6	923.1487	0.54	y52 +6
992.5535	1	992.5524	1.11	y8 +1
996.9192	5	996.9198	-0.60	b48-H2O +5
1000.5261	5	1000.5219	4.20	b48 +5
1017.5379	5	1017.5397	-1.77	a49 +5
1019.7376	5	1019.7334	4.12	b49-NH3 +5
1023.1398	5	1023.1387	1.08	b49 +5
1037.3532	5	1037.3534	-0.19	a50 +5
1042.9544	5	1042.9524	1.92	b50 +5
1075.5668	5	1075.5650	1.67	b51 +5
1088.1669	5	1088.1664	0.46	y51 +5
1093.6010	1	1093.6000	0.91	y9 +1
1101.1799	5	1101.1767	2.91	b52 +5
1107.5779	5	1107.5770	0.81	y52 +5
1132.3440	4	1132.3435	0.44	b43 +4
1150.6233	1	1150.6215	1.56	y10 +1
1186.1173	4	1186.1161	1.01	b46 +4
1207.6439	1	1207.6430	0.75	y11 +1
1218.1411	4	1218.1399	0.99	b47 +4
1250.4062	4	1250.4005	4.56	b48 +4

1322.6715	1	1322.6699	1.21	y12 +1

3.2.3 *S. aureus*

3.2.3.1 SAOUHSC_01135

UniProt Q2FZA4, SAOUHSC_01135
Seq: fMEGLFNAIKDTVTAAINNDGAKLGTSIVSIVENGVGLLGKLFGF
m/z monoisotopic = 1136.8509, +4; MW_{obs} = 4521.39 Da (2.0 ppm)



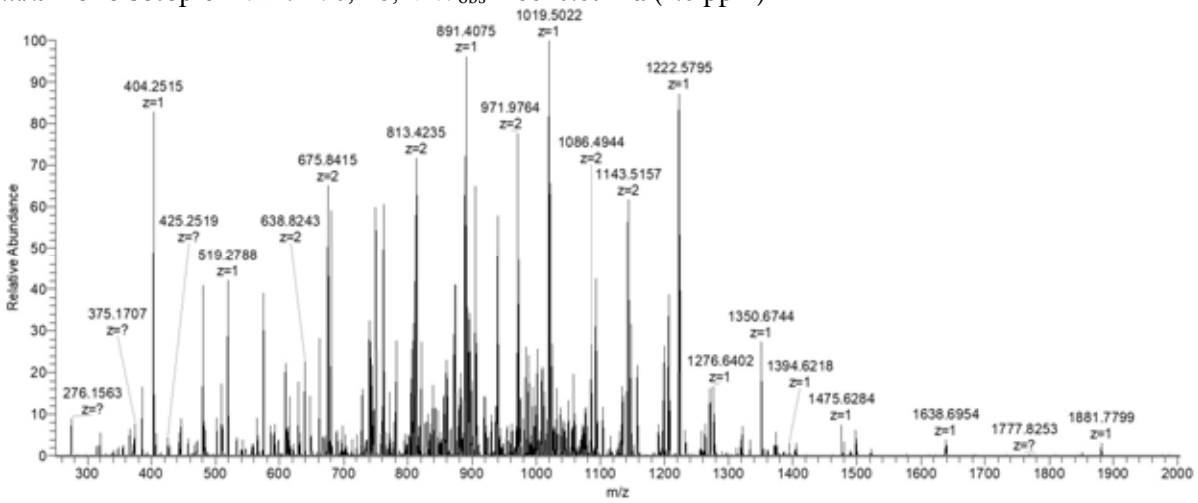
m/z meas	z	m/z calc	Δppm	assignment
346.1078	1	346.1067	-3.18	b3 +1
370.1771	1	370.1761	-2.70	y3 +1
459.1922	1	459.1908	-3.05	b4 +1
606.2608	1	606.2592	-2.64	b5 +1
611.3568	1	611.3552	-2.62	y5 +1
615.7965	2	615.7948	-2.76	b11-H2O +2
624.8017	2	624.8001	-2.56	b11 +2
665.3308	2	665.3290	-2.71	b12-H2O +2
668.3784	1	668.3766	-2.69	y6 +1
674.3360	2	674.3343	-2.52	b12 +2
715.8547	2	715.8529	-2.51	b13-H2O +2
720.3039	1	720.3021	-2.50	b6 +1
724.8600	2	724.8581	-2.62	b13 +2
751.3733	2	751.3714	-2.53	b14-H2O +2
781.4626	1	781.4607	-2.43	y7 +1
791.3413	1	791.3393	-2.53	b7 +1
795.8971	2	795.8952	-2.39	b15 +2
825.4683	2	825.4667	-1.94	y16 +2
826.1459	3	826.1440	-2.30	y25 +3

894.5471	1	894.5448	-2.57	y8 +1
902.5027	3	902.5006	-2.33	y27 +3
904.4255	1	904.4233	-2.43	b8 +1
940.5157	3	940.5149	-0.85	y28 +3
951.5685	1	951.5662	-2.42	y9 +1
978.2084	3	978.2096	1.23	y29 +3
1001.8905	3	1001.8886	-1.89	y30 +3
1016.8093	4	1016.8071	-2.13	y40 +4
1023.9958	2	1023.9937	-2.05	b19 +2
1034.5603	4	1034.5581	-2.10	b41-H2O +4
1039.0631	4	1039.0608	-2.21	b41 +4
1045.5833	2	1045.5802	-3.00	y21-NH3 +2
1050.6367	1	1050.6346	-2.00	y10 +1
1054.0944	2	1054.0935	-0.89	y21 +2
1059.3360	4	1059.3335	-2.36	y42 +4
1068.8318	4	1068.8292	-2.40	a42 +4
1071.3239	4	1071.3253	1.34	b42-H2O +4
1075.8307	4	1075.8279	-2.57	b42 +4
1107.6586	1	1107.6561	-2.26	y11 +1
1110.6375	2	1110.6356	-1.74	y22 +2
1152.0736	2	1152.0704	-2.81	b22 +2
1158.2599	3	1158.2591	-0.68	b34 +3
1164.2979	3	1164.2979	0.01	y35 +3
1174.6850	2	1174.6830	-1.73	y23 +2
1181.9530	3	1181.9503	-2.28	a35 +3
1191.2871	3	1191.2819	-4.36	b35 +3
1206.9983	3	1206.9962	-1.73	y36 +3
1221.7009	1	1221.6990	-1.56	y12 +1
1238.7129	2	1238.7123	-0.52	y25 +2
1244.6938	3	1244.6909	-2.32	y37 +3
1247.9863	3	1247.9838	-2.00	b37 +3
1248.5950	1	1248.5929	-1.68	b11 +1
1268.3724	3	1268.3700	-1.89	y38 +3
1285.6823	3	1285.6785	-2.95	b38 +3
1296.2278	2	1296.2258	-1.57	y26 +2
1306.3874	3	1306.3843	-2.37	y39 +3
1347.3845	3	1347.3839	-0.44	b40 +3
1347.6636	1	1347.6613	-1.71	b12 +1
1350.7443	1	1350.7416	-2.00	y13 +1
1353.2496	2	1353.2473	-1.73	y27 +2
1355.4120	3	1355.4071	-3.61	y40 +3
1379.0795	3	1379.0751	-3.18	b41-H2O +3
1385.0808	3	1385.0786	-1.58	b41 +3

1401.7590	2	1401.7554	-2.60	y28-NH3 +2
1410.2708	2	1410.2687	-1.52	y28 +2
1448.7128	1	1448.7090	-2.62	b13 +1
1449.8123	1	1449.8100	-1.59	y14 +1
1458.3034	2	1458.2975	-4.07	y29-NH3 +2
1466.8129	2	1466.8108	-1.46	y29 +2
1502.3315	2	1502.3293	-1.49	y30 +2
1519.7485	1	1519.7461	-1.58	b14 +1
1537.8516	2	1537.8479	-2.43	y31 +2
1562.8983	1	1562.8941	-2.69	y15 +1
1586.8387	2	1586.8316	-4.50	b31 +2
1588.3757	2	1588.3717	-2.54	y32 +2
1590.7851	1	1590.7832	-1.19	b15 +1
1649.9309	1	1649.9261	-2.91	y16 +1
1703.8703	1	1703.8673	-1.76	b16 +1
1748.9995	1	1748.9945	-2.86	y17 +1
1817.9147	1	1817.9102	-2.48	b17 +1
1931.9575	1	1931.9531	-2.28	b18 +1

3.2.3.2 SAOUHSC_01729

UniProt Q2FXV2, SAOUHSC_01729
Seq: SFMDKAKDAVEKFKNSDNEQVKNVKDKINEYTGSNNEEKKENEDKEK
m/z monoisotopic = 921.1190, +6; MW_{obs} = 5520.67 Da (2.0 ppm)



m/z meas	z	m/z calc	Δppm	assignment
276.1563	1	276.1554	-3.26	y2 +1
348.1386	1	348.1376	-2.87	b3-H2O +1
366.1492	1	366.1482	-2.73	b3 +1
373.1728	2	373.1718	-2.68	y6-NH3 +2

386.2409	1	386.2398	-2.85	y3-H2O +1
404.2515	1	404.2504	-2.72	y3 +1
446.2080	1	446.2063	-3.81	y7 +2
463.1660	1	463.1646	-3.02	b4-H2O +1
481.1765	1	481.1751	-2.91	b4 +1
501.2498	2	501.2485	-2.59	y8-H2O +2
510.2553	2	510.2538	-2.94	y8 +2
519.2788	1	519.2773	-2.89	y4 +1
533.2748	2	533.2735	-2.44	a10 +2
547.2722	2	547.2710	-2.19	b10 +2
565.2974	2	565.2960	-2.48	y9-H2O +2
574.3030	2	574.3013	-2.96	y9 +2
591.2609	1	591.2595	-2.37	b5-H2O +1
609.2718	1	609.2701	-2.79	b5 +1
611.7937	2	611.7923	-2.29	b11 +2
629.8190	2	629.8173	-2.70	y10-H2O +2
638.8243	2	638.8226	-2.66	y10 +2
648.3214	1	648.3199	-2.31	y5 +1
652.3143	1	652.3123	-3.07	a6 +1
662.2985	1	662.2967	-2.72	b6-H2O +1
666.8361	2	666.8345	-2.40	b12-H2O +2
675.8415	2	675.8397	-2.66	b12 +2
680.3091	1	680.3072	-2.79	b6 +1
729.0142	3	729.0125	-2.33	b19 +3
740.3706	2	740.3687	-2.57	b13-H2O +2
744.3542	1	744.3523	-2.55	y6-H2O +1
746.3639	4	746.3626	-1.74	y25 +4
749.3756	2	749.3740	-2.14	b13 +2
760.3677	2	760.3654	-3.02	y12 +2
762.3646	1	762.3628	-2.36	y6 +1
765.6966	3	765.6952	-1.83	b20-H2O +3
771.7006	3	771.6987	-2.46	b20 +3
777.3978	4	777.3961	-2.19	b27-H2O +4
778.3882	4	778.3864	-2.31	y26 +4
781.9008	4	781.8988	-2.56	b27 +4
798.7194	3	798.7180	-1.75	b21-H2O +3
800.0402	3	800.0383	-2.37	y20 +3
805.6691	4	805.6671	-2.48	b28-H2O +4
808.4043	1	808.4022	-2.60	b7 +1
810.1718	4	810.1698	-2.47	b28 +4
813.4235	2	813.4214	-2.58	b14 +2
817.3886	2	817.3868	-2.20	y13 +2
830.6677	4	830.6655	-2.65	y28-H2O +4

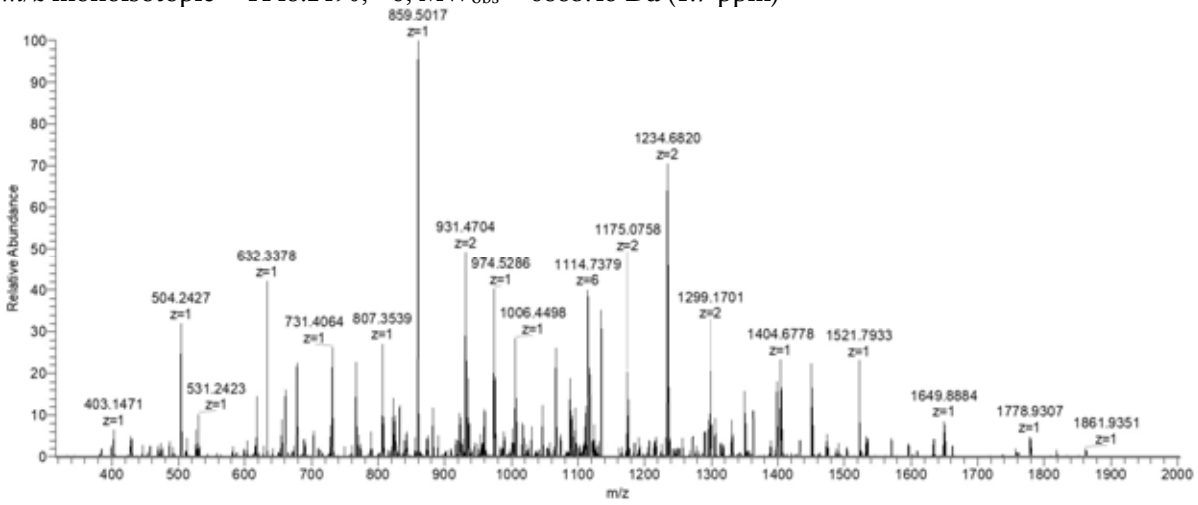
834.1783	4	834.1779	-0.48	b29-H2O +4
838.6825	4	838.6805	-2.38	b29 +4
841.0915	6	841.0891	-2.85	y43 +6
842.7387	3	842.7367	-2.37	y21 +3
847.4219	3	847.4198	-2.48	b22 +3
860.8271	5	860.8251	-2.32	y36 +5
860.9050	2	860.9028	-2.56	y14 +2
866.4408	4	866.4385	-2.65	b30-H2O +4
870.9444	4	870.9411	-3.79	b30 +4
873.3971	1	873.3949	-2.52	y7-H2O +1
875.2615	6	875.2591	-2.74	b45 +6
879.1002	6	879.0986	-1.82	y45-H2O +6
880.4106	2	880.4083	-2.61	y15-H2O +2
882.1023	6	882.1004	-2.15	y45 +6
885.4360	3	885.4341	-2.15	b23 +3
889.4157	2	889.4136	-2.36	y15 +2
891.4075	1	891.4054	-2.36	y7 +1
893.7665	6	893.7644	-2.35	b46-H2O +6
895.9414	4	895.9395	-2.12	y30 +4
896.7679	6	896.7662	-1.90	b46 +6
899.7698	6	899.7680	-2.00	b46+H2O +6
905.4208	1	905.4186	-2.43	b8-H2O +1
907.2067	4	907.2043	-2.65	b31-H2O +4
911.7093	4	911.7070	-2.52	b31 +4
930.9345	2	930.9321	-2.58	y16-H2O +2
936.9704	4	936.9689	-1.60	b32 +4
939.9398	2	939.9374	-2.55	y16 +2
943.6620	5	943.6601	-2.01	y40 +5
946.4562	4	946.4542	-2.11	y32 +4
952.8694	5	952.8680	-1.47	b41 +5
956.8026	3	956.8001	-2.61	y24 +3
961.1573	3	961.1553	-2.08	b25 +3
962.4700	2	962.4671	-3.01	b17-H2O +2
969.2812	5	969.2791	-2.17	y41 +5
971.4752	2	971.4724	-2.88	b17 +2
974.9678	4	974.9650	-2.87	y33 +4
979.8854	5	979.8844	-1.02	y42-H2O +5
983.4860	5	983.4865	0.51	y42 +5
988.8130	3	988.8109	-2.12	y25-H2O +3
993.4937	3	993.4941	0.40	b26-H2O +3
994.4687	1	994.4662	-2.51	b9 +1
994.8161	3	994.8144	-1.71	y25 +3
999.5001	3	999.4976	-2.50	b26 +3

1001.4922	1	1001.4898	-2.40	y8-H2O +1
1005.7046	5	1005.7002	-4.38	y43-NH3 +5
1006.9906	4	1006.9887	-1.89	y34 +4
1009.1077	5	1009.1055	-2.18	y43 +5
1012.4660	2	1012.4638	-2.17	y17-H2O +2
1019.5022	1	1019.5004	-1.77	y8 +1
1021.0902	5	1021.0852	-4.90	b44-NH3 +2
1021.4712	2	1021.4691	-2.06	y17 +2
1024.4924	5	1024.4905	-1.85	b44 +5
1028.4965	2	1028.4938	-2.63	b18 +2
1032.1131	5	1032.1109	-2.13	y44 +5
1037.5149	3	1037.5127	-2.12	y26 +3
1039.5033	4	1039.4992	-3.94	y35-NH3 +4
1042.1982	3	1042.1959	-2.21	b27 +3
1043.7586	4	1043.7558	-2.68	y35 +4
1050.1119	5	1050.1095	-2.29	b45 +5
1054.7181	5	1054.7169	-1.14	y45-H2O +5
1058.3209	5	1058.3190	-1.80	y45 +5
1065.5399	1	1065.5397	-0.19	a10 +1
1070.5385	3	1070.5355	-2.80	y27 +3
1072.3177	5	1072.3159	-1.68	b46-H2O +5
1073.8877	3	1073.8871	-0.56	b28-H2O +3
1075.9208	5	1075.9180	-2.60	b46 +5
1076.9874	2	1076.9851	-2.14	y18-H2O +2
1084.0125	2	1084.0099	-2.40	b19-H2O +2
1085.9926	2	1085.9904	-2.03	y18 +2
1093.0178	2	1093.0151	-2.47	b19 +2
1129.5881	1	1129.5848	-2.92	y9-H2O +1
1134.0091	2	1134.0066	-2.20	y19-H2O +2
1143.0144	2	1143.0118	-2.27	y19 +2
1147.5983	2	1147.5953	-2.61	y9 +1
1157.0468	2	1157.0444	-2.07	b20 +2
1190.5520	2	1190.5486	-2.86	y20-H2O +2
1197.5761	2	1197.5734	-2.25	b21-H2O +2
1199.5556	2	1199.5539	-1.42	y20 +2
1204.5697	1	1204.5667	-2.49	b11-H2O +1
1206.5812	2	1206.5786	-2.15	b21 +2
1222.5795	1	1222.5773	-1.80	b11 +1
1258.6306	1	1258.6274	-2.54	y10-H2O +1
1263.6054	2	1263.6014	-3.17	y21 +2
1270.6285	2	1270.6261	-1.89	b22 +2
1276.6402	1	1276.6379	-1.80	y10 +1
1312.1128	2	1312.1095	-2.52	y22-H2O +2

1321.1185	2	1321.1148	-2.80	y22 +2
1350.6744	1	1350.6722	-1.63	b12 +1
1385.1652	2	1385.1623	-2.09	y23 +2
1405.6849	1	1405.6805	-3.13	y11 +1
1479.7344	1	1479.7301	-2.91	b13-H2O +1
1497.7454	1	1497.7406	-3.20	b13 +1
1519.7276	1	1519.7235	-2.70	y12 +1

3.2.3.3 SAOUHSC_00845

UniProt Q2FZY9, UPF0337 protein SAOUHSC_00845
Seq:
ADESKFEQAKGNVKETVGNVTDNKNLENEGKEDKASGKAKEFVENAKEKATDFIDKVKGNKGE
m/z monoisotopic = 1148.2490, +6; MW_{obs} = 6883.45 Da (1.7 ppm)



m/z meas	**z**	**m/z calc**	**Δppm**	**assignment**
385.1363	1	385.1354	-2.34	b4-H2O +1
403.1471	1	403.1460	-2.73	b4 +1
430.2544	2	430.2534	-2.32	y8 +2
447.2208	1	447.2198	-2.24	y4 +1
487.2158	1	487.2147	-2.26	y5-NH3 +1
504.2427	1	504.2413	-2.78	y5 +1
513.2316	1	513.2304	-2.34	b5-H2O +1
531.2423	1	531.2409	-2.64	b5 +1
617.8446	2	617.8431	-2.43	y11 +2
632.3378	1	632.3362	-2.53	y6 +1
653.3087	2	653.3071	-2.45	b12 +2
660.3004	1	660.2988	-2.42	b6-H2O +1
678.3110	1	678.3093	-2.51	b6 +1
688.8454	2	688.8439	-2.18	a13 +2

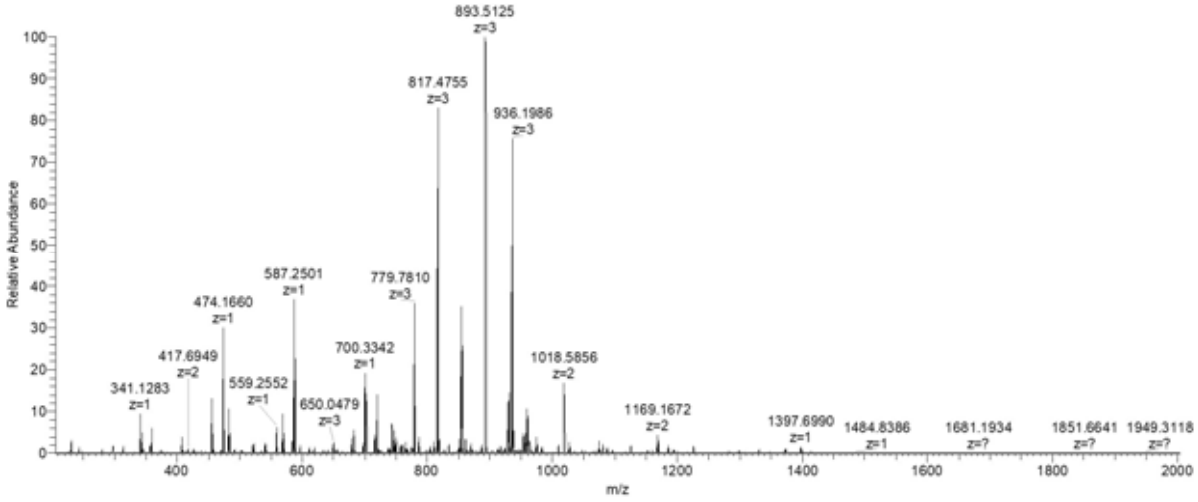
698.0401	3	698.0393	-1.15	y19 +3
702.8430	2	702.8413	-2.42	b13 +2
731.4064	1	731.4046	-2.46	y7 +1
757.8853	2	757.8835	-2.38	b14-H2O +2
766.8906	2	766.8888	-2.35	b14 +2
789.3432	1	789.3414	-2.28	b7-H2O +1
807.3539	1	807.3519	-2.48	b7 +1
822.4067	2	822.4048	-2.31	b15-H2O +2
825.4483	2	825.4465	-2.18	y15 +2
831.4120	2	831.4101	-2.29	b15 +2
841.4909	1	841.4890	-2.26	y8-H2O +1
859.5017	1	859.4996	-2.44	y8 +1
864.0992	3	864.0963	-3.36	b24 +3
872.9305	2	872.9287	-2.06	b16-H2O +2
880.9650	2	880.9625	-2.84	y16-H2O +2
881.9362	2	881.9339	-2.61	b16 +2
889.9697	2	889.9678	-2.13	y16 +2
896.1085	3	896.1071	-1.56	b25-H2O +3
902.1123	3	902.1106	-1.88	b25 +3
917.4730	2	917.4707	-2.51	a17 +2
922.4651	2	922.4629	-2.38	b17-H2O +2
924.7381	4	924.7364	-1.84	y34 +4
931.4704	2	931.4682	-2.36	b17 +2
935.4127	1	935.4105	-2.35	y8 +1
939.8068	3	939.8053	-1.60	b26 +3
954.0173	2	954.0153	-2.10	y17 +2
956.5183	1	956.5160	-2.40	y9-H2O +1
957.5018	1	957.5000	-1.88	y9-NH3 +1
959.9813	2	959.9789	-2.50	b18 +2
974.5286	1	974.5265	-2.15	y9 +1
982.8211	3	982.8195	-1.63	b27 +3
985.5118	4	985.5078	-4.06	y36 +4
988.4393	1	988.4371	-2.23	b9-H2O +1
989.5358	2	989.5338	-2.02	y18 +2
994.1969	3	994.1961	-0.80	y27 +3
1006.4498	1	1006.4476	-2.19	b9 +1
1017.0020	2	1017.0004	-1.57	b19 +2
1023.2084	3	1023.2068	-1.56	y28 +3
1023.2560	4	1023.2544	-1.56	b38 +4
1041.0152	4	1041.0137	-1.44	b39 +4
1046.5583	2	1046.5553	-2.87	y19 +2
1052.5394	2	1052.5371	-2.19	a20 +2
1063.8499	3	1063.8480	-1.79	b29 +3

1066.5369	2	1066.5346	-2.16	b20 +2
1073.0387	4	1073.0375	-1.12	b40 +4
1074.5500	4	1074.5502	0.19	y39 +4
1087.6125	1	1087.6106	-1.75	y10 +1
1089.5865	3	1089.5842	-2.11	y30 +3
1092.8951	6	1092.8915	-3.29	y60-NH3 +6
1095.7313	6	1095.7293	-1.83	y60 +6
1096.9631	5	1096.9614	-1.55	y50 +5
1102.0734	2	1102.0713	-1.91	y20-H2O +2
1111.0792	2	1111.0766	-2.34	y20 +2
1111.2342	6	1111.2329	-1.17	b61-H2O +6
1114.2365	6	1114.2346	-1.71	b61 +6
1117.2386	6	1117.2364	-1.97	b61+H2O +6
1120.9053	6	1120.9005	-4.28	b62-NH3 +6
1123.7405	6	1123.7382	-2.05	b62 +6
1130.9641	5	1130.9607	-3.01	b52 +5
1134.5450	1	1134.5426	-2.12	b10 +1
1135.0862	4	1135.0847	-1.32	y41 +4
1160.6134	2	1160.6108	-2.24	y21 +2
1165.5689	2	1165.5666	-1.97	b22-H2O +2
1170.9447	3	1170.9407	-3.42	y32 +3
1173.5564	1	1173.5535	-2.47	b11-H2O +1
1174.5749	2	1174.5719	-2.55	b22 +2
1190.8166	5	1190.8144	-1.85	y55 +5
1191.5671	1	1191.5640	-2.60	b11 +1
1205.9999	5	1205.9966	-2.74	b55 +5
1206.9131	3	1206.9100	-2.57	b33 +3
1213.0260	5	1213.0208	-4.29	y56 +5
1216.6705	1	1216.6684	-1.73	y11-H2O +1
1225.1417	2	1225.1397	-1.63	y22-H2O +2
1231.5955	2	1231.5933	-1.79	b23 +2
1232.6474	3	1232.6461	-1.05	y34 +3
1234.1473	2	1234.1450	-1.86	y22 +2
1234.6820	1	1234.6790	-2.43	y11 +1
1252.3852	4	1252.3799	-4.23	y46-NH3 +4
1256.6389	4	1256.6365	-1.91	y46 +4
1268.2460	5	1268.2430	-2.37	y58-NH3 +5
1271.6490	5	1271.6483	-0.55	y58 +5
1277.0528	5	1277.0482	-3.60	b58 +5
1289.6629	2	1289.6610	-1.47	y23-H2O +2
1295.6430	2	1295.6408	-1.70	b24 +2
1298.6684	2	1298.6663	-1.62	y23 +2
1305.6094	1	1305.6070	-1.84	b12 +1

1331.6978	1	1331.6954	-1.80	y12-H2O +1
1338.9250	4	1338.9262	0.90	y49 +4
1349.7082	1	1349.7060	-1.63	y12 +1
1352.6655	2	1352.6623	-2.37	b25 +2
1356.6937	3	1356.6889	-3.54	y37 +3
1362.7152	2	1362.7138	-1.03	y24 +2
1386.6670	1	1386.6648	-1.59	b13-H2O +1
1389.7256	2	1389.7191	-4.68	y25-NH3 +2
1398.2347	2	1398.2323	-1.72	y25 +2
1404.6778	1	1404.6754	-1.71	b13 +1
1432.3988	3	1432.3979	-0.63	y39 +3
1450.7562	1	1450.7536	-1.79	y13 +1
1473.7308	2	1473.7256	-3.53	b27 +2
1490.7940	2	1490.7906	-2.28	y27 +2
1503.7824	1	1503.7802	-1.46	y14-H2O +1
1521.7933	1	1521.7907	-1.71	y14 +1
1530.7517	2	1530.7471	-3.01	b28 +2
1532.7728	1	1532.7703	-1.63	b14 +1
1534.3100	2	1534.3066	-2.22	y28 +2
1569.8287	2	1569.8251	-2.29	y29 +2
1595.2729	2	1595.2684	-2.82	b29 +2
1633.8773	2	1633.8726	-2.88	y30 +2
1643.8076	1	1643.8024	-3.16	b15-H2O +1
1649.8884	1	1649.8857	-1.64	y15 +1
1661.8201	1	1661.8129	-4.33	b15 +1
1691.3880	2	1691.3861	-1.12	y31 +2
1760.9224	1	1760.9177	-2.67	y16-H2O +1
1778.9307	1	1778.9283	-1.35	y16 +1
1861.9351	1	1861.9290	-3.28	b17 +1

3.2.3.4 δ-haemolysin

UniProt Q6G7S2, δ-haemolysin
Seq: fMAQDIISTISDLVKWIIDTVNKFTKK
m/z monoisotopic = 759.6682, +4; MW_{obs} = 3034.64 Da (0.8 ppm)



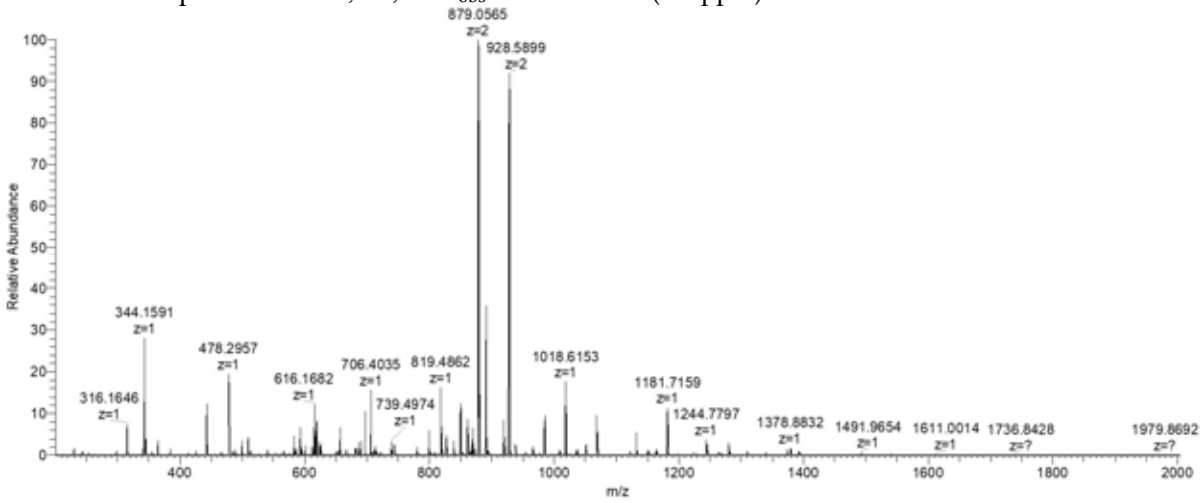
m/z meas	z	m/z calc	Δppm	assignment
231.0802	1	231.0798	-1.73	b2 +1
275.2083	1	275.2078	-1.82	y2 +1
359.1389	1	359.1384	-1.39	b3 +1
376.2560	1	376.2554	-1.59	y3 +1
383.2350	2	383.2345	-1.30	y6 +2
456.1554	1	456.1547	-1.53	b4-H2O +1
474.1660	1	474.1653	-1.48	b4 +1
483.2933	2	483.2926	-1.45	y8 +2
432.7694	2	432.7687	-1.62	y7 +2
523.3247	1	523.3239	-1.53	y4 +1
540.8067	2	540.8060	-1.29	y9 +2
559.2552	1	559.2545	-1.25	a5 +1
569.2395	1	569.2388	-1.23	b5-H2O +1
587.2501	1	587.2494	-1.19	b5 +1
611.7054	3	611.7047	-1.14	y15 +3
650.0479	3	650.0471	-1.23	y16 +3
679.0586	3	679.0577	-1.33	y17 +3
682.3238	1	682.3229	-1.32	b6-H2O +1
651.4197	1	651.4188	-1.38	y5 +1
697.6476	4	697.6468	-1.15	y24-H2O +4
700.3342	1	700.3334	-1.14	b6 +1
702.1504	4	702.1495	-1.28	y24 +4
715.6528	4	715.6521	-0.98	y25-NH3 +4
716.7533	3	716.7524	-1.26	y18 +3
719.9095	4	719.9088	-0.97	y25 +4
723.1419	4	723.1412	-0.97	b25 +4
737.9254	2	737.9245	-1.22	y12-H2O +2
744.4325	3	744.4315	-1.34	y19-H2O +3

746.9305	4	746.9298	-0.94	y12 +2
750.4357	3	750.4350	-0.93	y19 +3
765.4628	1	765.4618	-1.31	y6 +1
769.3561	1	769.3549	-1.56	b7-H2O +1
773.4429	3	773.4421	-1.03	y20-H2O +3
779.4465	3	779.4457	-1.03	y20 +3
787.3663	1	787.3655	-1.02	b7 +1
801.9727	2	801.9720	-0.87	y13-H2O +2
810.9781	2	810.9772	-1.11	y13 +2
817.1413	3	817.1403	-1.22	y21 +3
854.8358	3	854.8350	-0.94	y22 +3
860.5125	2	860.5114	-1.28	y14 +2
864.5313	2	864.5302	-1.27	y7 +1
870.4039	1	870.4026	-1.49	b8-H2O +1
893.1783	3	893.1773	-1.12	y23 +3
912.9832	2	912.9819	-1.42	b16-H2O +2
917.0541	2	917.0535	-0.65	y15 +2
929.8610	3	929.8600	-1.08	y24-H2O +3
935.8643	3	935.8635	-0.85	y24 +3
947.5685	1	947.5673	-1.27	y8-H2O +1
953.8680	3	953.8671	-0.94	y25-NH3 +3
957.8498	3	957.8490	-0.84	b25-H2O +3
915.1507	3	915.1507	0.00	b24-H2O +3
959.5435	3	959.5426	-0.94	y25 +3
963.8524	3	963.8525	0.10	b25 +3
965.5801	1	965.5778	-2.38	y8 +1
974.5678	2	974.5669	-0.92	y16 +2
983.4879	1	983.4866	-1.32	b9-H2O +1
1009.0786	2	1009.0777	-0.89	y17-H2O +2
1018.0838	2	1018.0830	-0.79	y17 +2
1001.4981	1	1001.4972	-0.90	b9 +1
1027.0385	2	1027.0374	-1.07	b18-H2O +2
1062.5955	1	1062.5942	-1.22	y9-H2O +1
1065.6208	2	1065.6197	-1.03	y18-H2O +2
1070.5196	1	1070.5187	-0.84	b10-H2O +1
1074.6258	2	1074.6250	-0.74	y18 +2
1080.6054	1	1080.6048	-0.56	y9 +1
1088.5302	1	1088.5292	-0.92	b10 +1
1116.1448	2	1116.1435	-1.16	y19-H2O +2
1125.1495	2	1125.1488	-0.62	y19 +2
1159.6602	2	1159.6596	-0.52	y20-H2O +2
1168.6657	2	1168.6648	-0.77	y20 +2
1185.5464	1	1185.5456	-0.67	b11-H2O +1

1193.6899	1	1193.6888	-0.92	y10 +1
1216.2025	2	1216.2016	-0.74	y21-H2O +2
1225.2075	2	1225.2069	-0.49	y21 +2
1281.7502	2	1281.7489	-1.01	y22 +2
1298.6304	1	1298.6297	-0.54	b12-H2O +1
1306.7746	1	1306.7729	-1.30	y11 +1
1397.6990	1	1397.6981	-0.64	b13-H2O +1
1415.7078	1	1415.7087	0.64	b13 +1
1492.8544	1	1492.8522	-1.47	y12 +1
1525.7951	1	1525.7931	-1.31	b14-H2O +1

3.2.3.5 PSMα4

UniProt P0C826, phenol soluble modulin α peptide 4
Seq: fMAIVGTIIKIIKAIIDIFAK
m/z monoisotopic = 733.7894, +3; MW_{obs} = 2198.35 Da (0.6 ppm)



m/z meas	**z**	**m/z calc**	**Δppm**	**assignment**
218.1499	1	218.1499	0.00	y2 +1
231.0797	1	231.0798	0.43	b2 +1
316.1687	1	316.1689	0.63	a3 +1
344.1636	1	344.1639	0.87	b3 +1
365.2181	1	365.2183	0.55	y3 +1
443.2320	1	443.2323	0.68	b4 +1
478.2861	2	478.2859	-0.42	b9 +2
478.3021	1	478.3024	0.63	y4 +1
500.2534	1	500.2537	0.60	b5 +1
509.8181	2	509.8184	0.59	y9 +2
566.3600	2	566.3604	0.71	y10 +2
567.3738	3	567.3743	0.88	y15 +3

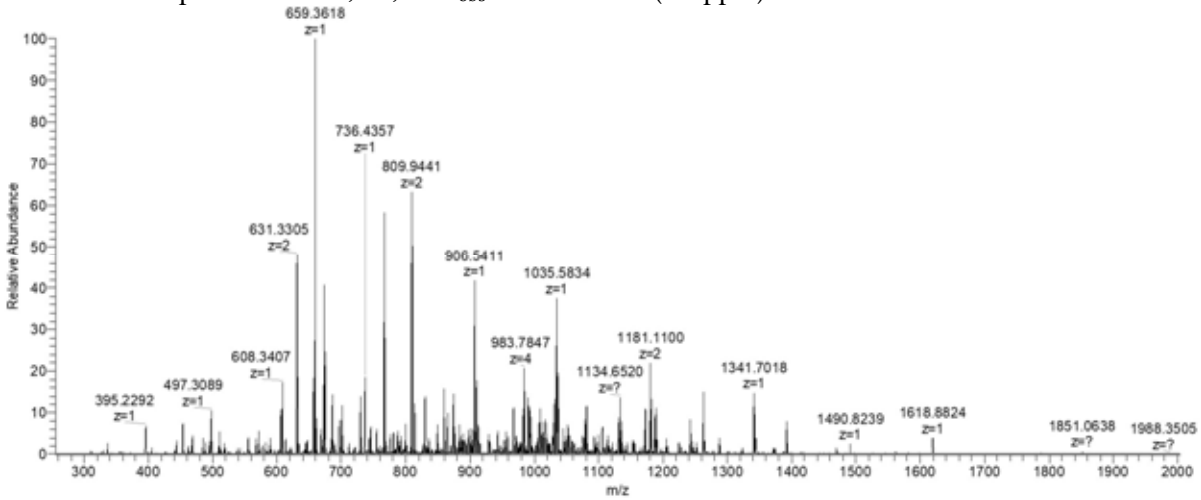
583.2904	1	583.2908	0.69	b6-H2O +1
586.3810	3	586.3815	0.85	y16 +3
593.3291	1	593.3293	0.34	y5 +1
601.3010	1	601.3014	0.67	b6 +1
619.4038	3	619.4043	0.81	y17 +3
622.9020	2	622.9025	0.80	y11 +2
651.0950	3	651.0955	0.77	y18-H2O +3
657.0986	3	657.0990	0.61	y18 +3
676.9381	2	676.9385	0.59	a13 +2
680.7774	3	680.7780	0.88	y19 +3
681.9301	2	681.9307	0.88	b13-H2O +2
686.9495	2	686.9500	0.73	y12 +2
690.9356	2	690.9360	0.58	b13 +2
696.3746	1	696.3749	0.43	b7-H2O +1
706.4129	1	706.4134	0.71	y6 +1
714.3850	1	714.3855	0.70	b7 +1
743.4914	2	743.4920	0.81	y13 +2
747.4777	2	747.4780	0.40	b14 +2
800.0336	2	800.0340	0.50	y14 +2
819.4972	1	819.4975	0.37	y7 +1
827.4690	1	827.4695	0.60	b8 +1
850.5573	2	850.5579	0.71	y15 +2
861.5331	2	861.5335	0.46	b16 +2
870.0629	2	870.0633	0.46	y16-H2O +2
879.0682	2	879.0686	0.46	y16 +2
890.5342	1	890.5346	0.45	y8 +1
918.0749	2	918.0756	0.76	b17 +2
919.5968	2	919.5975	0.76	y17-H2O +2
928.6024	2	928.6028	0.43	y17 +2
955.5638	1	955.5645	0.73	b9 +1
976.1388	2	976.1395	0.72	y18-H2O +2
985.1440	2	985.1448	0.81	y18 +2
991.6091	2	991.6098	0.71	b18 +2
1018.6291	1	1018.6295	0.39	y9 +1
1020.6629	2	1020.6634	0.49	y19 +2
1027.1277	2	1027.1283	0.58	b19 +2
1027.1277	2	1027.1283	0.58	b19 +2
1036.1326	2	1036.1336	0.97	b19+H2O +2
1068.6478	1	1068.6486	0.75	b10 +1
1113.7025	1	1113.7030	0.45	y10-H2O +1
1131.7129	1	1131.7136	0.62	y10 +1
1163.7213	1	1163.7221	0.69	b11-H2O +1
1181.7318	1	1181.7326	0.68	b11 +1

1226.7868	1	1226.7871	0.24	y11-H2O +1
1244.7975	1	1244.7977	0.16	y11 +1
1309.8273	1	1309.8276	0.23	b12 +1
1372.8923	1	1372.8926	0.22	y12 +1
1485.9762	1	1485.9767	0.34	y13 +1

3.2.4 *S. pneumoniae*

3.2.4.1 spr1385

UniProt Q8CYJ5, spr1385
Seq:
SLENKLEQATGAVKEGFGKVTGDSKTELEGAVEKTVAKAKDVVEDAKGAVEGAVEGLKNVFTK
m/z monoisotopic = 930.0688, +7; MW_{obs} = 6503.43 Da (0.3 ppm)



m/z meas	**z**	**m/z calc**	**Δppm**	**assignment**
330.1664	1	330.1660	1.21	b3 +1
395.2292	1	395.2289	0.76	y3 +1
444.2092	1	444.2089	0.68	b4 +1
453.7743	2	453.7740	0.66	y8 +2
494.2977	1	494.2973	0.81	y4 +1
507.2670	2	507.2667	0.59	b9 +2
509.2903	2	509.2900	0.59	y9-H2O +2
518.2958	2	518.2953	0.96	y9 +2
527.2829	1	527.2824	0.95	a5-NH3 +1
548.7855	2	548.7853	0.36	b10-H2O +2
554.2936	1	554.2933	0.54	b5-H2O +1
557.7907	2	557.7906	0.18	b10 +2
567.8299	2	567.8295	0.70	y10 +2
572.3042	1	572.3039	0.52	b5 +2

577.2962	2	577.2960	0.35	b11-H2O +2
582.9969	3	582.9965	0.69	y17 +3
586.3017	2	586.3013	0.68	b11 +2
603.3482	2	603.3481	0.17	y11 +2
606.6759	3	606.6755	0.66	y18 +3
621.8201	2	621.8199	0.32	b12 +2
631.8592	2	631.8588	0.63	y12 +2
645.0184	3	645.0178	0.93	y19 +3
685.3885	1	685.3879	0.88	b6 +1
718.4251	1	718.4246	0.70	y6-H2O +1
720.7270	3	720.7268	0.28	a21 +3
721.0557	3	721.0548	1.25	y21 +3
736.4357	1	736.4352	0.68	y6 +1
739.6578	4	739.6592	-1.89	y29-H2O +4
744.1623	4	744.1618	0.67	y29 +4
745.9147	2	745.9143	0.54	y14 +2
776.1860	4	776.1856	0.52	y30 +4
781.4333	2	781.4329	0.51	y15 +2
787.4082	3	787.4079	0.38	b23 +3
790.9184	2	790.9176	1.01	b15-H2O +2
792.4211	3	792.4199	1.51	y23 +3
799.9234	2	799.9229	0.63	b15 +2
808.4465	4	808.4462	0.37	y31 +4
809.9441	2	809.9436	0.62	y16 +2
814.4311	1	814.4305	0.74	b7 +1
826.2487	5	826.2505	-2.18	y40-H2O +5
829.8532	5	829.8526	0.72	y40 +5
833.2135	4	833.2133	0.24	y32 +4
835.1187	3	835.1183	0.48	y24 +3
840.1125	6	840.1140	-1.79	y49 +6
849.5198	1	849.5193	0.59	y7 +1
850.9725	4	850.9726	-0.12	y33 +4
853.1133	3	853.1134	-0.12	b25-H2O +3
858.7976	3	858.7973	0.35	y25 +3
859.1161	3	859.1169	-0.93	b25 +3
861.4608	6	861.4631	-2.67	y50 +6
865.2284	4	865.2280	0.46	y34 +4
869.6649	5	869.6645	0.46	b42 +5
873.9913	2	873.9911	0.23	y17 +2
877.9758	2	877.9745	1.48	y51 +6
880.6159	7	880.6135	2.73	y60-NH3 +7
883.0473	7	883.0458	1.70	y60 +7
884.4708	5	884.4718	-1.13	y43 +5

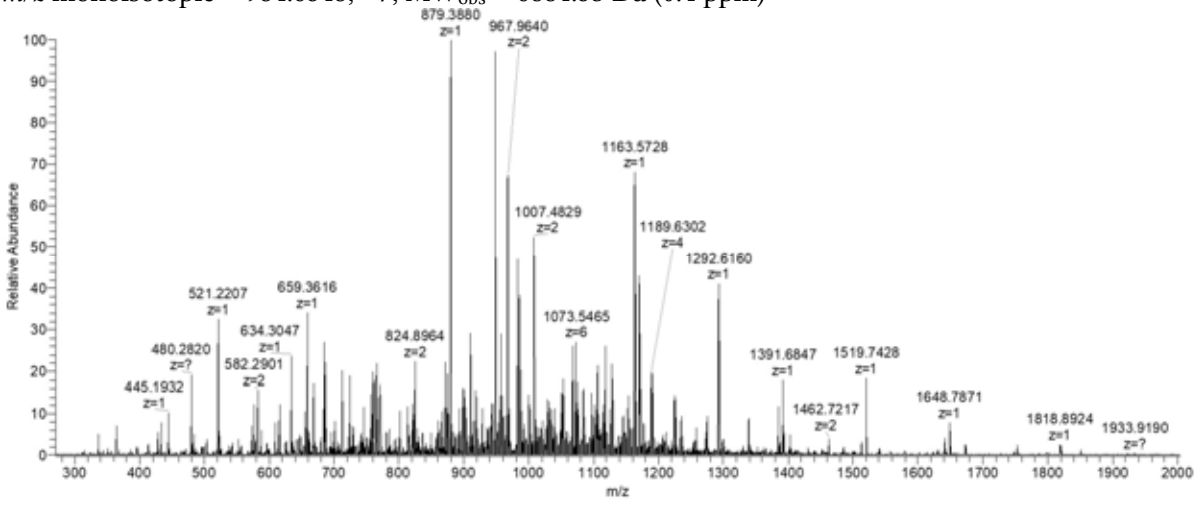
888.5303	1	888.5302	0.11	y8-H2O +1
889.4782	5	889.4782	0.00	b43 +5
892.7998	3	892.7995	0.34	b26 +3
895.9708	6	895.9713	-0.56	b53 +6
897.4895	4	897.4886	1.00	y35 +4
898.9088	7	898.9076	1.33	y61-H2O +7
901.4804	7	901.4805	-0.11	y61 +7
901.4961	3	901.4956	0.55	y26 +3
904.2856	5	904.2855	0.11	y44 +5
906.5411	1	906.5407	0.44	y8 +1
909.5101	2	909.5096	0.55	y18 +2
915.2882	5	915.2867	1.64	b44 +5
916.1582	6	916.1589	-0.76	y54 +6
942.4897	1	942.4891	0.64	b8 +1
946.5009	6	946.5038	-3.06	y56-NH3 +6
952.5000	5	952.4995	0.52	b46 +5
967.0236	2	967.0231	0.52	y19 +2
970.8484	6	970.8486	-0.21	y57 +6
973.5086	3	973.5084	0.21	b28 +3
978.1187	5	978.1185	0.20	b47 +5
982.1275	5	982.1268	0.71	y48 +5
983.2827	4	983.2822	0.51	y38 +4
986.8602	6	986.8582	2.03	y58-NH3 +6
989.6970	6	989.6960	1.01	y58 +6
991.8805	3	991.8800	0.50	y29 +3
994.5266	2	994.5260	0.60	b19 +2
1001.2763	4	1001.2812	-4.89	b39 +4
1003.7314	5	1003.7302	1.20	b49 +5
1004.5347	5	1004.5300	4.68	y49-NH3 +5
1007.9361	5	1007.9353	0.79	y49 +5
1011.0457	6	1011.0451	0.59	y59 +6
1013.5263	1	1013.5262	0.10	b9 +1
1015.3071	4	1015.3059	1.18	y39 +4
1016.5240	3	1016.5226	1.38	b29 +3
1017.5734	1	1017.5728	0.59	y9-H2O +1
1022.5405	2	1022.5391	1.37	y20-H2O +2
1027.2188	6	1027.2145	4.19	y60-NH3 +6
1030.0526	6	1030.0523	0.29	y60 +6
1031.5449	2	1031.5444	0.48	y20 +2
1033.5543	5	1033.5543	0.00	y50 +5
1035.5834	1	1035.5833	0.10	y9 +1
1037.0640	4	1037.0640	0.00	y40 +4
1044.0602	2	1044.0602	0.00	b20 +2

1048.5542	6	1048.5576	-3.24	y61-H2O +6
1051.5594	6	1051.5594	0.00	y61 +6
1053.3681	5	1053.3680	0.09	y51 +5
1057.5590	4	1057.5590	0.00	b41-H2O +4
1059.2097	3	1059.2088	0.85	b31 +3
1062.0630	4	1062.0617	1.22	b41 +4
1065.8171	4	1065.8207	-3.38	y41 +4
1067.5736	5	1067.5754	-1.69	y52 +5
1074.9655	5	1074.9641	1.30	b53 +5
1077.5936	3	1077.5925	1.02	y31 +3
1078.9795	5	1078.9797	-0.19	y53 +5
1080.0753	4	1080.0761	-0.74	y42 +4
1081.0788	2	1081.0786	0.19	y21 +2
1092.2317	3	1092.2316	0.09	b32 +3
1094.5846	5	1094.5840	0.55	b21 +2
1094.7798	5	1094.7778	1.83	b54 +5
1099.1899	5	1099.1892	0.64	y54 +5
1101.0855	4	1101.0813	3.81	y43-NH3 +4
1105.3385	4	1105.3380	0.45	y43 +4
1111.5925	4	1111.5959	-3.06	b43 +4
1113.3977	5	1113.3966	0.99	y55 +5
1114.5740	1	1114.5739	0.09	b10 +1
1120.5870	5	1120.5863	0.62	b55 +5
1123.0954	2	1123.0948	0.53	b22 +2
1130.1025	4	1130.1051	-2.30	y44 +4
1130.6118	4	1130.6128	-0.88	y22 +2
1134.6520	1	1134.6517	0.26	y10 +1
1135.2459	3	1135.2458	0.09	b33 +3
1139.0067	5	1139.0084	-1.49	y56 +5
1143.8568	4	1143.8565	0.26	b44 +4
1153.3030	3	1153.3015	1.30	y34 +3
1154.6129	5	1154.6074	4.76	b57 +5
1171.5981	1	1171.5953	2.39	b11 +1
1172.6136	4	1172.6133	0.26	b45 +4
1177.9454	3	1177.9441	1.10	b34 +3
1179.1208	4	1179.1172	3.05	a46-NH3 +4
1180.6082	2	1180.6083	-0.08	b23 +2
1188.1264	2	1188.1263	0.08	y23 +2
1190.3739	4	1190.3725	1.18	b46 +4
1196.3146	3	1196.3157	-0.92	y35 +3
1205.6898	1	1205.6888	0.83	y11 +1
1224.1245	2	1224.1243	0.16	b24 +2
1224.6246	1	1224.6219	2.20	b12 -H2O +1

1242.6331	1	1242.6325	0.48	b12 +1
1252.1757	2	1252.1738	1.52	y24 +2
1262.7109	1	1262.7103	0.48	y12 +1
1268.3339	3	1268.3285	4.26	b37 +3
1287.6947	2	1287.6923	1.86	y25 +2
1323.6919	1	1323.6903	1.21	b13-H2O +1
1341.7018	1	1341.7009	0.67	b13 +1
1387.2597	2	1387.2584	0.94	y27 +2
1391.7544	1	1391.7529	1.08	y13 +1
1469.7994	1	1469.7958	2.45	b14 +1
1490.8239	1	1490.8213	1.74	y14 +1
1561.8604	1	1561.8584	1.28	y15 +1
1580.8263	1	1580.8279	-1.01	b15-H2O +1
1598.8415	1	1598.8384	1.94	b15 +1
1618.8824	1	1618.8799	1.54	y16 +1

3.2.4.2 spr1626

UniProt Q8CYD7, spr1626
Seq:
SVEEKLNQAKGSIKEGVGKAIGDEKMEKEGAAEKVVSKVKEVAEDAKDAVEGAVEGVKNMLSGDD
K
m/z monoisotopic = 984.6548, +7; MW_{obs} = 6884.53 Da (0.4 ppm)



m/z meas	**z**	**m/z calc**	**Δppm**	**assignment**
316.1505	1	316.1503	0.63	b3 +1
377.1674	1	377.1667	1.86	y3 +1
434.1884	1	434.1882	0.46	y4 +1
445.1932	1	445.1929	0.67	b4 +1
521.2207	1	521.2202	0.96	y5 +1

573.2883	1	573.2879	0.70	b5 +1
582.2901	2	582.2899	0.34	y11 +2
634.3047	1	634.3042	0.79	y6 +1
646.8116	2	646.8112	0.62	y12 +2
661.7011	3	661.7005	0.91	b19 +3
668.3620	1	668.3614	0.90	b6-H2O +1
679.3777	3	679.3760	2.50	b20-H2O +3
685.3777	1	685.3795	-2.63	b20 +3
688.0091	3	688.0089	0.29	y20 +3
696.3458	2	696.3454	0.57	y13 +2
711.6883	3	711.6879	0.56	y21 +3
723.0749	3	723.0742	0.97	b21 +3
750.0306	3	750.0302	0.53	y22 +3
756.9228	2	756.9227	0.13	b14 +2
760.3751	2	760.3747	0.53	y15 +2
765.3451	1	765.3447	0.52	y7 +1
780.4244	3	780.4237	0.90	b23 +3
812.4389	2	812.4387	0.25	b15-H2O +2
815.8912	2	815.8907	0.61	y16-H2O +2
821.4446	2	821.4440	0.73	b15 +2
824.8964	2	824.8960	0.48	y16 +2
829.9330	4	829.9290	4.82	y32 +4
861.9500	4	861.9527	-3.13	y33 +4
866.1362	3	866.1362	0.00	b25 +3
874.4306	2	874.4302	0.46	y17 +2
879.3880	1	879.3877	0.34	y8 +1
884.4542	5	884.4544	-0.23	y42 +5
888.4823	5	888.4831	-0.90	b42 +5
893.2171	4	893.2174	-0.34	b34 +4
899.4881	2	899.4889	-0.89	b17 +2
902.6920	5	902.6905	1.66	b43 +5
909.9493	2	909.9487	0.66	y18 +2
910.2629	5	910.2629	0.00	y43 +5
913.4816	4	913.4819	-0.33	b35-H2O +4
917.9855	4	917.9845	1.09	b35 +4
921.0555	7	921.0561	-0.65	y62 +7
928.4738	1	928.4734	0.43	b8 +1
935.4592	3	935.4588	0.43	y27 +3
942.7527	4	942.7517	1.06	b36 +4
943.9876	4	943.9873	0.32	y37 +4
955.4961	7	955.4931	3.14	y64-NH3 +7
957.9247	7	957.9254	-0.73	y64 +7
960.6618	6	960.6621	-0.31	y56 +6

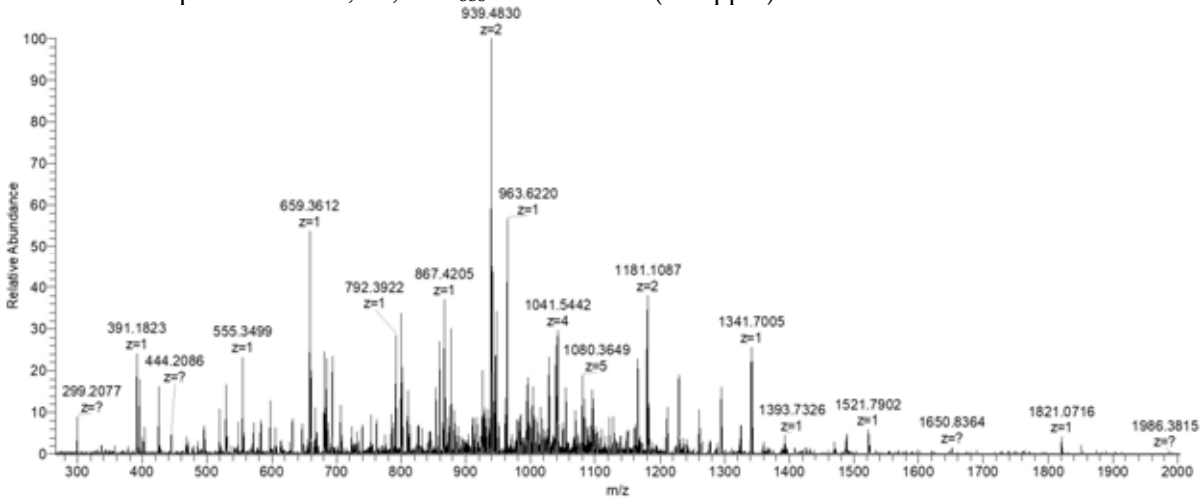
967.4625	2	967.4622	0.31	y19 +2
999.5106	1	999.5106	0.00	b9 +1
1007.4829	1	1007.4826	0.30	y9 +1
1021.3019	4	1021.3005	1.37	b39 +4
1027.5652	2	1027.5657	-0.49	b20 +2
1031.5101	2	1031.5097	0.39	y20 +2
1034.2037	6	1034.2011	2.51	y60 +6
1040.1911	3	1040.1906	0.48	y30 +3
1049.7405	5	1049.7381	2.29	y51 +5
1053.0502	6	1053.0484	1.71	y61 +6
1067.0283	2	1067.0283	0.00	y21 +2
1084.1088	2	1084.1077	1.01	b21 +2
1095.9046	6	1095.9047	-0.09	y63 +6
1099.7790	5	1099.7775	1.36	b53 +5
1104.9076	6	1104.9064	1.09	b64 +6
1106.5479	1	1106.5510	-2.80	y10 +1
1110.3517	4	1110.3520	-0.27	b42 +4
1114.5769	6	1114.5740	2.60	y64-NH3 +6
1117.4144	6	1117.4118	2.33	y64 +6
1119.5894	5	1119.5912	-1.61	b54 +5
1127.6055	1	1127.6055	0.00	b10 +1
1145.3998	5	1145.3997	0.09	b55 +5
1152.5905	5	1152.5931	-2.26	y56 +5
1156.8038	5	1156.8040	-0.17	b56 +5
1163.5728	1	1163.5725	0.26	y11 +1
1170.1321	2	1170.1319	0.17	b23 +2
1189.0635	2	1189.0630	0.42	y23 +2
1189.1292	4	1189.1287	0.42	b45 +4
1206.8905	4	1206.8879	2.15	b46 +4
1224.5828	2	1224.5816	0.98	y24 +2
1234.6548	2	1234.6532	1.30	b24 +2
1258.3135	3	1258.3140	-0.40	y37 +3
1274.1176	2	1274.1158	1.41	y25 +2
1292.6160	1	1292.6151	0.70	y12 +1
1338.6392	2	1338.6371	1.57	y26 +2
1384.7454	1	1384.7431	1.66	b13 +1
1391.6847	1	1391.6835	0.86	y13 +1
1402.6858	2	1402.6846	0.86	y27 +2
1462.7217	2	1462.7206	0.75	y14 +1
1512.8370	1	1512.8380	-0.66	b14 +1
1519.7428	1	1519.7421	0.46	y15 +1
1641.8868	1	1641.8806	3.78	b15 +1
1648.7871	1	1648.7847	1.46	y16 +1

1747.8544	1	1747.8531	0.74	y17 +1
1818.8924	1	1818.8902	1.21	y18 +1

3.2.5 *S. oralis*

3.2.5.1 HMPREF9950_0438

UniProt F9Q496, HMPREF9950_0438
Seq:
SLENKLDQATGAIKEGFGKVTGDSKTEAEGAVEKTVAKAKEVVEDAKGAVEGAVEGLKNSFKKED
m/z monoisotopic = 963.0736, +7; MW_{obs} = 6734.46 Da (-2.0 ppm)



m/z meas	z	m/z calc	Δppm	assignment
330.1660	1	330.1660	0.00	b3 +1
391.1823	1	391.1823	0.00	y3 +1
444.2086	1	444.2089	-0.68	b4 +1
519.2773	1	519.2773	0.00	y4 +1
572.3035	1	572.3039	-0.70	b5 +1
579.2934	2	579.2935	-0.17	b11 +2
583.3141	2	583.3142	-0.17	y10 +2
614.8117	2	614.8120	-0.49	b12 +2
666.3456	1	666.3457	-0.15	y5 +1
667.3770	1	667.3774	-0.60	b6-H2O +1
685.3878	1	685.3879	-0.15	b6 +1
690.3720	3	690.3724	-0.58	b20-H2O +3
693.0352	3	693.0357	-0.72	y20 +3
697.3694	2	697.3697	-0.43	y12 +2
731.3775	3	731.3780	-0.68	y21 +3
753.3769	1	753.3777	-1.06	y6 +1
761.3984	2	761.3990	-0.79	y14 +2

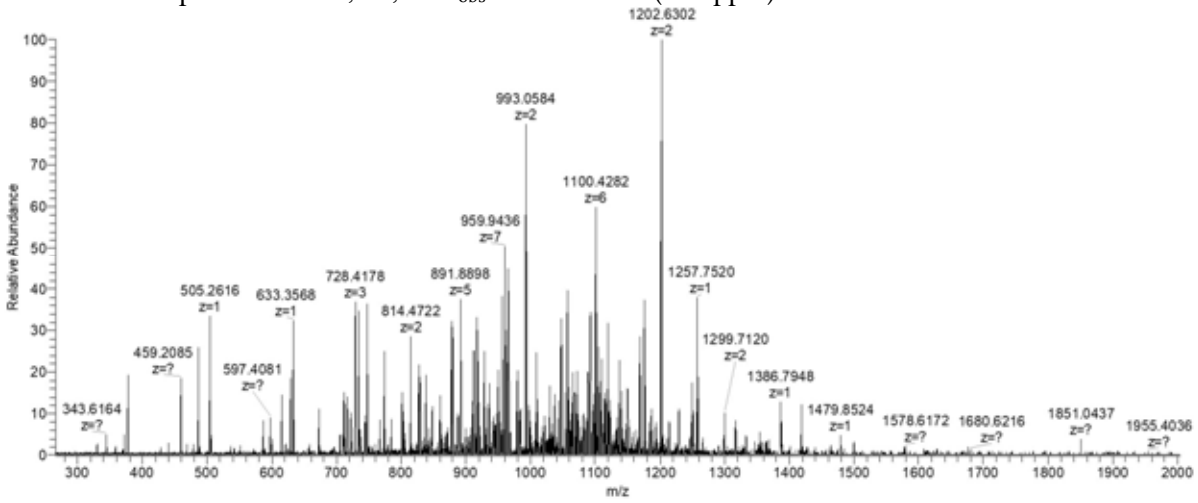
787.4076	3	787.4079	-0.38	b23 +3
799.9220	2	799.9229	-1.13	b15 +2
800.4135	1	800.4149	-1.75	b7 +1
812.4351	4	812.4359	-0.98	y31 +4
825.9199	2	825.9203	-0.48	y15 +2
840.4381	3	840.4378	0.36	y24 +3
844.4612	4	844.4596	1.89	y32 +4
859.1163	3	859.1169	-0.70	b25 +3
867.4205	1	867.4207	-0.23	y7 +1
875.4535	2	875.4545	-1.14	y16 +2
876.0618	5	876.0624	-0.68	y42 +5
883.4517	1	883.4520	-0.34	y25 +3
883.8689	5	883.8719	-3.39	b43 +5
892.7992	3	892.7995	-0.34	b26 +3
901.9675	1	901.9678	-0.33	b17 +2
910.9729	2	910.9731	-0.22	y17 +2
916.0528	7	916.0529	-0.11	y62 +7
923.2056	7	923.2032	2.60	b63-NH3 +7
925.6351	7	925.6355	-0.43	b63 +7
928.4731	1	928.4734	-0.32	b8 +1
930.4780	2	930.4785	-0.54	b18 +2
933.5009	4	933.5020	-1.18	y36 +4
934.4868	7	934.4875	-0.75	y63 +7
939.4830	2	939.4838	-0.85	y18 +2
941.6387	7	941.6378	0.96	b64-NH3 +7
944.0683	7	944.0702	-2.01	b64 +7
946.6427	7	946.6431	-0.42	b64+H2O +7
949.8287	3	949.8293	-0.63	y27 +3
977.5056	1	977.5051	0.51	y8-H2O +1
982.4822	1	982.4840	-1.83	b9-NH3 +1
983.5208	4	983.5219	-1.12	y38 +4
987.3503	6	987.3479	2.43	y58-NH3 +6
992.5261	3	992.5277	-1.61	y28 +3
994.5252	2	994.5260	-0.80	b19 +2
995.5140	1	995.5156	-1.61	y8 +1
998.1236	5	998.1239	-0.30	b49 +5
999.5101	1	999.5106	-0.50	b9 +1
1003.5310	2	1003.5313	-0.30	y19 +2
1009.3575	6	1009.3568	0.69	y59 +6
1015.7823	4	1015.7826	-0.30	y39 +4
1016.2066	3	1016.2067	-0.10	y29 +3
1025.2044	6	1025.2024	1.95	y60-H2O +6
1028.2036	6	1028.2041	-0.49	y60 +6

1030.0429	2	1030.0446	-1.65	y20-H2O +2
1039.0490	2	1039.0498	-0.77	y20 +2
1041.0435	4	1041.0445	-0.96	y40 +4
1044.0600	2	1044.0602	-0.19	b20 +2
1049.2291	3	1049.2295	-0.38	y30 +3
1050.7432	5	1050.7398	3.24	y51-NH3 +5
1054.1446	5	1054.1451	-0.47	y51 +5
1058.3910	6	1058.3911	-0.09	b62 +6
1068.5633	6	1068.5605	2.62	y62 +6
1079.7642	5	1079.7641	0.09	y52 +5
1082.9114	3	1082.9121	-0.65	y31 +3
1087.0670	6	1087.0658	1.10	y63-H2O +6
1090.0660	6	1090.0676	-1.47	y63 +6
1091.5751	1	1091.5732	1.74	y9-NH3 +1
1094.5832	2	1094.5840	-0.73	b21 +2
1094.8249	4	1094.8262	-1.19	y42 +4
1096.5628	2	1096.5633	-0.46	y21 +2
1100.5579	1	1100.5582	-0.27	b10 +1
1102.3800	5	1102.3809	-0.82	y53 +5
1104.5862	4	1104.5881	-1.72	b43 +4
1108.5994	1	1108.5997	-0.27	y9 +1
1121.2303	3	1121.2301	0.18	b33 +3
1125.6100	3	1125.6104	-0.36	y32 +3
1127.9882	5	1127.9926	-3.90	y55 +5
1148.2012	5	1148.2022	-0.87	y56 +5
1161.0836	2	1161.0846	-0.86	y22 +2
1165.6214	1	1165.6212	0.17	y10 +1
1180.6074	2	1180.6083	-0.76	b23 +2
1210.6188	2	1210.6188	0.00	y23 +2
1228.6166	1	1228.6168	-0.16	b12 +1
1260.1529	2	1260.1530	-0.08	y24 +2
1288.1718	2	1288.1717	0.08	b25 +2
1294.6636	1	1294.6638	-0.15	y11 +1
1324.6723	2	1324.6743	-1.51	y25 +2
1341.7005	1	1341.7009	-0.30	b13 +1
1388.7222	2	1388.7218	0.29	y26 +2
1393.7326	1	1393.7322	0.29	y12 +1
1424.2418	2	1424.2404	0.98	y27 +2
1464.7709	1	1464.7693	1.09	y13 +1
1469.7970	1	1469.7958	0.82	b14 +1
1521.7902	1	1521.7907	-0.33	y14 +1
1598.8404	1	1598.8384	1.25	b15 +1
1650.8364	1	1650.8333	1.88	y15 +1

3.2.6 *S. gordonii*

3.2.6.1 AWH02_03040

UniProt A0A0F5MKW3, AWH02_03040
Seq:
SLEEKFNQAKGSVKEGVGKLTGDTKTQAEGVAEKVASKAKEVAEDAKGAVEGAIEGVKNIIKKDDK
m/z monoisotopic = 980.2427, +7; MW_{obs} = 6854.65 Da (-0.6 ppm)



m/z meas	z	m/z calc	Δppm	assignment
330.1662	1	330.1660	0.61	b3 +1
359.1562	1	359.1561	0.28	y3-H2O +1
377.1667	1	377.1667	0.00	y3 +1
459.2085	1	459.2086	-0.22	b4 +1
487.2511	1	487.2511	0.00	y4-H2O +1
505.2616	1	505.2617	-0.20	y4 +1
551.3349	2	551.3350	-0.18	y9 +2
587.3036	1	587.3035	0.17	b5 +1
615.3461	1	615.3461	0.00	y5-H2O +1
620.3741	2	620.3746	-0.81	y11-H2O +2
629.3801	2	629.3799	0.32	y11 +2
633.3568	1	633.3566	0.32	y5 +1
673.0284	3	673.0286	-0.30	b19 +3
704.7395	3	704.7389	0.85	y20 +3
710.7235	3	710.7233	0.28	b20 +3
716.3613	1	716.3614	-0.14	b6-H2O +1
722.4143	3	722.4145	-0.28	y21-H2O +3
728.4178	3	728.4180	-0.27	y21 +3
734.3717	1	734.3719	-0.27	b6 +1

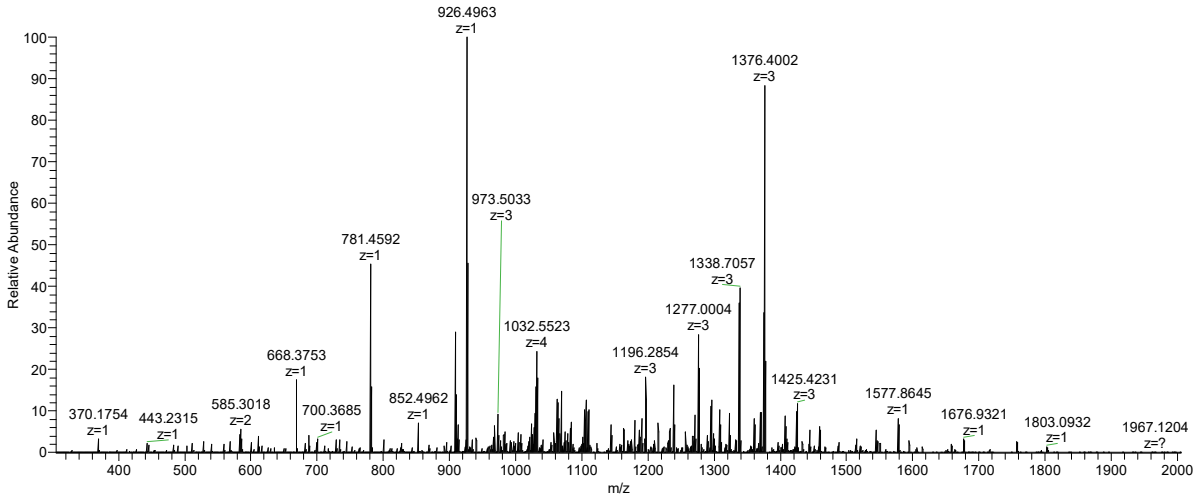
744.4054	3	744.4059	-0.67	b21 +3
746.4403	1	746.4407	-0.54	y6 +1
766.7601	3	766.7603	-0.26	y22 +3
773.9150	2	773.9148	0.26	b14 +2
801.7553	3	801.7553	0.00	b23 +3
814.4722	2	814.4725	-0.37	y15 +2
828.4638	4	828.4609	3.50	y32 +4
829.4308	2	829.4308	0.00	b15-H2O +2
835.4368	3	835.4379	-1.32	b24 +3
838.4364	2	838.4361	0.36	b15 +2
848.4149	1	848.4149	0.00	b7 +1
859.5248	1	859.5247	0.12	y7 +1
860.4851	4	860.4846	0.58	y33 +4
869.9886	2	869.9885	0.11	y16-H2O +2
878.1362	3	878.1362	0.00	b25 +3
887.8867	5	887.8867	0.00	y43-H2O +5
891.4897	5	891.4888	1.01	y43 +5
911.9832	4	911.9837	-0.55	b35 +4
916.4813	2	916.4811	0.22	b17 +2
928.5278	2	928.5280	-0.22	y17 +2
929.7430	4	929.7430	0.00	b36 +4
933.2225	7	933.2206	2.04	y63 +7
935.4904	5	935.4912	-0.86	b45 +5
942.9393	7	942.9386	0.74	b64 +7
947.6820	6	947.6829	-0.95	y56 +6
949.2247	7	949.2229	1.90	y64-NH3 +7
949.6985	5	949.6986	-0.11	b46 +5
951.6567	7	951.6552	1.58	y64 +7
955.5576	1	955.5571	0.52	y8-H2O +1
959.3716	7	959.3710	0.63	b65 +7
961.9443	7	961.9440	0.31	b65+H2O +7
964.0468	2	964.0466	0.21	y18 +2
992.5570	2	992.5573	-0.30	y19 +2
999.5462	4	999.5469	-0.70	y39 +4
1009.0387	2	1009.0393	-0.59	b19 +2
1042.7301	6	1042.7314	-1.25	y61-H2O +6
1045.7318	6	1045.7332	-1.34	y61 +6
1047.5103	1	1047.5106	-0.29	b9 +1
1047.5997	2	1047.5995	0.19	y20-H2O +2
1052.3207	4	1052.3209	-0.19	y41-H2O +4
1056.6041	2	1056.6048	-0.66	y20 +2
1059.1719	5	1059.1725	-0.57	y52-H2O +5
1062.7728	5	1062.7746	-1.69	y52 +5

1065.5826	2	1065.5813	1.22	b20 +2
1067.0832	6	1067.0823	0.84	y62 +6
1071.2555	3	1071.2559	-0.37	y31 +3
1072.1641	5	1072.1632	0.84	b53 +5
1080.7556	6	1080.7560	-0.37	b63 +6
1085.7541	6	1085.7517	2.21	y63-NH3 +6
1092.1227	2	1092.1233	-0.55	y21 +2
1096.9227	6	1096.9254	-2.46	b64-H2O +6
1099.9275	6	1099.9271	0.36	b64 +6
1107.0958	6	1107.0948	0.90	y64-H2O +6
1114.1070	4	1114.1092	-1.97	y43 +4
1119.0996	6	1119.0983	1.16	b65 +6
1131.9968	5	1131.9929	3.45	b56 +5
1137.0184	5	1137.0180	0.35	y56 +5
1139.9214	3	1139.9214	0.00	b33 +3
1149.6370	2	1149.6368	0.17	y22 +2
1151.8070	5	1151.8065	0.43	b57 +5
1169.1116	4	1169.1122	-0.51	b45 +4
1175.6052	1	1175.6055	-0.26	b10 +1
1186.8738	4	1186.8714	2.02	b46 +4
1193.1228	2	1193.1241	-1.09	b23-H2O +2
1202.1287	2	1202.1294	-0.58	b23 +2
1214.1569	2	1214.1581	-0.99	y23 +2
1249.6772	2	1249.6767	0.40	y24 +2
1257.7520	1	1257.7525	-0.40	y11 +1
1299.2110	2	1299.2109	0.08	y25 +2
1386.7948	1	1386.7951	-0.22	y12 +1
1418.7276	1	1418.7274	0.14	b13 +1
1499.8793	1	1499.8792	0.07	y13 +1

3.2.7 *Staphylococcus* sp.

3.2.7.1 PSMβ (de novo sequencing)

Phenol soluble modulin β (by homology)
Seq: fMQKLAEAIANAVQAAGNHDWAKLGTSIVGIVENGVGALGKIFGF
m/z monoisotopic = 1124.3430, +4; MW_{obs} = 4493.34 Da



m/z meas	z	m/z calc	Δppm	assignment
370.1754	1	370.1761	-1.89	y3 +1
416.1956	1	416.1962	-1.44	b3 +1
416.1956	1	416.1962	-1.44	b2 +1
483.2594	1	483.2602	-1.66	y4 +1
529.2792	1	529.2803	-2.08	b4 +1
549.7837	2	549.7842	-0.91	b10 +2
585.3018	2	585.3028	-1.71	b11 +2
600.3163	1	600.3174	-1.83	b5 +1
611.3540	1	611.3552	-1.96	y5 +1
634.8359	2	634.8370	-1.73	b12 +2
668.3753	1	668.3766	-1.95	y6 +1
711.3481	1	711.3494	-1.83	b6-H2O +1
729.3586	1	729.3600	-1.92	b6 +1
734.3834	2	734.3849	-2.04	b14 +2
764.4324	1	764.4341	-2.22	y7-NH3 +1
781.4592	1	781.4607	-1.92	y7 +1
800.3956	1	800.3971	-1.87	b7 +1
820.7512	3	820.7533	-2.56	b23 +3
852.4962	1	852.4978	-1.88	y8 +1
855.4340	2	855.4356	-1.87	b17 +2
895.4692	1	895.4706	-1.56	b8-H2O +1
909.5174	1	909.5193	-2.09	y9 +1
913.4795	1	913.4812	-1.86	b8 +1
923.9628	2	923.9651	-2.49	b18 +2
930.8139	3	930.8167	-3.01	a27 +3
934.1426	3	934.1448	-2.36	b27-H2O +3
940.1463	3	940.1483	-2.13	b27 +3
963.8370	3	963.8395	-2.59	a28 +3

966.5058	3	966.5077	-1.97	b9-H2O +1
967.1655	3	967.1676	-2.17	b28-H2O +3
973.1692	3	973.1711	-1.95	b28 +3
981.4765	2	981.4785	-2.04	b19 +2
984.5162	1	984.5183	-2.13	b9 +1
992.1759	3	992.1783	-2.42	b29 +3
999.2765	4	999.2796	-3.10	b40-H2O +4
1003.7802	4	1003.7823	-2.09	b40 +4
1008.5854	1	1008.5877	-2.28	y10 +1
1020.5386	3	1020.5414	-2.74	a30 +3
1023.8673	3	1023.8695	-2.15	b30-H2O +3
1025.0522	4	1025.0546	-2.34	a41 +4
1029.8707	3	1029.8730	-2.23	b30 +3
1032.0512	4	1032.0533	-2.03	b41 +4
1053.5611	3	1053.5642	-2.94	a31 +3
1056.8901	3	1056.8923	-2.08	b31-H2O +3
1061.8181	4	1061.8217	-3.39	a42 +4
1062.8936	3	1062.8958	-2.07	b31 +3
1064.3149	4	1064.3178	-2.72	b42-H2O +4
1065.6067	1	1065.6091	-2.25	y11 +1
1068.8178	4	1068.8204	-2.43	b42 +4
1074.5161	2	1074.5182	-1.95	b20 +2
1078.5700	4	1078.5731	-2.87	b43-H2O +4
1083.0732	4	1083.0758	-2.40	b43 +4
1098.5583	1	1098.5612	-2.64	b10 +1
1101.0289	2	1101.0315	-2.36	b21-H2O +2
1105.9074	3	1105.9100	-2.35	b32 +3
1110.0341	2	1110.0367	-2.34	b21 +2
1138.2477	3	1138.2488	-0.97	b33-NH3 +3
1143.9218	3	1143.9243	-2.19	b33 +3
1151.5856	1	1151.5878	-1.91	b11-H2O +1
1156.9246	3	1156.9279	-2.85	b34-H2O +3
1162.9289	3	1162.9315	-2.24	b34 +3
1169.5953	1	1169.5983	-2.56	b11 +1
1170.9514	3	1170.9546	-2.73	y35 +3
1174.0815	2	1174.0842	-2.30	b22 +3
1179.6495	1	1179.6521	-2.20	y12 +1
1180.9462	3	1180.9471	-0.76	a35-NH3 +3
1186.6202	3	1186.6226	-2.02	a35 +3
1189.9478	3	1189.9507	-2.44	b35-H2O +3
1194.6315	3	1194.6336	-1.76	y36 +3
1195.9506	3	1195.9543	-3.09	b35 +3
1208.9553	3	1208.9579	-2.15	b36-H2O +3

1214.9582	3	1214.9614	-2.63	b36 +3
1221.6186	2	1221.6210	-1.96	b23-H2O +2
1226.3204	3	1226.3248	-3.59	y37-H2O +3
1229.3058	3	1229.3088	-2.44	a37 +3
1230.6216	2	1230.6263	-3.82	b23 +2
1232.3256	3	1232.3283	-2.19	y37 +3
1232.6311	3	1232.6369	-4.71	b37-H2O +3
1238.6375	3	1238.6405	-2.42	b37 +3
1250.0009	3	1250.0038	-2.32	y38-H2O +3
1251.6370	1	1251.6402	-2.56	b12-NH3 +1
1256.0053	3	1256.0074	-1.67	y38 +3
1259.1359	2	1259.1370	-0.87	b24 +2
1261.3285	3	1261.3280	0.40	a38-NH3 +3
1267.2086	2	1267.2125	-3.08	y25 +2
1268.6646	1	1268.6667	-1.66	b12 +1
1270.3286	3	1270.3316	-2.36	b38-H2O +3
1276.3319	3	1276.3351	-2.51	b38 +3
1289.3354	3	1289.3388	-2.64	b39-H2O +3
1295.3391	3	1295.3423	-2.47	b39 +3
1299.0191	3	1299.0216	-1.92	y39 +3
1308.6923	1	1308.6947	-1.83	y13 +1
1322.6969	3	1322.7006	-2.80	y40 +3
1332.0332	3	1332.0371	-2.93	b40-H2O +3
1338.0370	3	1338.0406	-2.69	b40 +3
1360.3915	3	1360.3953	-2.79	y41 +3
1369.7283	3	1369.7318	-2.56	b41-H2O +3
1375.7316	3	1375.7353	-2.69	b41 +3
1389.7497	1	1389.7525	-2.01	y14-H2O +1
1393.2523	2	1393.2554	-2.23	y27 +2
1396.7222	1	1396.7253	-2.22	b13 +1
1400.7108	2	1400.7136	-2.00	b27-H2O +2
1407.7598	1	1407.7631	-2.34	y14 +1
1409.7164	2	1409.7189	-1.77	b27 +2
1419.0794	3	1419.0826	-2.25	b42-NH3 +3
1424.7540	3	1424.7581	-2.88	b42 +3
1443.7614	3	1443.7653	-2.70	b43 +3
1459.2494	2	1459.2531	-2.54	b28 +2
1467.7569	1	1467.7624	-3.75	b14 +1
1487.7606	2	1487.7638	-2.15	b29 +2
1505.7951	2	1505.7929	1.46	y30-NH3 +2
1514.3028	2	1514.3062	-2.25	y30 +2
1520.8437	1	1520.8471	-2.24	y15 +1
1544.3020	2	1544.3058	-2.46	b30 +2

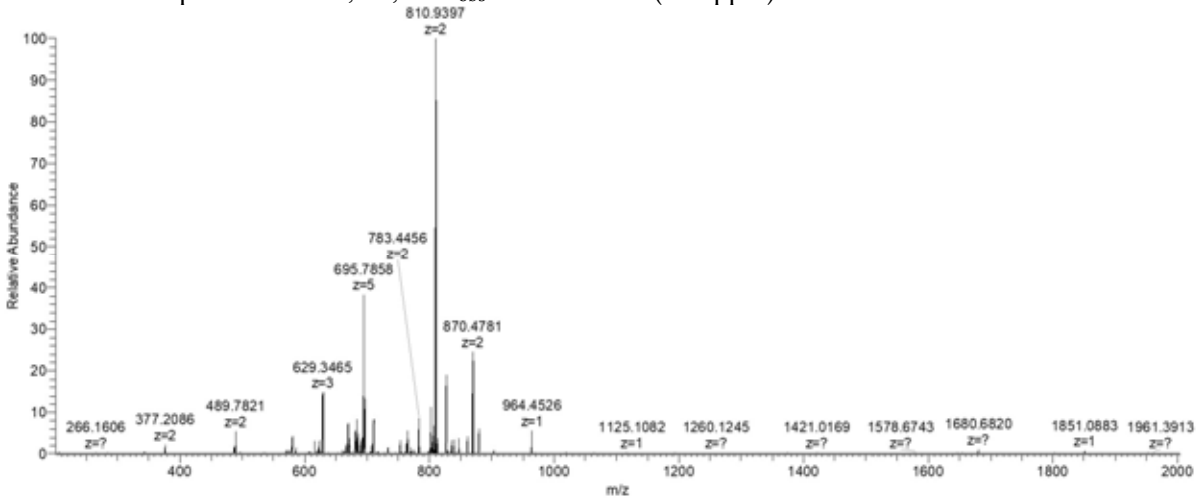
1549.8203	2	1549.8247	-2.84	y31 +2
1559.8550	1	1559.8580	-1.92	y16-H2O +1
1577.8645	1	1577.8686	-2.60	y16 +1
1593.8358	2	1593.8401	-2.70	b31 +2
1605.3390	2	1605.3407	-1.06	y32-NH3 +2
1613.8487	2	1613.8540	-3.28	y32 +2
1649.8474	2	1649.8481	-0.42	b32-NH3 +2
1658.3565	2	1658.3613	-2.89	b32 +2
1663.3875	2	1663.3882	-0.42	y33 +2
1676.9321	1	1676.9370	-2.92	y17 +1
1793.4219	2	1793.4278	-3.29	b35 +2

3.3 Chapter 5

3.3.1 *E. coli*

3.3.1.1 SRA

UniProt P68191, SRA
Seq: MKSNRQARHILGLDHKISNQRKIVTEGDKSSVVNNPTGRKRPAEK
m/z monoisotopic = 728.5430, +7; MW_{obs} = 5092.75 Da (-3.0 ppm)

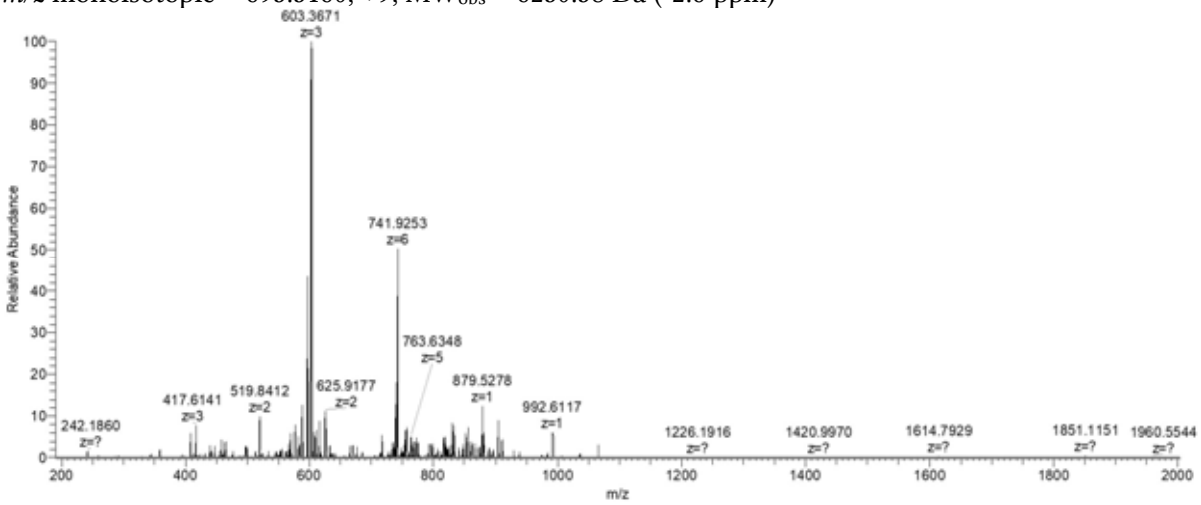


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B14	810.9423	1619.8701	1619.8678	0.0023	1.4
B15	879.4658	1756.9170	1756.9267	-0.0097	-5.5
B16	629.3452	1885.0136	1885.0217	-0.0081	-4.3
B17	667.0416	1998.1031	1998.1057	-0.0026	-1.3
B26	764.4204	3053.6524	3053.6781	-0.0257	-8.4
B28	807.4355	3225.7131	3225.7265	-0.0134	-4.2
B29	671.7661	3353.7943	3353.8214	-0.0271	-8.1

B29	839.4619	3353.8184	3353.8214	-0.0030	-0.9
B30	689.1733	3440.8303	3440.8535	-0.0232	-6.7
B44	707.6698	4946.6379	4946.6597	-0.0219	-4.4
Y12	684.3793	1366.7440	1366.7429	0.0011	0.8
Y14	783.4447	1564.8748	1564.8797	-0.0050	-3.2
Y15	826.9638	1651.9131	1651.9118	0.0013	0.8
Y16	870.4790	1738.9435	1738.9438	-0.0003	-0.2
Y16	580.6554	1738.9444	1738.9438	0.0006	0.3
Y17	623.3526	1867.0358	1867.0388	-0.0029	-1.6
Y19	680.7014	2039.0824	2039.0872	-0.0048	-2.3
Y29	802.9426	3207.7413	3207.7436	-0.0022	-0.7
Y30	834.9633	3335.8239	3335.8385	-0.0146	-4.4
Y31	695.5868	3472.8974	3472.8974	0.0000	0.0

3.3.1.2 L33

UniProt P0A7N9, L33
Seq: AKTIKITQTRSAIGRLPKHKATLLGLGLRRIGHTVEREDTPAIRGMINAVSFMVKVEE
m/z monoisotopic = 695.5160, +9; MW_{obs} = 6250.58 Da (-2.6 ppm)



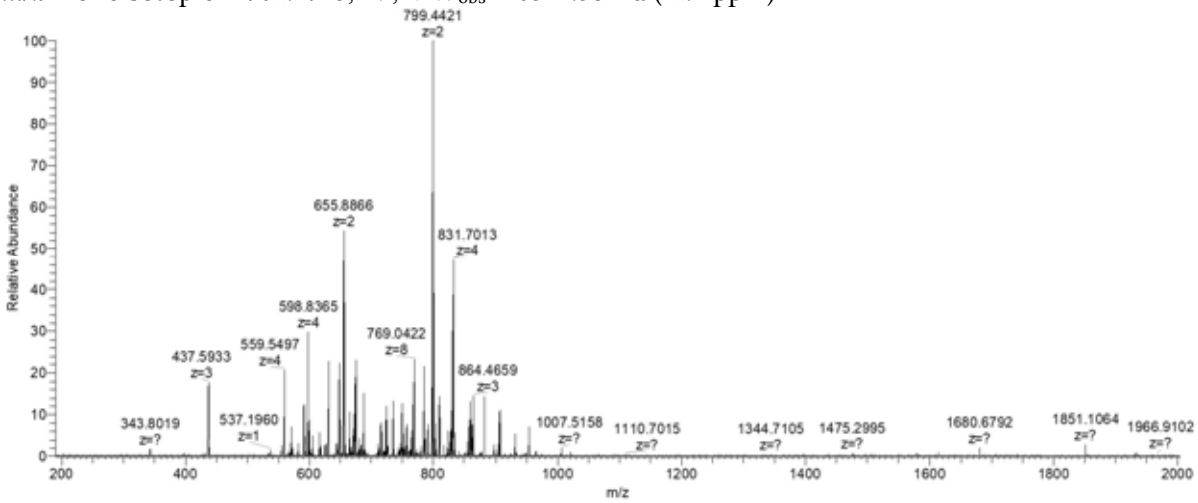
assignment	**m/z**	**MW_{obs}**	**MW_{theor}**	**ΔMW**	**Δppm**
B6	669.4026	668.3954	668.3969	-0.0016	-2.4
B7	797.4970	796.4897	796.4919	-0.0022	-2.8
B9	519.8423	1037.6699	1037.6709	-0.0010	-0.9
B10	576.3847	1150.7549	1150.7550	-0.0001	-0.1
B11	625.9179	1249.8212	1249.8234	-0.0022	-1.8
B11	417.6151	1249.8234	1249.8234	0.0000	0.0
B12	446.6239	1336.8500	1336.8554	-0.0054	-4.1
B12	669.4341	1336.8537	1336.8554	-0.0017	-1.3

B14	499.3129	1494.9167	1494.9246	-0.0078	-5.2
B15	776.9797	1551.9449	1551.9460	-0.0011	-0.7
B18	462.7736	1847.0654	1847.0741	-0.0087	-4.7
B18	616.6969	1847.0690	1847.0741	-0.0051	-2.8
B19	499.5386	1994.1254	1994.1425	-0.0170	-8.5
B19	665.7197	1994.1374	1994.1425	-0.0051	-2.5
B32	715.0046	3569.9867	3570.0158	-0.0291	-8.1
B34	763.4309	3812.1183	3812.1424	-0.0241	-6.3
B36	676.5599	4053.3156	4053.3214	-0.0058	-1.4
B36	811.6715	4053.3212	4053.3214	-0.0003	-0.1
B39	889.7039	4443.4832	4443.5117	-0.0285	-6.4
B39	741.5885	4443.4873	4443.5117	-0.0245	-5.5
B45	860.9876	5159.8820	5159.9199	-0.0380	-7.4
B45	738.1348	5159.8929	5159.9199	-0.0270	-5.2
B46	877.4995	5258.9535	5258.9883	-0.0348	-6.6
B52	749.9285	5991.3696	5991.4053	-0.0357	-6.0
Y2	260.1967	259.1895	259.1896	-0.0001	-0.5
Y4	459.3271	458.3198	458.3217	-0.0018	-4.0
Y5	588.3684	587.3611	587.3642	-0.0031	-5.3
Y6	716.4666	715.4594	715.4592	0.0001	0.2
Y7	440.2680	878.5214	878.5225	-0.0012	-1.3
Y8	992.6078	991.6005	991.6066	-0.0061	-6.2
Y8	496.8083	991.6021	991.6066	-0.0045	-4.5
Y15	603.3651	1807.0734	1807.0832	-0.0098	-5.4
Y15	603.3667	1807.0781	1807.0832	-0.0051	-2.8
Y18	733.4270	2197.2590	2197.2735	-0.0144	-6.6
Y20	610.6186	2438.4453	2438.4525	-0.0072	-3.0
Y34	819.6791	4093.3592	4093.3891	-0.0299	-7.3
Y35	852.2951	4256.4391	4256.4524	-0.0134	-3.1
Y36	734.9214	4403.4846	4403.5209	-0.0363	-8.2
Y36	881.7100	4403.5134	4403.5209	-0.0075	-1.7
Y40	793.6135	4755.6372	4755.6704	-0.0332	-7.0
Y41	805.4573	4826.7003	4826.7075	-0.0072	-1.5
Y43	834.4649	5000.7459	5000.7715	-0.0256	-5.1
Y52	755.6845	6037.4181	6037.4472	-0.0291	-4.8
Y53	771.6945	6165.4974	6165.5422	-0.0448	-7.3

3.3.1.3 L32

UniProt P0A7N4, L32
Seq: AVQQNKPTRSKRGMRRSHDALTAVTSLSVDKTSGEKHLRHHITADGYYRGRKVIAK

m/z monoisotopic = 702.2715, +9; MW_{obs} = 6311.38 Da (-2.1 ppm)

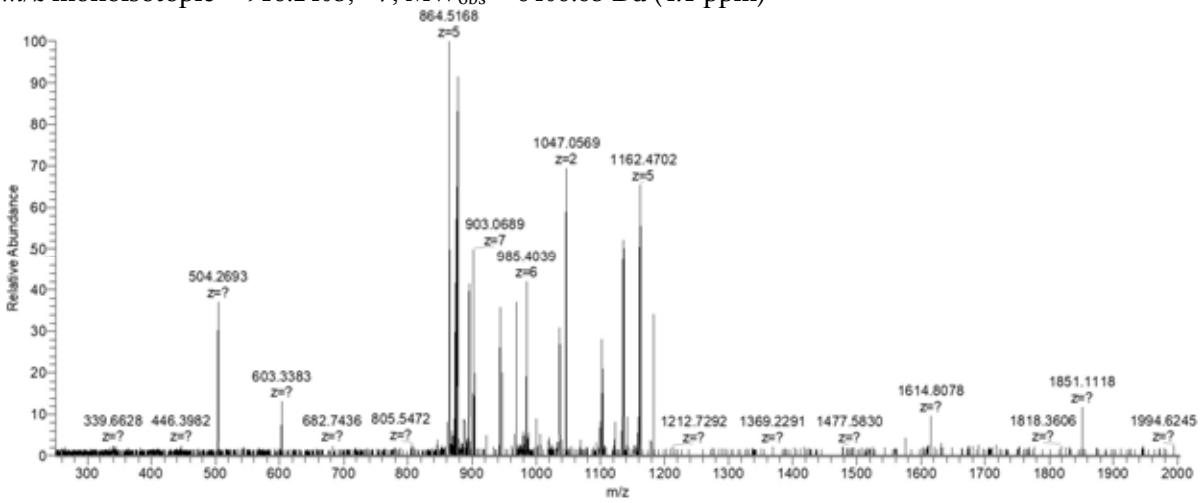


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B19	559.2992	2233.1677	2233.1723	-0.0046	-2.1
B23	648.3492	2589.3678	2589.3782	-0.0105	-4.0
B23	864.1349	2589.3829	2589.3782	0.0047	1.8
B24	673.1171	2688.4391	2688.4466	-0.0075	-2.8
B30	823.6888	3290.7260	3290.7378	-0.0118	-3.6
B41	767.9103	4601.4180	4601.4334	-0.0154	-3.3
B42	674.5053	4714.4864	4714.5174	-0.0310	-6.6
B45	834.6069	5001.5975	5001.6292	-0.0317	-6.3
Y11	655.8867	1309.7588	1309.7619	-0.0031	-2.4
Y11	437.5941	1309.7605	1309.7619	-0.0014	-1.1
Y13	748.9191	1495.8237	1495.8259	-0.0023	-1.5
Y14	799.4417	1596.8688	1596.8736	-0.0048	-3.0
Y15	855.9853	1709.9561	1709.9577	-0.0016	-0.9
Y15	570.9934	1709.9584	1709.9577	0.0008	0.4
Y19	752.0914	2253.2525	2253.2606	-0.0081	-3.6
Y20	797.7781	2390.3125	2390.3196	-0.0070	-2.9
Y20	598.5866	2390.3173	2390.3196	-0.0022	-0.9
Y21	630.6101	2518.4113	2518.4145	-0.0032	-1.3
Y25	724.1450	2892.5508	2892.5583	-0.0075	-2.6
Y26	756.1663	3020.6360	3020.6532	-0.0172	-5.7
Y26	605.1365	3020.6462	3020.6532	-0.0070	-2.3
Y27	784.9239	3135.6664	3135.6802	-0.0138	-4.4
Y28	809.6911	3234.7353	3234.7486	-0.0133	-4.1
Y29	831.4470	3321.7590	3321.7806	-0.0216	-6.5
Y30	859.7154	3434.8325	3434.8647	-0.0321	-9.4
Y31	881.4776	3521.8811	3521.8967	-0.0156	-4.4

Y32	906.7353	3622.9122	3622.9444	-0.0321	-8.9
Y32	725.5915	3622.9213	3622.9444	-0.0231	-6.4
Y54	768.6614	6141.2327	6141.2855	-0.0528	-8.6

3.3.1.4 L30

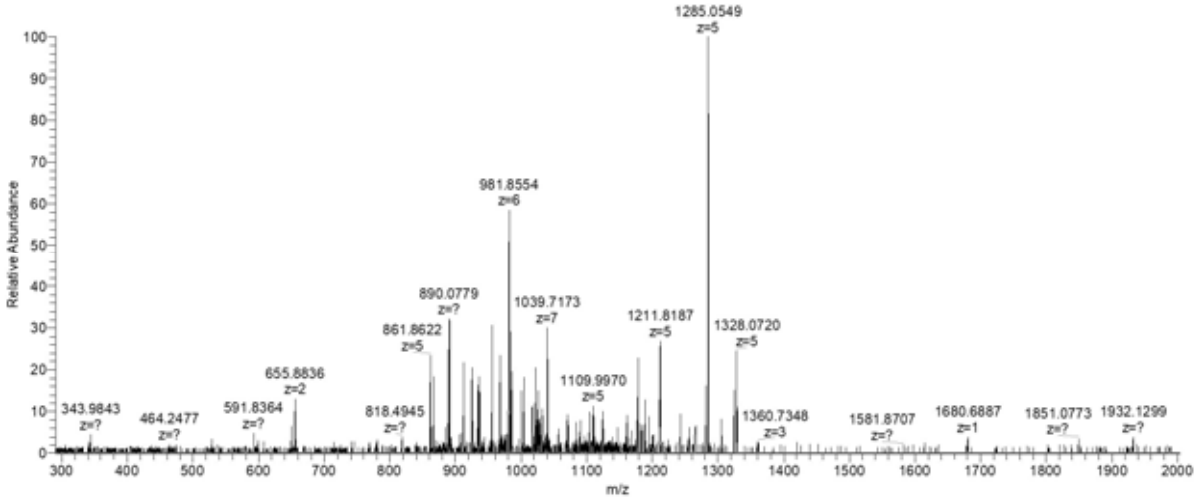
UniProt P0AG51, L30
Seq: AKTIKITQTRSAIGRLPKHKATLLGLGLRRIGHTVEREDTPAIRGMINAVSFMVKVEE
m/z monoisotopic = 916.2403, +7; MW_{obs} = 6406.63 Da (4.1 ppm)



assignment	**m/z**	**MW_{obs}**	**MW_{theor}**	**ΔMW**	**Δppm**
B39	864.1116	4315.5217	4315.5192	0.0025	0.6
B52	1135.6536	5673.2317	5673.2368	-0.0052	-0.9
B53	968.3857	5804.2707	5804.2773	-0.0067	-1.1
B53	1161.8660	5804.2935	5804.2773	0.0162	2.8
B54	984.9001	5903.3571	5903.3457	0.0113	1.9
B56	876.7961	6130.5219	6130.5091	0.0128	2.1
B57	895.2306	6259.5631	6259.5517	0.0114	1.8
Y19	1046.5511	2091.0876	2091.0856	0.0019	0.9

3.3.1.5 RaiA

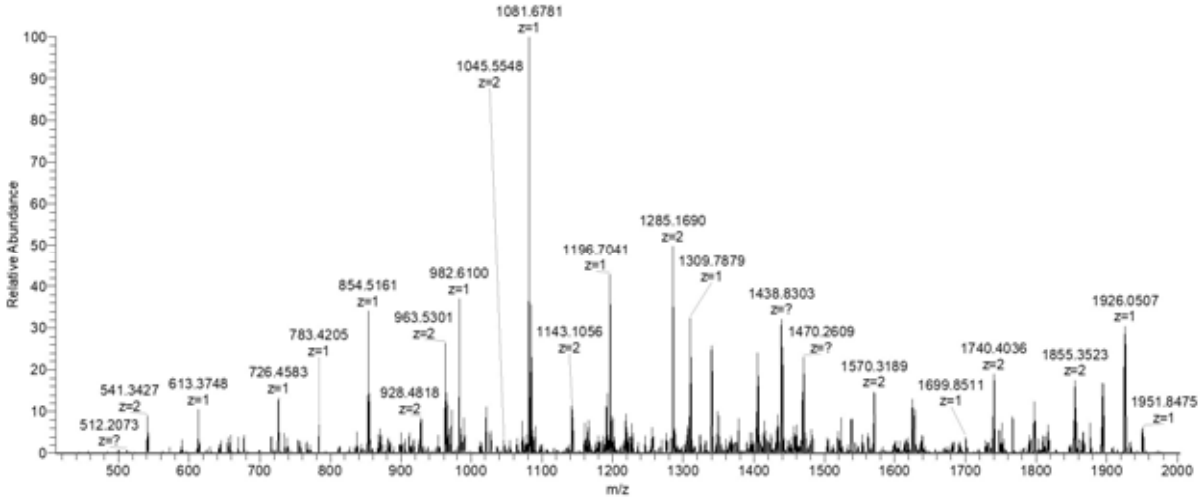
UniProt P0AD49, RaiA
Seq: TMNITSKQMEITPAIRQHVADRLAKLEKWQTHLINPHIILSKEPQGFVADATINTPNGVL
VASGKHEDMYTAINELINKLERQLNKLQHKGEARRAATSVKDANFVEEVEEE
m/z monoisotopic = 1054.7974, +12; MW_{obs} = 12645.48 Da (-8.3 ppm)



assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B35	1025.0399	4096.1306	4096.1826	-0.0520	-12.7
B37	867.0592	4330.2597	4330.2943	-0.0345	-8.0
B38	889.6748	4443.3377	4443.3783	-0.0406	-9.1
B39	912.2924	4556.4257	4556.4624	-0.0367	-8.1
B40	934.9077	4669.5021	4669.5465	-0.0444	-9.5
B43	1003.7408	5013.6678	5013.7160	-0.0483	-9.6
B48	924.6612	5541.9233	5541.9857	-0.0624	-11.3
B50	955.6690	5727.9703	5728.0497	-0.0794	-13.9
B55	1039.0477	6228.2423	6228.3092	-0.0669	-10.7
Y53	1211.0119	6050.0231	6050.0919	-0.0688	-11.4
Y57	1284.4531	6417.2291	6417.2774	-0.0483	-7.5

3.3.1.6 YgiW

UniProt P0ADU5, YgiW
Seq: AEQGGFSGPSATQSQAGGFQGPNGSVTTVESAKSLRDDTWVTLRGNIVERISDDLYVFKD
ASGTINVDIDHKRWNGVTVTPKDTVEIQGEVDKDWNSVEIDVKQIRKVNP
m/z monoisotopic = 1497.1216, +8; MW_{obs} = 11968.91 Da (-6.1 ppm)

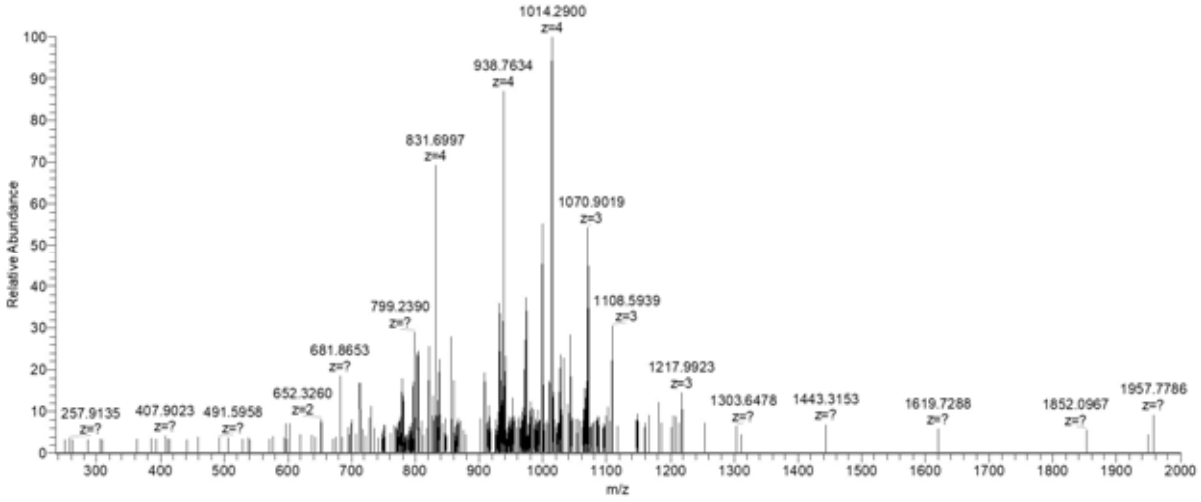


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B6	590.2551	589.2479	589.2496	-0.0018	-3.0
B7	677.2867	676.2795	676.2816	-0.0022	-3.2
B8	734.3088	733.3015	733.3031	-0.0016	-2.2
B10	918.3933	917.3860	917.3879	-0.0019	-2.0
B12	1090.4777	1089.4704	1089.4727	-0.0023	-2.1
B14	1305.5596	1304.5524	1304.5633	-0.0109	-8.4
B15	1433.6238	1432.6165	1432.6219	-0.0054	-3.8
B16	1504.6561	1503.6489	1503.6590	-0.0101	-6.7
B17	1561.6840	1560.6768	1560.6804	-0.0037	-2.4
B18	1618.7033	1617.6961	1617.7019	-0.0058	-3.6
B20	1893.8133	1892.8060	1892.8289	-0.0229	-12.1
B37	1796.8318	3591.6490	3591.6825	-0.0335	-9.3
B38	1854.3431	3706.6716	3706.7095	-0.0378	-10.2
B88	1567.9370	9401.5786	9401.6528	-0.0743	-7.9
B107	1456.1042	11640.7753	11640.8125	-0.0372	-3.2
B108	1468.4815	11739.7941	11739.8809	-0.0867	-7.4
B109	1482.7391	11853.8549	11853.9238	-0.0689	-5.8
Y5	613.3765	612.3692	612.3707	-0.0016	-2.5
Y6	726.4585	725.4512	725.4548	-0.0036	-4.9
Y7	854.5170	853.5098	853.5134	-0.0036	-4.2
Y9	541.3437	1080.6728	1080.6767	-0.0039	-3.7
Y10	1196.7077	1195.7004	1195.7037	-0.0033	-2.7
Y11	655.3993	1308.7840	1308.7877	-0.0038	-2.9
Y11	1309.7923	1308.7850	1308.7877	-0.0027	-2.1
Y12	1438.8333	1437.8260	1437.8303	-0.0043	-3.0
Y13	1537.8966	1536.8893	1536.8987	-0.0095	-6.2
Y14	1624.9266	1623.9193	1623.9308	-0.0115	-7.1

Y14	812.9684	1623.9223	1623.9308	-0.0085	-5.2
Y15	869.9904	1737.9663	1737.9737	-0.0074	-4.2
Y16	963.0259	1924.0373	1924.0530	-0.0158	-8.2
Y17	1020.5364	2039.0582	2039.0800	-0.0217	-10.7
Y18	1084.5912	2167.1678	2167.1749	-0.0071	-3.3
Y19	1142.0988	2282.1831	2282.2019	-0.0187	-8.2
Y20	1191.6307	2381.2469	2381.2703	-0.0234	-9.8
Y21	1256.1630	2510.3114	2510.3129	-0.0014	-0.6
Y22	1284.6653	2567.3160	2567.3343	-0.0183	-7.1
Y23	1348.7029	2695.3912	2695.3929	-0.0017	-0.6
Y24	937.1579	2808.4518	2808.4770	-0.0252	-9.0
Y24	1405.2350	2808.4554	2808.4770	-0.0216	-7.7
Y25	1469.7632	2937.5118	2937.5195	-0.0077	-2.6
Y26	1519.2829	3036.5512	3036.5880	-0.0367	-12.1
Y27	1569.8155	3137.6165	3137.6356	-0.0191	-6.1
Y27	1046.8797	3137.6171	3137.6356	-0.0185	-5.9
Y28	1627.3338	3252.6530	3252.6626	-0.0095	-2.9
Y30	1739.8934	3477.7722	3477.8103	-0.0381	-10.9
Y30	1160.2684	3477.7833	3477.8103	-0.0270	-7.8
Y31	1790.4252	3578.8358	3578.8580	-0.0222	-6.2
Y40	932.2868	4656.3977	4656.4412	-0.0434	-9.3
Y40	1165.1077	4656.4016	4656.4412	-0.0396	-8.5
Y42	1222.1356	4884.5132	4884.5522	-0.0389	-8.0
Y57	1305.4654	6522.2906	6522.3458	-0.0552	-8.5
Y72	1378.0443	8262.2220	8262.2777	-0.0557	-6.7
Y73	1197.7522	8377.2141	8377.3046	-0.0905	-10.8
Y73	1397.2203	8377.2784	8377.3046	-0.0262	-3.1
Y89	1432.3019	10019.0624	10019.1368	-0.0743	-7.4
Y90	1440.4478	10076.0834	10076.1583	-0.0749	-7.4
Y91	1458.7425	10204.1467	10204.2168	-0.0701	-6.9
Y103	1412.5825	11292.6019	11292.7055	-0.1036	-9.2
Y107	1456.1042	11640.7753	11640.8489	-0.0736	-6.3

3.3.1.7 YejL

UniProt P0AD24, YejL
Seq:
PQISRYSDEQVEQLLAELLNVLEKHKAPTDLSLMVLGNMVTNLINTSIAPAQRQAIANSFARALQSSIN
EDKAH
m/z monoisotopic = 906.4584, +9; MW_{obs} = 8154.38 Da (+2.04 Da)

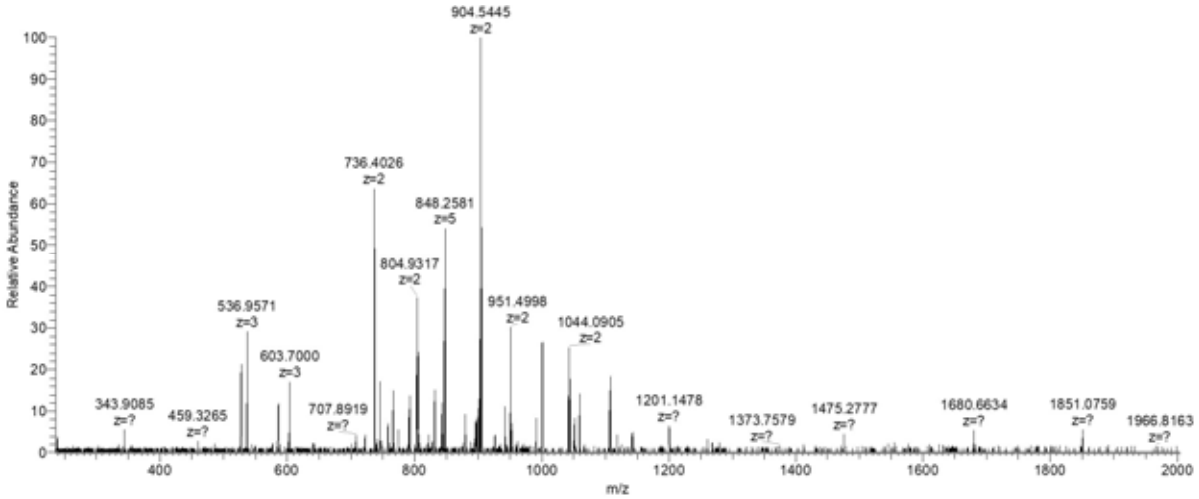


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B13	780.8784	1559.7423	1559.7216	0.0207	13.3
B14	837.4185	1672.8224	1672.8057	0.0167	10.0
B33	940.5181	3758.0433	3757.9890	0.0543	14.4
B34	778.8167	3889.0471	3889.0295	0.0176	4.5
B34	973.2768	3889.0780	3889.0295	0.0485	12.5
B35	998.0455	3988.1528	3988.0979	0.0549	13.8
B38	855.4672	4272.2995	4272.2463	0.0532	12.4
Y25	908.4769	2722.4088	2722.3899	0.0190	7.0
Y26	932.1634	2793.4683	2793.4270	0.0413	14.8
Y27	969.8554	2906.5445	2906.5111	0.0335	11.5
Y30	1070.5623	3208.6649	3208.6337	0.0312	9.7
Y31	1108.2602	3321.7587	3321.7178	0.0409	12.3
Y31	831.4483	3321.7640	3321.7178	0.0462	13.9
Y35	938.2564	3748.9965	3748.9608	0.0356	9.5
Y37	999.5234	3994.0646	3994.0443	0.0204	5.1

3.3.2 *P. aeruginosa*

3.3.2.1 L33

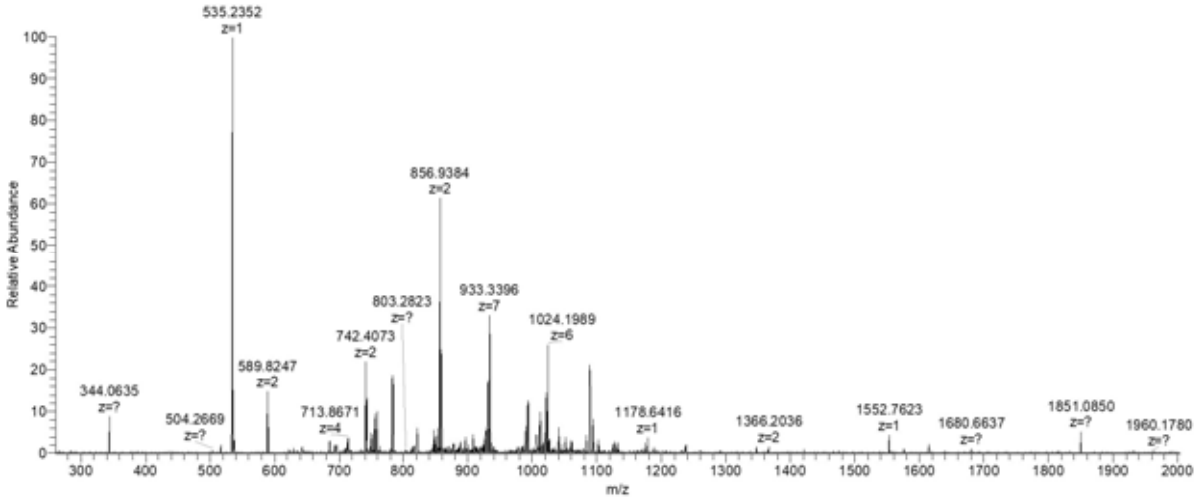
UniProt Q9HTN9, L33
Seq: MRELIRLVSSAGTGHFYTTDKNKRTKPEKIEIKKYDPVVRQHVIYKEAKIK
m/z monoisotopic = 864.0533, +7; MW_{obs} = 6041.32 Da (-4.9 ppm)



assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B14	736.4039	1470.7932	1470.7977	-0.0045	-3.1
B15	804.9336	1607.8527	1607.8566	-0.0038	-2.4
B16	585.9793	1754.9161	1754.9250	-0.0089	-5.1
B20	746.0414	2235.1023	2235.1106	-0.0083	-3.7
B33	766.6179	3828.0531	3828.0832	-0.0301	-7.9
B34	792.2377	3956.1524	3956.1782	-0.0258	-6.5
B36	847.8510	4234.2185	4234.2684	-0.0499	-11.8
Y15	603.3617	1807.0633	1807.0832	-0.0199	-11.0
Y15	904.5405	1807.0664	1807.0832	-0.0168	-9.3
Y17	1043.5888	2085.1631	2085.1735	-0.0104	-5.0
Y18	1107.6354	2213.2562	2213.2684	-0.0122	-5.5

3.3.2.2 L32

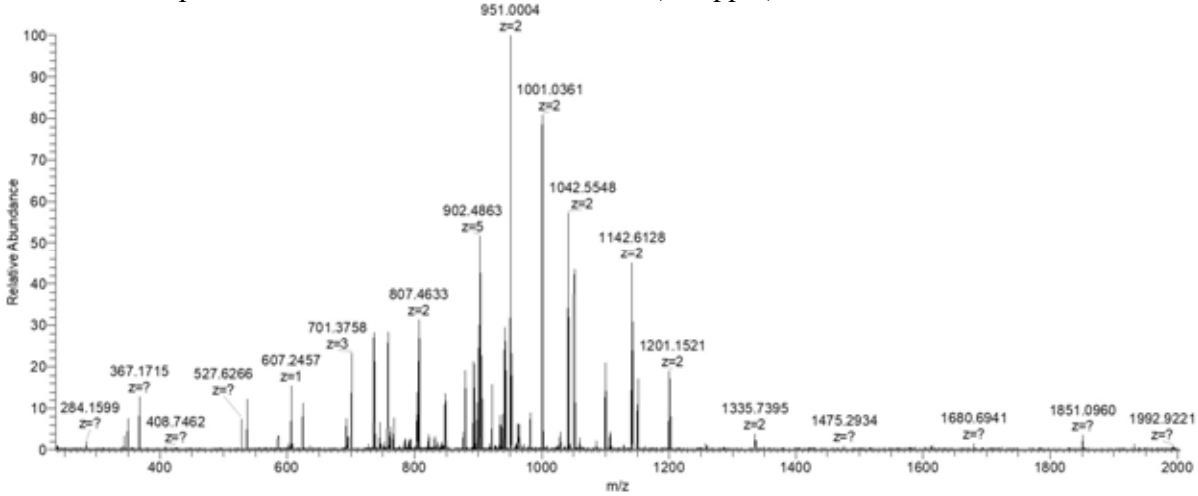
UniProt Q9HZN4, L32
Seq: AVQQNKKSRSARDMRRSHDALESNALSVEKSTGEVHLRHHVSPDGFYRGRKVVDKGSDE
m/z monoisotopic = 953.9173, +7; MW_{obs} = 6670.37 Da (-1.2 ppm)



assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B13	742.4063	1482.7980	1482.8015	-0.0035	-2.3
B19	756.0623	2265.1650	2265.1621	0.0029	1.3
B29	1093.8913	3278.6522	3278.6650	-0.0129	-3.9
B30	852.6957	3406.7538	3406.7600	-0.0061	-1.8
B44	992.7043	4958.4850	4958.5142	-0.0291	-5.9
B58	932.9102	6523.3201	6523.3252	-0.0051	-0.8
Y5	535.2366	534.2293	534.2285	0.0007	1.4
Y15	856.9401	1711.8657	1711.8641	0.0015	0.9
Y29	1088.8728	3263.5964	3263.6184	-0.0219	-6.7

3.3.2.3 CsrA

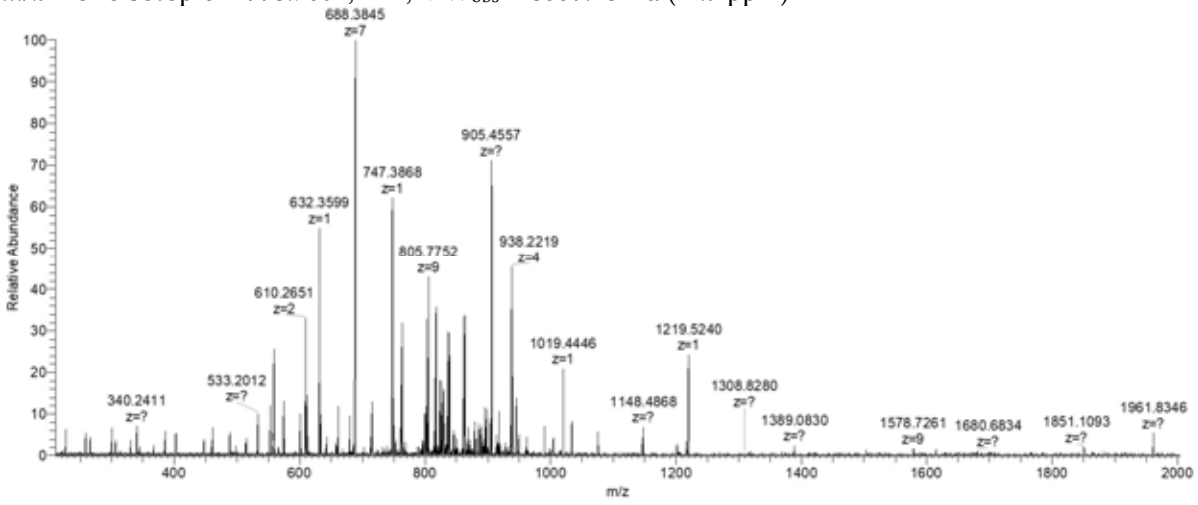
UniProt O69078, CsrA
Seq: MLILTRRVGETLMVGDDVTVTVLGVKGNQVRIGVNAPKEVAVHREEIYQRIQKEKDQEPN
H
m/z monoisotopic = 864.0915, +8; MW_{obs} = 6904.67 Da (-2.6 ppm)



assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B13	757.9316	1513.8487	1513.8473	0.0014	0.9
B14	807.4647	1612.9148	1612.9157	-0.0009	-0.5
B16	893.4883	1784.9621	1784.9641	-0.0020	-1.1
B17	950.9957	1899.9768	1899.9910	-0.0143	-7.5
B18	1000.5361	1999.0576	1999.0594	-0.0018	-0.9
B19	1051.0558	2100.0971	2100.1071	-0.0100	-4.8
B20	1100.5943	2199.1741	2199.1755	-0.0014	-0.6
B21	767.7470	2300.2193	2300.2232	-0.0039	-1.7
B21	1151.1222	2300.2299	2300.2232	0.0067	2.9
B22	1200.6514	2399.2882	2399.2916	-0.0034	-1.4
B58	935.0794	6538.5046	6538.5267	-0.0221	-3.4
B60	965.2323	6749.5754	6749.6224	-0.0470	-7.0
Y38	879.4654	4392.2907	4392.3162	-0.0256	-5.8
Y39	902.0867	4505.3969	4505.4003	-0.0034	-0.7
Y40	921.8963	4604.4452	4604.4687	-0.0235	-5.1
Y41	942.1055	4705.4910	4705.5164	-0.0254	-5.4
Y42	961.9190	4804.5588	4804.5848	-0.0260	-5.4

3.3.2.4 PA4739

UniProt Q9HV60, PA4739
Seq: ANDTMQKTEEAVSDTWITSKVKSSLIANKNVSGVDIKVETNKGVVSLSGNVKS
DAERDLAIETAKGIKGVKAVSADGLKSVE
m/z monoisotopic = 778.9604, +11; MW_{obs} = 8557.48 Da (-2.7 ppm)

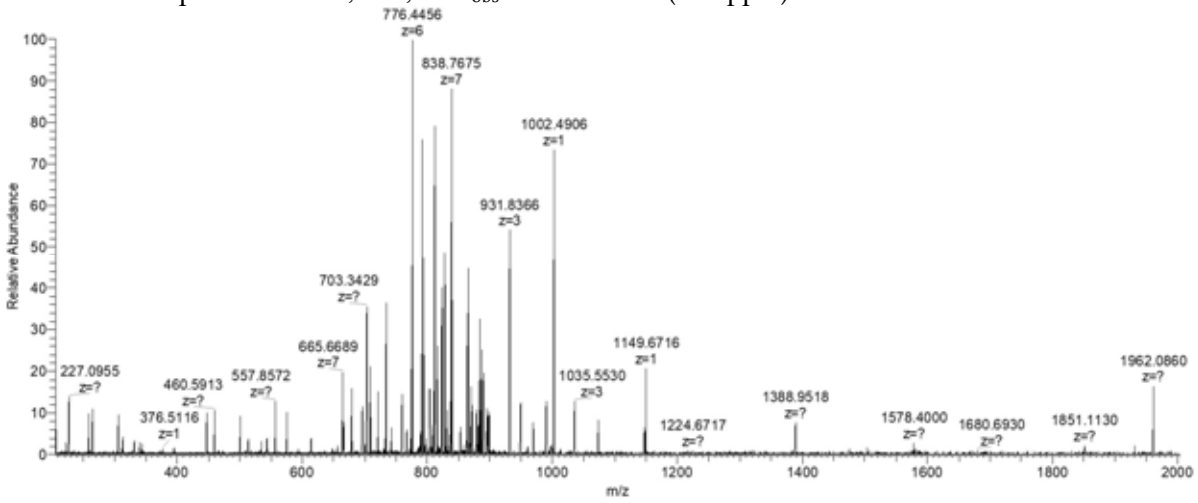


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
Y6	632.3614	631.3542	631.3541	0.0001	0.1
Y7	747.3874	746.3802	746.3810	-0.0008	-1.1

Y9	905.4585	904.4512	904.4502	0.0010	1.1
Y17	829.4767	1656.9389	1656.9410	-0.0021	-1.2
Y17	553.3212	1656.9416	1656.9410	0.0007	0.4
Y28	714.6509	2854.5744	2854.5763	-0.0019	-0.7
Y37	749.4048	3741.9878	3742.0112	-0.0234	-6.3
Y40	800.4377	3997.1523	3997.1694	-0.0171	-4.3
Y45	762.4180	4568.4644	4568.4660	-0.0016	-0.3
Y47	802.6103	4809.6179	4809.6450	-0.0271	-5.6
Y47	688.0975	4809.6314	4809.6450	-0.0136	-2.8
Y51	862.3056	5167.7902	5167.7939	-0.0037	-0.7
Y70	905.9901	7239.8623	7239.9204	-0.0582	-8.0
Y70	805.4407	7239.9013	7239.9204	-0.0192	-2.7
Y71	816.4465	7338.9531	7338.9889	-0.0357	-4.9
Y72	824.3432	7410.0237	7410.0260	-0.0022	-0.3

3.3.2.5 HU-β

UniProt P05384, HU-β
Seq: MLILTRRVGETLMVGDDVTVTVLGVKGNQVRIGVNAPKEVAVHREEIYQRIQKEKDQEPN
H
m/z monoisotopic = 757.7571, +12; MW_{obs} = 9080.99 Da (-5.1 ppm)

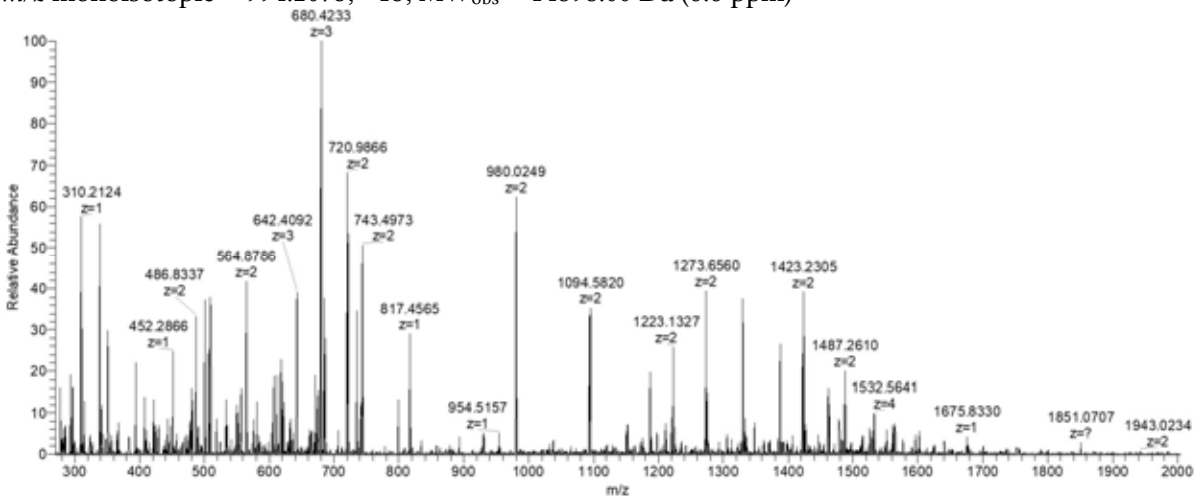


assignment	**m/z**	**MW_{obs}**	**MW_{theor}**	**ΔMW**	**Δppm**
B9	1002.4909	1001.4836	1001.4852	-0.0016	-1.6
B27	898.4806	2692.4201	2692.4217	-0.0016	-0.6
B28	931.4994	2791.4765	2791.4901	-0.0136	-4.9
B33	831.1950	3320.7510	3320.7649	-0.0139	-4.2
Y14	708.4039	1414.7933	1414.7932	0.0001	0.1
Y45	776.1112	4650.6238	4650.6349	-0.0112	-2.4

Y45	665.3841	4650.6378	4650.6349	0.0028	0.6
Y46	679.5349	4749.6931	4749.7034	-0.0103	-2.2
Y46	792.6235	4749.6976	4749.7034	-0.0057	-1.2
Y47	695.6858	4862.7495	4862.7874	-0.0379	-7.8
Y47	811.4688	4862.7690	4862.7874	-0.0184	-3.8
Y48	827.9822	4961.8498	4961.8558	-0.0060	-1.2
Y57	823.9004	5760.2521	5760.2794	-0.0273	-4.7
Y58	733.6678	5861.2839	5861.3270	-0.0431	-7.4
Y58	838.3393	5861.3239	5861.3270	-0.0031	-0.5
Y60	864.9193	6047.3844	6047.4275	-0.0430	-7.1
Y62	899.5096	6289.5163	6289.5541	-0.0378	-6.0

3.3.2.6 L11

UniProt Q9HWC5, L11
Seq:
MAKKIQAYIKLQVKAGQANPSPPVGPALGQHGVNIMEFCKAFNAKTQGQEPGLPTPVIITVYSDRSF
TFETKSTPAAVLLKKAAGITSGSARPNSQKVGTVTRAQLEEIAKTKQADLTAADLDAAVRTIAGSAR
SMGLNVEGV
m/z monoisotopic = 994.2075, +15; MW_{obs} = 14898.00 Da (6.6 ppm)

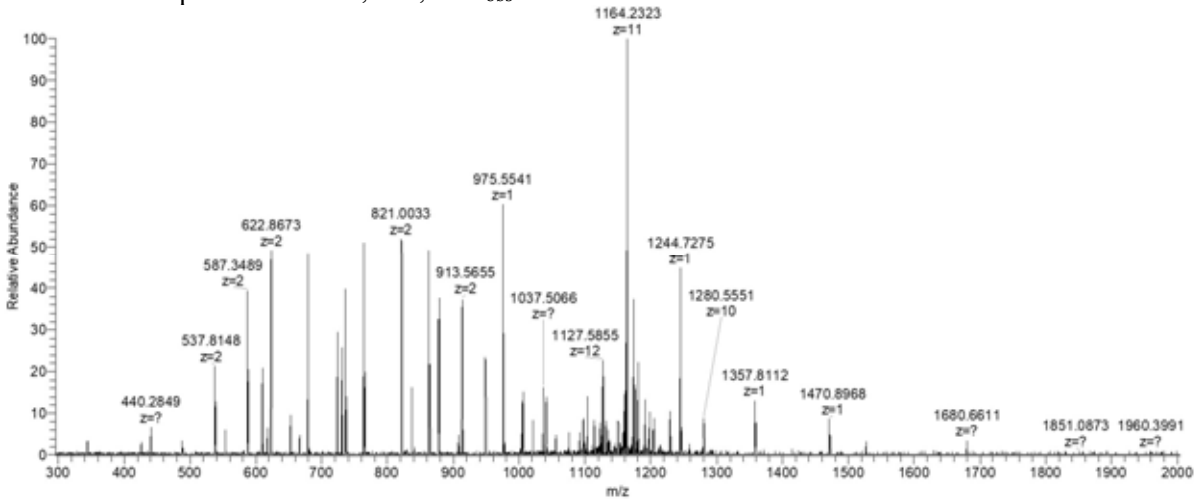


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
Y22	1094.0795	2186.1445	2186.1477	-0.0032	-1.4
Y23	1151.5916	2301.1687	2301.1746	-0.0059	-2.6
Y24	1187.1059	2372.1973	2372.2118	-0.0145	-6.1
Y25	1222.6258	2443.2370	2443.2489	-0.0119	-4.9
Y26	1273.1604	2544.3062	2544.2965	0.0096	3.8
Y27	1329.6951	2657.3756	2657.3806	-0.0050	-1.9
Y28	1387.2001	2772.3856	2772.4075	-0.0220	-7.9
Y29	1422.7346	2843.4546	2843.4447	0.0100	3.5

Y30	1486.7542	2971.4939	2971.5032	-0.0093	-3.1
Y31	1550.7989	3099.5833	3099.5982	-0.0149	-4.8
Y61	1531.5652	6122.2319	6122.2141	0.0178	2.9
Y62	1563.5791	6250.2875	6250.3091	-0.0216	-3.5

3.3.2.7 Pilin

UniProt P04739, fimbrial protein (pilin)
Seq:
FTLIELMIVVAIIGILAAIAIPQYQNYVARSEGASALATINPLKTTVEESLSRGIAGSKIKIGTTASTATET
YVGVEPDANKLGVIAVAIEDSGAGDITFTFQTGTSSPKNATKVITLNRTADGVWACKSTQDPMFTPK
GCDN
m/z monoisotopic = 1071.3510, +14; MW_{obs} = 14983.79 Da



assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B6	731.4312	730.4239	730.4265	-0.0026	-3.5
B7	862.4730	861.4658	861.4670	-0.0012	-1.4
B8	975.5574	974.5501	974.5511	-0.0010	-1.0
B9	537.8156	1073.6166	1073.6195	-0.0029	-2.7
B10	587.3502	1172.6859	1172.6879	-0.0019	-1.7
B10	1173.6950	1172.6877	1172.6879	-0.0002	-0.1
B11	622.8681	1243.7217	1243.7250	-0.0033	-2.6
B11	1244.7313	1243.7240	1243.7250	-0.0010	-0.8
B12	1357.8101	1356.8029	1356.8091	-0.0062	-4.6
B12	679.4099	1356.8053	1356.8091	-0.0038	-2.8
B13	1470.8920	1469.8847	1469.8931	-0.0084	-5.7
B13	735.9516	1469.8886	1469.8931	-0.0045	-3.1
B14	764.4631	1526.9116	1526.9146	-0.0029	-1.9
B15	821.0012	1639.9879	1639.9986	-0.0108	-6.6
B16	877.5465	1753.0785	1753.0827	-0.0042	-2.4
B17	913.0595	1824.1044	1824.1198	-0.0154	-8.4

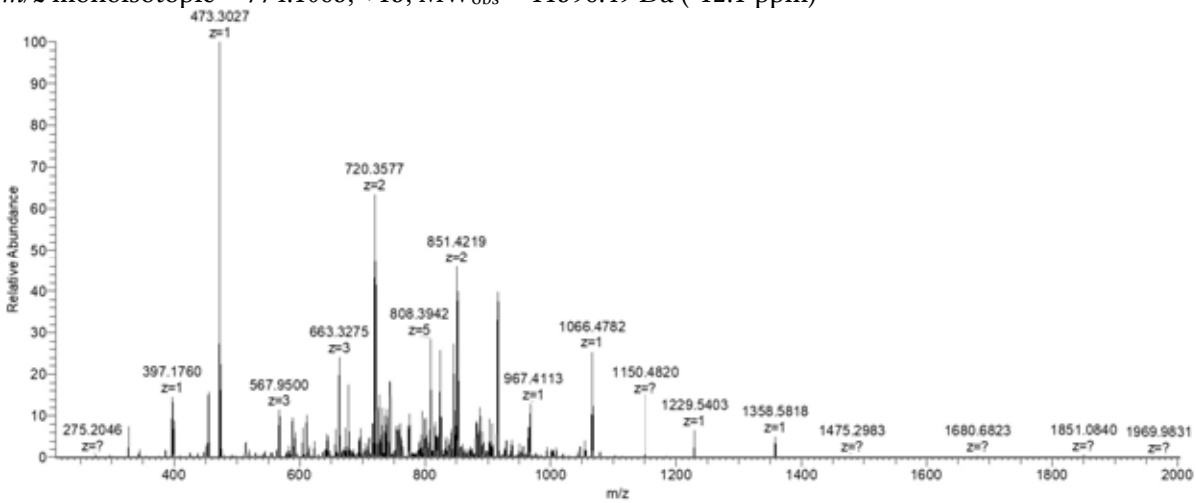
B18	948.5790	1895.1435	1895.1569	-0.0134	-7.1
B19	1005.1265	2008.2385	2008.2410	-0.0025	-1.2
B20	1040.6380	2079.2613	2079.2781	-0.0167	-8.0

3.4 Chapter 6

3.4.1 *S. cerevisiae*

3.4.1.1 HSP12 (15+)

UniProt P22943, HSP12
Seq: SDAGRKGFGEKASEALKPDSQKSYAEQGKEYITDKADKVAGKVQPEDNKGVFQGVHDSA
EKGKDNAEGQGESLADQARDYMGAAKSKLNDAVEYVSGRVHGEEDPTKK
m/z monoisotopic = 774.1065, +15; MW_{obs} = 11596.49 Da (-12.1 ppm)

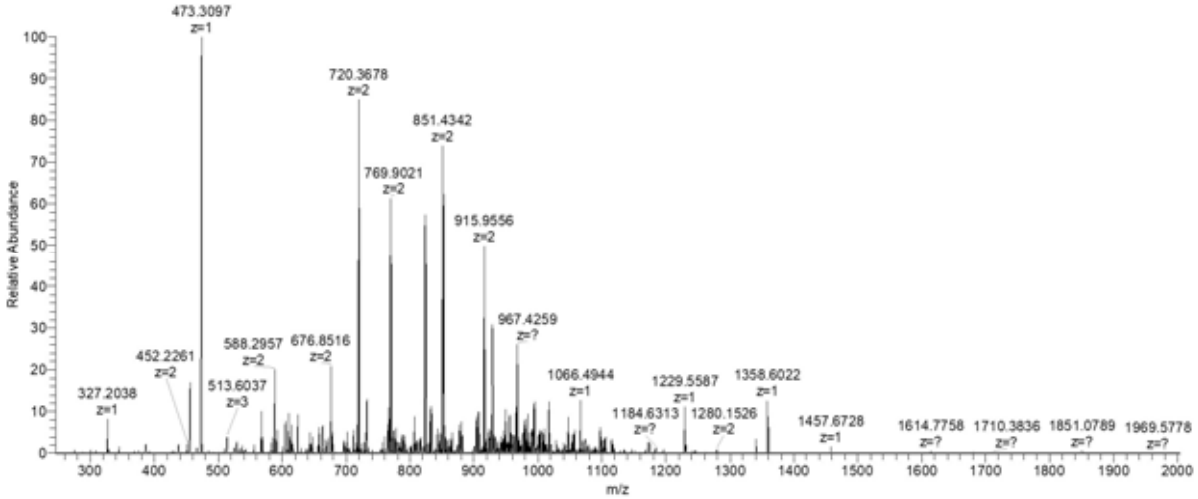


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B6	657.3259	656.3186	656.3242	-0.0056	-8.5
B11	588.2871	1174.5597	1174.5731	-0.0133	-11.3
B12	623.8046	1245.5946	1245.6102	-0.0156	-12.5
B14	731.8429	1461.6712	1461.6848	-0.0135	-9.3
B16	823.9030	1645.7914	1645.8060	-0.0145	-8.8
B17	887.9445	1773.8745	1773.9009	-0.0264	-14.9
B17	592.2993	1773.8760	1773.9009	-0.0249	-14.1
B19	662.9945	1985.9617	1985.9806	-0.0189	-9.5
B22	583.2912	2329.1359	2329.1662	-0.0303	-13.0
B34	929.4386	3713.7254	3713.7808	-0.0555	-14.9
B34	743.7523	3713.7254	3713.7808	-0.0555	-14.9
B37	672.3235	4027.8971	4027.9399	-0.0427	-10.6
B38	832.2025	4155.9761	4156.0348	-0.0588	-14.1
B38	693.6709	4155.9816	4156.0348	-0.0532	-12.8

B42	752.8735	4511.1972	4511.2568	-0.0595	-13.2
B57	881.1440	6160.9573	6161.0365	-0.0792	-12.9
B57	685.5584	6160.9600	6161.0365	-0.0765	-12.4
B67	720.0527	7190.4547	7190.5092	-0.0545	-7.6
B69	738.5566	7375.4934	7375.5892	-0.0959	-13.0
B71	757.1623	7561.5506	7561.6533	-0.1027	-13.6
Y4	473.3032	472.2959	472.3009	-0.0050	-10.6
Y5	588.3299	587.3226	587.3279	-0.0053	-9.0
Y6	717.3678	716.3605	716.3705	-0.0100	-13.9
Y8	452.2188	902.4231	902.4345	-0.0114	-12.6
Y12	676.8428	1351.6710	1351.6844	-0.0134	-9.9
Y13	720.3605	1438.7065	1438.7164	-0.0099	-6.9
Y14	513.5979	1537.7718	1537.7848	-0.0130	-8.4
Y15	851.4191	1700.8236	1700.8482	-0.0245	-14.4
Y15	567.9499	1700.8279	1700.8482	-0.0203	-11.9
Y16	915.9405	1829.8664	1829.8908	-0.0244	-13.3
Y17	643.9868	1928.9385	1928.9592	-0.0206	-10.7
Y17	965.4774	1928.9402	1928.9592	-0.0190	-9.8
Y18	667.6650	1999.9732	1999.9963	-0.0231	-11.5
Y23	640.3198	2557.2501	2557.2772	-0.0271	-10.6
Y33	609.1321	3648.7489	3648.7954	-0.0465	-12.8
Y36	790.5859	3947.8932	3947.9435	-0.0503	-12.7
Y37	807.9920	4034.9237	4034.9755	-0.0518	-12.8
Y39	845.2039	4220.9831	4221.0396	-0.0565	-13.4
Y61	725.1223	6517.0353	6517.1228	-0.0875	-13.4
Y61	815.6374	6517.0414	6517.1228	-0.0814	-12.5
Y71	757.8652	7568.5794	7568.6890	-0.1096	-14.5

3.4.1.2 HSP12 (13+)

UniProt P22943, HSP12
Seq:
SDAGRKGFGEKASEALKPDSQKSYAEQGKEYITDKADKVAGKVQPEDNKGVFQGVHDSAEKGKDN
AEGQGESLADQARDYMGAAKSKLNDAVEYVSGRVHGEEDPTKK
m/z monoisotopic = 893.0568, +13; MW_{obs} = 11596.64 Da (1.3 ppm)



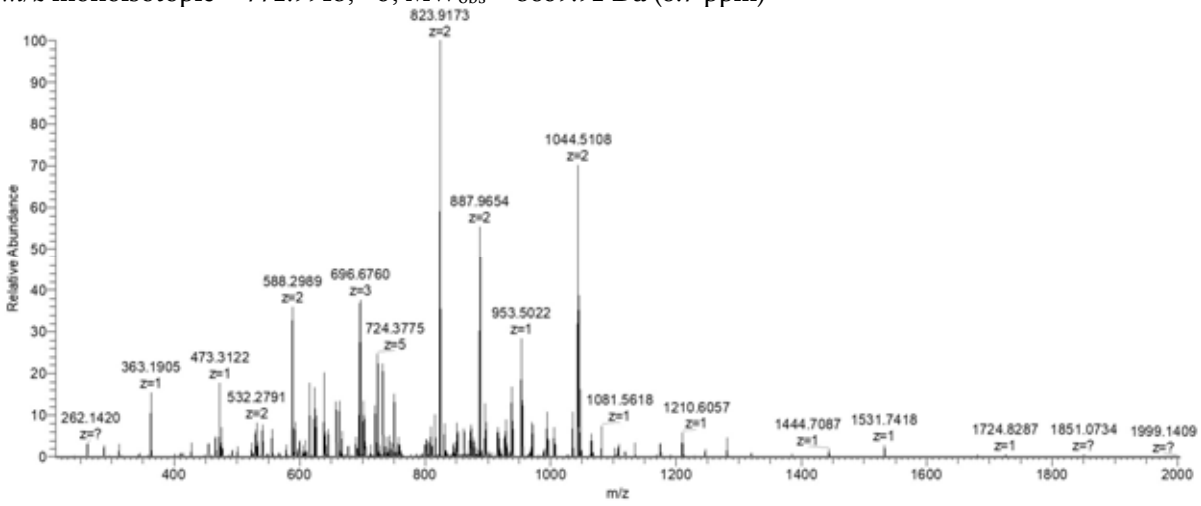
assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B6	657.3343	656.3270	656.3242	0.0029	4.3
B9	918.4470	917.4397	917.4355	0.0042	4.5
B10	524.2493	1046.4841	1046.4781	0.0060	5.8
B11	588.2956	1174.5767	1174.5731	0.0037	3.1
B12	623.8123	1245.6100	1245.6102	-0.0001	-0.1
B13	667.3323	1332.6501	1332.6422	0.0079	6.0
B15	767.3721	1532.7296	1532.7219	0.0077	5.0
B16	823.9103	1645.8060	1645.8060	0.0001	0.0
B17	592.3109	1773.9110	1773.9009	0.0101	5.7
B19	993.9975	1985.9804	1985.9806	-0.0002	-0.1
B19	663.0049	1985.9928	1985.9806	0.0122	6.2
B21	734.6993	2201.0759	2201.0712	0.0047	2.1
B24	860.7611	2579.2614	2579.2615	-0.0001	-0.1
B29	774.1378	3092.5222	3092.5162	0.0060	1.9
B30	806.4005	3221.5727	3221.5588	0.0139	4.3
B35	769.3800	3841.8634	3841.8758	-0.0124	-3.2
B35	961.4798	3841.8900	3841.8758	0.0142	3.7
B36	783.5881	3912.9042	3912.9129	-0.0087	-2.2
B37	806.5949	4027.9383	4027.9399	-0.0016	-0.4
B37	1007.9984	4027.9647	4027.9399	0.0248	6.2
B38	832.2176	4156.0517	4156.0348	0.0169	4.1
B40	866.2357	4326.1422	4326.1403	0.0018	0.4
B42	903.2558	4511.2424	4511.2568	-0.0143	-3.2
B44	948.6839	4738.3831	4738.3837	-0.0006	-0.1
B44	790.7384	4738.3865	4738.3837	0.0028	0.6
B46	828.4167	4964.4564	4964.4791	-0.0227	-4.6
B46	993.9001	4964.4641	4964.4791	-0.0150	-3.0

B47	1016.9070	5079.4987	5079.5060	-0.0073	-1.4
B47	847.5934	5079.5166	5079.5060	0.0106	2.1
B55	985.8344	5908.9630	5908.9507	0.0123	2.1
B56	864.7273	6046.0405	6046.0096	0.0309	5.1
B57	881.1579	6161.0541	6161.0365	0.0176	2.9
B61	940.4706	6576.2431	6576.2432	-0.0001	0.0
B63	846.1768	6761.3563	6761.3596	-0.0033	-0.5
B64	860.5554	6876.3847	6876.3866	-0.0018	-0.3
B69	922.9585	7375.6097	7375.5892	0.0205	2.8
B83	1106.0383	8840.2484	8840.2195	0.0288	3.3
B93	990.5847	9895.7743	9895.7806	-0.0063	-0.6
B94	915.4497	10058.8669	10058.8440	0.0229	2.3
B94	1006.8950	10058.8768	10058.8440	0.0328	3.3
Y2	275.2087	274.2014	274.2005	0.0009	3.5
Y4	473.3102	472.3029	472.3009	0.0020	4.2
Y5	588.3384	587.3311	587.3279	0.0032	5.5
Y6	717.3817	716.3744	716.3705	0.0039	5.5
Y8	903.4466	902.4394	902.4345	0.0049	5.4
Y13	720.3683	1438.7221	1438.7164	0.0057	3.9
Y14	513.6046	1537.7921	1537.7848	0.0073	4.7
Y14	769.9039	1537.7933	1537.7848	0.0085	5.5
Y15	851.4348	1700.8550	1700.8482	0.0069	4.0
Y15	567.9594	1700.8562	1700.8482	0.0081	4.7
Y16	610.9715	1829.8927	1829.8908	0.0019	1.1
Y16	915.9552	1829.8959	1829.8908	0.0051	2.8
Y17	965.4903	1928.9660	1928.9592	0.0068	3.5
Y18	1001.0104	2000.0063	1999.9963	0.0100	5.0
Y20	1115.5423	2229.0700	2229.0661	0.0039	1.7
Y21	781.7245	2342.1517	2342.1502	0.0015	0.6
Y21	1172.0841	2342.1537	2342.1502	0.0035	1.5
Y23	853.4365	2557.2878	2557.2772	0.0106	4.2
Y33	730.7667	3648.7971	3648.7954	0.0018	0.5
Y34	941.9611	3763.8153	3763.8223	-0.0070	-1.9
Y35	959.7241	3834.8673	3834.8594	0.0079	2.0
Y37	1009.7509	4034.9744	4034.9755	-0.0012	-0.3
Y37	808.0064	4034.9958	4034.9755	0.0203	5.0
Y39	1056.2681	4221.0432	4221.0396	0.0037	0.9
Y39	845.2168	4221.0476	4221.0396	0.0080	1.9
Y41	1102.5393	4406.1282	4406.1196	0.0086	1.9
Y47	837.7411	5020.4029	5020.3856	0.0173	3.4
Y48	859.0873	5148.4804	5148.4806	-0.0001	0.0

Y53	813.5367	5687.7061	5687.6782	0.0279	4.9
Y53	948.9585	5687.7073	5687.6782	0.0292	5.1
Y61	815.6471	6517.1186	6517.1228	-0.0041	-0.6
Y64	858.2875	6858.2415	6858.2451	-0.0036	-0.5
Y74	876.8829	7882.8810	7882.8480	0.0331	4.2

3.4.1.3 HSP12 N-terminal

UniProt P22943, HSP12 N-terminal fragment
Seq: SDAGRKGFGEKASEALKPDSQKSYAEQGKEYITDK
m/z monoisotopic = 772.9913, +5; MW_{obs} = 3859.92 Da (8.7 ppm)

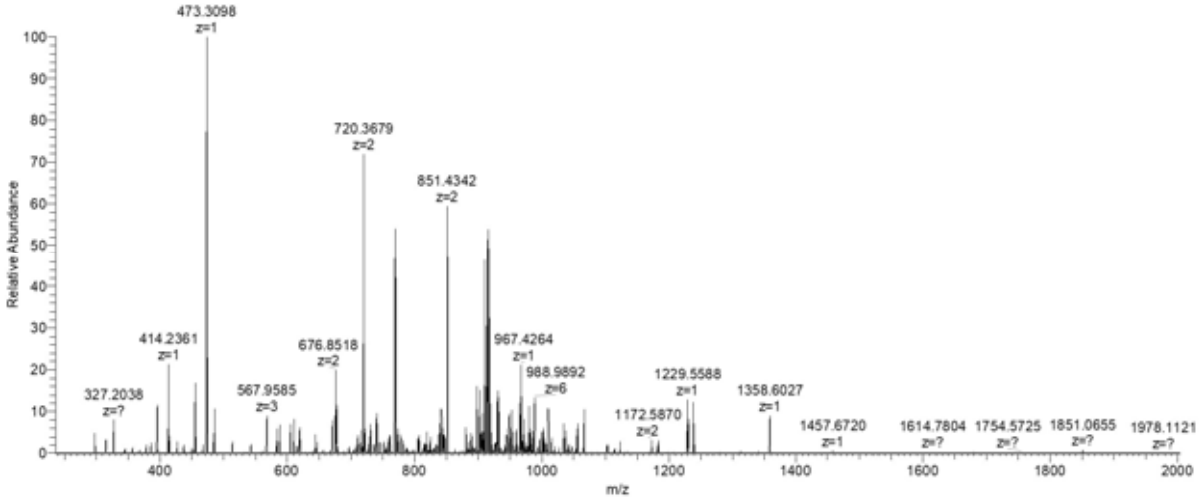


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B6	657.3339	656.3267	656.3242	0.0025	3.8
B8	861.4259	860.4186	860.4140	0.0046	5.3
B9	918.4491	917.4418	917.4355	0.0063	6.9
B10	1047.4899	1046.4826	1046.4781	0.0045	4.3
B10	524.2492	1046.4838	1046.4781	0.0057	5.4
B11	588.2969	1174.5793	1174.5731	0.0062	5.3
B13	667.3320	1332.6494	1332.6422	0.0072	5.4
B14	731.8525	1461.6905	1461.6848	0.0057	3.9
B16	823.9147	1645.8149	1645.8060	0.0089	5.4
B17	887.9578	1773.9010	1773.9009	0.0000	0.0
B17	592.3107	1773.9102	1773.9009	0.0093	5.2
B18	624.6609	1870.9608	1870.9537	0.0071	3.8
B19	993.9965	1985.9784	1985.9806	-0.0022	-1.1
B19	663.0045	1985.9917	1985.9806	0.0111	5.6
B21	734.6988	2201.0745	2201.0712	0.0032	1.5
B21	1101.5476	2201.0807	2201.0712	0.0094	4.3

B25	884.4416	2650.3029	2650.2987	0.0043	1.6
B26	927.4562	2779.3466	2779.3412	0.0054	1.9
B27	727.8612	2907.4157	2907.3998	0.0159	5.5
B30	806.3974	3221.5605	3221.5588	0.0017	0.5
B32	875.4356	3497.7135	3497.7062	0.0072	2.1
B34	929.4529	3713.7824	3713.7808	0.0016	0.4
B34	743.7644	3713.7857	3713.7808	0.0049	1.3
Y2	262.1408	261.1335	261.1325	0.0011	4.1
Y3	363.1888	362.1815	362.1801	0.0014	3.7
Y4	476.2728	475.2656	475.2642	0.0014	2.9
Y5	639.3380	638.3308	638.3275	0.0032	5.0
Y7	896.4763	895.4690	895.4651	0.0039	4.4
Y8	953.4975	952.4902	952.4865	0.0037	3.9
Y9	541.2830	1080.5514	1080.5451	0.0062	5.8
Y10	605.8037	1209.5928	1209.5877	0.0051	4.2
Y11	641.3221	1280.6296	1280.6248	0.0048	3.7
Y11	1281.6419	1280.6346	1280.6248	0.0098	7.6
Y12	722.8554	1443.6963	1443.6882	0.0081	5.6
Y13	1531.7405	1530.7332	1530.7202	0.0130	8.5
Y14	830.4198	1658.8251	1658.8151	0.0099	6.0
Y16	625.6412	1873.9018	1873.9058	-0.0040	-2.1
Y16	937.9630	1873.9115	1873.9058	0.0058	3.1
Y17	995.4738	1988.9331	1988.9327	0.0004	0.2
Y18	1043.9980	2085.9814	2085.9855	-0.0040	-1.9
Y18	696.3386	2085.9939	2085.9855	0.0084	4.0
Y19	739.0354	2214.0845	2214.0804	0.0041	1.8
Y19	1108.0512	2214.0879	2214.0804	0.0075	3.4
Y21	800.4097	2398.2073	2398.2016	0.0058	2.4
Y24	896.1147	2685.3222	2685.3133	0.0089	3.3
Y25	704.3614	2813.4166	2813.4083	0.0084	3.0
Y33	724.1744	3615.8355	3615.8168	0.0186	5.2
Y34	747.1736	3730.8316	3730.8438	-0.0122	-3.3

3.4.1.4 HSP12 C-terminal

UniProt P22943, HSP12 C-terminal fragment
Seq:
ADKVAGKVQPEDNKGVFQGVHDSAEKGKDNAEGQGESLADQARDYMGAAKSKLNDAVEYVSGR
VHGEEDPTKK
m/z monoisotopic = 862.6493, +9; MW_{obs} = 7754.78 Da (3.2 ppm)

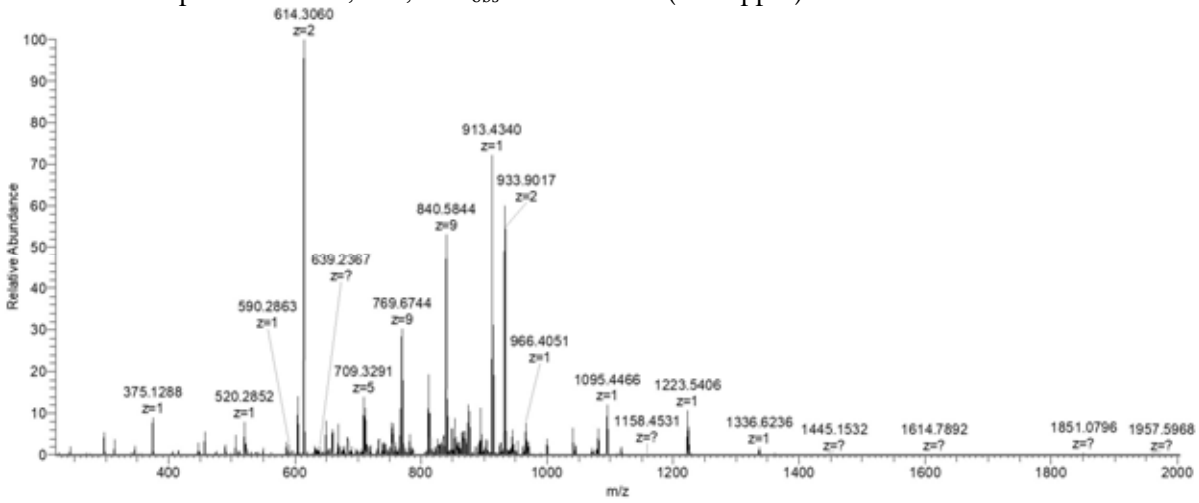


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B3	315.1672	314.1599	314.1590	0.0009	2.9
B4	414.2366	413.2293	413.2274	0.0019	4.5
B5	485.2742	484.2669	484.2645	0.0024	4.9
B7	670.3919	669.3847	669.3809	0.0037	5.6
B9	897.5197	896.5124	896.5079	0.0045	5.0
B12	619.8259	1237.6373	1237.6302	0.0071	5.7
B12	1238.6450	1237.6377	1237.6302	0.0075	6.1
B14	740.8961	1479.7776	1479.7681	0.0095	6.4
B20	1034.5404	2067.0663	2067.0749	-0.0085	-4.1
B21	735.7167	2204.1283	2204.1338	-0.0055	-2.5
B22	774.0630	2319.1671	2319.1607	0.0064	2.8
B28	730.8775	2919.4809	2919.4838	-0.0029	-1.0
B29	759.6350	3034.5108	3034.5108	0.0000	0.0
B31	805.9081	3219.6032	3219.5908	0.0124	3.9
B32	838.1667	3348.6378	3348.6334	0.0044	1.3
B34	884.4347	3533.7097	3533.7134	-0.0037	-1.0
B34	707.7523	3533.7249	3533.7134	0.0115	3.2
B36	744.9646	3719.7866	3719.7775	0.0091	2.5
B37	762.3683	3806.8052	3806.8095	-0.0043	-1.1
B44	916.2486	4576.2066	4576.1813	0.0253	5.5
B46	975.0646	4870.2866	4870.2852	0.0014	0.3
B47	986.4693	4927.3100	4927.3066	0.0034	0.7
B48	1000.6756	4998.3415	4998.3437	-0.0022	-0.4
B49	1014.8827	5069.3773	5069.3809	-0.0035	-0.7
B53	921.9568	5525.6969	5525.6869	0.0100	1.8
B55	823.1169	5754.7673	5754.7567	0.0106	1.8
B56	971.9728	5825.7930	5825.7938	-0.0009	-0.1

B56	833.2662	5825.8128	5825.7938	0.0189	3.2
B57	988.4830	5924.8544	5924.8623	-0.0078	-1.3
B57	847.4169	5924.8676	5924.8623	0.0053	0.9
B58	1009.9876	6053.8818	6053.9048	-0.0230	-3.8
B59	1037.1699	6216.9759	6216.9682	0.0077	1.2
B59	889.1474	6216.9807	6216.9682	0.0125	2.0
B60	903.2994	6316.0449	6316.0366	0.0083	1.3
B64	960.3332	6715.2816	6715.2596	0.0221	3.3
B65	979.9117	6852.3312	6852.3185	0.0127	1.9
B69	911.3146	7282.4586	7282.4521	0.0066	0.9
B72	952.0884	7608.6487	7608.6475	0.0012	0.2
Y4	473.3091	472.3018	472.3009	0.0009	1.8
Y6	717.3795	716.3722	716.3705	0.0018	2.5
Y8	903.4435	902.4363	902.4345	0.0018	1.9
Y12	676.8553	1351.6960	1351.6844	0.0116	8.6
Y13	720.3703	1438.7260	1438.7164	0.0096	6.7
Y14	513.6046	1537.7920	1537.7848	0.0072	4.7
Y14	769.9038	1537.7930	1537.7848	0.0082	5.3
Y15	851.4346	1700.8547	1700.8482	0.0065	3.8
Y15	567.9593	1700.8561	1700.8482	0.0079	4.6
Y16	610.9698	1829.8876	1829.8908	-0.0031	-1.7
Y16	915.9550	1829.8954	1829.8908	0.0047	2.6
Y17	965.4900	1928.9655	1928.9592	0.0063	3.3
Y17	643.9966	1928.9680	1928.9592	0.0088	4.6
Y33	730.7645	3648.7860	3648.7954	-0.0094	-2.6
Y35	959.7239	3834.8664	3834.8594	0.0069	1.8
Y37	1009.7506	4034.9733	4034.9755	-0.0022	-0.6
Y39	1056.2715	4221.0568	4221.0396	0.0172	4.1
Y41	1102.5429	4406.1425	4406.1196	0.0229	5.2
Y59	897.4361	6275.0015	6274.9849	0.0166	2.6
Y61	932.0226	6517.1073	6517.1228	-0.0155	-2.4
Y61	815.6494	6517.1373	6517.1228	0.0145	2.2
Y64	980.7561	6858.2415	6858.2451	-0.0035	-0.5
Y68	909.8177	7270.4837	7270.4885	-0.0048	-0.7
Y69	918.6996	7341.5388	7341.5256	0.0132	1.8
Y70	931.0812	7440.5912	7440.5940	-0.0028	-0.4
Y71	841.9701	7568.6658	7568.6890	-0.0232	-3.1

3.4.1.5 SIP18 (11+)

UniProt P50263, SIP18
Seq:
SNMMNKFAEKLQGNDDSHQKGKNAKSSNKERDDMNMDMGMGHDQSEGGMKMGHDQSGTK
MNAGRGIANDWKTYENMKK
m/z monoisotopic = 799.0788, +11; MW_{obs} = 8778.79 Da (-11.3 ppm)

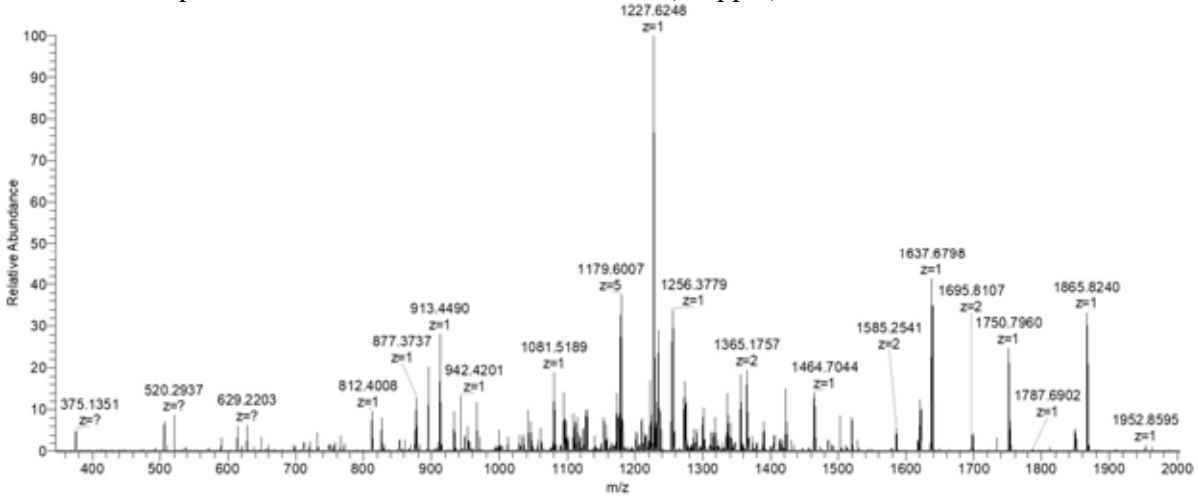


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B2	244.0899	243.0826	243.0855	-0.0029	-11.9
B3	375.1295	374.1222	374.1260	-0.0038	-10.1
B8	966.4085	965.4012	965.4099	-0.0087	-9.0
B9	1095.4469	1094.4396	1094.4525	-0.0129	-11.8
B10	1223.5449	1222.5376	1222.5475	-0.0098	-8.0
B11	1336.6280	1335.6207	1335.6315	-0.0108	-8.1
B11	668.8177	1335.6209	1335.6315	-0.0106	-8.0
B12	732.8434	1463.6723	1463.6901	-0.0178	-12.2
B15	875.8905	1749.7665	1749.7814	-0.0149	-8.5
B18	1045.4446	2088.8747	2088.8993	-0.0246	-11.8
B25	711.8340	2843.3070	2843.3443	-0.0373	-13.1
B32	732.9414	3659.6708	3659.7169	-0.0461	-12.6
B33	755.9487	3774.7072	3774.7438	-0.0366	-9.7
B37	854.1761	4265.8441	4265.8947	-0.0505	-11.8
B48	1071.4486	5352.2068	5352.2805	-0.0737	-13.8
B55	773.4505	6179.5459	6179.6224	-0.0765	-12.4
B69	945.0290	7552.1734	7552.2741	-0.1006	-13.3
B69	840.1382	7552.1786	7552.2741	-0.0954	-12.6
B70	860.8145	7738.2647	7738.3534	-0.0887	-11.5
Y4	520.2857	519.2785	519.2839	-0.0054	-10.5
Y5	649.3276	648.3203	648.3265	-0.0062	-9.6

Y7	457.2210	912.4274	912.4375	-0.0101	-11.1
Y8	521.2681	1040.5217	1040.5325	-0.0107	-10.3
Y8	1041.5309	1040.5237	1040.5325	-0.0088	-8.5
Y9	614.3076	1226.6006	1226.6118	-0.0112	-9.1
Y19	743.0193	2226.0360	2226.0674	-0.0314	-14.1
Y24	682.8292	2727.2878	2727.3221	-0.0343	-12.6
Y27	760.1041	3036.3872	3036.4294	-0.0421	-13.9
Y28	634.4937	3167.4321	3167.4699	-0.0378	-11.9
Y32	886.1579	3540.6025	3540.6483	-0.0457	-12.9
Y34	752.3425	3756.6760	3756.7229	-0.0468	-12.5
Y37	828.3699	4136.8131	4136.8673	-0.0542	-13.1
Y40	877.3868	4381.8978	4381.9507	-0.0529	-12.1
Y41	753.1633	4512.9363	4512.9912	-0.0549	-12.2
Y41	903.5961	4512.9442	4512.9912	-0.0470	-10.4
Y59	821.2365	6561.8335	6561.9280	-0.0944	-14.4
Y66	915.3951	7315.1023	7315.1958	-0.0934	-12.8
Y68	756.6299	7556.2259	7556.3384	-0.1125	-14.9
Y69	769.4422	7684.3493	7684.4334	-0.0841	-10.9
Y75	841.4731	8404.6584	8404.7598	-0.1015	-12.1

3.4.1.6 SIP18 (7+)

UniProt P50263, SIP18
Seq:
SNMMNKFAEKLQGNDDSHQKGKNAKSSNKERDDMNMDMGMGHDQSEGGMKMGHDQSGTK
MNAGRGIANDWKTYENMKK
m/z monoisotopic = 1255.1402, +7; MW_{obs} = 8778.93 Da (5.1 ppm)

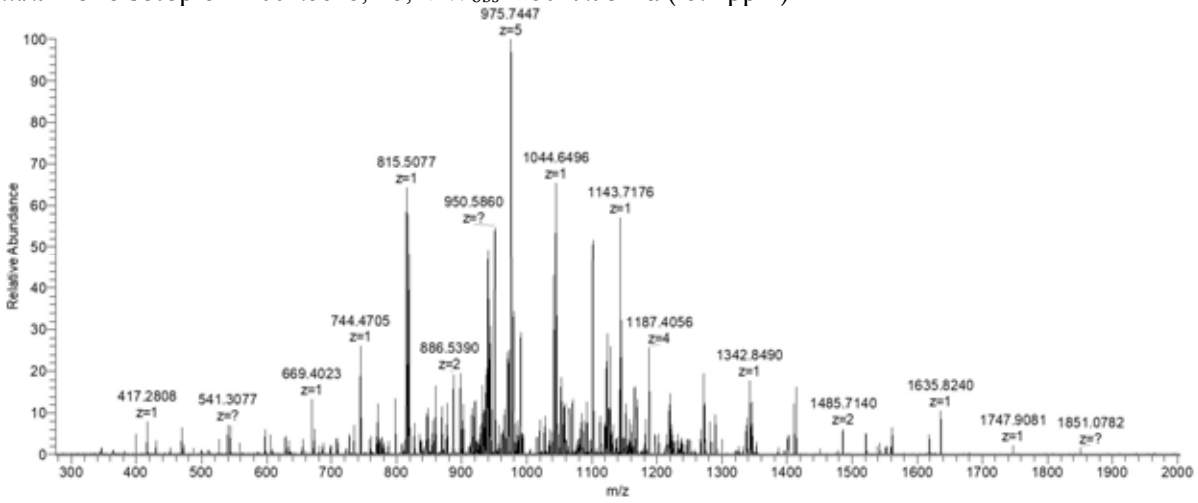


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm

B7	895.3822	894.3750	894.3728	0.0022	2.4
B8	966.4252	965.4179	965.4099	0.0080	8.3
B9	1095.4675	1094.4602	1094.4525	0.0077	7.0
B10	1223.5652	1222.5579	1222.5475	0.0105	8.6
B11	1336.6516	1335.6443	1335.6315	0.0128	9.6
B12	1464.7084	1463.7011	1463.6901	0.0110	7.5
B15	875.9046	1749.7947	1749.7814	0.0133	7.6
B15	1750.8035	1749.7962	1749.7814	0.0148	8.4
B16	933.4157	1864.8169	1864.8084	0.0086	4.6
B16	1865.8264	1864.8191	1864.8084	0.0107	5.8
B18	1045.4601	2088.9057	2088.8993	0.0064	3.0
B19	1109.4879	2216.9613	2216.9579	0.0034	1.5
B37	1422.9761	4265.9065	4265.8947	0.0118	2.8
B76	1215.9636	8504.6945	8504.6854	0.0091	1.1
B77	1234.2622	8632.7843	8632.7803	0.0040	0.5
Y6	812.4017	811.3944	811.3898	0.0046	5.6
Y7	913.4526	912.4453	912.4375	0.0078	8.6
Y8	1041.5463	1040.5390	1040.5325	0.0066	6.3
Y9	1227.6259	1226.6186	1226.6118	0.0069	5.6
Y9	614.3184	1226.6222	1226.6118	0.0105	8.5
Y10	1342.6502	1341.6429	1341.6387	0.0042	3.1
Y14	1697.8468	1696.8395	1696.8243	0.0152	9.0
Y16	955.9748	1909.9351	1909.9468	-0.0117	-6.1
Y18	1048.5237	2095.0328	2095.0269	0.0059	2.8
Y19	1114.0499	2226.0852	2226.0674	0.0178	8.0
Y23	1300.6449	2599.2752	2599.2635	0.0117	4.5
Y24	1364.6649	2727.3153	2727.3221	-0.0068	-2.5
Y25	1422.1852	2842.3558	2842.3490	0.0067	2.4
Y27	1519.2285	3036.4424	3036.4294	0.0130	4.3
Y28	1056.8376	3167.4911	3167.4699	0.0212	6.7
Y28	1584.7571	3167.4997	3167.4699	0.0298	9.4
Y32	1181.2219	3540.6438	3540.6483	-0.0044	-1.2
Y36	1334.2780	3999.8123	3999.8084	0.0039	1.0
Y37	1035.2223	4136.8600	4136.8673	-0.0074	-1.8
Y41	1129.2609	4513.0146	4512.9912	0.0234	5.2
Y62	1153.3550	6914.0862	6914.0775	0.0087	1.3
Y72	1339.6091	8031.6109	8031.5815	0.0294	3.7

3.4.1.7 EF1A

UniProt P02994, EF1A
Seq: FSEYPPLGRFAVRDMRQTVAVGVIKSVDKTEKAAKVTKAAQKAAKK
m/z monoisotopic = 1004.5645, +5; MW_{obs} = 5017.78 Da (-3.1 ppm)

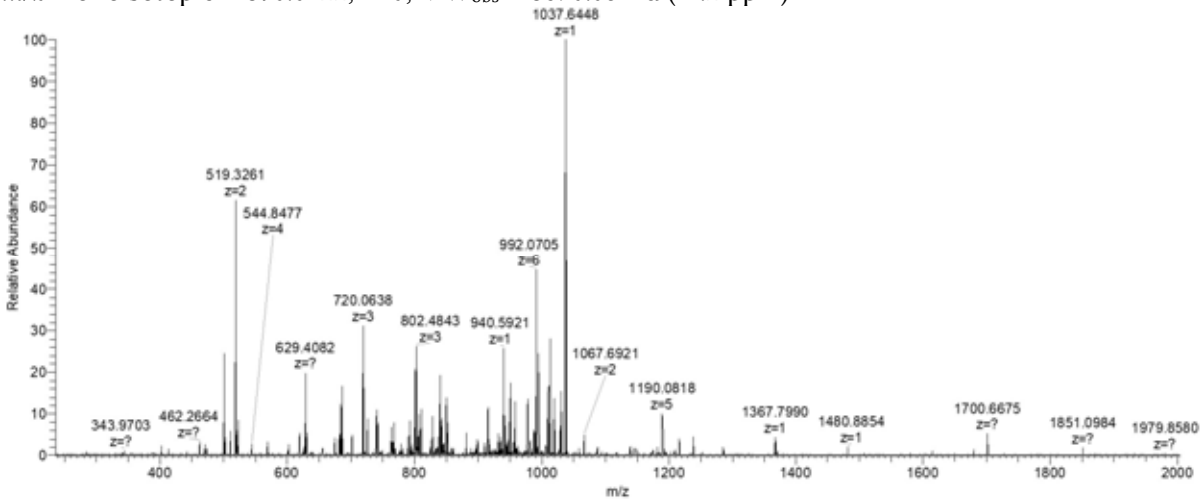


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B39	1069.5893	4274.3279	4274.3361	-0.0082	-1.9
B40	1087.3452	4345.3518	4345.3732	-0.0214	-4.9
B41	1119.3602	4473.4118	4473.4318	-0.0200	-4.5
B42	1151.3868	4601.5181	4601.5267	-0.0086	-1.9
B43	1169.1445	4672.5490	4672.5638	-0.0148	-3.2
B43	935.5183	4672.5553	4672.5638	-0.0085	-1.8
B44	1186.9003	4743.5721	4743.6010	-0.0289	-6.1
B44	949.7263	4743.5951	4743.6010	-0.0059	-1.2
B45	975.3390	4871.6586	4871.6959	-0.0373	-7.7
B45	1218.9230	4871.6628	4871.6959	-0.0331	-6.8
Y3	346.2442	345.2369	345.2376	-0.0007	-1.9
Y7	744.4699	743.4626	743.4654	-0.0028	-3.7
Y8	815.5080	814.5008	814.5025	-0.0017	-2.1
Y9	943.5985	942.5913	942.5974	-0.0062	-6.5
Y11	1143.7231	1142.7158	1142.7135	0.0023	2.0
Y12	1271.8141	1270.8068	1270.8085	-0.0017	-1.3
Y13	1342.8489	1341.8416	1341.8456	-0.0040	-3.0
Y14	1413.8852	1412.8780	1412.8827	-0.0047	-3.4
Y14	707.4486	1412.8826	1412.8827	-0.0001	0.0
Y15	771.4943	1540.9741	1540.9777	-0.0036	-2.3
Y16	836.0174	1670.0203	1670.0202	0.0000	0.0
Y18	950.5847	1899.1548	1899.1629	-0.0081	-4.3
Y20	1057.6328	2113.2510	2113.2582	-0.0072	-3.4

Y21	1101.1464	2200.2782	2200.2903	-0.0120	-5.5
Y22	1165.1933	2328.3721	2328.3852	-0.0131	-5.6
Y22	777.1328	2328.3767	2328.3852	-0.0085	-3.7
Y23	1221.7343	2441.4540	2441.4693	-0.0153	-6.3
Y23	814.8312	2441.4718	2441.4693	0.0025	1.0
Y32	1128.6578	3382.9516	3382.9809	-0.0293	-8.7
Y32	846.7491	3382.9671	3382.9809	-0.0138	-4.1
Y33	875.5004	3497.9725	3498.0079	-0.0354	-10.1
Y42	1123.9026	4491.5813	4491.5951	-0.0138	-3.1
Y43	931.9339	4654.6333	4654.6584	-0.0251	-5.4
Y45	975.1434	4870.6806	4870.7330	-0.0524	-10.8

3.4.1.8 L38

UniProt P49167, L38
Seq: FSEYPPLGRFAVRDMRQTVAVGVIKSVDKTEKAAKVTKAAQKAAKK
m/z monoisotopic = 870.0122, +10; MW_{obs} = 8690.05 Da (-1.2 ppm)

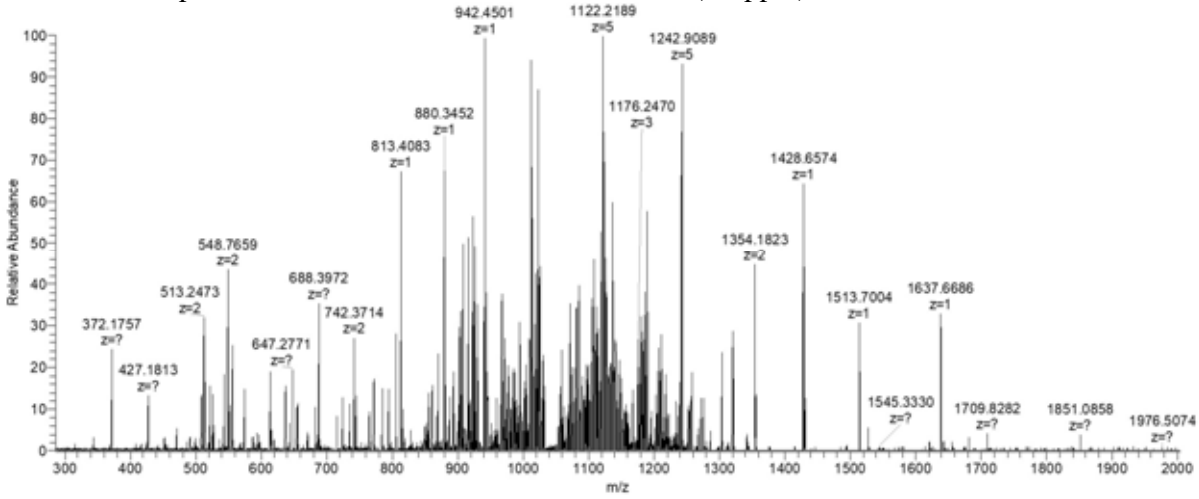


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B6	686.3472	685.3399	685.3395	0.0004	0.6
B28	810.7200	3238.8507	3238.8513	-0.0006	-0.2
B29	842.7386	3366.9253	3366.9463	-0.0210	-6.2
B52	994.5645	5961.3432	5961.3908	-0.0475	-8.0
B53	1013.4179	6074.4635	6074.4748	-0.0113	-1.9
Y4	501.3138	500.3065	500.3071	-0.0005	-1.1
Y5	629.4064	628.3991	628.4020	-0.0030	-4.7
Y8	940.5956	939.5883	939.5865	0.0018	1.9
Y9	519.3269	1036.6392	1036.6393	-0.0001	-0.1
Y9	1037.6488	1036.6415	1036.6393	0.0022	2.1

Y11	619.3840	1236.7535	1236.7554	-0.0019	-1.5
Y12	683.4127	1364.8109	1364.8140	-0.0031	-2.3
Y19	702.1148	2103.3225	2103.3255	-0.0031	-1.5
Y20	725.7942	2174.3608	2174.3626	-0.0018	-0.8
Y21	764.1348	2289.3827	2289.3896	-0.0069	-3.0
Y23	630.1351	2516.5112	2516.5166	-0.0053	-2.1
Y31	842.7386	3366.9253	3366.9602	-0.0349	-10.4

3.4.1.9 ZEO1

UniProt Q08245, ZEO1
Seq:
SEIQNKAETAAQDVQQKLEETKESLQNKGQEVKEQAEASIDNLKNEATPEAEQVKKEEQNIADGVE
QKKTEAANKVEETKKQASAAVSEKKETKKEGGFLKKLNRKIASIFN
m/z monoisotopic = 1042.0490, +12; MW_{obs} = 12492.50 Da (2.8 ppm)

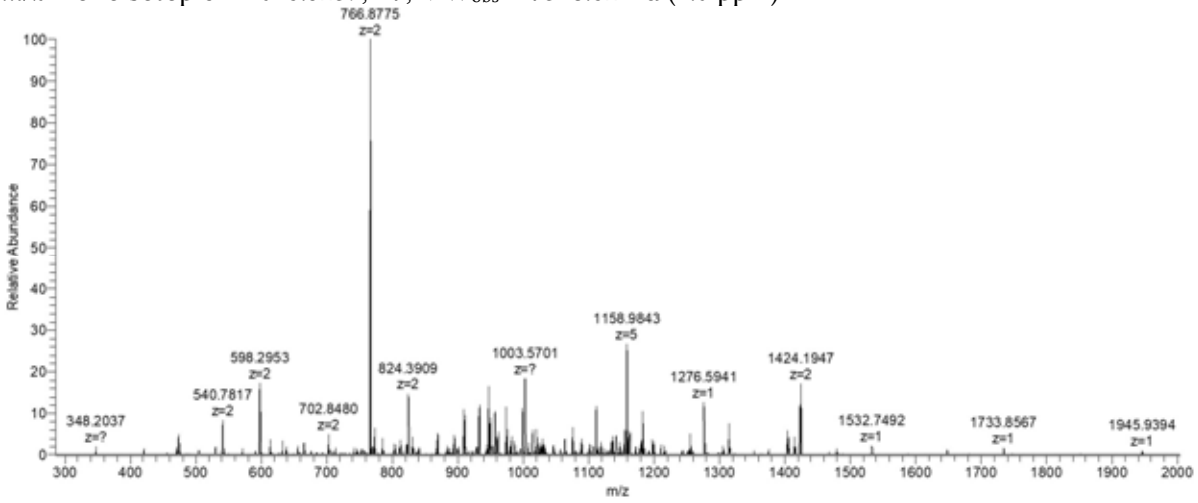


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B7	813.4077	812.4005	812.4028	-0.0024	-2.9
B9	522.2539	1042.4933	1042.4931	0.0002	0.2
B10	1114.5370	1113.5297	1113.5302	-0.0005	-0.4
B11	1185.5773	1184.5700	1184.5673	0.0027	2.3
B13	1428.6630	1427.6558	1427.6528	0.0029	2.0
B20	1142.0541	2282.0937	2282.1026	-0.0089	-3.9
B28	1071.2075	3210.6007	3210.6004	0.0003	0.1
B31	1175.9120	3524.7140	3524.7230	-0.0090	-2.5
B33	938.9705	3751.8529	3751.8864	-0.0335	-8.9
B34	971.2313	3880.8959	3880.9290	-0.0331	-8.5
B35	1003.2483	4008.9640	4008.9876	-0.0236	-5.9
B41	1149.8167	4595.2375	4595.2474	-0.0099	-2.1

B43	1206.5961	4822.3552	4822.3744	-0.0192	-4.0
B71	1140.1345	7973.8904	7973.9377	-0.0473	-5.9
Y16	903.5361	1805.0577	1805.0675	-0.0099	-5.5
Y17	968.0612	1934.1079	1934.1101	-0.0023	-1.2
Y20	1146.6731	2291.3316	2291.3477	-0.0161	-7.0
Y21	1211.1973	2420.3799	2420.3903	-0.0104	-4.3
Y23	893.2010	2676.5811	2676.5802	0.0009	0.3
Y30	1098.2864	3291.8373	3291.8666	-0.0293	-8.9
Y31	1140.9782	3419.9127	3419.9252	-0.0125	-3.7
Y32	888.0073	3548.0000	3548.0202	-0.0201	-5.7
Y34	945.2966	3777.1571	3777.1628	-0.0057	-1.5
Y36	1009.8166	4035.2374	4035.2480	-0.0105	-2.6
Y41	904.7075	4518.5011	4518.5285	-0.0274	-6.1
Y44	976.3474	4876.7006	4876.7137	-0.0132	-2.7
Y49	1084.6020	5417.9736	5417.9997	-0.0262	-4.8
Y49	904.0043	5417.9821	5417.9997	-0.0177	-3.3
Y50	1107.6051	5532.9893	5533.0267	-0.0374	-6.8
Y51	1121.8089	5604.0080	5604.0638	-0.0557	-9.9
Y64	1019.1196	7126.7862	7126.8265	-0.0403	-5.7
Y64	1188.8094	7126.8128	7126.8265	-0.0137	-1.9
Y68	1078.4260	7541.9309	7541.9968	-0.0659	-8.7

3.4.1.10 FMP16

UniProt Q12497, FMP16
Seq:
PHLKNNQDEAEKKEQGLFDSNKKRLDTLEHGKNPDYKQPGMEDLKKKGDDARIEQNRPDDGVY
m/z monoisotopic = 1046.5239, +7; MW_{obs} = 7318.62 Da (1.0 ppm)

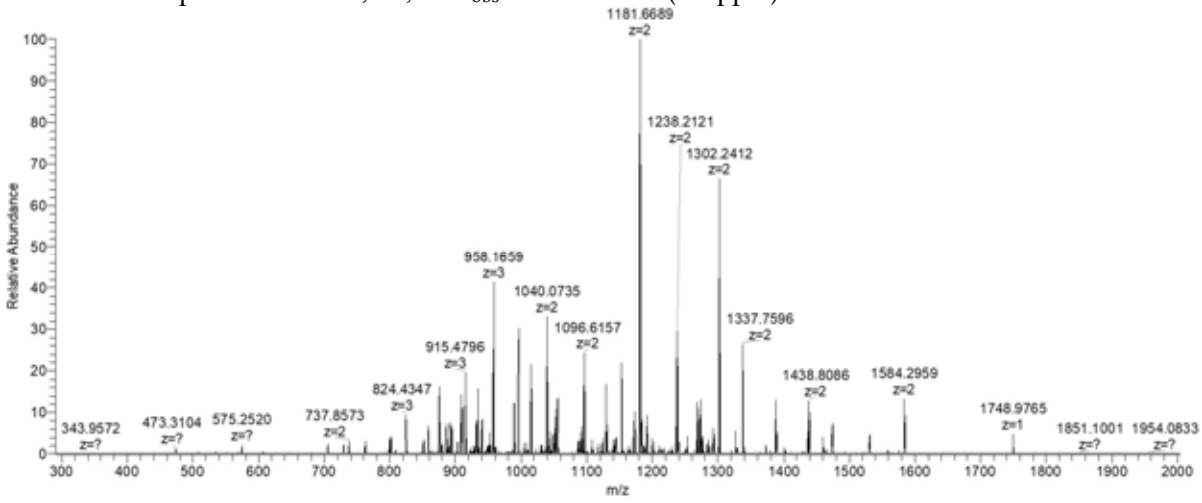


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B3	348.2037	347.1964	347.1957	0.0007	2.0
B4	476.2997	475.2924	475.2907	0.0017	3.6
B8	947.4728	946.4655	946.4621	0.0034	3.6
B9	1076.5119	1075.5046	1075.5047	-0.0001	0.0
B11	638.8012	1275.5879	1275.5844	0.0035	2.7
B11	1276.5957	1275.5884	1275.5844	0.0041	3.2
B12	702.8491	1403.6836	1403.6793	0.0043	3.0
B12	1404.6976	1403.6903	1403.6793	0.0110	7.8
B13	766.8975	1531.7804	1531.7743	0.0061	4.0
B14	831.4197	1660.8249	1660.8169	0.0081	4.9
B15	895.4450	1788.8754	1788.8754	0.0000	0.0
B16	923.9561	1845.8976	1845.8969	0.0007	0.4
B17	980.4968	1958.9791	1958.9810	-0.0019	-1.0
B19	1111.5466	2221.0787	2221.0763	0.0024	1.1
B29	1136.2490	3405.7253	3405.7276	-0.0024	-0.7
B30	886.7002	3542.7718	3542.7866	-0.0147	-4.2
B31	900.9546	3599.7892	3599.8080	-0.0188	-5.2
B32	1243.6443	3727.9111	3727.9030	0.0082	2.2
B32	932.9852	3727.9116	3727.9030	0.0087	2.3
B33	961.4962	3841.9558	3841.9459	0.0099	2.6
B35	1014.5138	4054.0261	4054.0256	0.0005	0.1
B37	1087.3032	4345.1838	4345.1839	-0.0001	0.0
B38	895.6527	4473.2272	4473.2425	-0.0152	-3.4
B38	1119.3244	4473.2684	4473.2425	0.0259	5.8
B46	1075.3525	5371.7261	5371.7007	0.0254	4.7
B47	1100.9638	5499.7825	5499.7957	-0.0132	-2.4
B49	1135.3698	5671.8126	5671.8441	-0.0315	-5.6
B50	1158.3760	5786.8436	5786.8710	-0.0274	-4.7
B59	981.9249	6866.4235	6866.4182	0.0053	0.8
B60	998.3511	6981.4068	6981.4451	-0.0383	-5.5
B62	1020.6500	7137.4989	7137.5350	-0.0361	-5.1
Y6	665.2799	664.2726	664.2704	0.0022	3.3
Y9	1063.4830	1062.4757	1062.4730	0.0027	2.6
Y11	1305.6161	1304.6088	1304.5997	0.0091	7.0
Y13	766.8793	1531.7440	1531.7379	0.0061	4.0
Y14	824.3913	1646.7681	1646.7648	0.0033	2.0
Y15	881.9063	1761.7981	1761.7918	0.0063	3.6
Y16	910.4120	1818.8094	1818.8132	-0.0039	-2.1
Y17	974.4625	1946.9104	1946.9082	0.0022	1.1
Y19	1102.5596	2203.1046	2203.0981	0.0065	2.9

Y20	773.0711	2316.1914	2316.1822	0.0092	4.0
Y21	1216.6191	2431.2236	2431.2091	0.0145	5.9
Y24	1375.1616	2748.3087	2748.3137	-0.0049	-1.8
Y25	949.4586	2845.3539	2845.3664	-0.0126	-4.4
Y25	1423.7002	2845.3859	2845.3664	0.0195	6.9
Y28	1089.1997	3264.5773	3264.5833	-0.0060	-1.8
Y30	870.1709	3476.6547	3476.6630	-0.0083	-2.4
Y30	1159.8928	3476.6564	3476.6630	-0.0066	-1.9
Y31	1197.9030	3590.6871	3590.7059	-0.0188	-5.2
Y31	898.6810	3590.6949	3590.7059	-0.0110	-3.1
Y33	944.9638	3775.8262	3775.8223	0.0039	1.0
Y34	979.2232	3912.8635	3912.8812	-0.0177	-4.5
Y44	1020.5130	5097.5287	5097.5326	-0.0039	-0.8
Y48	1106.9505	5529.7163	5529.7334	-0.0172	-3.1
Y50	1158.3760	5786.8436	5786.8346	0.0090	1.6
Y51	1183.9913	5914.9202	5914.9296	-0.0094	-1.6
Y51	986.8295	5914.9333	5914.9296	0.0038	0.6
Y52	1008.1771	6043.0189	6043.0245	-0.0056	-0.9
Y59	1141.5642	6843.3415	6843.3182	0.0233	3.4
Y61	1181.7515	7084.4652	7084.4972	-0.0320	-4.5

3.4.1.11 OM14 C-terminal

UniProt P38325, OM14
Seq: LCNGYANHNARFLKGKPNSTVLATTAGALGLLTLDGIISKKYYSRYDKK
m/z monoisotopic = 1069.7826, +5; MW_{obs} = 5343.88 Da (4.0 ppm)

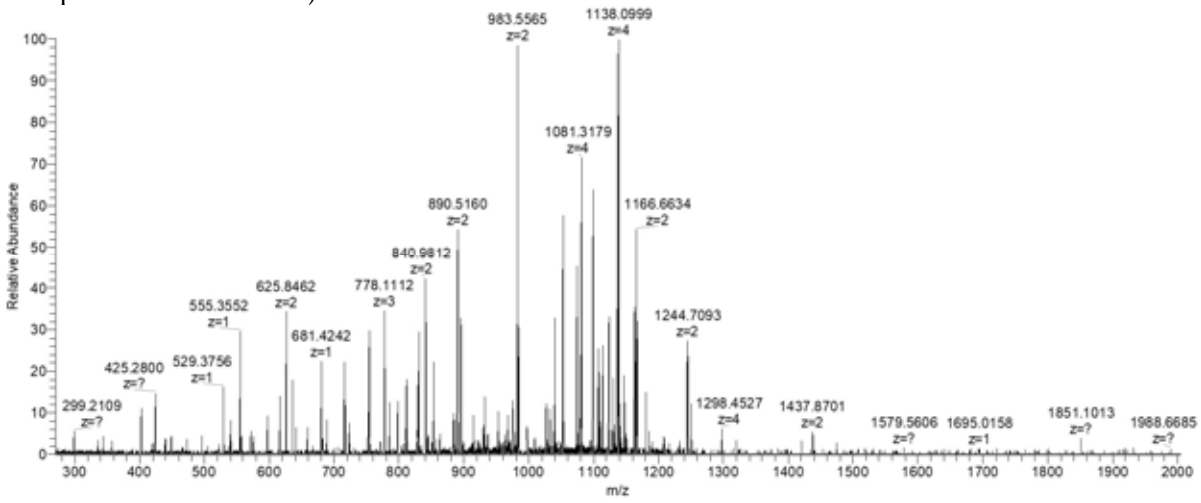


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B13	737.8585	1473.7025	1473.6935	0.0090	6.1

B14	801.9083	1601.8021	1601.7885	0.0136	8.5
B16	894.4658	1786.9171	1786.9049	0.0121	6.8
B21	1143.5838	2285.1531	2285.1487	0.0044	1.9
B21	762.7266	2285.1579	2285.1487	0.0092	4.0
B22	1200.1236	2398.2327	2398.2328	0.0000	0.0
B22	800.4203	2398.2391	2398.2328	0.0063	2.6
B23	824.0961	2469.2665	2469.2699	-0.0034	-1.4
B23	1235.6430	2469.2715	2469.2699	0.0016	0.7
B24	1286.1653	2570.3161	2570.3176	-0.0015	-0.6
B24	857.7833	2570.3280	2570.3176	0.0104	4.0
B26	915.1393	2742.3961	2742.4024	-0.0063	-2.3
B26	1372.2145	2742.4144	2742.4024	0.0121	4.4
B27	934.1512	2799.4319	2799.4238	0.0081	2.9
B28	957.8307	2870.4702	2870.4609	0.0093	3.2
B28	1436.2494	2870.4843	2870.4609	0.0234	8.1
B29	995.5256	2983.5550	2983.5450	0.0100	3.3
B30	1014.5261	3040.5565	3040.5665	-0.0100	-3.3
B31	1052.2280	3153.6623	3153.6505	0.0118	3.7
B32	1089.9154	3266.7245	3266.7346	-0.0101	-3.1
B33	1123.5979	3367.7719	3367.7823	-0.0104	-3.1
B47	1268.4245	5069.6687	5069.6549	0.0138	2.7
B48	1040.5641	5197.7843	5197.7498	0.0344	6.6
Y8	1122.5581	1121.5508	1121.5506	0.0002	0.2
Y14	1748.9778	1747.9706	1747.9621	0.0085	4.9
Y14	874.9930	1747.9715	1747.9621	0.0094	5.4
Y15	932.5035	1862.9924	1862.9890	0.0034	1.8
Y16	989.0428	1976.0711	1976.0731	-0.0020	-1.0
Y17	1039.5699	2077.1253	2077.1208	0.0046	2.2
Y18	1096.1158	2190.2171	2190.2048	0.0122	5.6
Y19	1152.6507	2303.2868	2303.2889	-0.0020	-0.9
Y20	1181.1675	2360.3205	2360.3103	0.0102	4.3
Y21	1237.7092	2473.4039	2473.3944	0.0095	3.8
Y22	1273.2288	2544.4431	2544.4315	0.0116	4.6
Y23	1301.7339	2601.4532	2601.4530	0.0002	0.1
Y24	1337.2504	2672.4862	2672.4901	-0.0038	-1.4
Y25	1387.7812	2773.5479	2773.5378	0.0101	3.7
Y26	1438.3016	2874.5887	2874.5854	0.0033	1.1
Y27	1473.8181	2945.6217	2945.6226	-0.0008	-0.3
Y28	1530.3755	3058.7364	3058.7066	0.0298	9.7
Y33	1186.6539	3556.9400	3556.9504	-0.0104	-2.9

3.4.1.12 UPF0495

UniProt A5Z2X5, UPF0495 protein YPR010C-A
Seq: NTAKKTSGGYKIPVELTPLFLAVGVALCSGTYFTYKKLRTDETLRLTGNPEL SSLDEVLAKDKD
m/z monoisotopic = 997.3978, +7; MW_{obs} = 6974.73 Da (4.9 ppm)
cleavage of initial 8 amino acids not predicted in database; **Δm mode active** (takes into account unexpected modifications)



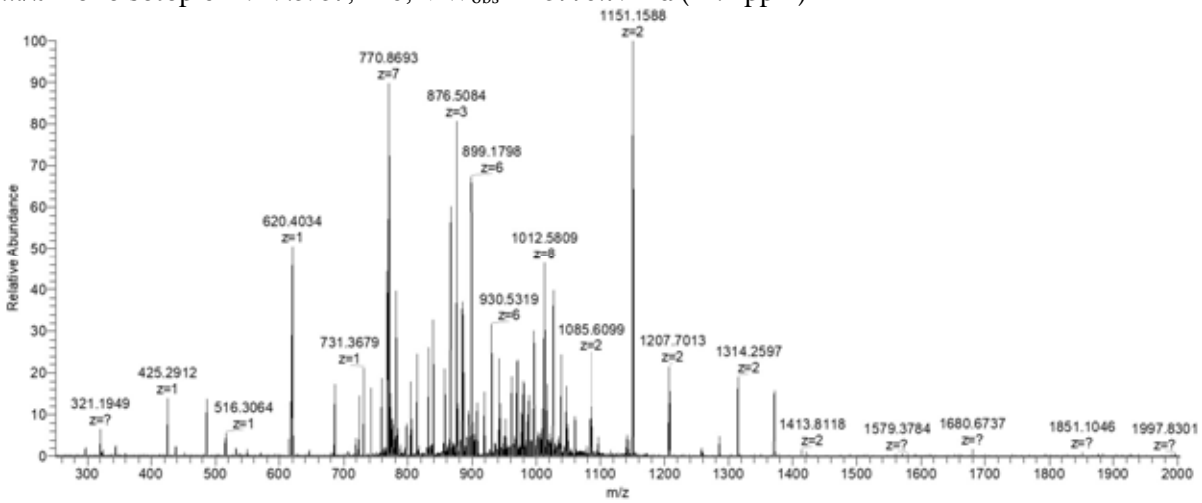
assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm	Δm
B28	1074.1005	2146.1865	3067.7256	-0.0448	-14.6	TRUE
B29	1130.6432	2259.2718	3180.8096	-0.0435	-13.7	TRUE
B30	1166.1611	2330.3077	3251.8468	-0.0447	-13.7	TRUE
B30	777.7790	2330.3153	3251.8468	-0.0372	-11.4	TRUE
B32	829.8048	2486.3926	3407.9366	-0.0497	-14.6	TRUE
Y36	1026.7910	4103.1349	4103.1062	0.0287	7.0	
Y37	1052.5470	4206.1589	4206.1154	0.0435	10.3	
Y38	1080.8109	4319.2144	4319.1995	0.0150	3.5	
Y39	1098.5713	4390.2561	4390.2366	0.0196	4.5	
Y40	1123.3435	4489.3448	4489.3050	0.0399	8.9	
Y41	1137.5994	4546.3684	4546.3264	0.0420	9.2	
Y42	1162.3650	4645.4307	4645.3948	0.0359	7.7	

3.4.2 *C. glabrata*

3.4.2.1 H2A.1

UniProt Q6FWM7, histone H2A.1
Seq:
SGGKGGKAGSAAKASQTRSAKAGLTFPVGRVHRLLRRGNYAQRIGSGAPVYLTAVLEYLAAEILELA GNAARDNKKSRIIPRHLQLAIRNDDELNKLLGNVTIAQGGVLPNIHQNLLPKKSAKPSASQEL

m/z monoisotopic = 919.3937, +15; MW_{obs} = 13775.79 Da (14.4 ppm)



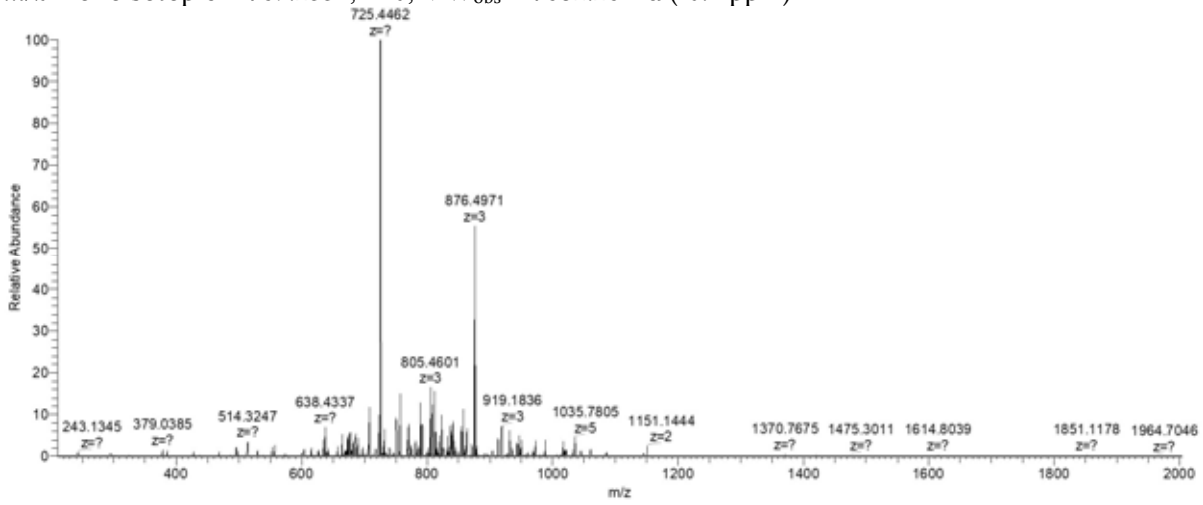
assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B50	838.9740	5027.8003	5027.7618	0.0385	7.7
B50	719.2685	5027.8287	5027.7618	0.0669	13.3
B51	866.1554	5190.8889	5190.8251	0.0637	12.3
B51	742.5644	5190.8999	5190.8251	0.0747	14.4
B52	758.7179	5303.9743	5303.9092	0.0651	12.3
B52	885.0052	5303.9874	5303.9092	0.0783	14.8
B53	773.1534	5405.0226	5404.9569	0.0657	12.2
B54	783.3025	5476.0667	5475.9940	0.0728	13.3
B55	797.4520	5575.1131	5575.0624	0.0507	9.1
B57	832.0463	5817.2734	5817.1890	0.0843	14.5
B60	881.6437	6164.4547	6164.3735	0.0812	13.2
Y18	988.5577	1975.1009	1975.0850	0.0159	8.0
Y21	1150.6562	2299.2978	2299.2648	0.0331	14.4
Y22	1207.1931	2412.3717	2412.3489	0.0228	9.5
Y22	805.1334	2412.3783	2412.3489	0.0294	12.2
Y25	1313.7569	2625.4992	2625.4602	0.0390	14.9
Y71	961.3024	7682.3613	7682.2611	0.1002	13.0
Y72	975.4374	7795.4411	7795.3452	0.0960	12.3
Y73	995.8176	7958.4825	7958.4085	0.0740	9.3
Y75	1026.0875	8200.6422	8200.5351	0.1070	13.1
Y76	1038.4661	8299.6709	8299.6036	0.0673	8.1
Y77	1047.3486	8370.7309	8370.6407	0.0902	10.8

3.4.2.2 H2A.1 C-terminal

UniProt Q6FWM7, histone H2A.1 C-terminal fragment

Seq:
AAEILELAGNAARDNKKSRIIPRHLQLAIRNDDELNKLLGNVTIAQGGVLPNIHQNLLPKKSAKPSAS
QEL

m/z monoisotopic = 769.2331, +10; MW_{obs} = 7682.26 Da (-0.4 ppm)

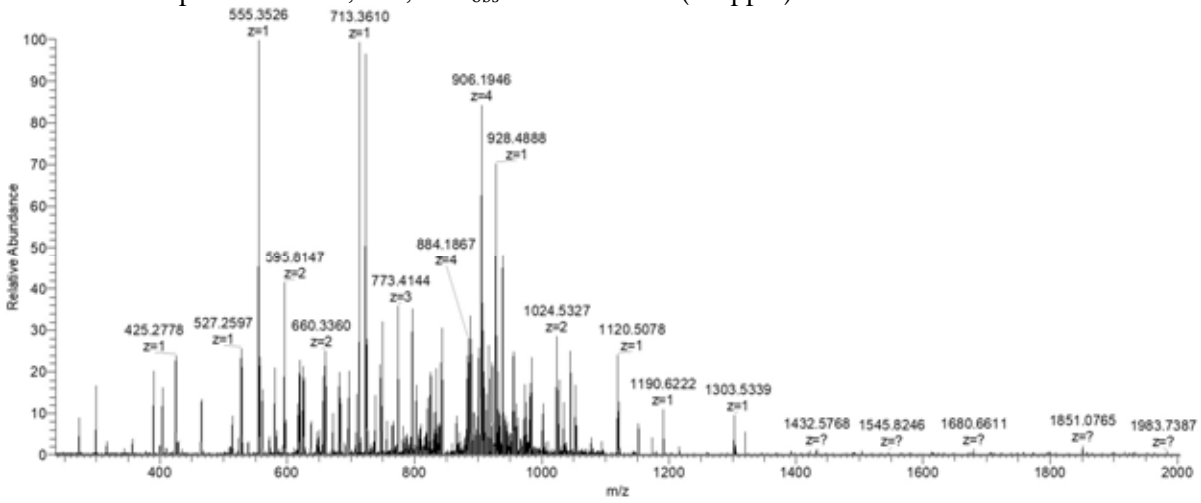


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B42	774.9349	4643.5656	4643.5735	-0.0078	-1.7
B43	949.9316	4744.6218	4744.6212	0.0007	0.1
B43	791.7814	4744.6447	4744.6212	0.0235	5.0
B44	972.5460	4857.6935	4857.7052	-0.0117	-2.4
B44	810.6259	4857.7120	4857.7052	0.0068	1.4
B45	822.4638	4928.7391	4928.7423	-0.0033	-0.7
B45	705.1152	4928.7557	4928.7423	0.0133	2.7
B47	731.5527	5113.8182	5113.8224	-0.0042	-0.8
B47	853.3109	5113.8215	5113.8224	-0.0009	-0.2
B48	862.8127	5170.8327	5170.8438	-0.0111	-2.2
B48	1035.1759	5170.8429	5170.8438	-0.0009	-0.2
B70	840.0279	7551.1855	7551.1665	0.0190	2.5
Y7	731.3601	730.3528	730.3497	0.0031	4.2
Y10	1017.5237	1016.5165	1016.5138	0.0026	2.6
Y13	685.8881	1369.7617	1369.7565	0.0052	3.8
Y18	988.5485	1975.0824	1975.0850	-0.0027	-1.4
Y19	1045.0884	2088.1623	2088.1691	-0.0068	-3.3
Y21	1150.6400	2299.2654	2299.2648	0.0006	0.2
Y22	604.0932	2412.3436	2412.3489	-0.0053	-2.2
Y23	838.1474	2511.4203	2511.4173	0.0031	1.2
Y24	857.1537	2568.4392	2568.4387	0.0005	0.2
Y25	876.1611	2625.4616	2625.4602	0.0014	0.5
Y25	657.3750	2625.4708	2625.4602	0.0106	4.0
Y26	689.3879	2753.5225	2753.5188	0.0037	1.4

Y27	707.1445	2824.5490	2824.5559	-0.0068	-2.4

3.4.2.3 HSP12 (13+)

UniProt Q6FPF6, HSP12
Seq:
SDAGRKNFSDKLNEGLTPDSQKSTWDKGKEFVTDETDKLAGKFQGEENKGVAQGMHDSAQKGAD
EANAESYADTAREYMDAAKSKLNDAVEYVSKSVHGGEK
m/z monoisotopic = 853.6437, +13; MW_{obs} = 11084.27 Da (4.8 ppm)

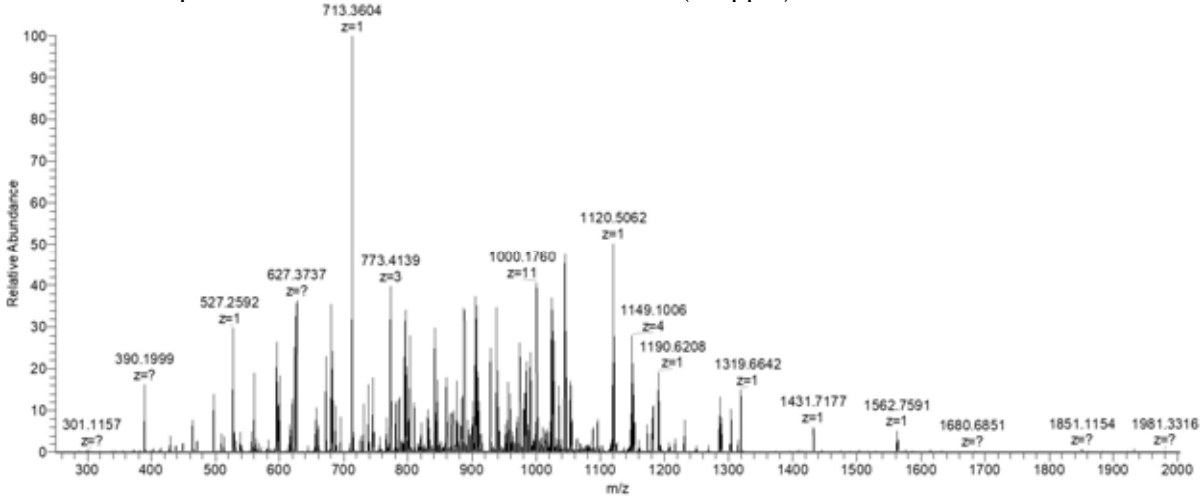


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
Y20	725.7208	2174.1407	2174.1331	0.0076	3.5
Y22	580.0599	2316.2106	2316.2073	0.0033	1.4
Y29	637.5235	3182.5811	3182.5666	0.0145	4.6
Y29	796.6551	3182.5914	3182.5666	0.0249	7.8
Y30	825.4061	3297.5953	3297.5935	0.0018	0.5
Y31	843.1654	3368.6325	3368.6306	0.0019	0.6
Y32	883.9310	3531.6948	3531.6940	0.0009	0.2
Y32	707.3494	3531.7104	3531.6940	0.0165	4.7
Y34	937.9575	3747.8009	3747.7686	0.0323	8.6
Y35	955.7112	3818.8158	3818.8057	0.0101	2.7
Y36	984.2196	3932.8492	3932.8486	0.0005	0.1
Y37	1001.9768	4003.8781	4003.8857	-0.0077	-1.9
Y38	1034.2413	4132.9363	4132.9283	0.0080	1.9
Y45	959.0613	4790.2703	4790.2365	0.0338	7.1
Y46	982.0633	4905.2803	4905.2634	0.0169	3.4
Y47	841.3967	5042.3364	5042.3224	0.0141	2.8
Y53	931.9362	5585.5734	5585.5699	0.0035	0.6
Y62	944.3070	6603.0978	6603.0578	0.0400	6.1

Y63	954.4483	6674.0871	6674.0949	-0.0078	-1.2
Y65	865.4149	6915.2614	6915.2739	-0.0126	-1.8
Y73	982.4639	7851.6531	7851.6452	0.0079	1.0
Y76	908.2129	8164.8506	8164.8566	-0.0059	-0.7
Y76	817.4927	8164.8541	8164.8566	-0.0025	-0.3
Y80	962.5701	8654.0650	8654.0425	0.0225	2.6
Y83	900.7311	8997.2384	8997.2281	0.0103	1.1
Y86	932.0458	9310.3853	9310.3555	0.0298	3.2
Y99	898.4353	10769.1366	10769.1134	0.0233	2.2

3.4.2.4 HSP12 (12+)

UniProt Q6FPF6, HSP12
Seq:
SDAGRKNFSDKLNEGLTPDSQKSTWDKGKEFVTDETDKLAGKFQGEENKGVAQGMHDSAQKGAD
EANAESYADTAREYMDAAKSKLNDAVEYVSKSVHGGEK
m/z monoisotopic = 924.6934, +12; MW_{obs} = 11084.23 Da (1.2 ppm)



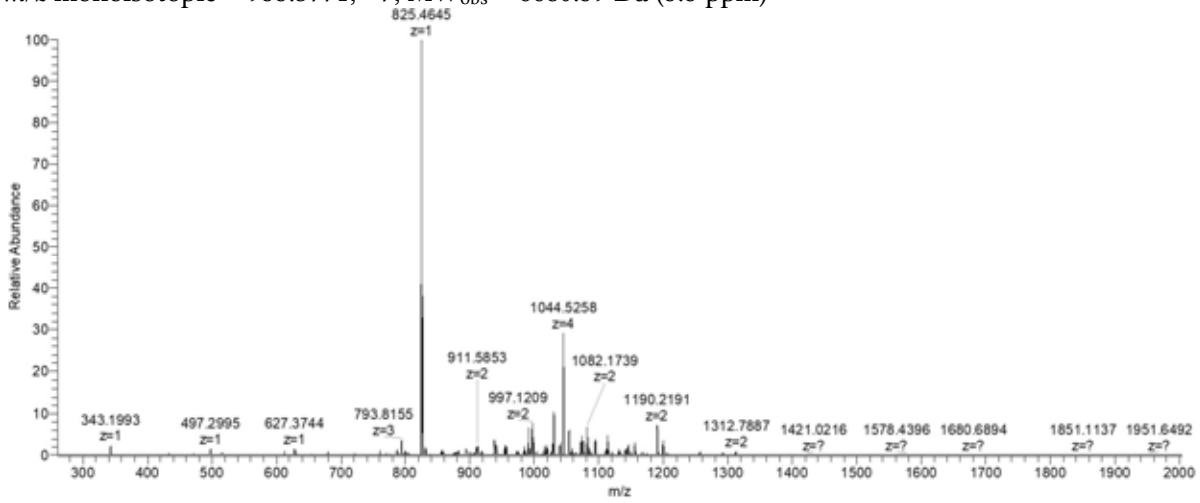
assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B10	560.7584	1119.5023	1119.4945	0.0078	7.0
B10	1120.5109	1119.5036	1119.4945	0.0091	8.1
B11	624.8066	1247.5987	1247.5894	0.0093	7.4
B12	681.3478	1360.6811	1360.6735	0.0076	5.6
B13	738.3699	1474.7252	1474.7164	0.0088	6.0
B14	802.8908	1603.7671	1603.7590	0.0081	5.0
B15	831.4021	1660.7896	1660.7805	0.0091	5.5
B16	887.9445	1773.8745	1773.8645	0.0100	5.6
B17	938.4681	1874.9216	1874.9122	0.0093	5.0
B19	1044.5034	2086.9922	2086.9919	0.0003	0.2

B21	1152.0489	2302.0833	2302.0825	0.0007	0.3
B26	974.1281	2919.3626	2919.3635	-0.0009	-0.3
B27	1016.8276	3047.4611	3047.4584	0.0027	0.9
B29	809.1498	3232.5701	3232.5748	-0.0048	-1.5
B64	869.9154	6951.2653	6951.2917	-0.0264	-3.8
Y5	527.2591	526.2518	526.2499	0.0019	3.5
Y6	626.3276	625.3203	625.3184	0.0019	3.1
Y7	713.3624	712.3551	712.3504	0.0047	6.6
Y8	841.4597	840.4525	840.4453	0.0071	8.5
Y9	928.4850	927.4778	927.4774	0.0004	0.4
Y10	1027.5561	1026.5488	1026.5458	0.0030	3.0
Y11	1190.6187	1189.6115	1189.6091	0.0024	2.0
Y11	595.8153	1189.6161	1189.6091	0.0070	5.9
Y12	1319.6624	1318.6551	1318.6517	0.0034	2.6
Y12	660.3361	1318.6577	1318.6517	0.0060	4.5
Y14	745.3906	1488.7666	1488.7572	0.0094	6.3
Y15	802.8908	1603.7671	1603.7842	-0.0171	-10.6
Y16	859.9230	1717.8314	1717.8271	0.0043	2.5
Y18	980.5088	1959.0031	1959.0061	-0.0030	-1.5
Y19	1024.0298	2046.0451	2046.0381	0.0070	3.4
Y19	683.0230	2046.0473	2046.0381	0.0091	4.5
Y20	725.7160	2174.1262	2174.1331	-0.0070	-3.2
Y29	796.6513	3182.5763	3182.5666	0.0097	3.0
Y31	843.1635	3368.6249	3368.6306	-0.0057	-1.7
Y32	883.9286	3531.6853	3531.6940	-0.0087	-2.5
Y33	905.6906	3618.7334	3618.7260	0.0074	2.1
Y35	955.7079	3818.8026	3818.8057	-0.0031	-0.8
Y36	984.2192	3932.8476	3932.8486	-0.0010	-0.3
Y37	1001.9762	4003.8759	4003.8857	-0.0098	-2.5
Y38	1034.2369	4132.9187	4132.9283	-0.0097	-2.3
Y41	1095.0113	4376.0159	4376.0138	0.0021	0.5
Y45	959.0580	4790.2536	4790.2365	0.0171	3.6
Y46	982.0597	4905.2620	4905.2634	-0.0015	-0.3
Y47	1009.4707	5042.3171	5042.3224	-0.0053	-1.0
Y47	841.3948	5042.3251	5042.3224	0.0028	0.5
Y58	1024.8123	6142.8304	6142.8144	0.0160	2.6
Y63	954.4513	6674.1083	6674.0949	0.0134	2.0
Y64	970.6046	6787.1812	6787.1790	0.0022	0.3
Y64	849.4060	6787.1898	6787.1790	0.0108	1.6
Y65	865.4128	6915.2440	6915.2739	-0.0299	-4.3
Y67	892.4305	7131.3856	7131.3486	0.0371	5.2

Y68	908.5592	7260.4154	7260.3911	0.0243	3.3
Y80	962.5667	8654.0344	8654.0425	-0.0081	-0.9
Y83	900.7285	8997.2124	8997.2281	-0.0157	-1.7

3.4.2.5 COX7A

UniProt Q6FWE8, cytochrome c oxidase subunit 7A + propeptide
Seq: SALAPITGTLKKRIITDIVIGFSLGGVMASYWWWGFHKNVIDRREAFYADLAEKKKAEN
m/z monoisotopic = 955.3771, +7; MW_{obs} = 6680.59 Da (5.5 ppm)

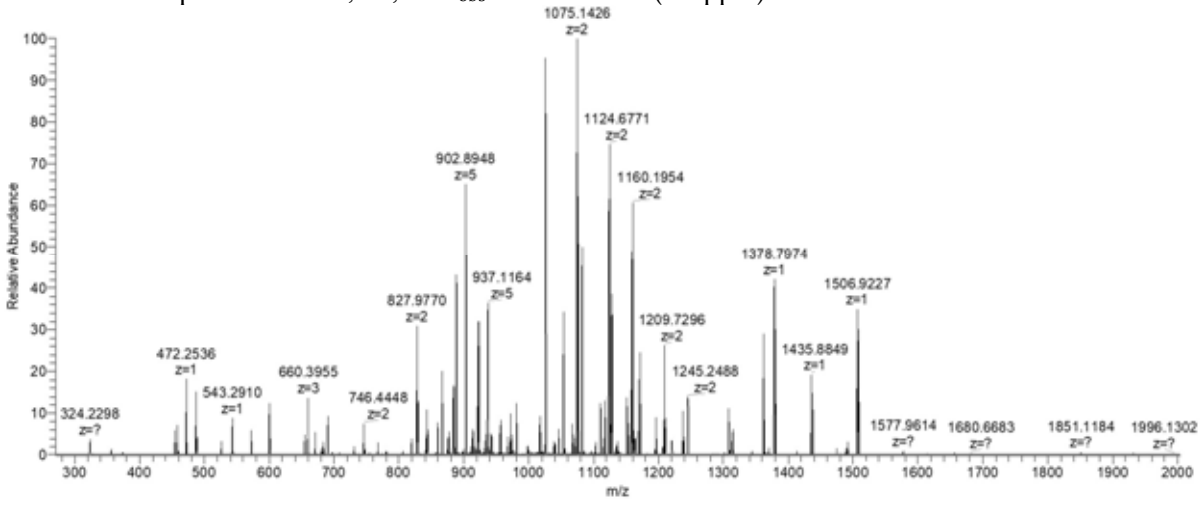


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B4	343.1987	342.1914	342.1903	0.0011	3.2
B17	890.5425	1779.0704	1779.0618	0.0086	4.8
B19	996.6150	1991.2154	1991.2142	0.0012	0.6
B20	1053.1588	2104.3031	2104.2983	0.0047	2.3
B21	1081.6643	2161.3140	2161.3198	-0.0058	-2.7
B22	1155.2068	2308.3991	2308.3882	0.0109	4.7
B23	1198.7204	2395.4262	2395.4202	0.0060	2.5
B24	837.1811	2508.5214	2508.5043	0.0171	6.8
B25	856.1864	2565.5375	2565.5257	0.0117	4.6
B26	1312.2861	2622.5577	2622.5472	0.0105	4.0
B26	875.1947	2622.5624	2622.5472	0.0152	5.8
B27	908.2177	2721.6312	2721.6156	0.0156	5.7
Y9	515.8022	1029.5898	1029.5818	0.0080	7.7
Y17	680.3642	2038.0708	2038.0595	0.0113	5.5
Y28	877.7009	3506.7743	3506.7636	0.0108	3.1
Y29	918.4677	3669.8419	3669.8269	0.0150	4.1
Y30	940.2248	3756.8700	3756.8589	0.0111	3.0
Y32	990.7453	3958.9519	3958.9365	0.0154	3.9

Y33	1015.5095	4058.0088	4058.0049	0.0039	1.0
Y34	1029.7636	4115.0251	4115.0264	-0.0013	-0.3
Y35	1044.0200	4172.0509	4172.0478	0.0031	0.7
Y36	1072.2920	4285.1388	4285.1319	0.0069	1.6
Y36	858.0392	4285.1595	4285.1319	0.0276	6.4
Y37	1094.0547	4372.1897	4372.1639	0.0258	5.9
Y37	875.4463	4372.1950	4372.1639	0.0310	7.1
Y38	904.8531	4519.2294	4519.2323	-0.0030	-0.7
Y38	1130.8190	4519.2470	4519.2323	0.0147	3.3
Y39	916.2581	4576.2539	4576.2538	0.0001	0.0
Y39	1145.0726	4576.2614	4576.2538	0.0076	1.7
Y40	1173.3413	4689.3360	4689.3379	-0.0019	-0.4
Y40	938.8801	4689.3639	4689.3379	0.0260	5.6
Y55	1057.4070	6338.3985	6338.3618	0.0367	5.8
Y56	1069.2436	6409.4178	6409.3989	0.0188	2.9
Y57	1088.0896	6522.4937	6522.4830	0.0107	1.6

3.4.2.6 COX7

UniProt B4UN07, uncharacterized protein CAGL0H02491g (COX7)
Seq: ANRIIELQKLFQSSTKPIWWRHPRSAFYLYPYYGLLAVAVVAPLLYIPNAVRGIKAPKNN
m/z monoisotopic = 990.1314, +7; MW_{obs} = 6923.87 Da (3.8 ppm)

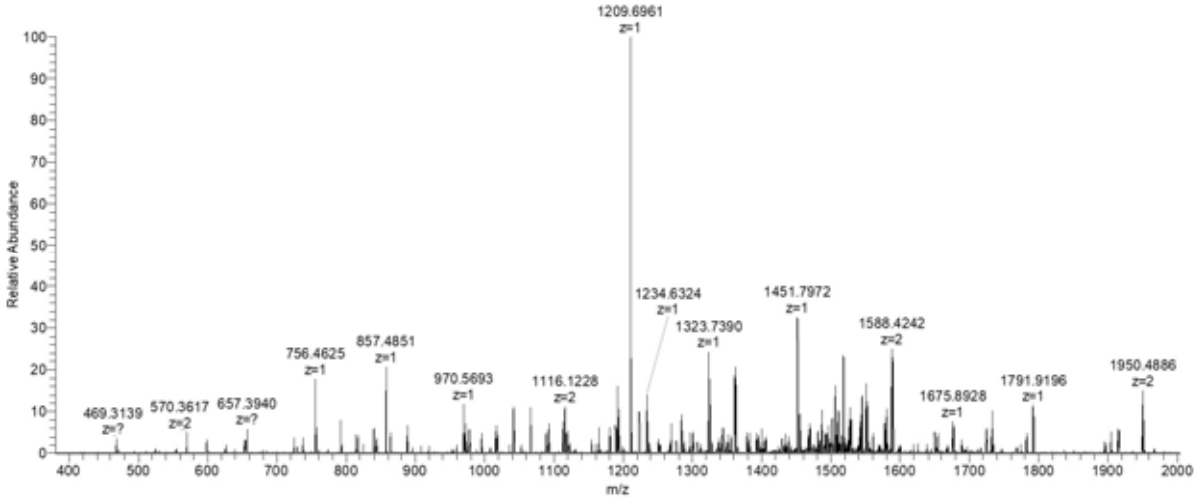


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B30	933.5097	3730.0095	3730.0048	0.0047	1.3
B33	1039.3031	4153.1834	4153.1842	-0.0008	-0.2
B34	843.0496	4210.2117	4210.2057	0.0060	1.4
B34	1053.5611	4210.2153	4210.2057	0.0096	2.3
B35	1081.8278	4323.2823	4323.2898	-0.0075	-1.7

B35	865.6691	4323.3091	4323.2898	0.0193	4.5
B36	1110.1023	4436.3802	4436.3738	0.0064	1.4
B36	888.2837	4436.3821	4436.3738	0.0083	1.9
B37	1127.8582	4507.4037	4507.4109	-0.0072	-1.6
B37	902.4926	4507.4264	4507.4109	0.0155	3.4
B38	1152.6270	4606.4791	4606.4793	-0.0002	-0.1
B38	922.3036	4606.4816	4606.4793	0.0023	0.5
B39	936.5140	4677.5337	4677.5164	0.0173	3.7
B39	1170.3920	4677.5389	4677.5164	0.0225	4.8
B40	956.3246	4776.5865	4776.5849	0.0017	0.4
B40	1195.1575	4776.6011	4776.5849	0.0162	3.4
B41	1219.9225	4875.6608	4875.6533	0.0075	1.5
B42	1237.6820	4946.6987	4946.6904	0.0084	1.7
Y4	472.2538	471.2465	471.2441	0.0024	5.0
Y5	543.2907	542.2835	542.2813	0.0022	4.1
Y13	689.9045	1377.7945	1377.7841	0.0104	7.6
Y13	1378.8024	1377.7951	1377.7841	0.0110	8.0
Y14	746.4430	1490.8714	1490.8681	0.0033	2.2
Y15	827.9777	1653.9408	1653.9314	0.0094	5.7
Y16	884.5197	1767.0248	1767.0155	0.0093	5.3
Y18	660.0604	1977.1593	1977.1523	0.0070	3.5
Y19	1025.1042	2048.1938	2048.1894	0.0044	2.1
Y19	683.7401	2048.1985	2048.1894	0.0091	4.4
Y20	1074.6404	2147.2663	2147.2578	0.0085	3.9
Y21	1124.1747	2246.3349	2246.3263	0.0086	3.8
Y22	1159.6901	2317.3657	2317.3634	0.0023	1.0
Y23	1209.2252	2416.4358	2416.4318	0.0040	1.7
Y24	1244.7430	2487.4714	2487.4689	0.0025	1.0

3.4.2.7 Thioredoxin

UniProt Q6FND5, thioredoxin
Seq:
VKQITSVSEFDSAIAVDKLVVVDFFATWCGPCKMIAPMIEKFAAEYSTADFYKLDVDELPEVAQKNE
VSAMPTLVLFKNGKEVAKVVGANPAAIKQAIANNV
m/z monoisotopic = 1380.0925, +8; MW_{obs} = 11032.68 Da (-0.3 ppm)



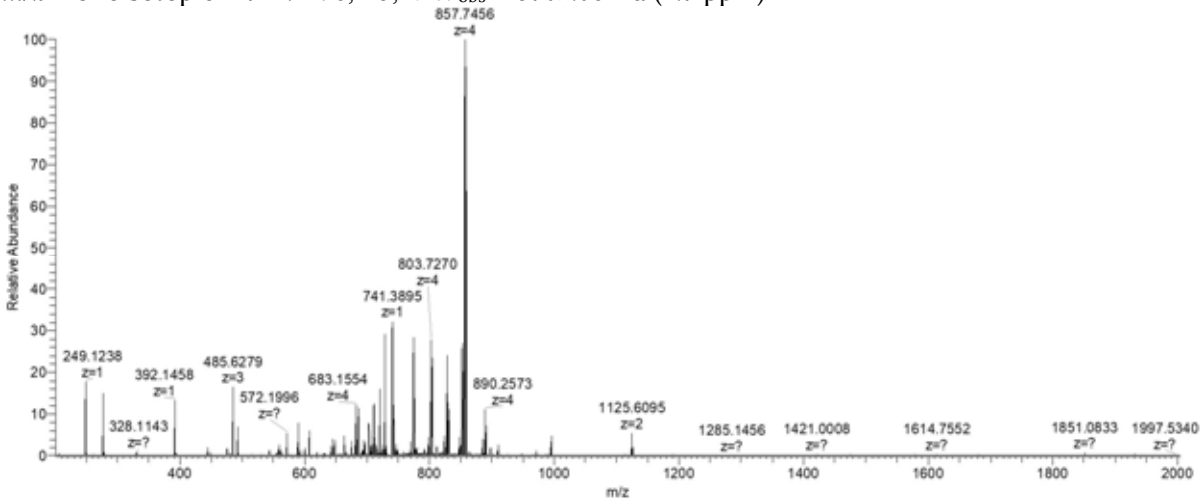
assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B7	756.4629	755.4557	755.4541	0.0015	2.0
B9	972.5417	971.5344	971.5287	0.0057	5.9
B10	1119.6067	1118.5994	1118.5972	0.0023	2.0
B11	1234.6360	1233.6287	1233.6241	0.0046	3.7
B13	1392.7072	1391.7000	1391.6932	0.0067	4.8
B14	1505.7784	1504.7711	1504.7773	-0.0062	-4.1
B15	1576.8253	1575.8180	1575.8144	0.0036	2.3
B16	1675.8958	1674.8885	1674.8828	0.0057	3.4
B17	1790.9182	1789.9110	1789.9098	0.0012	0.7
B19	1016.5498	2031.0851	2031.0888	-0.0037	-1.8
B20	1066.0823	2130.1501	2130.1572	-0.0071	-3.3
B21	1115.6138	2229.2131	2229.2256	-0.0125	-5.6
B22	1165.1584	2328.3022	2328.2940	0.0082	3.5
B23	1222.6658	2443.3171	2443.3209	-0.0039	-1.6
B24	1296.2116	2590.4086	2590.3894	0.0192	7.4
B26	1405.2505	2808.4864	2808.4949	-0.0084	-3.0
B90	1404.4287	9823.9499	9823.9970	-0.0471	-4.8
Y8	857.4851	856.4778	856.4767	0.0011	1.3
Y9	970.5695	969.5622	969.5607	0.0015	1.6
Y11	1112.6485	1111.6412	1111.6349	0.0062	5.6
Y12	1209.6903	1208.6830	1208.6877	-0.0046	-3.8
Y13	1323.7407	1322.7334	1322.7306	0.0028	2.1
Y14	1394.7827	1393.7755	1393.7677	0.0077	5.5
Y16	1550.8751	1549.8679	1549.8576	0.0102	6.6
Y17	1649.9427	1648.9355	1648.9260	0.0094	5.7
Y24	1188.6734	2375.3322	2375.3284	0.0038	1.6
Y26	1326.2604	2650.5062	2650.4918	0.0144	5.4

Y31	1587.9224	3173.8302	3173.8288	0.0014	0.4
Y32	1653.4456	3304.8766	3304.8693	0.0073	2.2
Y34	1732.4802	3462.9458	3462.9384	0.0074	2.1
Y35	1782.0106	3562.0066	3562.0068	-0.0002	-0.1
Y37	1903.5489	3805.0833	3805.0923	-0.0091	-2.4
Y37	1269.3726	3805.0960	3805.0923	0.0037	1.0
Y43	1486.8293	4457.4660	4457.4467	0.0192	4.3

3.4.3 *C. neoformans*

3.4.3.1 H3 C-terminal

UniProt J9VMZ0, histone H3 C-terminal fragment
Seq: FEDTNLAAIHAKRVTIQPKDLQLARRLRGERS
m/z monoisotopic = 741.4196, +5; MW_{obs} = 3702.06 Da (1.7 ppm)

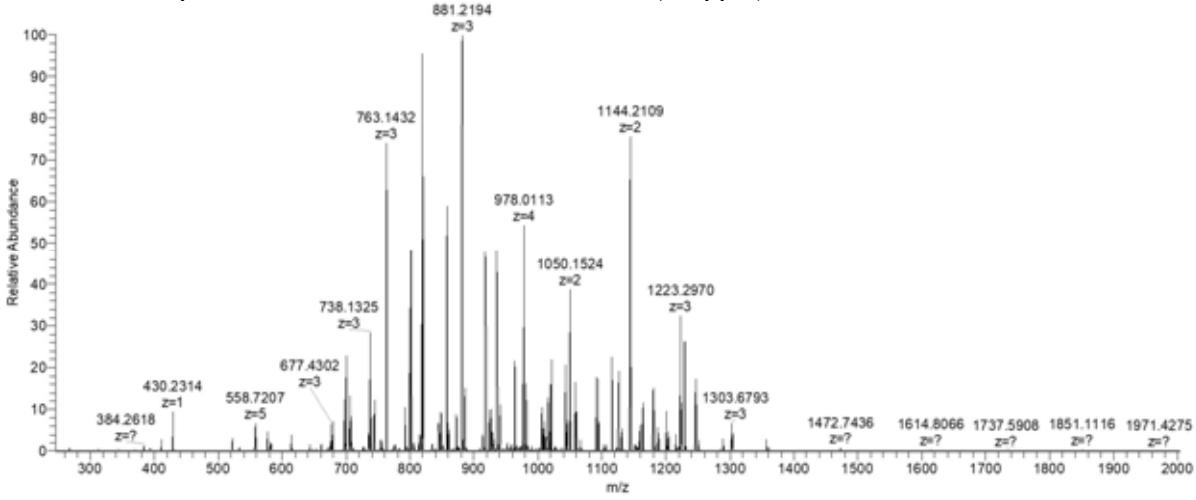


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B2	277.1184	276.1111	276.1110	0.0001	0.5
B3	392.1464	391.1391	391.1379	0.0011	2.9
B4	493.1951	492.1878	492.1856	0.0022	4.5
B20	1125.1085	2248.2024	2248.1964	0.0060	2.7
B31	720.4061	3596.9940	3597.0127	-0.0187	-5.2
Y12	485.6289	1453.8648	1453.8590	0.0059	4.0
Y22	648.6383	2590.5241	2590.5255	-0.0014	-0.6
Y23	682.9050	2727.5908	2727.5844	0.0064	2.3
Y24	711.1735	2840.6648	2840.6685	-0.0037	-1.3
Y25	728.9366	2911.7172	2911.7056	0.0116	4.0
Y26	995.2528	2982.7366	2982.7427	-0.0061	-2.1
Y26	746.6917	2982.7379	2982.7427	-0.0049	-1.6

Y27	774.9636	3095.8251	3095.8268	-0.0017	-0.5
Y28	803.4734	3209.8643	3209.8697	-0.0054	-1.7
Y29	663.1886	3310.9067	3310.9174	-0.0107	-3.2
Y29	828.7400	3310.9308	3310.9174	0.0134	4.1
Y30	686.1963	3425.9452	3425.9443	0.0008	0.2
Y30	857.4967	3425.9579	3425.9443	0.0136	4.0
Y31	889.7541	3554.9874	3554.9869	0.0005	0.1
Y31	712.0056	3554.9914	3554.9869	0.0045	1.3

3.4.3.2 COX7A

UniProt J9VVE3, cytochrome c oxidase subunit 7A
Seq: PVAPVVGKLRKRLITDLTASIGIGLAGAYTFWYTVHLPMVKKRDDYYLRLEQAKSS
m/z monoisotopic = 901.6492, +7; MW_{obs} = 6304.49 Da (3.5 ppm)

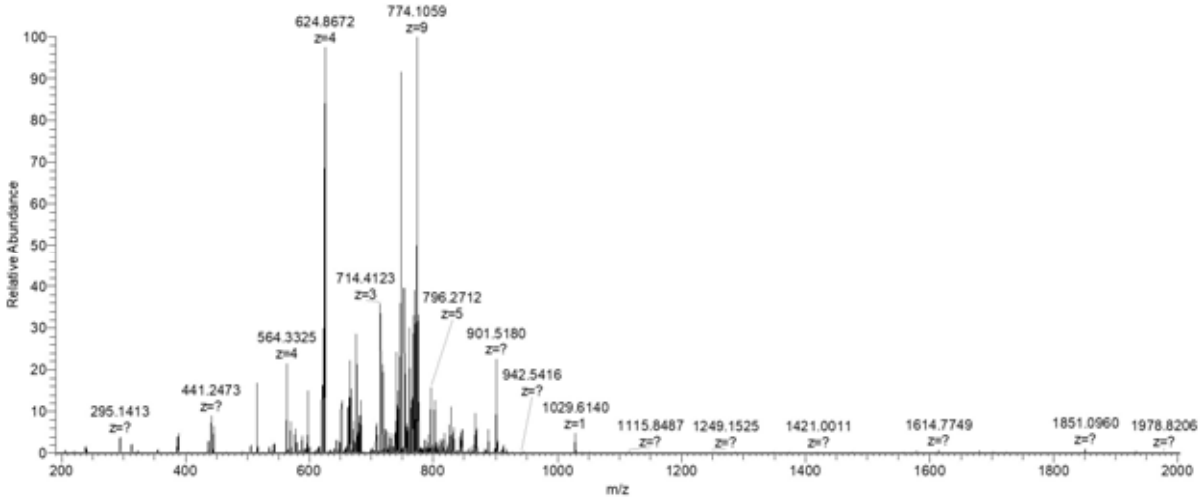


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B16	582.0387	1743.0942	1743.0883	0.0059	3.4
B16	872.5591	1743.1037	1743.0883	0.0155	8.9
B19	1015.1352	2028.2559	2028.2571	-0.0012	-0.6
B20	706.1061	2115.2966	2115.2891	0.0075	3.5
B20	1058.6591	2115.3036	2115.2891	0.0144	6.8
B21	1115.1941	2228.3737	2228.3732	0.0005	0.2
B21	743.8021	2228.3844	2228.3732	0.0112	5.0
B22	1143.7140	2285.4134	2285.3947	0.0188	8.2
B23	1200.2500	2398.4854	2398.4787	0.0066	2.8
B24	614.8821	2455.4994	2455.5002	-0.0008	-0.3
B24	1228.7577	2455.5009	2455.5002	0.0007	0.3
B25	857.2052	2568.5938	2568.5842	0.0095	3.7
B26	660.9116	2639.6171	2639.6213	-0.0042	-1.6

B27	675.1715	2696.6569	2696.6428	0.0141	5.2
B28	923.5704	2767.6892	2767.6799	0.0093	3.4
Y11	679.3600	1356.7054	1356.7038	0.0016	1.2
Y19	1164.1194	2326.2242	2326.2103	0.0138	5.9
Y21	859.7945	2576.3616	2576.3533	0.0083	3.2
Y23	926.4964	2776.4672	2776.4694	-0.0021	-0.8
Y24	980.8509	2939.5309	2939.5327	-0.0018	-0.6
Y25	1042.8785	3125.6135	3125.6120	0.0015	0.5
Y26	819.1753	3272.6719	3272.6804	-0.0085	-2.6
Y26	1091.9027	3272.6863	3272.6804	0.0059	1.8
Y27	844.4407	3373.7335	3373.7281	0.0054	1.6
Y27	1125.5858	3373.7357	3373.7281	0.0075	2.2
Y28	885.2066	3536.7974	3536.7914	0.0059	1.7
Y28	1179.9422	3536.8048	3536.7914	0.0134	3.8
Y29	1203.6184	3607.8333	3607.8286	0.0047	1.3
Y30	917.2201	3664.8514	3664.8500	0.0014	0.4
Y30	1222.6292	3664.8657	3664.8500	0.0157	4.3
Y31	934.9792	3735.8878	3735.8871	0.0006	0.2
Y31	1246.3117	3735.9132	3735.8871	0.0261	7.0
Y32	963.2526	3848.9812	3848.9712	0.0100	2.6
Y33	1303.0008	3905.9807	3905.9926	-0.0119	-3.1
Y33	977.5092	3906.0077	3905.9926	0.0150	3.8
Y34	1005.7752	4019.0716	4019.0767	-0.0051	-1.3
Y34	804.8255	4019.0912	4019.0767	0.0145	3.6
Y35	1020.0334	4076.1044	4076.0982	0.0063	1.5
Y53	1007.2275	6037.3215	6037.3131	0.0084	1.4
Y54	1019.0683	6108.3663	6108.3502	0.0161	2.6

3.4.3.3 L39

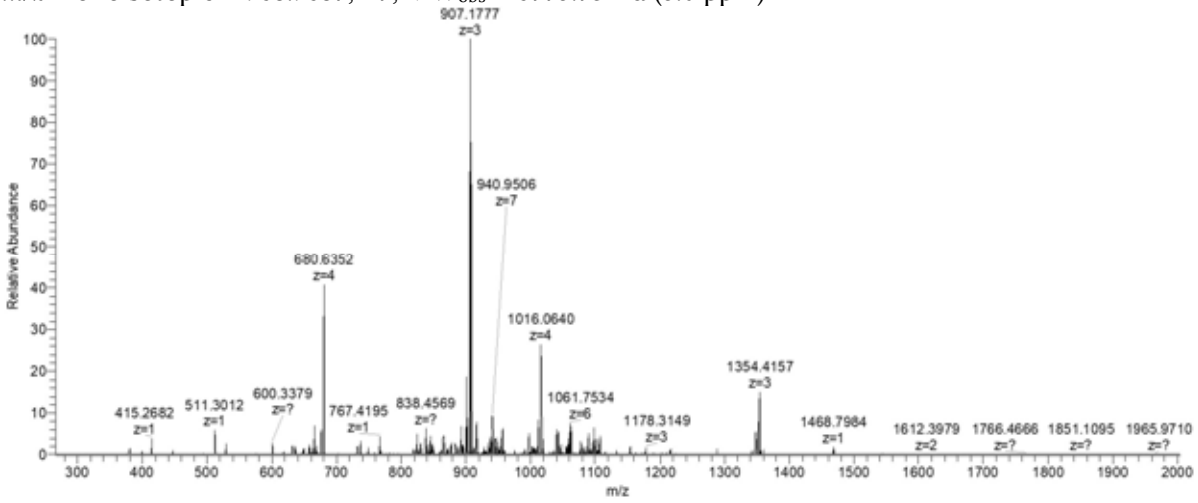
UniProt J9VPZ8, L39
Seq: PSQKTFIVKQKLAKKARQNRPLPQWFRLKTDNKIQYNAKRRHWRRTKLNL
m/z monoisotopic = 692.8529, +9; MW_{obs} = 6226.61 Da (4.3 ppm)



assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B3	313.1520	312.1447	312.1434	0.0013	4.2
B8	901.5203	900.5130	900.5069	0.0061	6.8
B11	643.3875	1284.7605	1284.7554	0.0051	4.0
B20	588.6124	2350.4204	2350.4073	0.0131	5.6
B31	623.0323	3732.1499	3732.1580	-0.0080	-2.1
B31	747.4406	3732.1666	3732.1580	0.0087	2.3
B33	663.3907	3974.3008	3974.2958	0.0050	1.2
B33	795.8690	3974.3084	3974.2958	0.0125	3.2
B34	682.2409	4087.4020	4087.3799	0.0221	5.4
Y15	671.3875	2011.1406	2011.1452	-0.0047	-2.3
Y16	714.0757	2139.2052	2139.2038	0.0014	0.6
Y17	564.0794	2252.2884	2252.2879	0.0005	0.2
Y17	751.7709	2252.2909	2252.2879	0.0030	1.3
Y18	596.1052	2380.3919	2380.3828	0.0090	3.8
Y19	624.6169	2494.4383	2494.4258	0.0125	5.0
Y19	832.4870	2494.4393	2494.4258	0.0135	5.4
Y39	824.6466	4941.8358	4941.8283	0.0075	1.5
Y40	725.2840	5069.9372	5069.9233	0.0139	2.7
Y41	867.3412	5198.0034	5197.9819	0.0215	4.1
Y43	776.0265	5425.1348	5425.1452	-0.0104	-1.9
Y44	792.1860	5538.2513	5538.2293	0.0220	4.0
Y45	813.1960	5685.3208	5685.2977	0.0231	4.1
Y46	827.6285	5786.3486	5786.3454	0.0032	0.5
Y46	724.3021	5786.3587	5786.3454	0.0133	2.3
Y47	845.9319	5914.4721	5914.4404	0.0317	5.4
Y49	767.2009	6129.5491	6129.5310	0.0182	3.0

3.4.3.4 L29

UniProt J9VST2, L29
Seq: AKSKNHTAHNQNKKAHRNKIQRPKTNKYHSLKGVDPKFRRNARYAAQGSAKAIREAKASA
m/z monoisotopic = 968.9687, +7; MW_{obs} = 6775.73 Da (5.0 ppm)

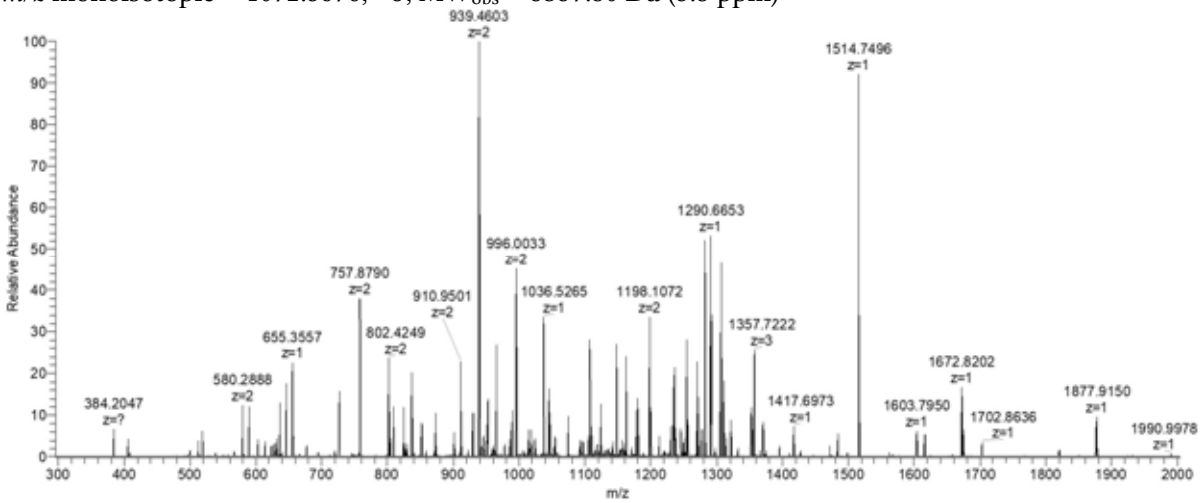


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B6	666.3718	665.3645	665.3609	0.0036	5.4
B7	767.4225	766.4153	766.4086	0.0067	8.7
B8	838.4604	837.4531	837.4457	0.0074	8.8
B27	1053.9268	3158.7587	3158.7510	0.0077	2.4
B29	865.7287	3458.8856	3458.8733	0.0124	3.6
B30	887.4888	3545.9262	3545.9053	0.0209	5.9
B31	915.7563	3658.9962	3658.9893	0.0069	1.9
B32	947.7767	3787.0778	3787.0843	-0.0065	-1.7
B35	1353.7413	4058.2019	4058.2011	0.0008	0.2
B35	1015.5579	4058.2024	4058.2011	0.0013	0.3
B57	933.6603	6528.5712	6528.5792	-0.0080	-1.2
B57	1089.1047	6528.5846	6528.5792	0.0054	0.8
B58	943.8071	6599.5988	6599.6164	-0.0175	-2.7
B59	956.2435	6686.6535	6686.6484	0.0051	0.8
Y8	845.4884	844.4811	844.4766	0.0045	5.3
Y9	916.5240	915.5167	915.5138	0.0030	3.2
Y21	1095.5978	2189.1811	2189.1777	0.0034	1.6
Y25	906.8423	2717.5051	2717.4949	0.0102	3.7
Y25	680.3844	2717.5085	2717.4949	0.0135	5.0
Y28	997.2206	2988.6399	2988.6118	0.0281	9.4
Y30	1077.6076	3229.8010	3229.7908	0.0102	3.2
Y31	830.2124	3316.8204	3316.8228	-0.0024	-0.7
Y31	1106.6185	3316.8338	3316.8228	0.0110	3.3

Y32	864.4813	3453.8963	3453.8817	0.0146	4.2
Y33	905.2444	3616.9483	3616.9450	0.0033	0.9
Y38	1047.3347	4185.3098	4185.2783	0.0314	7.5
Y52	990.7150	5938.2465	5938.2504	-0.0039	-0.7
Y53	1002.5573	6009.2999	6009.2875	0.0124	2.1
Y54	1019.4029	6110.3735	6110.3352	0.0383	6.3
Y55	1042.2425	6247.4116	6247.3941	0.0175	2.8
Y56	1061.2518	6361.4670	6361.4370	0.0300	4.7
Y57	1082.5976	6489.5422	6489.5320	0.0102	1.6

3.4.3.5 COX7C C-terminal

UniProt J9W384, cytochrome c oxidase subunit 7c C-terminal fragment
Seq: VHFENVVNHTLPTDVTNKYFLAAKIITFGVLGFGTPFYAAYFHLNKSG
m/z monoisotopic = 1072.5676, +5; MW_{obs} = 5357.80 Da (5.8 ppm)

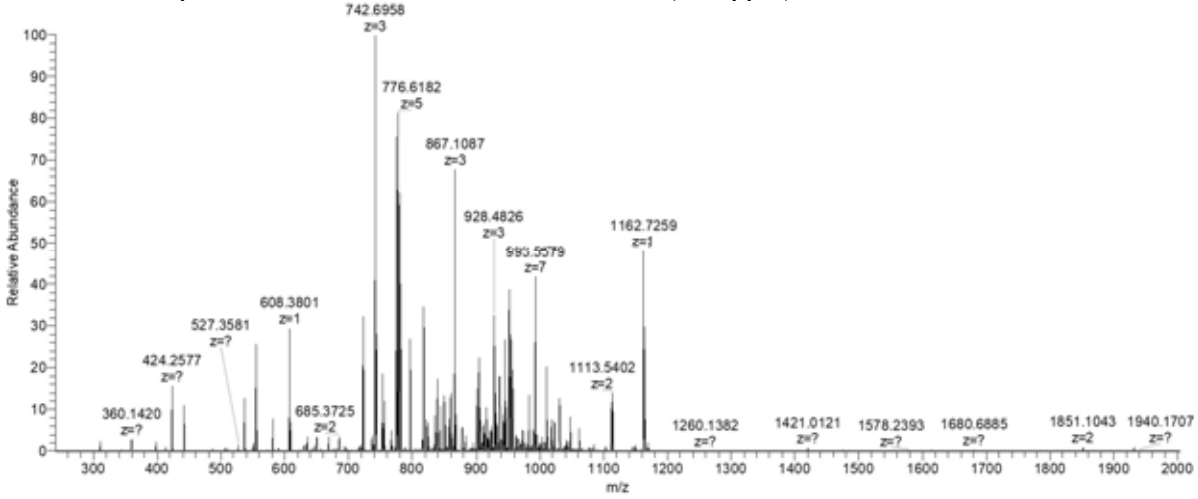


assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
B27	1022.8916	3065.6530	3065.6338	0.0192	6.3
B30	1123.9392	3368.7957	3368.7921	0.0036	1.1
B31	1161.6350	3481.8832	3481.8761	0.0071	2.0
B32	1180.6460	3538.9163	3538.8976	0.0187	5.3
B33	1229.6606	3685.9601	3685.9660	-0.0059	-1.6
B35	962.0192	3844.0475	3844.0351	0.0124	3.2
Y6	655.3553	654.3480	654.3449	0.0031	4.7
Y8	965.4899	964.4827	964.4767	0.0060	6.2
Y9	1036.5305	1035.5233	1035.5138	0.0095	9.2
Y10	1107.5654	1106.5581	1106.5509	0.0072	6.5
Y11	1270.6293	1269.6220	1269.6142	0.0078	6.2
Y13	1514.7463	1513.7391	1513.7354	0.0037	2.4

Y13	757.8799	1513.7452	1513.7354	0.0098	6.5
Y14	808.4048	1614.7951	1614.7831	0.0121	7.5
Y15	1672.8172	1671.8100	1671.8045	0.0055	3.3
Y15	836.9152	1671.8159	1671.8045	0.0114	6.8
Y16	910.4440	1818.8734	1818.8729	0.0005	0.3
Y17	938.9547	1875.8948	1875.8944	0.0004	0.2
Y17	938.9577	1875.9009	1875.8944	0.0065	3.5
Y17	1876.9155	1875.9082	1875.8944	0.0138	7.4
Y18	995.5015	1988.9885	1988.9784	0.0100	5.0
Y19	1045.0355	2088.0564	2088.0469	0.0096	4.6
Y21	1147.0788	2292.1431	2292.1367	0.0064	2.8
Y22	1197.6032	2393.1918	2393.1844	0.0074	3.1
Y23	1254.1441	2506.2737	2506.2685	0.0052	2.1
Y24	1310.6821	2619.3496	2619.3525	-0.0029	-1.1
Y37	1018.0391	4068.1273	4068.1189	0.0084	2.1
Y37	1357.0528	4068.1365	4068.1189	0.0176	4.3
Y42	1159.1125	4632.4208	4632.4209	-0.0001	0.0
Y44	1212.3926	4845.5415	4845.5322	0.0093	1.9
Y45	1244.6573	4974.6000	4974.5748	0.0252	5.1
Y46	1281.4215	5121.6571	5121.6432	0.0139	2.7

3.4.3.6 ATPase

UniProt J9VK09, F-type H-transporting ATPase subunit F
Seq:
SLAGLIPPKIATPGAVSSGTTSARTAQVIDFYSKLPKGPKPASERVGGIKGRFFEGKNASGAPVVATLA
ALFLIGYTIDYNMHLKHHKHGHH
m/z monoisotopic = 878.9375, +11; MW_{obs} = 9657.23 Da (10.2 ppm)



assignment	m/z	MW_{obs}	MW_{theor}	ΔMW	Δppm
Y15	635.3162	1902.9268	1902.9172	0.0095	5.0
Y16	668.9962	2003.9669	2003.9649	0.0020	1.0
Y17	723.3525	2167.0356	2167.0282	0.0073	3.4
Y18	1113.0368	2224.0590	2224.0497	0.0094	4.2
Y21	866.7720	2597.2941	2597.2862	0.0079	3.0
Y22	904.4680	2710.3823	2710.3703	0.0120	4.4
Y23	928.1458	2781.4155	2781.4074	0.0081	2.9
Y24	951.8251	2852.4534	2852.4445	0.0089	3.1
Y25	989.5221	2965.5446	2965.5286	0.0160	5.4
Y26	1023.2023	3066.5851	3066.5762	0.0089	2.9
Y80	944.7264	8493.4718	8493.4613	0.0106	1.2
Y80	850.3600	8493.5268	8493.4613	0.0656	7.7
Y86	911.0935	9100.8618	9100.8306	0.0312	3.4
Y90	946.5159	9455.0859	9455.0573	0.0286	3.0