Liquid Extraction Surface Analysis Mass Spectrometry of Microbes: Towards Point-of-Care Diagnosis for Wound Infection

by

Jana Havlikova

A thesis submitted to the University of Birmingham for the degree of doctor of philosophy

School of Chemistry
College of Engineering and Physical Sciences
University of Birmingham

January 2022

Abstract

Trauma is one of the leading causes of death of people under age 49, while complications associated with the wound infection are the primary cause of death in the first day of injury. Microbial infections are mostly caused by bacteria and fungi commonly present in the hospital environment and these are associated with antibiotic and antifungal resistance. Rapid and accurate microbial identification is therefore of high importance. Current diagnostics takes hours or even days to obtain the result. This work uses liquid extraction surface analysis mass spectrometry (LESA MS) for top-down (TD) analysis of proteins in clinically important microbes including the ESKAPE pathogens (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa* and *Enterobacter spp.*) and yeast *Candida glabrata*. First, the ESKAPE microbes and their two close clinical isolates were grown on simple substrates and subjected to LESA MS. Searches against individual protein databases resulted in identification of 24 proteins from 37 MS/MS mass spectra, while searching against multiple databases determined identification success rate at protein and species level to be 79%. A biofilm model of *P. aeruginosa* and *C. glabrata* was cultured, resulting in identification of three *P. aeruginosa* proteins. Next, a LESA MS workflow was developed for analysis of wounded *in vitro* three-dimensional living skin models inoculated with four of the ESKAPE species and *C. glabrata*, resulting in assignment of seven human, nine bacterial and one yeast proteins. The LESA MS workflow was subsequently applied to *ex vivo* human skin grafts including those inoculated with *S. aureus*, and two human skin proteins were identified from both types of samples. Lastly, LESA MS was coupled to state-of-art cylindrical high-field asymmetric waveform ion mobility spectrometry (FAIMS) and mass spectrometer and a high-throughput approach for TD identification of proteins in bacteria was developed. In total, the high-throughput TD LESA FAIMS

analysis of *Escherichia coli* K12 and four ESKAPE species allowed identification of 179 proteins and 277 proteoforms. The results presented in this thesis suggest that LESA MS and LESA FAIMS MS are tools capable of relatively fast extraction of proteins from microbes cultured on a variety of substrates with their subsequent tandem MS identification, showing potential of LESA as a future *in situ* diagnostic.

To Margita and Jozef

Acknowledgements

I would like to thank the Sci-Phy Centre for Doctoral Training for giving me this wonderful opportunity to start a PhD at the University of Birmingham and EPSRC for funding. A huge thank you belongs to all my supervisors – Prof. Helen Cooper, Prof. Iain Styles and Prof. Robin May. Thank you for all your advices, guidance and help. Helen, thank you for all the opportunities I had during my PhD and your patience with me. I have learnt a lot from you and I truly cannot express my gratitude in words.

I would like to thank the members of Cooper group and research groups on the Biosciences 5^{th} floor for help in the lab and, importantly, all the good chat we had over a cup of tea or cake (or biscuits of course). Special thanks to Rian and Klaudia for introducing me into the world of LESA and microbes and thank you Eva and Carys for always being there for me. I would also like to thank members of HAPI lab for help with my (sometimes weird) experiments. Another thanks goes to everyone in Sci-Phy for all the fun we had, with special thanks to Tasha, Menisha, Sophie and the most special thanks to my best friend Larry who helped me to become not only a better scientist, but also a better person. Thank you for your support during difficult times. Thank you to all of the Staff House Gang lovely people for all the beers, laughter and (sometimes) scientific conversations on a Friday night, with a special thanks to Rebeca who was always up for a good dance. Thanks to Maria for all the chats in our native language and for cheering me up when needed.

Last but not least, I would like to thank my family for their support during this PhD journey. Thank you for everything you have done for me. Your love and care always pushed me forward and I am forever grateful for that. A huge thanks belongs to Paulinka, Matko and Lukasko for the jolly moments accompanied with brutal child honesty and for showing me that the present moment is all that matters.

Contents

List of Figures

List of Tables

List of Abbreviations

3D three-dimensional

AGC automatic gain control

BUP bottom-up proteomics

c-FAIMS cylindrical high-field asymmetric waveform ion mobility spectrometry

CCS collision cross section

CF compensation field

CID collision induced dissociation

CV compensation voltage

DC direct current

DESI desorption electrospray ionisation

DF dispersion field

DV dispersion voltage

ESI electrospray ionisation

ETD electron transfer dissociation

FAIMS high-field asymmetric waveform ion mobility spectrometry

FDR false discovery rate

FTICR fourier transform ion cyclotron resonance

HCD higher-energy collision dissociation

IMS ion mobility spectrometry

IRM ion routing multipole

LBA lysogeny broth agar

LBB lysogeny broth broth

LC liquid chromatography

LESA liquid extraction surface analysis

LIT linear ion trap

MALDI matrix-assisted laser desorption ionisation

MS mass spectrometry

MS/MS tandem mass spectrometry

MSI mass spectrometry imaging

MW molecular weight

nanoESI nanoelectrospray ionisation

p-FAIMS planar high-field asymmetric waveform ion mobility spectrometry

PTM post-translational modification

REIMS rapid evaporative ionisation mass spectrometry

S/N signal-to-noise

SNP single nucleotide polymorphism

TD top-down

TOF time-of-flight

YPD yeast extract peptone dextrose

Chapter 1

Introduction

The work in this thesis describes top-down (TD) liquid extraction surface analysis mass spectrometry (LESA MS) as a tool for analysis of clinically important microbes. The workflow established in previous studies was improved and applied to the group of bacteria known as the ESKAPE pathogens (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa* and *Enterobacter spp.*) and yeast *Candida glabrata*, all commonly found in hospital environment and known for their resistance to antibiotic and antifungal treatment. The microbial species were studied by LESA MS directly as single-species cultures from simple and complex substrates including *in vitro* three-dimensional living skin models and *ex vivo* human skin grafts or mixed-species biofilms grown in a simple culturing medium. A state-of-art mass spectrometer coupled to a new high-field asymmetric waveform ion mobility spectrometry (FAIMS) device allowed implementation of a high-throughput LESA FAIMS MS workflow, significantly improving the number of identified proteins and proteoforms from each of the species investigated.

This Chapter will provide an introduction to MS principles and techniques used in this work, and the project aims.

1.1 Introduction to mass spectrometry

MS is a tool used for identification of compounds based on measurement of the mass-to-charge (m/z) ratios of its constituents [1]. The main purpose of mass spectrometry (MS) is to "ionise the analyte, separate the ions according to their m/z and detect the ions based on their m/z abundance" [2, 3]. MS measures the m/z of ions,

either positively or negatively charged, in the gas phase. The first work on separating ions by their m/z was done by J. J. Thompson in 1913 [4–6]. Research in the last century led to continuous growth of the MS field and MS itself found applications across many scientific areas.

A mass spectrometer consists of an ionisation source, mass analyser(s), detector and a computer for data collection and data analysis. A sample of interest is ionised and introduced into the mass spectrometer, where m/z of the ions is measured and the ion numbers at each m/z are registered by the detector [7]. Collected data are presented as mass spectra containing information about the ion's signal abundance (y-axis) versus m/z (x-axis).

1.1.1 Ionisation

Ionisation is a process of creating ions from neutral molecules, necessary for further m/z analysis of the ions in the mass analyser. Numerous ionisation methods exist. Electron ionisation (EI) and chemical ionisation (CI) are suitable for gas phase ionisation of volatile and thermally stable analytes [8], while biomolecules require softer ionisation techniques such as matrix-assisted laser desorption ionisation (MALDI) [9] and electrospray ionisation (ESI) [10]. Both MALDI and ESI were introduced in the 1980s and are considered as important milestones in MS development – these techniques allowed the study of biomolecules without their extensive fragmentation or degradation and analysis of compounds with high molecular masses. The work in this thesis makes use of ESI and nanoelectrospray (nanoESI) and these are further described below.

1.1.1.1 Electrospray ionisation

ESI was introduced by the group of Fenn *et al.* [10, 11] and is considered groundbreaking as it allowed production of intact ions of large species such as proteins and polymers [10]. The main purpose of ESI is to transfer ions from solution to the gas phase at atmospheric conditions. A voltage of a few kV (typically in range 1 – 4 kV) is applied between a heated

capillary delivering the sample solution and a counter-electrode, resulting in production of either positively or negatively charged droplets [10, 12]. When a strong electric field is applied, the drop at the end of the capillary containing liquid elongates under the pressure of the accumulated charge [8]. At the critical electric field strength, the cone elongates into a cone-jet, also called Taylor cone [13] (see Fig. 1.1), and charged droplets are released towards the counter-electrode. The droplets produced shrink due to the solvent evaporation until the accumulated charge on the surface reaches a critical value, known as Rayleigh limit. At this point, the surface tension is overcome by electrostatic (Coulombic) repulsion and droplet fission occurs (Fig. 1.1). To aid droplet desolvation and transfer of ions to the gas phase, sheath gas (such as nitrogen) is applied alongside the heated ESI capillary. ESI produces multiply charged ions; these are highly beneficial for analysis of analytes with high molecular weight (MW) – higher charge states allow the large analyte to be detected on mass spectrometers with a limited m/z range and facilitate dissociation of ions in tandem MS analysis (see section 1.1.3) [14].

The ESI flow rate is typically 5 – 20 µL/min [1], however sensitivity may be decreased for higher flow rates. Wilm and Mann implemented changes to ESI and developed nanoelectrospray (nanoESI) [15]. To introduce the sample into the mass spectrometer, nanoESI typically uses gold-coated pulled borosilicate capillaries with very fine tips of size approximately 1 – 2 µm and flow rates 20 – 40 nL/min [15, 16], however other nanoESI variations exist (see section 1.1.1.2). The main advantages of nanoESI compared to standard ESI is higher tolerance to salts present in the liquid, smaller diameter droplets and low sample consumption, ideal for analysis of biological samples.

Generally, there are three mechanisms describing the formation of gas phase ions [12, 14] (Fig. 1.1):

1. Charge residue model (CRM) – the droplet from the last fission process contains only one analyte molecule with excess charge (proposed by Dole in 1968).

2. Ion evaporation model (IEM) – evaporation of ions from the droplet occurs after last fission when the accumulated electric field on the surface reached a critical value

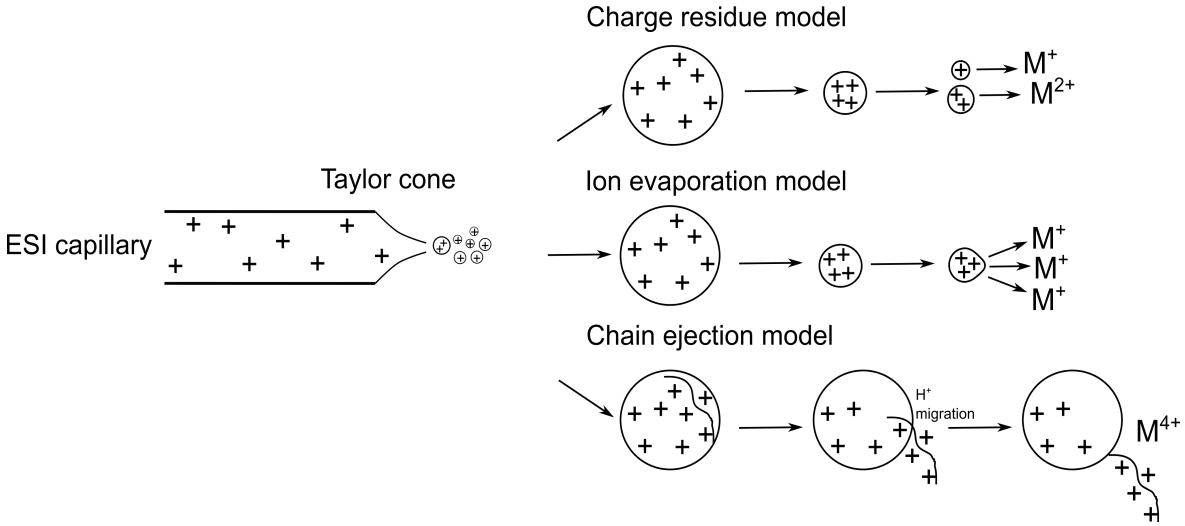


Figure 1.1: A schematic representation of ESI. Taylor cone is formed at the end of the ESI capillary, elongates and droplets are released. Formation of gas phase ions is described by the three mechanisms – charge residue model, ion evaporation model and chain ejection model.

(proposed by Iribarne and Thomson in 1976).

3. Chain ejection model (CEM) – electrostatic and hydrophobic factors cause unfolded polymer chains to move towards the droplet surface, where ejection from the droplet occurs (proposed by Konermann in 2013).

CRM describes release of large species such as peptides, globular proteins, native folded proteins or nucleic acid duplexes, while IEM is a good representation of ionisation of small molecules (analytes with low MW) and small inorganic ions [14, 17]. CEM explains that unfolded proteins under denaturing conditions are much more highly charged as a result of the ESI mechanism [18]. Unfolded proteins change their properties from hydrophilic to hydrophobic as the whole amino acid chain is exposed. The unfolded protein travels to the droplet surface and is gradually expelled while undergoing charge equilibration with the droplet (via mobile H^+ proton attached to the protruding tail) (Fig. 1.1) [18].

1.1.1.2 Liquid extraction surface analysis

Liquid extraction surface analysis (LESA) MS is a surface sampling technique which is coupled to nanoESI and was developed by Kertesz and Van Berkel [19, 20]. Analytes are

extracted *via* formation of a liquid microjunction between the solid surface and a pipette tip containing an extraction solvent system. The LESA process is described in detail in Fig. 1.2. LESA, now commercially available as Advion Triversa Nanomate platform (Advion, Ithaca, NY), introduces the sample into mass spectrometer *via* chip-based nanoESI (Fig. 1.2). The chip consists of 400 nozzles of diameter 2.5 – 5.5 μm etched in a silicon wafer [21]. A conductive pipette tip containing the sample comes into contact with one of the nozzles and low amounts of sample may be delivered over a long period of time [21]. For each extraction a new nozzle is used to prevent sample carryover. The independent sampling and ionisation process is beneficial especially for adding steps such as digestion or separation, e.g., ion mobility MS or liquid chromatography (LC)-MS. The whole LESA robotic system is controlled with dedicated software also developed by Advion.

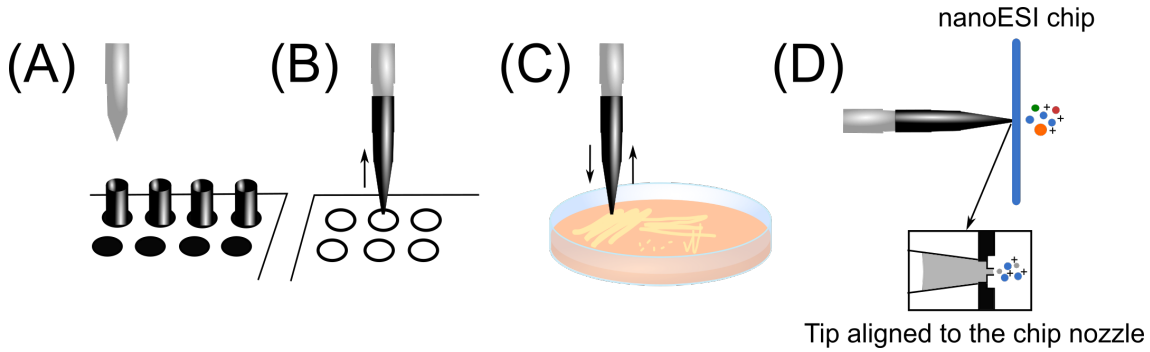


Figure 1.2: A schematic description of the LESA process. (A) The robotic arm picks up a pipette tip from the tip rack. (B) The robotic arm with the new pipette tip aspirates few μL of extraction solvent system from the solvent well of the microtiter plate. (C) The robotic pipette is relocated to the position above the sample, dispenses the extraction solvent system which is subsequently re-aspirated back into the tip with the analytes of interest. (D) The sample is introduced into the mass spectrometer *via* chip-based nanoESI. The inset shows tip alignment to the chip nozzle in greater detail.

In the last decade, LESA has been shown as a suitable technique for extracting small molecules [20, 22–24], lipids [25–29], denatured proteins [30–33], native proteins and protein complexes [34–38]. Quantitative studies employing LESA MS were performed on small molecules [22, 24, 39] and proteins [40]. One of the major strengths of LESA is the *in situ* analysis of proteins which have been detected from a variety of biological surfaces

including dried blood spots [30, 41, 42], tissue sections [31, 35], bacterial colonies [32, 43], yeast [44] and *in vitro* skin models [33]. Another advantage of LESA over other ambient ionisation techniques, such as desorption electrospray ionisation (DESI) or Flowprobe, is sensitivity – the extracted analytes are concentrated in a small droplet, delivered over a long time, suitable for MS/MS experiments. LESA can also be used for MS imaging (MSI), although the spatial resolution of 1 mm can be achieved due to software limitation, and hypothetically up to 400 µm which corresponds to the inner diameter of the pipette tip when "contact" LESA is applied [45].

1.1.2 Mass analysers

Once the analyte of interest is ionised and in the gas phase, the ions are separated based on their m/z in the mass analyser. Multiple types of mass analysers exist and their use depends on the analysis and sample type. A single mass analyser is capable of mass analysis but typically a combination of two or more is highly beneficial and present in many mass spectrometers. In the text below, three commonly used mass analysers – a linear quadrupole, a linear ion trap and an Orbitrap – will be described as these were used in the experimental work presented in this thesis.

1.1.2.1 Quadrupole

A linear quadrupole consists of four parallel cylindrical or hyperbolical rods. An oscillating electric field is used to separate ions according to their m/z [8]. A potential of opposite polarities is applied to the each pair of rods:

$$\begin{aligned} \phi_0 &= (U + V cos\ \omega t) \\ -\phi_0 &= -(U + V cos\ \omega t) \end{aligned} \tag{1.1}$$

consisting of DC (U) and RF ($Vcos\ \omega t$) components (Fig. 1.3) [1]. Ions oscillate in the quadrupolar field and oscillations consists of many different frequencies [46]. The ion motion is described by Mathieu equations and solutions to these equations are either

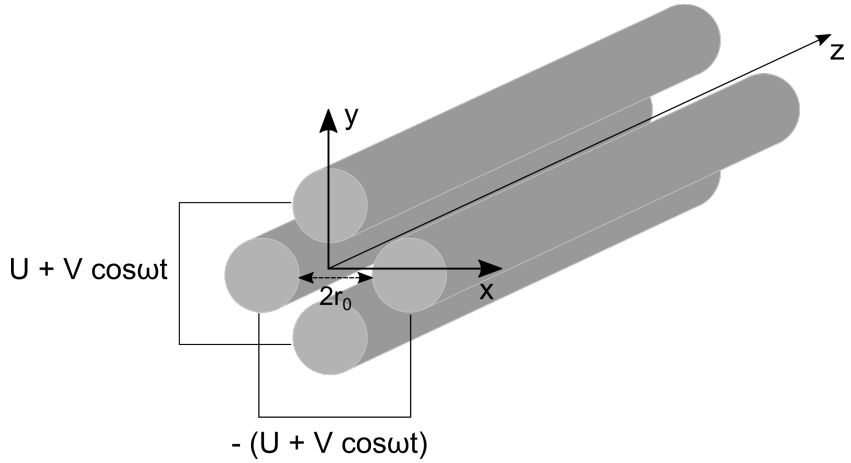


Figure 1.3: A schematic representation of a linear quadrupole mass analyser.

"stable" or "unstable" [46]. For ions to pass through the quadrupole, the ion trajectories must be stable in both x and y planes within $2r_0$ distance of electrodes (Fig. 1.3) and the stability path for the ion of particular m/z is defined by the U/V ratio. This way, the ions of selected m/z are transmitted and other ions with unstable trajectories strike one of the rods and cannot be detected.

Quadrupoles can be used as mass filters, mass analysers and ion guides. Higher sensitivity and low resolution is utilised for instruments such as triple-quadrupoles. For the instrumentation described in the section 1.1.2.4, linear quadrupoles are used as mass filter and ion guides, where a different analyser is used for performing the mass analysis.

1.1.2.2 Linear ion trap

The working principle of the linear ion trap (LIT) is similar to the linear quadrupole. A LIT consists of four parallel hyperbolical rods aligned around a central axis (Fig. 1.4) where a potential of opposite polarities (Equation 1.1) is applied to the each pair of the rods. The main difference between the LIT and quadrupole is that an additional DC potential is applied at the end sections in order to trap ions along the z-axis. The ions are accelerated into the trap and accumulated [1]. Buffer gas such as nitrogen or helium is present in the trap. Collisional cooling takes place, the ions lose their kinetic energies which helps them to align along the z-axis. Subsequently, the ions are ejected radially via

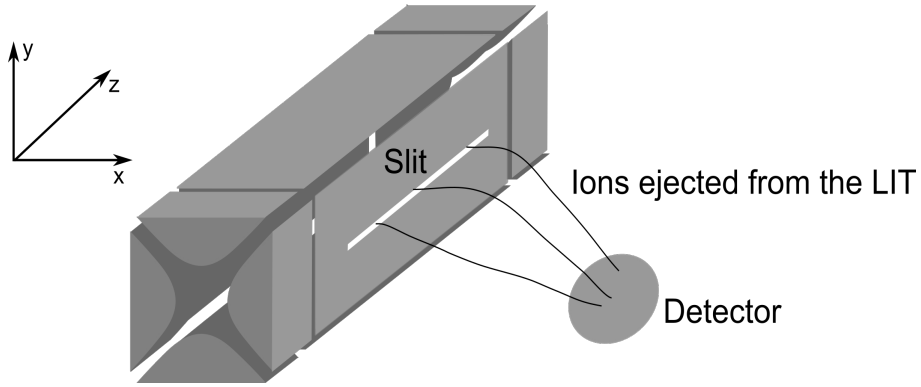


Figure 1.4: A schematic of a linear ion trap with ions ejected radially from the trap.

slits on the two opposite rods (Fig. 1.4) [8, 47]. LITs are used for ion isolation, tandem MS analysis and as mass analysers.

1.1.2.3 Orbitrap

The Orbitrap mass analyser was developed by Makarov and first introduced in 1999 [48]. The working principle of the orbitrap lies in trapping ions in an electrostatic field. The electrodes are specifically designed to produce a quadro-logarithmic potential distribution:

$$U(r,z) = \frac{k}{2}\left(z^2 - \frac{r^2}{2}\right) + \frac{k}{2}(R_m)^2 \cdot ln\left[\frac{r}{R_m}\right] + C \tag{1.2}$$

where r and z are cylindrical coordinates, k represents field curvature, R_m is characteristic radius and C is a constant [49]. The trapped ions form stable trajectories around the central spindle-shaped electrode. The trajectories consist of rotational, radial and axial motion. Frequencies of both rotational and radial motion depend on the initial radius R, while the frequency of the axial motion (along the z-axis)

$$\omega = \sqrt{\frac{e}{(m/z)}k} \tag{1.3}$$

where e is elementary charge, is independent of all initial ion velocities and coordinates [50]. The axial frequency is therefore used for calculation of m/z.

Ions are sent into the Orbitrap analyser in ion packets collected in a curved linear ion

trap termed C-trap. In the C-trap, ions are cooled down by presence of nitrogen gas. After a DC signal is generated, the ion packet is injected off-axis through the C-trap slit and additional focusing lenses into the orbitrap (Fig. 1.5). Differential pumping ensures that a high vacuum of in the Orbitrap chamber is achieved. After injection, the ions start oscillating around the central electrode and the image current is measured on the outer orbitrap electrodes (Fig. 1.5). The difference between the image currents is detected with a differential amplifier. The signal is digitised and subsequently transformed by using fast Fourier transform from time to frequency (m/z) domain (Fig. 1.5). A resolving power of above one million can be achieved [50].

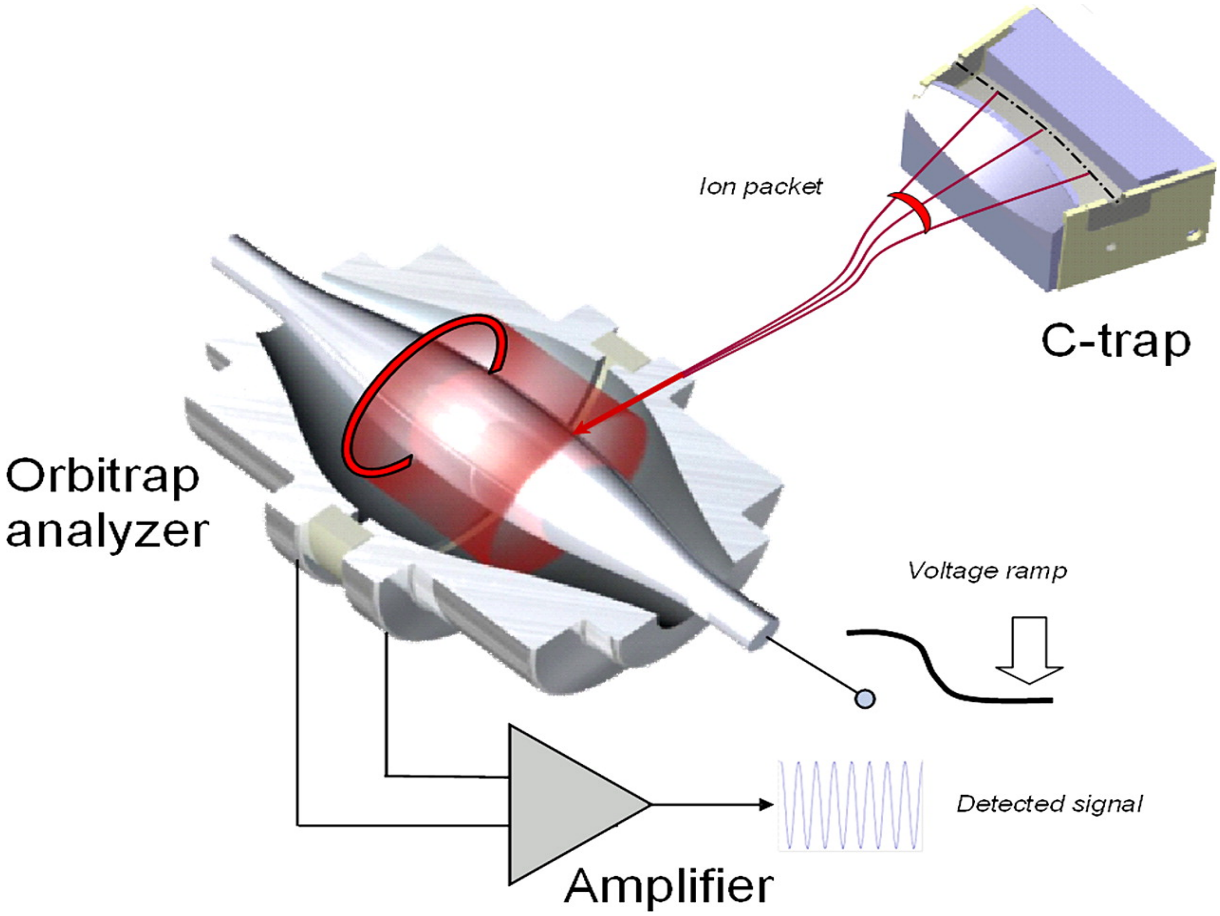


Figure 1.5: A schematic representation of an ion packet injected from C-trap into the orbitrap mass analyser. The ions oscillate around the central electrode and induce current which is detected by the differential amplifier. Reprinted with permission from [51]. Copyright 2013 American Chemical Society.

Fragmentation of ions occurs within the Orbitrap, however the analyser itself is not

suitable for tandem MS and is solely used for high resolution mass analysis. To overcome the limitations of MS/MS analysis, Orbitrap analysers have been coupled to front-end LIT or quadrupole. This instrument was commercialised as the hybrid LTQ Orbitrap and introduced for the first time in 2005 [51]. The main advantage of coupling with the LIT was its speed, sensitivity, MS^n capability, ability to perform electron transfer dissociation (ETD) and compatibility with LC-MS proteomics experiments [52]. Importantly, the instrument makes use of automatic gain control (AGC), where a pre-scan in the LIT allows the ion current within the specified mass range to be determined and subsequently only a specific number of ions (so-called AGC target) are present in the next scan [50,52]. The AGC feature together with the controlled injection time is crucial for quantitative analyses, control of the number of ions in packets and a decrease of space charge effects.

1.1.2.4 Orbitrap mass spectrometers

The next generation of LTQ Orbitraps introduced a stacked ring ion guide (S-lens), a dual-pressure LIT with faster scanning rate and a cell for higher-energy collision dissociation (HCD) fragmentation technique [50,53]. The S-lens was placed at the front to improve the ion transmission and ion beam focusing [53]. A new generation of Orbitrap analysers with higher electric field and thus improved resolving power was introduced as an Orbitrap Elite in 2011 (Thermo Fisher Scientific, Bremen, Germany), and the fragmentation technique ETD was made available [54]. A schematic of the Orbitrap Elite mass spectrometer is presented in Fig. 1.6. A different branch of the Orbitrap-based instruments, the Q Exactive (Thermo Fisher Scientific, Bremen, Germany) combined quadrupole, HCD cell and the Orbitrap mass analyser (Fig. 1.7). The Q Exactive was widely accepted for 'omics' applications for its high sensitivity, increased duty cycle and fast fragmentation in the HCD cell [50, 55].

Later in 2013, the first tribrid instrument incorporating quadrupole, LIT and the Orbitrap mass analyser was presented [58]. The main advantage of such instrument is an expanded option of MS/MS fragmentation techniques, sensitivity, higher mass range,

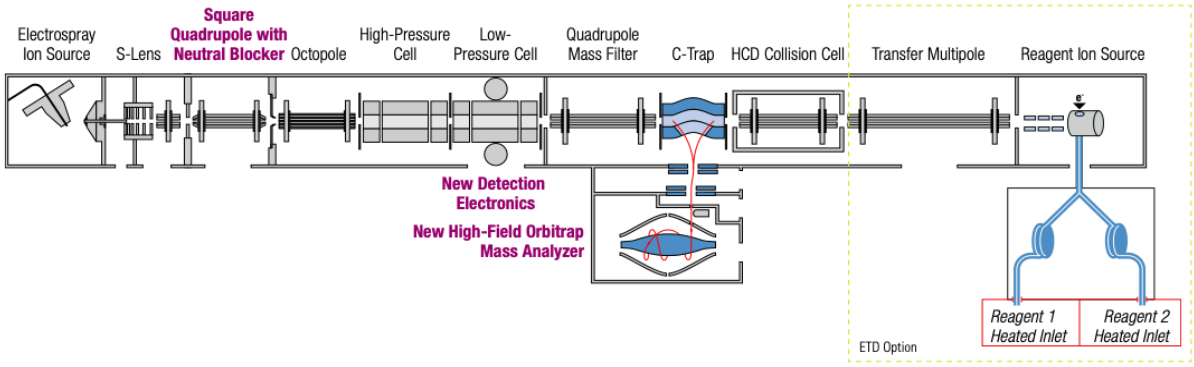


Figure 1.6: A schematic figure of the Orbitrap Elite LTQ instrument. Reprinted with permission [56].

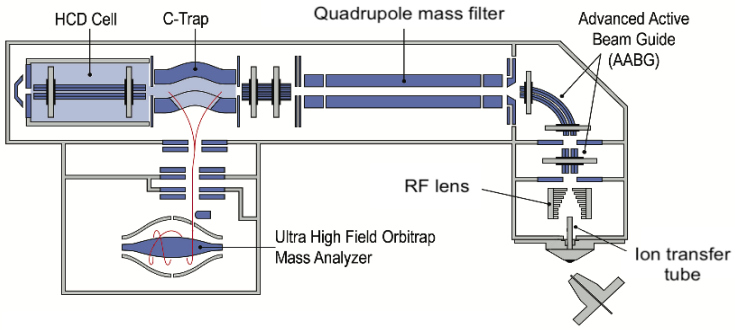


Figure 1.7: A schematic figure of the Orbitrap Q Exactive HF instrument. Reprinted with permission [57].

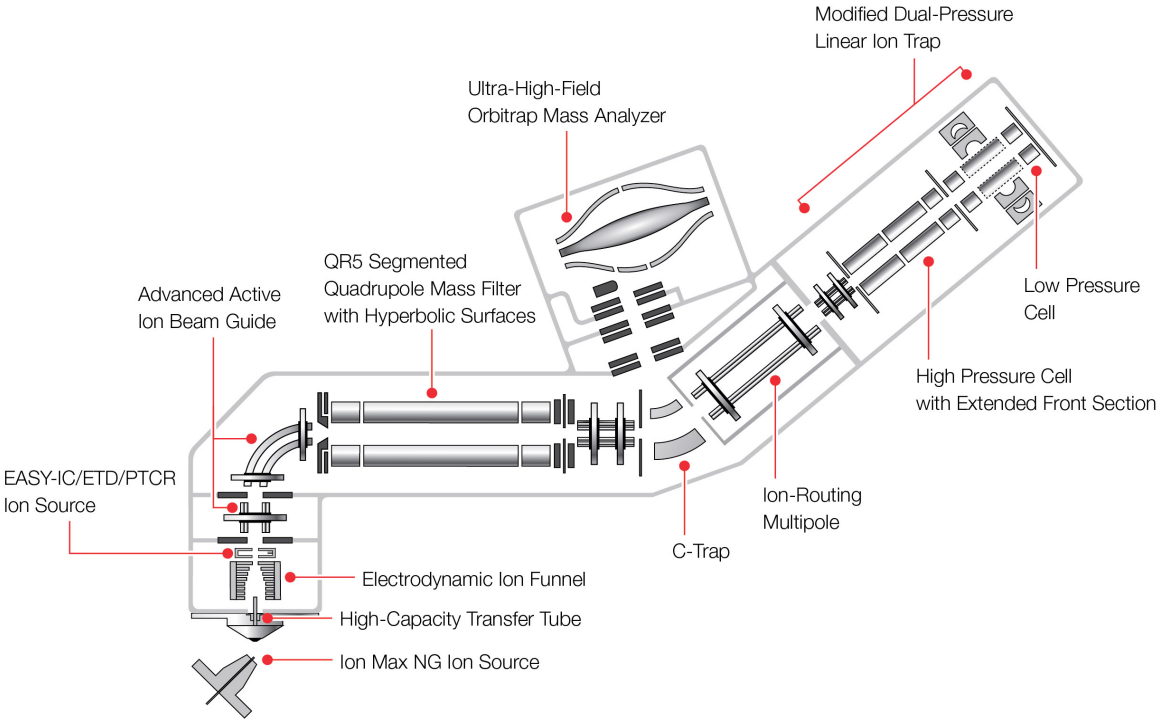


Figure 1.8: A schematic figure of the Orbitrap Eclipse tribrid instrument. Reprinted with permission [60].

fast parallel and coordinated operation of multiple mass analysers [50, 58]. The schematic figure of the Orbitrap Eclipse (Thermo Fisher Scientific, San Jose, CA) tribrid is presented in Fig. 1.8. This mass spectrometer includes ETD reagent ion source at the front-end (as opposed to the Orbitrap Elite, for comparison see Fig. 1.6 and Fig. 1.8). ETD can be combined into EThcD or ETciD with other fragmentation techniques, with further addition of ultra-violet photodissociation (UVPD) [59]. The HCD cell was exchanged for an ion routing multipole (IRM), a multipurpose ion storage device which is used for routing ions into the mass analysers [58] or HCD fragmentation. The mass range can be extended up to 8000 m/z (compared to maximum of 4000 m/z on the Elite) which is suitable for analysis of antibodies, drug-antibody conjugates, native proteins and protein complexes [52].

1.1.3 Tandem mass spectrometry

Tandem MS (also referred to as MS/MS or MS^n) is a technique by which structural information about an ion is determined from its fragments. Typically, a precursor ion with a specific m/z is isolated and subsequently activated, i.e., the ions are energised which promotes fragmentation and the m/z of the fragments are recorded. The activation mechanism determines the type of products or fragments formed [61]. In the context of proteins and peptides, fragmentation nomenclature depends on which backbone bond or side chain is cleaved as shown in Fig. 1.9 [62]. If a signal for fragments is detected in a mass spectrum and correlated to the fragments of a certain protein/peptide, the amino acid sequence can be calculated [63]. Three types of activation methods are briefly described here: collision-induced dissociation (CID, occasionally reported in literature as CAD – collision activated dissociation), higher-energy collision dissociation (HCD) and electron transfer dissociation (ETD).

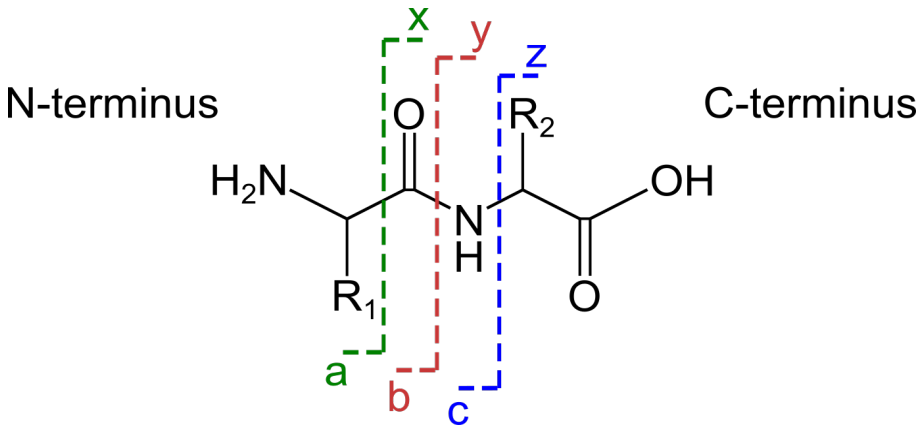


Figure 1.9: Fragmentation nomenclature of peptides. CID and HCD yield N-terminal b and C-terminal y fragments by cleaving the amide CO—NH bond. ETD uses a different activation mechanism which cleaves the Cα—N bond, thus generating N-terminal c and C-terminal z fragments.

CID is a low energy activation method, referred to as resonant or slow-heating method, where ions collide with molecules of an inert gas such as helium or nitrogen. Once the collisions take place, some of the initial kinetic energy is transferred into internal (vibrational) energy, resulting in the breakage of weakest bonds, which is CO—NH amide

bond on the peptide backbone (Fig. 1.9). CID yields N-terminal b fragments and C-terminal y fragments (Fig. 1.9). Fragments with neutral losses (NH_3 or H_2O) are commonly observed in CID mass spectra. The fragmentation pattern of peptides and proteins depends on the charge state selected for fragmentation and is also related to the number of mobile protons [62]. The mobile proton model [64] explains that after ion activation, the proton from the basic sites of the amino acid residues is transferred to peptide linkages (either side of the amide bond), thus initiating fragmentation [65]. Ion activation by CID takes place in a collision cell. Certain limitations of this technique have been observed including limited energy input that restricts dissociation of larger proteins. Another drawback when identifying proteins and their proteoforms (protein variants, see section 1.3) is labile post-translational modification (PTM) losses (such as phosphorylation) during the ion activation process [66] and lack of disulfide bond cleavage. Nevertheless, CID is one of the most robust and widely implemented fragmentation method available on practically every mass spectrometer and used in numerous proteomics studies over the last years [62, 67].

HCD – a variation of beam-type CID activation – was developed by Makarov and co-workers [68] and is available on the Orbitrap mass spectrometers used in this work (see also section 1.1.2.4). An octopole cell was aligned to the C-trap device (at the back-end of the instrument), where ions were transferred from the C-trap and collisions of ions with nitrogen gas took place [68] (see also Fig. 1.6). Most recently, the configuration of the instrument was changed with HCD taking place in the IRM (see Fig. 1.8). HCD, unlike CID in the LIT, does not suffer from the low mass cutoff as the ion activation is independent of the ion trapping [67, 68]. While both CID and HCD produce b fragments and y fragments (Fig. 1.9), there is a difference observed between these two methods. CID in a trapping instrument results in one fragmentation event per precursor, while the fragment ions in a HCD cell may undergo further collisions, producing more fragments resulting in differences between the ion intensities in CID and HCD mass spectra [69].

ETD is a method which uses ion-ion chemistry to generate c-type and z-type fragments

(Fig. 1.9) [70]. The underlying mechanism involves transfer of an electron from a radical ion to a protonated peptide, resulting in cleavage of backbone Cα—N bond [71]. The main advantage of ETD over CID and HCD is that it preserves labile PTMs such as phosphorylation, glycosylation, nitrosylation or sulfonation [71, 72]. ETD can be used for fragmentation alone but in recent years, supplemental activation was added to ETD aiming to improve the protein sequence coverage. In the new instruments, the charge reduced precursors are allowed to undergo further fragmentation by CID (ETciD) in the LIT or HCD (EThcD) in the IRM. The c, z, b and y fragments are complementary which results in higher sequence coverage and better characterisation of protein PTMs.

1.2 Ion mobility spectrometry

MS itself is a powerful tool for analysis of biological samples. These can be, however, very complex and often require additional separation prior to MS analysis. MS is therefore coupled to various separation techniques including ion mobility spectrometry (IMS) which separates ions in the gas phase. Further text in this section will describe general working principles of IMS, and high-field waveform ion mobility spectrometry (FAIMS) which was used in the work presented in this thesis.

The main definition characterises IMS as a technique for studying behaviour of ions exposed to electric field in a region filled with carrier gas [73, 74]. The main advantages of IMS is analysis of gaseous samples at ambient pressure and temperature, fast screening capability and high sensitivity, highly beneficial for analysis of chemical vapours [75, 76]. Advances in IMS led to its application to a variety of biological samples from small molecules (metabolites, lipids) to large protein assemblies [77].

The mobility of ions K is described as the movement of ions with velocity v_d in the electric field E:

$$K = \frac{v_d}{E} \tag{1.4}$$

The ions travel through a gas, therefore K also depends on various experimental

parameters, such as gas density which is directly related to gas pressure p and temperature T [74]. K is therefore expressed in experimental conditions as reduced mobility K_0:

$$K_0 = K \cdot \frac{p}{p_0} \cdot \frac{T_0}{T} \tag{1.5}$$

where p_0 is standard pressure and T_0 is standard temperature. K_0 depends on T, nature of the buffer gas and E/N (N corresponds to the number of molecules per unit volume) [74, 75]. Mobility of ions is assumed to be independent of E/N ($K_0 = K_0(0)$) for low electric fields, however for high electric fields dependence of E/N occurs [73, 74, 78]. E/N ratio can also be expressed in Townsends, where $1\,\mathrm{Td} = 10^{-21}\,\mathrm{V{\cdot}m^2}$ [77].

IMS may be used to determine ion-neutral collision cross sections (CCS) from structural models of ions [74]. CCS represents a rotationally averaged collision area of an ion (charged analyte) and a neutral (ambient air, N_2, He) [79]. CCS values of the ion-neutral pair are analyte specific, therefore are useful not only for analyte identification but also for structural biology, i.e., to differentiate between two conformers of the same protein.

Linear methods for measuring ion mobility (independent of E/N) include:

1. Drift-tube IM-MS (DTIMS) – ions move in a uniform weak electric field that is propagated through a drift region.

2. Travelling wave IM-MS (TWIMS) – an oscillating electric field is applied that creates a wave pushing ions through the drift region.

3. Trapped IM-MS (TIMS) – ions are carried in the gas flow and the static electric field is applied opposed to the direction of ions.

4. Differential mobility analysers (DMA) – ions travel between two electrodes and the gas flow is applied perpendicularly to the trajectory of ions.

All of the described IMS techniques are capable to determine CCS values. FAIMS is a non-linear method (dependent on E/N) and is discussed in more detail in the text below.

1.2.1 High-field asymmetric waveform ion mobility spectrometry

FAIMS is a technique where ions are carried in a gas flow and an oscillating electric field is applied perpendicularly to the ion trajectory (Fig. 1.10). In contrast with the IMS described above (DTIMS, TWIMS, TIMS and DMA), FAIMS cannot be used to directly determine the CCSs, however, FAIMS has gained popularity in the recent years and is used as a separation device improving sensitivity and significantly reducing background noise (hence improving signal-to-noise (S/N)).

To introduce the fundamentals of FAIMS – the equation 1.6 expresses the mobility of ions at high electric fields:

$$K_0\left(\frac{E}{N}\right) = K_0(0)\left[1 + \alpha_2\left(\frac{E}{N}\right)^2 + \alpha_4\left(\frac{E}{N}\right)^4 + ...\right] \tag{1.6}$$

where the function $\alpha(E/N)$ represents the dependence of ion mobility on electric field (at ambient temperature and constant gas pressure) and α are coefficients for the dependence of mobility of an ion on the electric field strength [75, 80]. The main principle of FAIMS lies in the differing mobilities of ions in high and low electric fields. Ideally, a square waveform signal is applied (Fig. 1.10A), the amplitude of which is referred to as the dispersion voltage (DV) (giving rise to the dispersion field (DF)). The waveform is periodic, asymmetric and perpendicular to the ion path, while the overall net electric field applied during one waveform equals zero [73] (see Fig. 1.10A). In such conditions, the ions follow a sawtooth-shaped trajectory and eventually collide with one of the electrodes (Fig. 1.10B). A DC voltage is superposed to correct for the ion drift [81], known as the compensation voltage (CV) (giving rise to the compensation field (CF)). Scanning across the CV voltages can selectively transmit ions through the FAIMS device (Fig. 1.10B).

Two geometries of FAIMS devices exist (Fig. 1.10): planar (p-FAIMS) and cylindrical (c-FAIMS) and both types can be coupled to ESI and nanoESI ion sources. The p-FAIMS

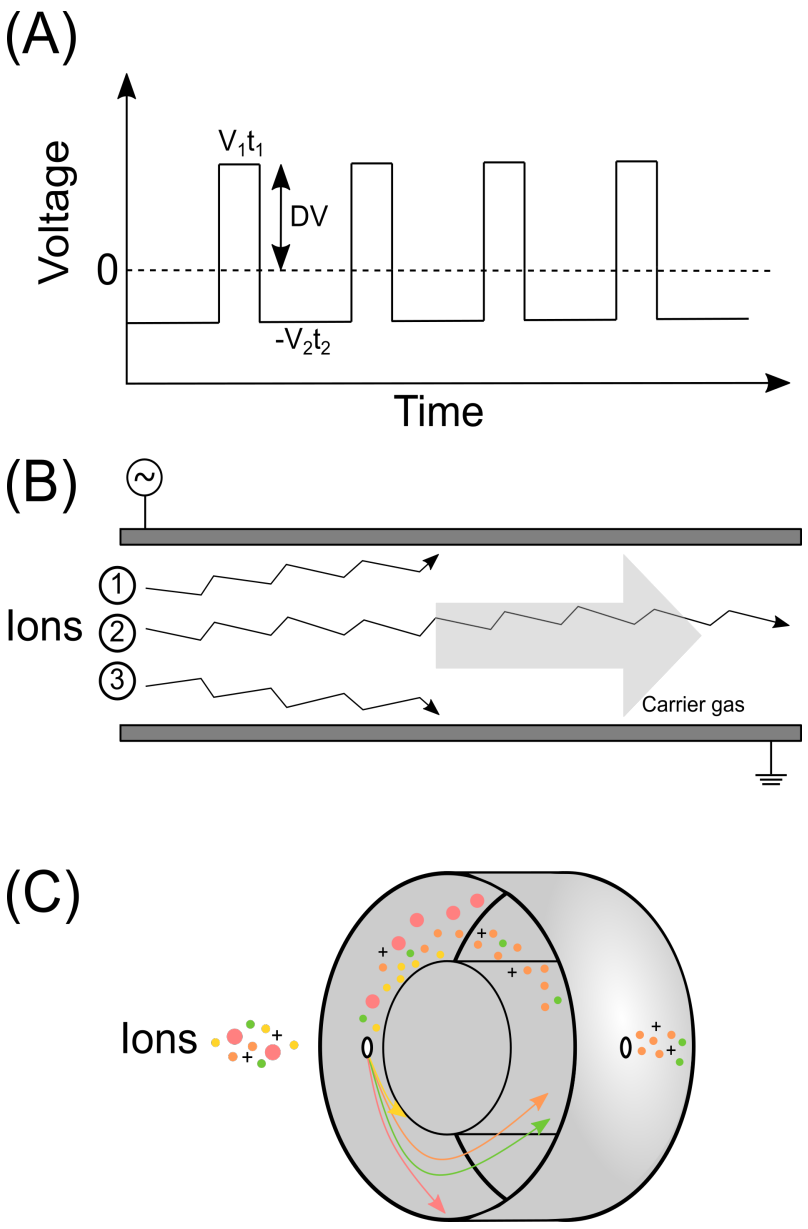


Figure 1.10: The main principle of FAIMS. (A) A square waveform is applied (DV or DF) such that the overall net electric field during one waveform equals zero. (B) Movement of three ions between the FAIMS electrodes. Ions follow a sawtooth-shaped trajectory and application of a "correct" CV voltage results in selective ion transmission (in this case ion 2). This figure represents planar FAIMS geometry. (C) Cylindrical FAIMS geometry. The ions are separated in the gap between the two coaxial cylinders. Red, orange, yellow and green dots represent different ions and their corresponding trajectories (see arrows) inside the FAIMS device.

device consists of two planar electrodes, through which ions are carried in a buffer gas (Fig. 1.10B). Introduction of FAIMS chips resulted in development of miniaturised p-FAIMS devices [82]. Ultra-FAIMS device (Owlstone, Cambridge, UK) is a type of miniature p-FAIMS which uses a silicon microchip mounted onto a a printed circuit board [80, 83].

The cylindrical FAIMS (c-FAIMS) geometry uses two coaxial cylinders in which ions are separated in space between the cylinders (Fig. 1.10C). The waveform and CV voltages are applied to the inner electrode [81]. One of the advantages of c-FAIMS over p-FAIMS is that the sensitivity increases with increasing DV [84]. This cannot be achieved on a p-FAIMS device as increase in the field strength would cause the ions to collide with the electrodes (hence decreasing the ion transmission). The most recent prototype of c-FAIMS is marketed as FAIMS Pro (Thermo Fisher Scientific, San Jose, CA). The FAIMS Pro uses a decreased gap between the electrodes compared to the previous c-FAIMS model and N_2 only as a carrier gas [84]. These modifications resulted in improvement of peak capacity and S/N [84, 85], advantageous for separation of complex mixtures.

1.2.1.1 FAIMS in analysis of peptides and intact proteins

Peptides and proteins were analysed by FAIMS for the first time by Purves and Guevremont [86–88]. It was shown that short glycine peptide chains were separated at different CVs based on the peptide chain length and it was observed that FAIMS is capable of discrimination between the cytochrome *c* charge states [86] and ubiquitin conformers [87]. FAIMS analysis of pig haemoglobin tryptic digest showed that very low correlation is observed between ion m/z and CV, making FAIMS an ideal tool for ion separation prior to MS analysis [88]. The presented benefits suggested that FAIMS is a technique orthogonal to MS suitable for separation of many analyte classes including peptides and proteins. FAIMS, particularly useful for improvement of S/N, found its application in analysis of complex biological samples. Many works exploit FAIMS coupled to ambient ionisation techniques such as desorption electrospray ionisation (DESI), Flowprobe and LESA for direct analysis of proteins from a variety of biological substrates and these will

be discussed in the text below.

DESI is a technique which uses a stream of charged solvent to desorb analytes from sample surface [89]. DESI MS was coupled to p-FAIMS chip-based device for analysis of mouse kidney tissue, where 11 protein species were detected [90]. Furthermore, the optimised DESI-FAIMS workflow was applied to MSI of mouse brain and human normal and cancerous ovarian tissue sections. It was shown that the spatial distribution of proteins is reproducible across the mouse brain tissue samples. The comparison of normal and cancerous tissues resulted in differences in relative abundances of protein ions, where the S100A6 protein was more abundant in the cancerous tissue. Overall, a significant improvement of S/N of detected proteins was observed together with the higher image contrast and quality while performing DESI-FAIMS MSI.

Flowprobe (also referred to as liquid microjunction surface sampling probe (LMJ-SSP)), is a type of surface sampling technique using a continuous flow of solvent capable of raster sampling. Flowprobe coupled to chip-based p-FAIMS was used for MSI of rat brain and allowed detection of 84 protein species while 67 were detected by the Flowprobe alone [91]. The use of Flowprobe-DESI MS increased the S/N of protein species 7-fold compared to Flowprobe MS. Flowprobe-DESI MSI of human normal and ovarian cancer tissue sections showed that the relative abundance of three out of four selected proteins was higher in the tumor region when compared to the necrotic region and normal tissue.

LESA MS coupled to the chip-based device (p-FAIMS geometry) was successfully employed to investigate proteins from multiple substrate types including dried blood spots [92], tissue sections [93,94] and bacterial colonies [93,95]. In these studies, FAIMS allowed separation of lipids from the haemoglobin subunits [92], increased the protein S/N, reduced background noise, separated various classes of molecular species and increased the number of detected proteins when compared to the no-FAIMS experiment [93–95]. The FAIMS Pro device (c-FAIMS) coupled to LESA MS dramatically improved the number of proteins detected from rat brain (7-fold), testes and kidney (10-fold) tissue sections when compared

to p-FAIMS [96]. More recently, LESA c-FAIMS MS was explored in combination with native MS and demonstrated that intact protein complexes are transmitted through the FAIMS Pro device without disruption of the non-covalent interactions [37] while also improving S/N and protein MS^n analysis [38].

All of the above-mentioned studies show that FAIMS (both geometries) significantly reduces background noise (thus improving S/N) and its capability to separate ions from complex biological samples without extensive sample pre-fractionation or preparation. The numbers of detected proteins species are always higher with the FAIMS incorporated in the workflow than with the FAIMS absent. There are, however, some drawbacks commonly observed during the FAIMS experiments. P-FAIMS devices enhance S/N but also reduce signal intensity, often precluding the tandem MS analysis and protein identification. This has been observed in multiple studies [90, 91, 95] and may be considered problematic especially when the proteins of interest are detected only with the FAIMS device "on". Despite the fact that FAIMS is an effective separation device, ion suppression and matrix effects should still be considered as these may hinder detection of the analyte of interest [91, 97]. Furthermore, FAIMS separates ions based on their structural properties [87], but both p- and c-FAIMS cannot be used to determine CCS values. Thus, any information about the protein tertiary and quaternary structure or stoichiometry (if protein complexes are studied) cannot be delivered [37]. Despite the disadvantages, the use of FAIMS in proteomics workflow is still considered to be highly beneficial, particularly for direct analysis of complex biological samples.

1.3 Top-down mass spectrometry

Top-down (TD) MS is a branch of mass spectrometry which focuses on the analysis of intact proteins and proteoforms [98]. The term proteoform was first introduced by the group of Kelleher in 2013 [99], unifying the terminology for different protein variants which are products of gene splice variants, amino acid mutations, post-translational

modifications (PTMs) and post-translational cleavage (Fig. 1.11A) [100, 101]. Recently, a five level system for proteoform classification was published, describing four types of ambiguity (PTM localisation, PTM identification, amino acid sequence and gene) in proteoform identification [102]. Level 1 represents no ambiguity (all ambiguity types known), while level 5 contains information only about the observed mass (none of the ambiguity types known). TD MS was introduced by McLafferty and co-workers in 1999 [103]. In the "traditional" bottom-up proteomics (BUP) approach, a protein sample is digested by an enzyme into smaller peptides before the analysis, therefore the information about the sample PTMs, mutations or single nuclear polymorphisms (SNPs) may be lost (Fig. 1.11). In TD MS, the intact protein of interest undergoes MS/MS analysis and the protein fragments recorded in the tandem mass spectra are subjected to a database search (Fig. 1.11B). It is clear that intact protein analysis can provide a deeper insight into PTMs and offers information not available by the BUP approach, while no digestion is required, hence reducing the sample preparation time [104]. Most recently, TD MS has been shown as a valuable tool for analysis of native-like proteins and protein complexes, allowing their structure and non-covalent interactions to be studied [105]. Moreover, relative proteoform quantification can also be achieved [106].

TD MS is still not as widespread as BUP due to its inherent challenges. Initially, the major limitation for TD MS was instrumentation. Precise determination of different proteoforms and potential mass shifts in protein sequences requires high-resolution mass spectrometers and the only such instruments available at the time of TD MS introduction were Fourier-transform ion cyclotron resonance (FTICR) instruments. FTICR mass spectrometers are expensive, less user-friendly (require specialised training), utilise super-conductive magnets (therefore are also high-maintenance) while achieving high mass resolving power and mass accuracy [51, 106, 107]. Despite disadvantages, FTICR is still preferred for analysis of large proteins, their in-depth proteoform characterisation and labile PTMs mapping [106]. The expansion of the TD field and more widespread usage of TD MS began with the introduction of commercial high-resolution Orbitrap instruments

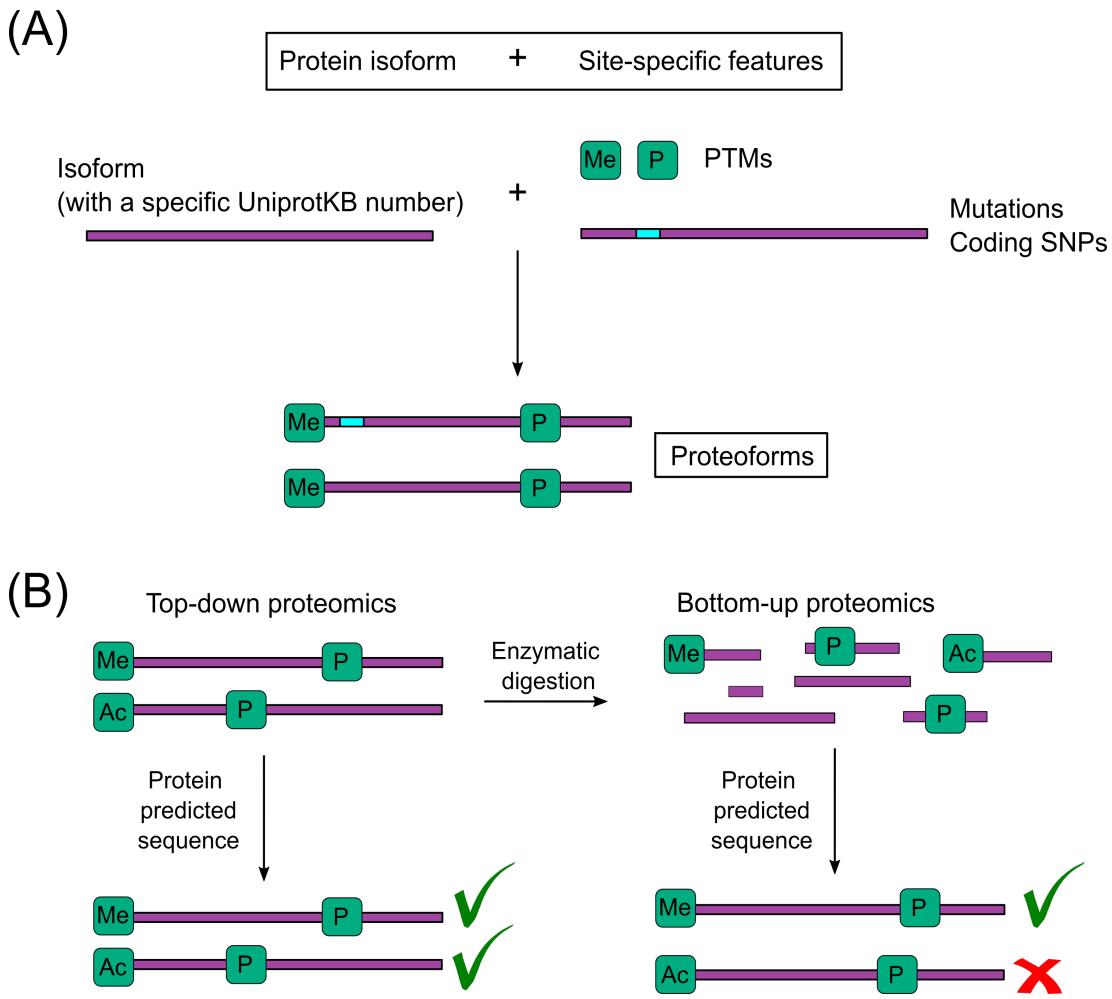


Figure 1.11: A schematic description of proteoforms and a comparison of TD MS and BUP approaches. (A) The available databases are gene-centric, i.e. contain information about protein sequences. Proteoforms are formed from those sequences by addition of PTMs, occurrence of mutations or single nucleotide polymorphisms (SNPs), while this information is not always known and recorded in the database. (B) Proteoforms are studied in their intact form by the TD MS workflow (left) which allows more precise PTM characterisation compared to the BUP approach (right), where proteoforms are enzymatically digested into smaller peptides and the predicted sequence is inferred based on the identified peptide sequences. In this process, PTMs or sequence variants information may be either lost (during the digestion) or misassigned (see bottom right).

which were more affordable, compact and combined multiple types of powerful mass analysers (see sections 1.1.2.3 and 1.1.2.4). As well as the progress of instrumentation, introduction of new fragmentation techniques such as electron-capture dissociation (ECD) [108], ultraviolet photodissociation (UVPD) [109], electron ionisation dissociation (EID) [110,111], surface-induced dissociation (SID) [112], ETD and HCD (see also section 1.1.3) contributed significantly to the TD analysis of proteins, highlighting TD MS as a robust technique for proteoform characterisation. In theory, TD MS can achieve 100% protein sequence coverage and elucidate the complete proteoform PTM profile [98].

TD MS has been applied to various sample types including purified intact proteins, protein mixtures or complex biological samples. To reduce sample complexity, offline separation can be coupled to online separation techniques such as liquid chromatography (LC), capillary electrophoresis (CE), ion exchange chromatography, size exclusion chromatography or ion mobility for optimal sample preparation [101, 113]. A comprehensive publication describing various sample preparation protocols for intact TD protein analysis was recently published by the Consortium for Top-Down Proteomics [114]. A simple decision tree allows a correct protocol to be found for sample cleanup, preparation and MS analysis for technically any protein sample analysed under either denaturing or native conditions [114].

While BUP is the standard high-throughput approach used for technically any sample type, the applicability of TD MS for such analyses proved challenging. Human samples are very complex and contain a number of proteoforms for each protein sequence due to various sources contributing to the overall proteoform diversity [115, 116]. High-throughput TD MS approaches have been applied to more simple microbial samples. Recently, there has been an initiative to characterise the whole human proteome [117]. The aim of that work is to map all sources of proteoform variations and thus better understand human health and disease. High-throughput TD MS was reported for the first time by Tran *et al.* [118]. In that work, more than 3093 proteoforms as products of 1045 genes were identified from HeLa cells, while proteins with MW up to 105 kDa were detected. The analysis

included extensive fractionation with subsequent LC MS nanocapillary separation. Since then, several studies applied the TD MS approach to characterisation of proteoforms from human samples. A protocol was developed by Toby and co-workers [116] for TD analysis of human peripheral blood mononuclear cells (PBMCs) with the aim of unifying the clinical sample preparation process and its TD analysis. The protocol describes extensive sample preparation including pre-fractionation, LC separation, considerations for MS analysis and data analysis. Despite these efforts confirming that TD MS is an outstanding tool suitable for analysis of proteoforms, there is still an urgent need for improvement in accuracy, reproducibility, sample analysis time and the applicability in clinical settings [119].

While it is clear that a remarkable improvement has been observed in the TD MS and that this approach is capable of high-throughput performance, data analysis still remains a challenge. Multiple TD MS database search engines have been developed over the years, namely ProSight [120], Mascot TD [121], MSAlign+ [122], TopPIC [123], pTOP [124], Informed Proteomics [125] and Proteoform Suite [126]. All of the software focuses on assignment of terminal fragments, however the information about internal fragments is usually lost. Loo and co-workers [127] introduced software capable of identification of internal fragments, hence improving the sequence coverage and protein modification sites. ProSightPC was developed as the first tool for TD data by Kelleher and co-workers [120,128,129]. The process of a typical TD search is as follows: before initiating the database search (the database is usually downloaded from UniProtKB website, www.uniprot.org), the tandem mass spectra are deconvoluted, i.e., the monoisotopic masses and charge states of the observed protein species are determined, and a peak list is created. Subsequently, matching of the observed fragments against the theoretical fragments from a selected database begins. Three types of searches are available: biomarker (for truncated proteins), absolute mass (the whole protein sequence) and sequence tag (searching for a specific sequence within database). A delta-mass (Δm) mode allows detection of mass shifts and possible amino acid mutations. To evaluate the correctness of the search results, ProSightPC uses a scoring system [128, 130]. Once

the search is completed, a list of identified proteoforms (if any) is generated and the proteoforms are usually validated manually. The National Resource for Translational and Developmental Proteomics at Northwestern also offers a license-free software ProSight Lite which aids the precise PTM localisation, also used for targeted proteoform identifications. To address the high-throughput data analysis, Thermo Scientific has incorporated the ProSightPC search nodes into the Proteome Discoverer search engine (commonly used for BUP data), thus enabling false discovery rate (FDR) determination. FDR is a measure of acceptance of how many incorrect matches may be present in the dataset [131] or, in other words, a rate of false positives in accepted hits [132]. FDR calculation is not trivial and needs a different approach than for the BUP data, while improvements for calculating the FDR are still under development [101, 133].

1.3.1 LESA MS in top-down protein analysis

In recent years, a focus has been placed on development of MS tools for direct *in situ* analysis of intact proteins [38]. Ambient surface sampling techniques including DESI and liquid microjunction-based techniques nanoDESI, Flowprobe, MassSpec Pen [134] or LESA were introduced. Each of the presented techniques have their advantages and disadvantages, while applied to analysis of small molecules, lipids and proteins. The focus here is placed on LESA (see also section 1.1.1.2), which offers several benefits to intact protein analysis such as low sample consumption and its pre-concentration in a small droplet, possibility of coupling to liquid separation and gas phase separation techniques or, in particular, no restriction in terms of the surface sampled. Investigation of dried blood spots (DBS) by TD LESA MS resulted in detection of α- and β-haemoglobin subunits, eight haemoglobin variants (based on mass shifts) [30, 41] and a possible indication of β-thalassemia major (based on the absence of the β-globin chain) [42]. The direct analysis of intact proteins from blood spots presents a possibility of clinical testing of newborns by TD LESA MS approach [30]. Numerous studies were published using TD LESA MS for direct analysis of intact proteins from thin tissue sections in which both denaturing and

native extraction solvent systems have been used. TD LESA MS was applied to study non-alcoholic steatohepatitis (NASH) tissue and was capable to detect a putative NASH biomarker fatty acid binding protein (FABP) containing a Thr→Ala substitution [31]. Moreover, BUP and TD MS analysis of NASH tissue were compared. Although BUP allowed identification of a much higher number of proteins, the study concludes that TD MS is more suitable for distinguishing the FABP variants [31]. Mouse liver and brain were also analysed [93, 94] while employing FAIMS (see also section 1.2.1.1). Native analysis of the same tissue types revealed presence of proteins not observed by denaturing LESA MS [35]. Most recently, native TD LESA MS detected and identified proteins up to 47 kDa by using an improved detergent-based extraction solvent system [45]. The use of TD LESA MS for analysis of living bacterial colonies was also successfully demonstrated (see text below).

1.4 Identification of microorganisms by mass spectrometry

Bacterial samples were studied by MS for the first time in the 1970s [135–137]. The technique used for bacterial characterisation was pyrolysis MS which uses heat to decompose material that is subsequently analysed by gas chromatography MS (Py-GC-MS). The work of Simmonds [135] investigated correlation of pyrolysates of *Micrococcus luteus* and *Bacillus subtilis* (var. *niger*) with those of meteoritic and fossil organic matters as a part of experiments for Martian mission in 1975. It was concluded that Py-GC-MS yields fragments indicative of biomolecules, hence suggesting evidence of life. Meuzelaar and Kistemaker [136] showed that fingerprints mass spectra of *Neisseria sicca* are reproducible, but the mass range studied was only up to 50 m/z. Anhalt and Fenselau [137] used lower pyrolysis temperature (compared to the previous studies) to analyse seven lyophilised bacterial species including both Gram-positive and Gram-negative microbes in the mass range up to 800 m/z. The differences between the

bacterial species were clearly observed in the mass spectra and this work is considered a breakthrough in MS analysis of bacteria.

The introduction of soft ionisation techniques allowed analysis of biomolecules without disrupting their structure, thus allowing more information to be obtained about chemical composition. One technique in particular, matrix-assisted laser desorption ionisation time-of-flight (MALDI TOF) MS, overtook the MS field for bacterial identification. MALDI is based on absorption of laser light by a solid sample layer covered with an appropriate matrix [1]. The ionisation process occurs under vacuum conditions where the sample is rapidly heated by the laser light and ionisation eventually occurs [1, 8]. MALDI is typically coupled to a TOF mass analyser which separates ions in a field-free region according to their velocities and the time that it takes for the ions to move from the source to the detector. The first demonstration of MALDI TOF MS for bacterial characterisation was reported in 1996 [138–140]. Currently, the use of MALDI TOF MS for analysis and microbial identification is well-established, with the use of dedicated software for spectral fingerprinting [141]. The technique is also FDA approved for use in clinical settings and is therefore considered as a gold standard MS approach for microbial characterisation.

Despite the numerous advantages of MALDI TOF MS, some drawbacks have been observed in bacterial and fungal analysis including sample preparation requirements and the fact that the sampling process takes place under vacuum conditions, precluding analysis of live colonies. Ambient ionisation MS techniques including rapid evaporative ionisation MS (REIMS), DESI, nanoDESI, Flowprobe and LESA MS overcome these limitations. REIMS uses alternating electric field causing thermal disintegration of microbial cells. During this process an aerosol is produced (containing the analytes of interest) and introduced into mass spectrometer [142]. REIMS can be used for direct analysis of living colonies while classification and characterisation of bacteria and fungi was achieved [142–144]. Species with the highest abundance presented phospholipids [143], while 28 bacterial and five *Candida* species were classified at the species level with 95.9% and 98.8% correctness, respectively. DESI MS has been used for the

detection of bacteria from swabs and the spectral features observed suggested that MS-based methods are suitable for medical applications [145, 146]. Furthermore, the examination of swabs by DESI MS was capable of detecting metabolite differences between *Lactobacillus*-dominating and *Lactobacillus*-depleted vaginal microbiome (suppression of *Lactobacillus* spp. may indicate a host immune response) [146]. Different papers describe the DESI MS analysis of fungal metabolites imprinted on various substrate types [147,148]. Imprints of fungal monocultures and co-cultures were also analysed, while focusing on metabolite detection [149]. Quantification of amino acids from *Cordyceps* fungus extracts by DESI MS showed potential to be implemented in food quality screening [150]. Liquid microjunction techniques such as nanoDESI MS and Flowprobe have been used for microbial identification [151–153]. NanoDESI is a surface sampling technique which desorbs analytes into a solvent droplet formed between two capillaries. The primary capillary delivers the solvent and the desorbed analytes, transferred into the secondary capillary, are subsequently introduced into mass spectrometer *via* nanoESI [151]. NanoDESI MS was used for characterisation of glycolipids and metabolites from a *Synechococcus* colony grown on agar with high salt content [151]. Another study using nanoDESI MS was focused on bacteria growing on agar with no additional sample preparation and two different types of bacteria were successfully detected in a Petri dish containing a mixed biofilm [152]. Flowprobe was used for direct metabolic analysis of a variety of colonies growing on agar substrate [153].

1.4.1 LESA MS as a tool for direct top-down analysis of microbial proteins

Despite the success of microbial analysis by the described MS techniques and their applications, complications have been observed [153, 154] and most previous studies have focused on analysis of small molecules rather than proteins (except for MALDI TOF, which detects intact proteins). Cooper and co-workers have been developing LESA MS as a tool for TD analysis of proteins directly from bacteria growing on solid substrates

[32,43,95]. This approach is different when compared to the MALDI TOF MS diagnostic approach and the other ambient MS techniques because it focuses on identification of bacterial proteins rather than spectral matching or metabolic profiling.

The initial work by Randall *et al.* [43] described the newly-developed LESA MS approach, where the proteins are extracted when the pipette tip comes into contact (as opposed to the surface sampling of tissue sections) with the living bacterial colony of *Escherichia coli* K12. The optimised workflow resulted in detection of approximately 150 peaks corresponding to around 60 proteins in MW range 5 – 32 kDa. Seven precursors were selected for CID fragmentation resulting in identification of six *E. coli* K12 proteins. The next study extended the TD analysis to multiple different bacterial species including Gram-negative and Gram-positive species. This work by Kocurek *et al.* [32] provided a deeper insight into the LESA colony sampling which resulted in selection of two extraction solvent systems with different acetonitrile content for Gram-negative and Gram-positive species. It was shown that the mass spectra of Gram-negative *E. coli* K12, *E. coli* BL21, *Pseudomonas aeruginosa* PS1054 contain predominantly intracellular proteins, while those of Gram-positive *Staphylococcus aureus* MSSA476 contain mostly highly abundant peaks of secreted peptides. The Gram-positive *Streptococcus pneumoniae* D39, *Streptococcus oralis* and *Streptococcus gordonii* were selected as an additional challenge for LESA MS and it was demonstrated that LESA MS was capable of differentiating between these species. The effect of refrigeration on the colonies of *E. coli* K12, *P. aeruginosa* PS1054 and *S. aureus* MSSA476 was examined and revealed that the number of protein peaks and their abundance changes depending on culturing conditions. This study identified in total 39 proteins by TD LESA MS, reporting 95% identification success rate.

Furthermore, incorporation of FAIMS in the workflow reduced background noise from the sample matrix, and significantly improved S/N and the number of proteins observed in the LESA mass spectra [93,95]. The initial work by Sarsby *et al.* [93] identified one protein by LESA FAIMS MS/MS (chip-based p-FAIMS device) that was not observed in the

previous studies. Kocurek *et al.* [95] used the same instrumentation for a more in-depth investigation of *E. coli* K12, *P. aeruginosa* PS1054 and *S. aureus* MSSA476. Eleven proteins were newly-identified from *E. coli* and *P. aeruginosa* colonies. FAIMS was able to separate overlapping protein peaks in the *P. aeruginosa* mass spectra. LESA FAIMS MS analysis of *S. aureus* did not result in any new protein ID assignments, however, reduction of signal abundance of secreted peptides was achieved and many previously undetected peptide and protein peaks emerged in the FAIMS mass spectra.

More recent work by Kocurek *et al.* [44] applied the TD LESA MS approach to analysis of fungal proteins. Here, an electroporation device was developed which delivered high voltage pulses (up to 3 kV) into yeast colonies causing their lysis. The lysed cells/colonies, subsequently subjected to LESA, allowed extraction of proteins as well as their detection in the mass spectra from the three investigated yeast species *Saccharomyces cerevisiae* (baker's yeast) and clinically relevant *Candida glabrata* and *Cryptococcus neoformans*. It was highlighted that any of the previous attempts to analyse yeast proteins by LESA MS alone was unsuccessful, while incorporation of electroporator in the workflow resulted in identification of 30 yeast proteins (from all three species). The great promise of this approach is to be used not only for analysis of yeast samples, but also other bacterial species, currently posed as a challenge for LESA MS. The work presented in this thesis is a direct continuation of these published studies.

1.5 Project aims

The ESKAPE pathogens represent six clinically relevant bacterial species *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* spp., of which two are Gram-positive (*E. faecium, S. aureus*) and the remaining four are Gram-negative [155]. The ESKAPE microbes are responsible for most of the nosocomial (hospital-acquired) infections [156,157] and their antibiotic resistance is constantly rising [158]. In fact, WHO reports at

least 700000 deaths annually due to infections by the drug-resistant strains, affecting both developed and developing countries [159]. In addition, nosocomial fungal infections also pose a serious risk to immunocompromised patients, while the most commonly involved species represent *Candida, Aspergillus, Mucorales, Fusarium* and *Scedosporium* [160]. The antifungal treatment resistance is on the rise and the treatment is costly [160,161]. Current diagnostics of fungal infections require higher specificity, selectivity and often take too long to perform – the delay in pathogen identification ultimately results in higher mortality rates [162, 163]. Development of improved tools for rapid and accurate identification for these microorganisms and hence tailored treatment of patients is therefore of high importance.

The aim of this project is to take steps towards development of a fast *in situ* tool for analysis of clinically relevant microbial species based on LESA MS, which does not require any specific sample preparation and is performed under ambient conditions, for identification of bacteria from skin wounds. Analysis may be therefore potentially carried out at the bedside directly from a patient's injury site. With LESA MS, it may be possible (in the future) to identify the pathogenic microbes and determine the immune response of the host (inflammatory biomarkers from patient). Currently, microbial examination takes hours or days to produce a result, while analysis performed by LESA MS is relatively fast (within minutes). The full development of a working *in situ* tool is a longer term project, while this project aims:

1. To investigate the potential of LESA MS for analysis of all of the ESKAPE pathogens growing on simple substrates while growing only one species per plate. The aim was to see whether LESA MS was capable of distinguishing the ESKAPE pathogens grown on simple substrates as a first step towards direct analysis of patients' wounds. Furthermore, it was necessary to focus on development of tools capable of detecting proteins from both microbes and substrate for further application of LESA MS method (Chapter 3).

2. To understand LESA MS of mixed-species biofilms (Chapter 3).

3. To develop a method for identification of ESKAPE pathogens growing on substrates of increasing complexity including blood agar and *in vitro* three-dimensional skin model ("Labskin"). This experiment included studying of wounded and inoculated/infected skin models (Chapter 4).

4. To develop a method for identification of yeast growing on *in vitro* 3D skin model based on use of the home-built electroporator (Chapter 4).

5. To develop a method for LESA MS analysis of *ex vivo* intact, wounded and inoculated/infected human skin samples (Chapter 5).

6. To develop and optimise a high-throughput workflow for identification of proteins from ESKAPE pathogens by TD LESA MS and TD LESA FAIMS MS (Chapter 6).

Chapter 2

Materials and methods

2.1 Materials

2.1.1 Chemicals

Analytical grade water, acetonitrile, ethanol, methanol and formic acid were purchased from Fisher Scientific (Loughborough, UK). Bacteriological agar was purchased from Appleton Woods (Birmingham, UK), sodium chloride, tryptone soya agar and defibrinated horse blood were purchased from Fisher Scientific (Loughborough, UK), yeast extract and dehydrated brain heart infusion were purchased from VWR (Lutterworth, UK), mycological peptone, glucose and peptone was purchased from Sigma Aldrich (Gillingham, UK).

2.1.2 Biological material

Freeze dried *S. aureus* NCTC13435 was obtained from Public Health England (Porton Down, UK) via Innovenn (Sand Hutton, UK). Bacterial samples of *E. faecium* E745, *E. faecalis* V583 and *K. pneumoniae* KP257 were obtained from Willem van Schaik (Institute of Microbiology and Infection (IMI), University of Birmingham), *S. aureus* MSSA476 and *P. aeruginosa* PS1054 were obtained from Mark Webber (clinical isolates library of the Queen Elizabeth Hospital, Birmingham), *A. baumannii* AYE and AC02 were obtained from Jessica Blair (IMI, University of Birmingham) and *E. cloacae* S11 was obtained from Allan McNally (IMI, University of Birmingham). *E. coli* K-12 and *C. glabrata* (in-house strain) was obtained from Robin May (IMI, University of Birmingham).

In vitro 3D skin models – "Labskin" were purchased from Innovenn. *Ex vivo* human skin samples were collected from Human Biomaterials Resource Centre (HBRC) of the University of Birmingham under project approval 16-253.

2.2 Methods

2.2.1 Preparation of microbial samples

Preparation of microbial samples and experiments involving microbial samples were carried out in laboratory of Containment level 2.

2.2.1.1 Preparation of liquid media and agar plates

All ingredients for each type of broth and blood agar base are listed in Table 2.1. The ingredients were weighed out and dissolved in 1 L of purified water (18 MΩ). For preparation of solid LB, BHI and YPD media, 15 g of bacteriological agar was added to the broth. All media solutions were autoclaved at 121 °C. Broth was stored in a glass bottle on a laboratory bench at room temperature prior use. Agar plates of two sizes were prepared – 60 mm and 100 mm. For 60 mm plates, approximately 7 mL of media was poured in the plate and approximately 15 mL of media for 100 mm plate. Plates with solid agar were sealed with parafilm and stored at 4 °C prior use.

Preparation of blood agar required dissolving of the base in 950 mL of purified water and autoclaving at 121 °C. The base medium was cooled down to 45 – 50 °C and subsequently 50 mL of defibrinated horse blood was added aseptically. Agar was poured into 60 mm plates. Solid plates were sealed with parafilm and stored at 4 °C.

2.2.1.2 Preparation of liquid cultures and cultures growing on agar plates

Liquid cultures were prepared by adding approximately 5 – 10 mL of liquid broth to a 50 mL tube. An inoculation loop of size 1 µL was used to scrape bacteria or yeast from the surface of a frozen glycerol stock or a colony grown on agar plate and resuspended

Table 2.1: Ingredients for preparation of growth media for microbial cultures.

Type of media	**Ingredient**	**Amount needed for 1 L**
Lysogeny broth (LB)	peptone	10 g
	yeast extract	5 g
	sodium chloride	10 g
Brain heart infusion (BHI)	brain heart infusion	37 g
Yeast extract peptone dextrose (YPD)	yeast extract	10 g
	mycological peptone	10 g
	glucose	20 g
Blood agar base	tryptone soya agar	40 g

in the liquid broth. To provide sufficient aeration, the lid of the tube was not tightly screwed. Liquid cultures were incubated up to 18 hours in a shaking incubator (200 rpm) at 37 °C. On the following day, 1 µL (bacteria) or 10 µL (yeast) of the microbial liquid culture was spotted onto agar plates. The plates were inverted and incubated at 37 °C (bacteria) or 30 °C (yeast) for 24 hours (bacteria) or up to 48 hours (yeast).

2.2.1.3 Preparation of glycerol stocks

Liquid cultures of the desired bacterial or yeast species were prepared. On the following day (up to 18 hours of incubation), 500 µL of the liquid culture was dispensed into a cryovial and mixed with 500 µL of 50% (v/v) sterile-filtered glycerol/water (purified, 18 MΩ) solution. The stock was stored at –80 °C.

2.2.1.4 Counting of microbial colonies

Agar plates with bacteria or yeast were prepared. An inoculation loop (10 µL) was used to scrape microbial colony from the agar plate surface and microbial cells were resuspended in sterile deionised water. Suspensions were diluted to achieve optical density (OD) of 0.2 (at 600 nm). 10 µL of suspension of OD 0.2 and 4 further suspension dilutions (1:10, 1:100, 1:1000 and 1:10000) of each species were dispensed on an agar plate and cultured

in a static incubator at 37 °C. Grown colonies were counted after 24 hours of incubation. The experiment was performed in triplicates on 3 different days.

2.2.1.5 Preparation of biofilms

On the first day, liquid cultures of *P. aeruginosa* PS1054 (in LB broth) and *C. glabrata* (in YPD broth) were incubated in a shaking incubator (200 rpm) up to 18 hours at 37 °C.

On day 2, liquid cultures were collected. *C. glabrata* was washed two times in analytical grade water and resuspended in LB broth. Both cultures were diluted such that the culture of *P. aeruginosa* achieved OD 0.2 (at 600 nm) and the number of cells/mL for *C. albicans* reached 1 x 10^7. First, 2 mL of *C. glabrata* was added to a tissue culture treated Petri dish of size 30 mm. Subsequently, 100 µL of *P. aeruginosa* was dispensed into the tissue culture plate and 1 mL of LB broth was added for the optimal culture growth. Biofilm cultures were incubated at static conditions at 37 °C.

After 2 hours of incubation, medium with unattached cells was removed and replaced with 3 mL of a fresh LB broth. Medium was replaced again after 24 hours and biofilms were collected after 48 hours of incubation. Medium was discarded prior to LESA MS analysis.

2.2.2 *In vitro* 3D skin model sample preparation

Upon arrival, Labskin samples were transferred from the 12-well plate with transfer agar medium to a new, sterile 12-well plate with fresh proprietary Labskin medium (Innovenn, Sand Hutton, UK) and incubated overnight at 37 °C, 5% CO_2 and >95% relative humidity. On the following day, the medium was replaced and the Labskin samples were wounded with a scalpel blade, such that the scalpel did not penetrate the membrane at the bottom of the well insert. The wounded samples were inoculated with desired bacterial species and yeast with a calculated infectious dose (see Chapter 4).

2.2.3 *Ex vivo* human skin sample preparation

One day before the scheduled surgery, a plastic container with low glucose DMEM medium without L-glutamine (Sigma Aldrich, Gillingham, UK), supplied with 1% Penicillin/ Streptomycin (P/S) (Sigma Aldrich, Gillingham, UK) was stored at 4 °C in HBRC (University of Birmingham). On the day of surgery, skin samples were placed into the container with the medium specified above and stored at 4 °C prior collection. The collected samples were transferred to laboratory on ice in a sealed box.

The piece of skin (approximately 6 x 6 cm) was disinfected with 70% ethanol and cleaned to remove any traces of blood. If present, fat was removed from the dermis and the skin sample was cut by scalpel and scissors to pieces of size 1.5 x 1.5 cm. Pieces of skin were sutured onto cell strainers with a surgical needle (Ethicon Mersilk), with 5 – 6 sutures per sample to maintain the skin's natural tension.

Cell strainers with sutured samples were immediately transferred into a 6-well sterile cell culture plate. Samples inoculated with bacteria were cultured in DMEM medium with high glucose, without L-glutamine, supplied with 10% of foetal bovine serum (FBS) (Sigma Aldrich, Gillingham, UK). For non-inoculated samples, 1% P/S supplement was kept in the medium. All samples were incubated at 37 °C, 5 % CO_2 and >95% relative humidity and medium was changed every day.

2.2.4 LESA MS analysis

LESA MS analysis was performed by using an Advion Triversa Nanomate (Advion, Ithaca, NY) coupled to a Thermo Orbitrap Elite, a Thermo Q Exactive HF and a Thermo Orbitrap Eclipse mass spectrometers (all Thermo Fisher Scientific, Bremen, Germany). The use of this technique in combination with Orbitrap mass spectrometers was described previously [32, 43, 44]. The extraction solvent system used for analysis was 60% acetonitrile, 35% water and 5% formic acid. Agar plates (60 mm diameter) with bacterial colonies were placed adjacent to the half of the 96-well microtiter plate. The

robotic Triversa Nanomate pipette aspirated 3 μL of extraction solvent from the microtiter plate. After relocation of the robotic pipette to a position above the plate, 2 μL of the solvent system was dispensed on the colony while touching the colony surface. After sampling, 2.5 μL of the extraction solvent system with analytes was re-aspirated back into the pipette tip and introduced into the mass spectrometer *via* chip-based nano-ESI at gas pressure 0.3 psi and a tip voltage 1.75 kV. The robotic Nanomate system was controlled with the advanced user interface (AUI) of the Chipsoft software 8.3.1 (Advion, Ithaca, NY).

Data presented in the experimental chapters were acquired on Orbitrap Elite unless otherwise stated. The mass spectra were acquired for at least 3 min in full scan mode, positive ion mode, in mass range 600 – 2000 m/z, at resolution 120 000 at 400 m/z. Top-down MS/MS analysis of proteins was performed by using CID in the ion trap using helium gas at 35% normalised collision energy. When fragmentation was carried out, each scan comprised 30 co-added microscans. Any deviations from this method will be described in the corresponding chapter.

2.2.5 Data analysis and protein identification

Top-down identification of proteins was carried out in ProSight 4.1 software (Thermo Fisher Scientific, Bremen, Germany). Databases containing the whole proteome of the species of interest were downloaded from the Uniprot website in XML format. Databases included *Homo sapiens* (UP000005640, 71599 entries), *Equus caballus* (UP000002281, 44485 entries), *Escherichia coli* K-12 (UP000000625, 4391 entries), *Enterococcus faecium* (UP000095783, 2983 entries), *Enterococcus faecalis* ATCC 700802/V583 (UP000001415, 3240 entries), *Staphylococcus aureus* NCTC 8325 (UP000008816, 2889 entries), *Klebsiella pneumoniae* ATCC 700721 (UP000000265, 5126 entries), *Acinetobacter baumannii* strain AYE (UP000002446, 3652 entries), *Pseudomonas aeruginosa* ATCC 15692/PA01 (UP000002438, 5563 entries), *Enterobacter cloacae* ATCC 13407 (UP000002363, 5412 entries) and *Candida glabrata* (UP000002428, 5200 entries). MS/MS spectra were

imported into ProSightPC software in "import profile" mode and subsequently deconvoluted with THRASH algorithm and default settings of S/N 3. Absolute mass search mode parameters accounted for all posttranslational modifications and included delta-mass (Δm) option on for locating any possible mutations or mass shifts. The search window was 1000 Da, initially with mass fragment tolerance $\pm$15 ppm and minimum matching fragments set to 4. All of the identified protein sequences were checked with the Sequence Gazer function of the ProSightPC software, followed by manual manual peak assignment where the fragment tolerance was narrowed to $\pm$5 ppm.

Chapter 3

Top-down LESA MS of ESKAPE pathogens and mixed microbial biofilm

The work presented in this chapter was published in Havlikova *et al.* [164] (see Appendix A). Portions of the text were re-used in agreement with the open access publishing policy.

3.1 Background

Previously, LESA MS has been applied to the analysis of proteins in *Escherichia coli* K12 [43]. Later work identified 39 proteins from multiple species including two of the ESKAPE pathogens *Staphylococcus aureus* and *Pseudomonas aeruginosa* [32]. The aim of the work presented in this chapter was to extend LESA MS to the remaining four ESKAPE pathogens *Enterococcus faecium* E745 (and its close relative *Enterococcus faecalis* V583, commonly found in the hospital environment), *Klebsiella pneumoniae* KP257, *Acinetobacter baumannii* (the reference strain AYE and a clinical isolate AC02) and *Enterobacter cloacae* S11.

A key consideration if LESA MS is to find use as a diagnostic tool for microbial identification is a universal sampling approach (i.e. a single solvent system suitable for all Gram-positive and Gram-negative species). Initial work focused on optimisation of the solvent system to enable successful protein extractions from all of the ESKAPE pathogens, before applying MS/MS for TD protein identification. In total, 24 proteins

were identified from 37 MS/MS mass spectra. Subsequently, protein identification from one database searches was compared to searches against multiple databases, and success of LESA diagnosis was associated.

Next, LESA MS was applied to the study of biofilms comprising the ESKAPE pathogen *P. aeruginosa* and the yeast *Candida glabrata*. Microbial species are rarely present in a planktonic ("free-floating") state in the clinical environment because they form biofilms. Biofilms are found at surfaces (e.g. air-liquid interfaces, medical implants, plant or mammalian tissue) and comprise microorganisms attached to each other and encased in an extracellular polymeric substance (EPS) – a combination of lipids, polysaccharides and proteins [165–168]. Such conditions are highly advantageous for growth of different bacterial species and yeast together while forming polymicrobial biofilms often resistant to antibiotic treatment, immunologic and chemical attacks [165]. *P. aeruginosa* and *Candida* are regularly present as commensals, however any change of conditions may cause their invasive growth and potentially lead to death in compromised individuals. *P. aeruginosa* and *Candida* are known for their specific antagonistic relationship and mutual growth suppression [169]. Despite the antagonism, mixed *P. aeruginosa* and *C. albicans* infection enhances *P. aeruginosa* pathogenicity [170]. Interaction at the level of antimicrobial resistance shows that co-colonisation of *P. aeruginosa* and *C. albicans* protects the *P. aeruginosa* and results in dissemination of the infection [171]. Early diagnostics and mixed biofilm identification in such cases is therefore highly desired.

Here, preliminary results from LESA MS analysis of *P. aeruginosa* and *C. glabrata* mixed biofilm sample are presented, where the microbes were cultured in a simple LBB liquid medium. Only protein peaks corresponding to *P. aeruginosa* PS1054 were detected and identified in the mixed biofilm sample.

3.2 Materials and methods

Chemicals and origin of microbial samples used for the experiments are listed in the section 2.1. Preparation of bacterial samples followed protocol described in 2.2 and two biological replicates were prepared for each species. For each biological replicate, between 1–11 technical replicates were performed. Biofilms were prepared according to 2.2.1.5. LESA MS analysis using the Orbitrap Elite only is described in section 2.2.4. Data analysis is described in section 2.2.5.

Deconvolution of the mass spectra was performed by using the Xtract algorithm in selected mass range 600–2000 m/z. The output mass was selected to neutral, charge range was set to 3–50. Orbitrap was selected as a mass analyser type, isotope table was set to protein and relative absolute threshold to 0% with minimum of 1 detected charge state per protein.

BLAST (Basic Local Alignment Search Tool) searches were performed as follows: for each protein identified, the protein sequence was downloaded from Uniprot in FASTA format and searched against the nonredundant protein sequences (nr) database using the BLASTP algorithm with default parameters (blast.ncbi.nlm.nih.gov).

3.3 Results and discussion

3.3.1 Optimization of extraction solvent system for LESA MS analysis

In the earlier work, two different extraction solvent systems containing acetonitrile, water and formic acid were employed for the LESA MS analysis of Gram-positive (50:45:5) and Gram-negative (40:59:1) bacteria [32]. The higher acetonitrile and formic acid content for the Gram-positive species ensures the extraction of proteins, possibly due to cell lysis or increased ionisation efficiency [32]. Utilisation of two solvents is, however,

impractical when analysing multiple and unknown bacterial species. In addition, initial LESA MS experiments with the 50:45:5 extraction solvent system of Gram-positive *Enterococci* species resulted in observation of no proteins peaks for *E. faecalis* V583 and only a few, low abundance protein peaks for *E. faecium* E745 (Fig. 3.1a and 3.1b). Subsequent experiments were focused on *E. faecalis* V583 as a model organism for optimisation of extraction solvent system composition. Solvents comprising various ratios of ethanol, acetonitrile, water and formic acid were investigated (Fig. 3.2). The most suitable ratio for extracting proteins from colonies of *E. faecalis* V583 was found to be 60:35:5 acetonitrile:water:formic acid (Fig. 3.1a and 3.2), where 12 proteins were extracted during the optimisation process. One protein was observed for the 50:10:25:5 acetonitrile:ethanol:water:formic acid and two proteins for each 50:45:5 and 60:33:7 acetonitrile:water:formic acid solvent systems. Interestingly, there appears to be an upper limit to the acetonitrile content that allows proteins to be detected. At 80% acetonitrile content, no protein peaks were detected, likely a result of poor protein solubility. The optimised solvent system also resulted in detection of 23 proteins from *E. faecium* E745 (Fig. 3.1b) – compared to 10 proteins detected with the initially used 50:45:5 acetonitrile:water:formic acid content. The optimised LESA extraction solvent was subsequently tested for the remaining three ESKAPE species (Fig. 3.1c) as well as the previously studied *S. aureus* MSSA476 and *P. aeruginosa* PS1054 (Fig. 3.3). The successful LESA extraction from *S. aureus* MSSA476 resulted in presence of four out of five previously identified proteins in the mass spectra [32]. The analysis of *P. aeruginosa* PS1054 allowed extraction of four out of 15 identified proteins published previously [32]. It is important to note that the study of Kocurek *et al.* shows data obtained from bacterial colonies sampled after various incubation and storage conditions. Therefore, presence of the identical number of previously identified proteins (or identical proteins) in the mass spectra for both MSSA476 and PS1054 was not expected. The results show that 60:35:5 acetonitrile:water:formic acid is a suitable solvent system for extracting proteins from both Gram-positive and Gram-negative ESKAPE microorganisms.

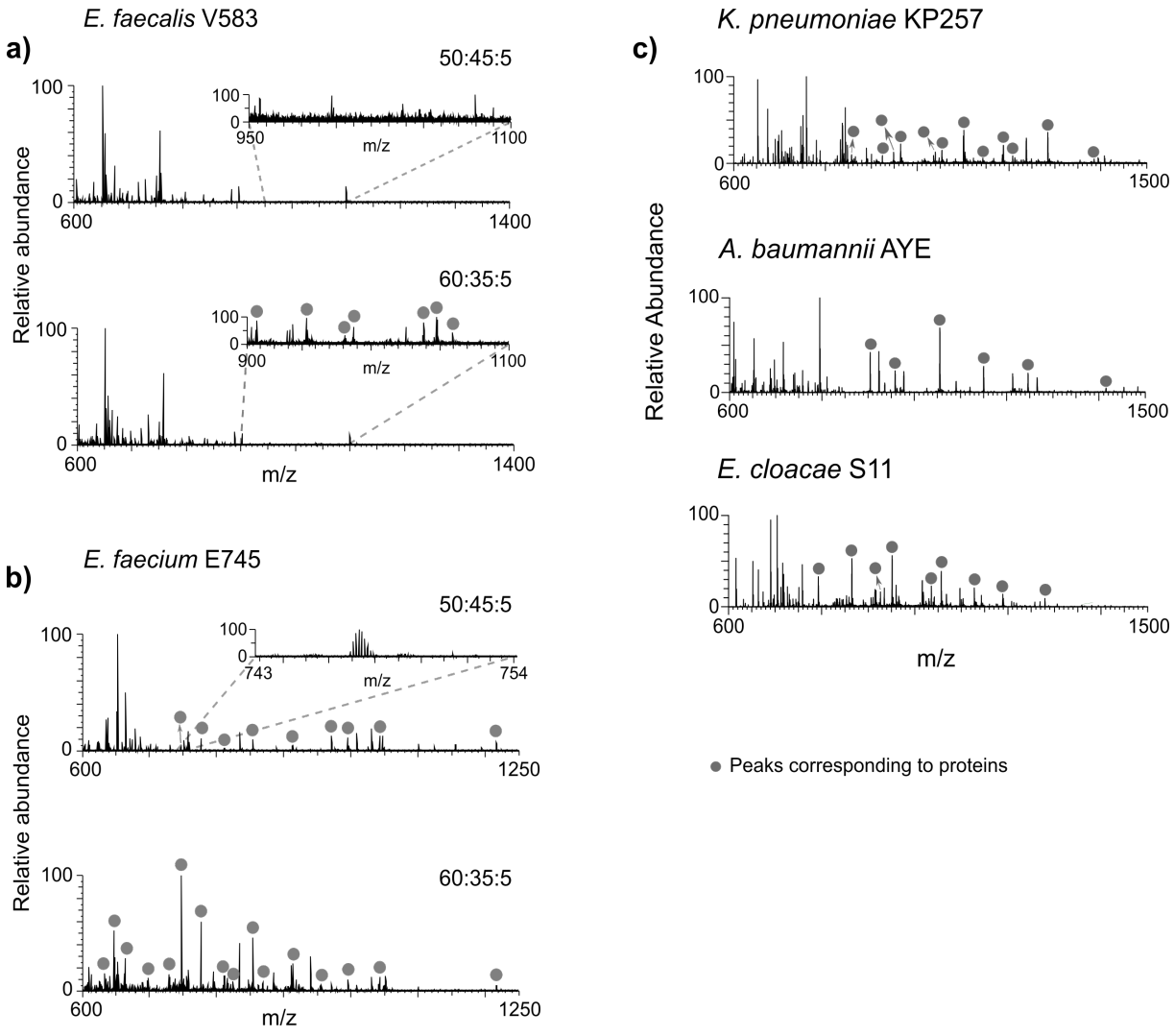


Figure 3.1: LESA mass spectra of bacterial colonies. Comparison of extraction solvent systems, acetonitrile:water:formic acid 50:45:5 and 60:35:5 for a) *E. faecalis* V583 and b) *E. faecium* E745. c) LESA MS analysis of *K. pneumoniae* KP257, *A. baumannii* AYE and *E. cloacae* S11 with the 60:35:5 extraction solvent system.

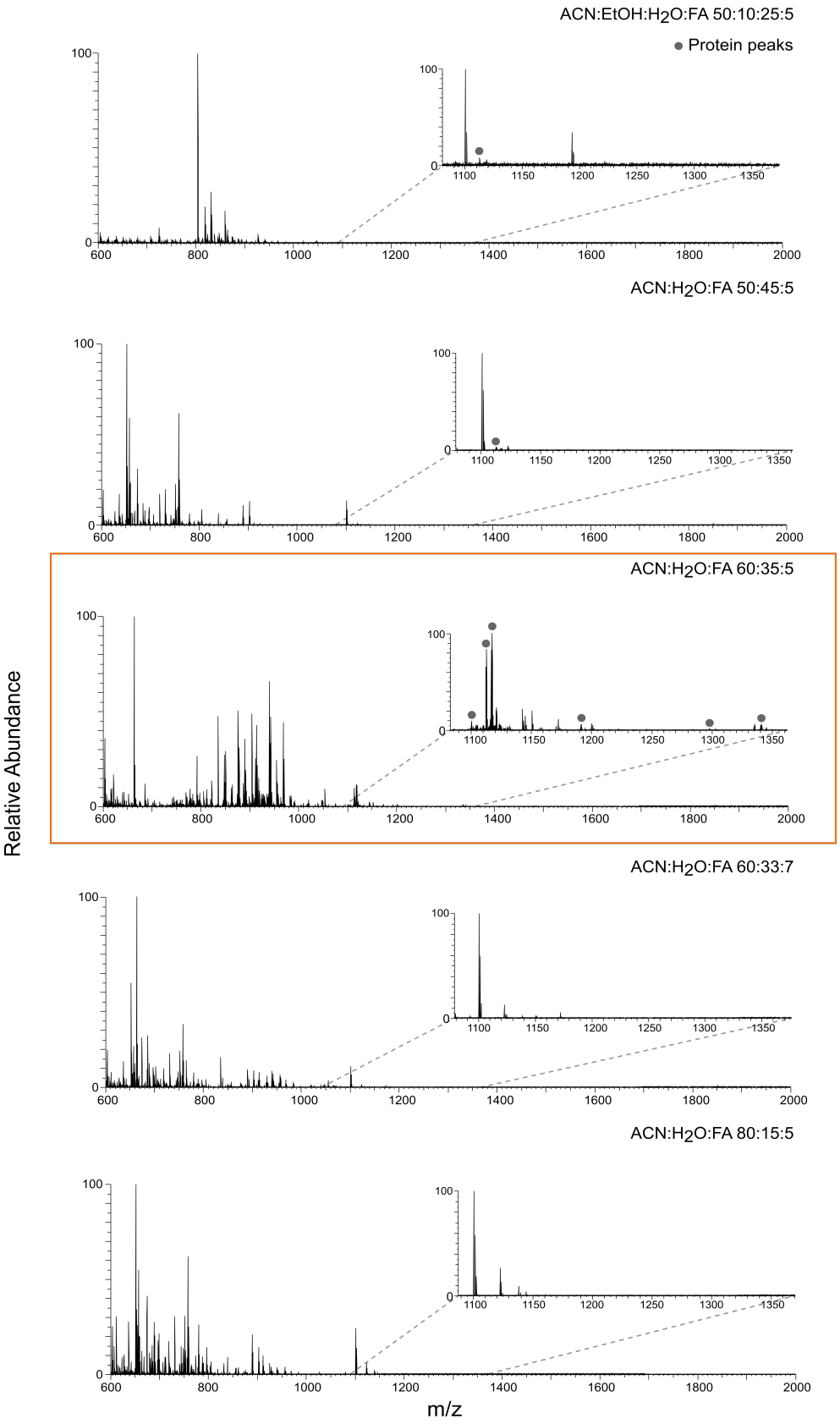


Figure 3.2: LESA mass spectra of *E. faecalis* V583 sampled with 5 different acetonitrile-based extraction solvent systems. The solvent system resulting in the detection of the greatest number of protein peaks in the mass spectra is highlighted.

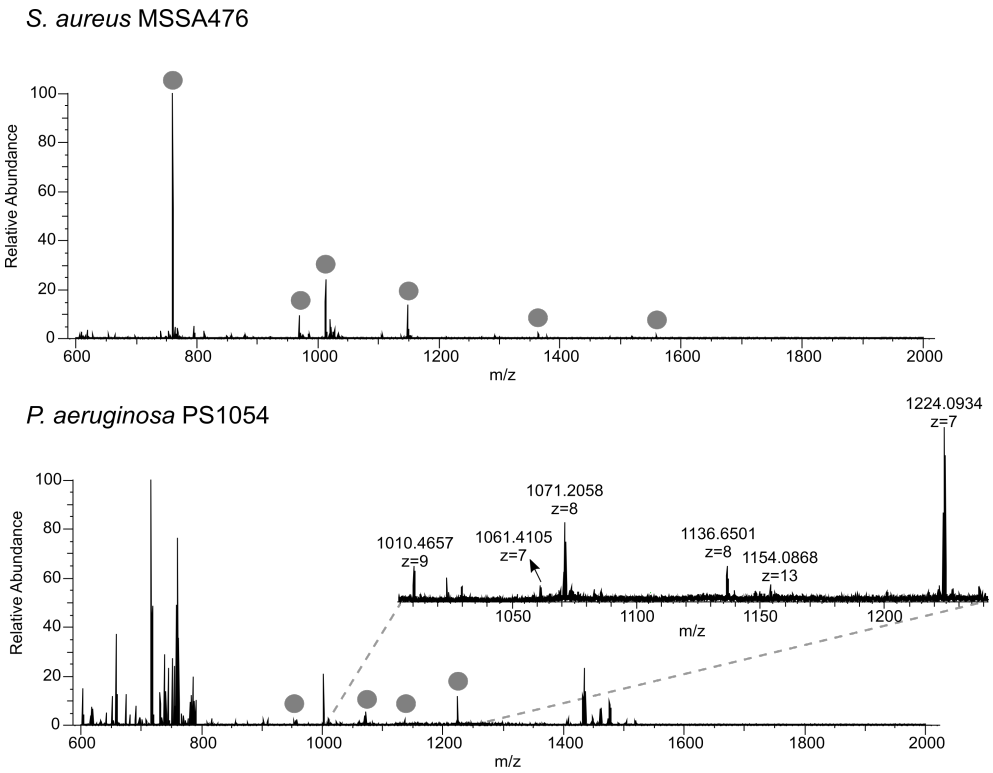


Figure 3.3: Mass spectra of *P. aeruginosa* PS1054 and *S. aureus* MSSA476 after LESA extraction with the optimised extraction solvent system 60:35:5 acetonitrile:water:formic acid.

3.3.2 LESA MS analysis of ESKAPE pathogens

For each of the species studied, the most abundant protein peaks were selected for fragmentation by CID. The resulting MS/MS spectra were searched against the corresponding individual bacterial protein database. TD LESA MS analysis of the ESKAPE species resulted in identification of 24 proteins from 37 MS/MS spectra. MS/MS spectra and fragment assignments for all proteins can be found in Appendix B.1.

3.3.2.1 *Enterococci*

For *E. faecium* E745, five proteins were identified after CID fragmentation of 11 MS/MS spectra (Fig. 3.4): 50S ribosomal protein L29, 30S ribosomal protein S20, and three uncharacterised proteins HMPREF0351_11703, HMPREF0351_11270 and HMPREF0351_12038. Four out of five proteins were observed in both biological replicates (Fig. 3.5). According to the Uniprot database, the predicted HMPREF0351_11703 protein sequence includes a signal peptide. Data suggest no signal peptide cleavage and formylation of the N-terminal methionine (Fig. 3.4). The second uncharacterised protein HMPREF0351_11270 (Fig. 3.4) was identified with formylated methionine at the N-terminus which has not been yet recorded in the Uniprot database. For the third uncharacterised protein HMPREF0351_12038, a mutation R→Q was detected at either position 38 or position 43, however no fragments were observed in this region to allow unambiguous localisation (Fig. 3.4).

A close relative of *E. faecium* is *E. faecalis*, a clinically important species associated with infective endocarditis, biofilm formation and antimicrobial resistance [172]. This microbe was investigated due to its abundance in the hospital environment and increasing antibiotic resistance [173]. LESA MS/MS of five *E. faecalis* V583 intact precursors resulted in identification of four proteins (Fig. 3.6): DNA-binding protein HU, 50S ribosomal protein L29, UPF0337 protein EF_1180 and uncharacterised protein (gene EF_0665). The UPF0337 protein EF_1180 is a protein inferred from homology and

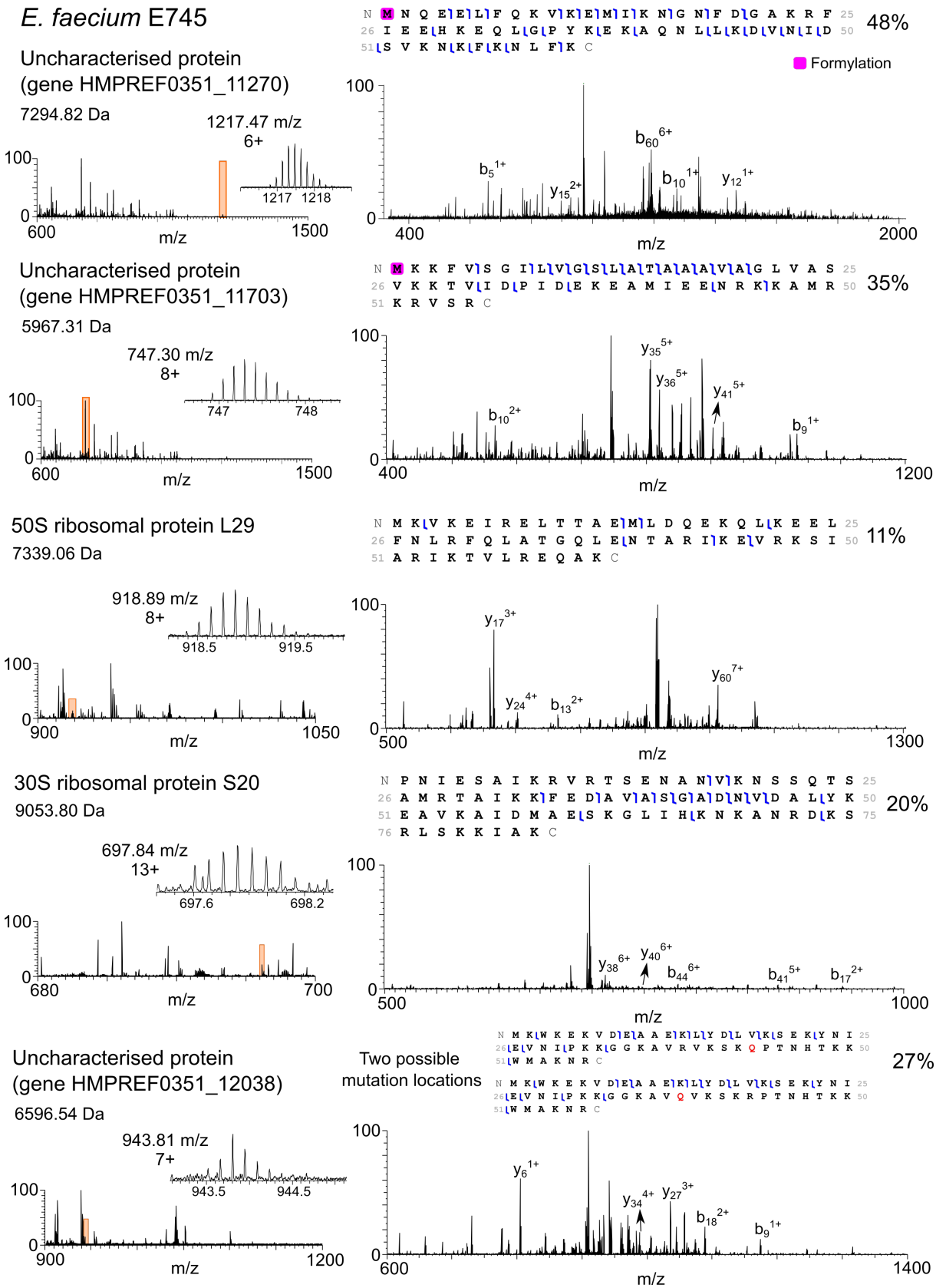


Figure 3.4: Representative MS/MS mass spectra of five newly-identified proteins for the *E. faecium* E745.

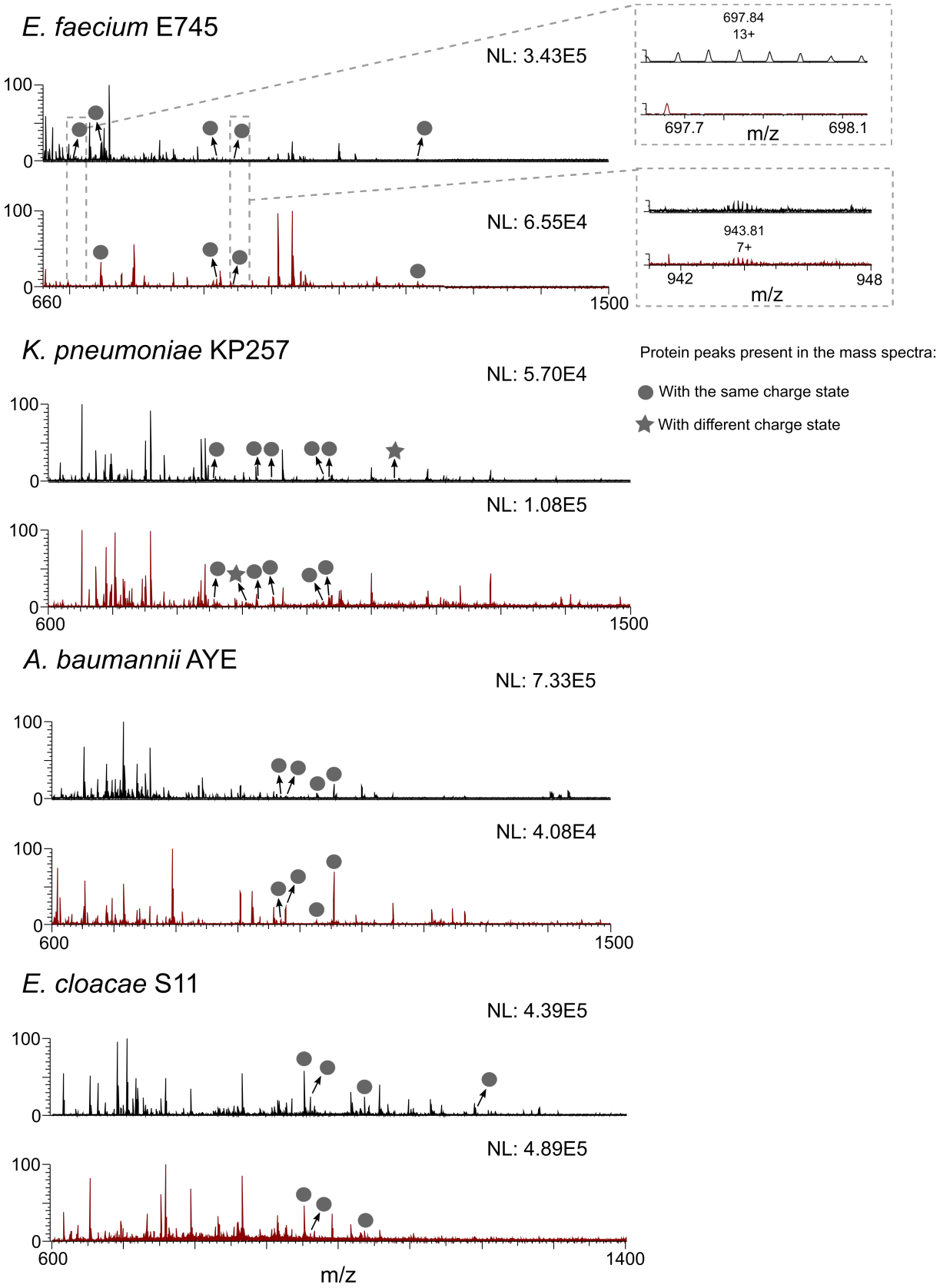


Figure 3.5: A comparison of representative mass spectra of biological replicates of the four investigated ESKAPE species.

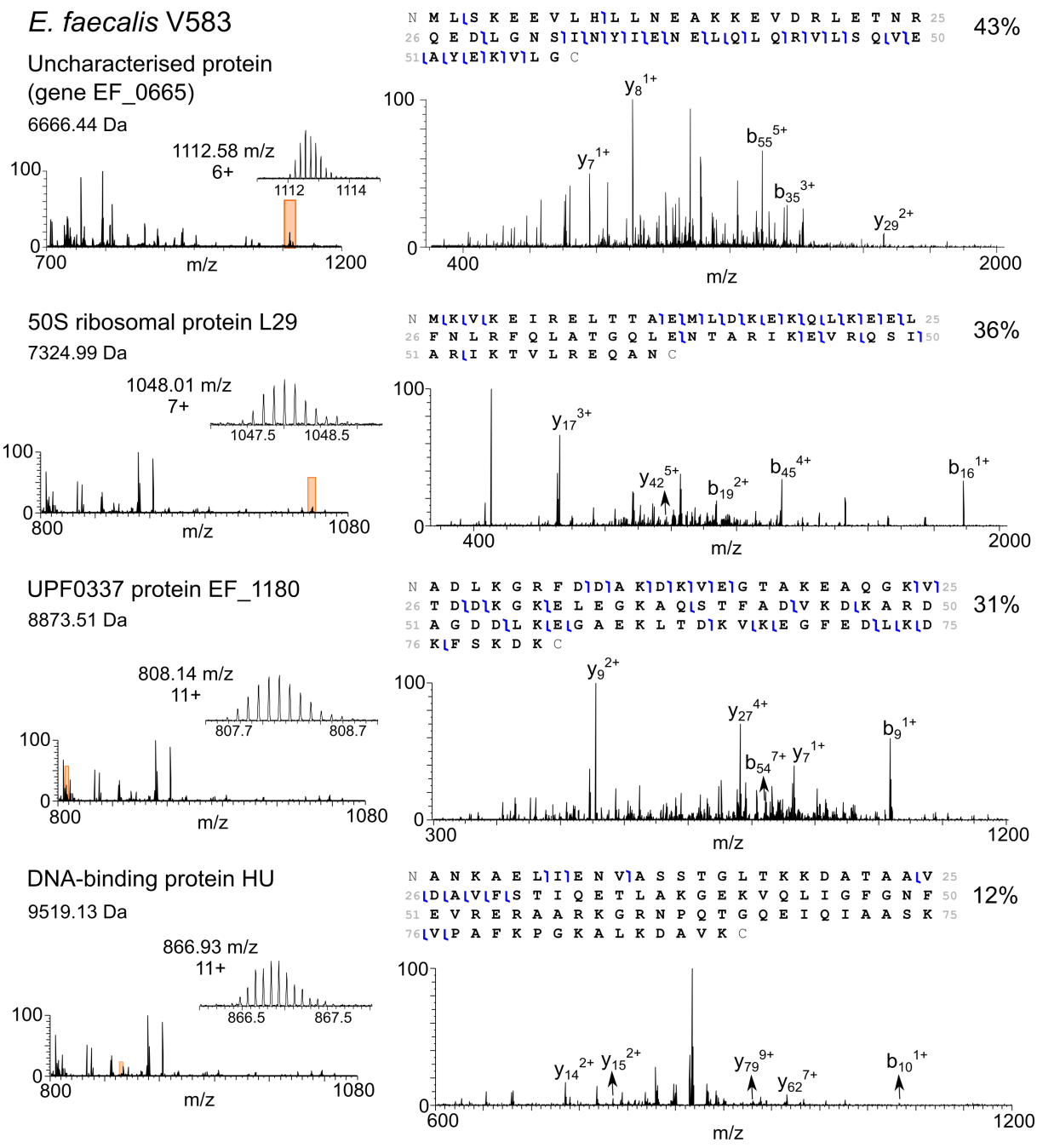


Figure 3.6: Representative MS/MS mass spectra of five newly-identified proteins for the *E. faecalis* V583.

belongs to the bacterial general stress response protein CsbD family (Fig. 3.6). The uncharacterised protein (gene EF_0665) (Fig. 3.6) is a predicted protein. Comparison of the LESA mass spectra obtained from *E. faecalis* V583 and *E. faecium* E745 revealed that the 50S ribosomal protein L29 was observed for both species (see Fig. 3.7). The sequence differs by N→K substitution at the position 62, resulting in a mass difference

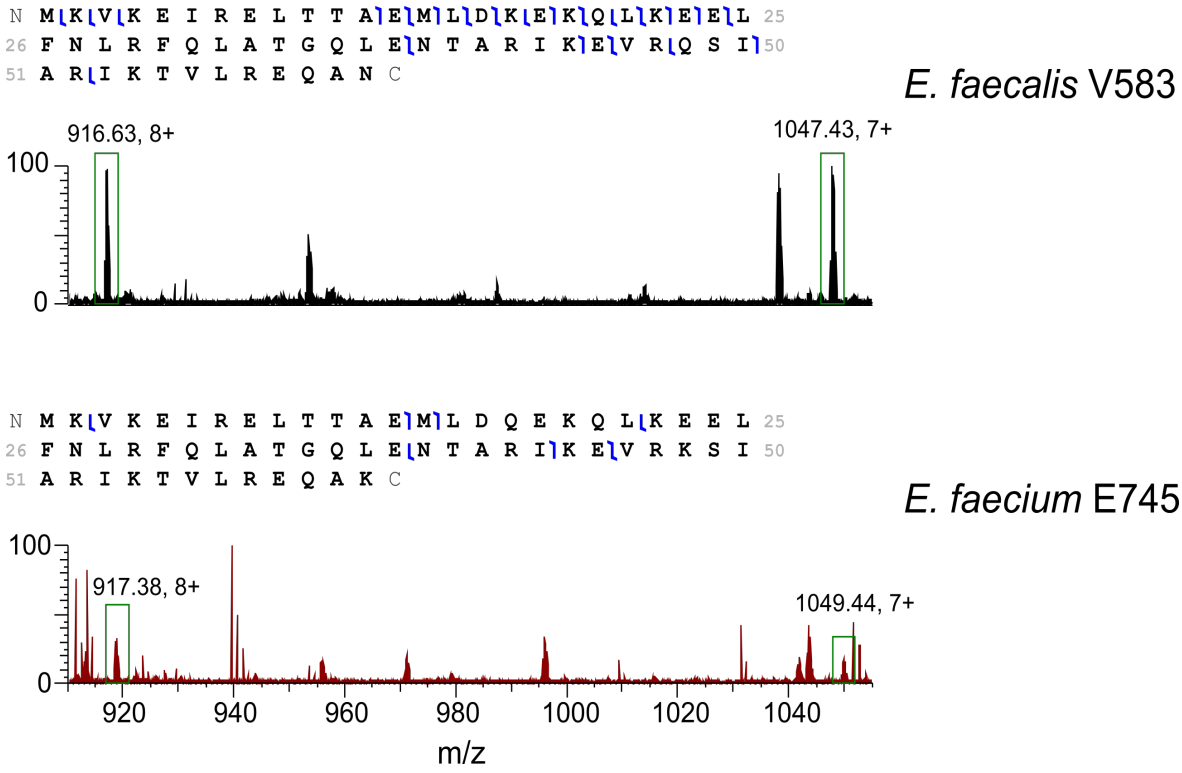


Figure 3.7: *E. faecalis* and *E. faecium* LESA mass spectra showing the difference between the 50S ribosomal proteins L29.

of 14.07 Da. This observation may be potentially useful as a diagnostic for differentiation between these two microorganisms.

3.3.2.2 *Klebsiella pneumoniae*

LESA MS/MS of *K. pneumoniae* KP257 resulted in identification of six out of nine precursor proteins. All six proteins were observed in both of the two biological replicates. The identified proteins (Fig. 3.8) included DNA-binding protein HU-α, KPN_00497, two ribosomal proteins (50S ribosomal protein L29 and 30S ribosomal protein S16), one uncharacterised protein (gene yciG) and CsbD domain-containing protein. A mutation of the protein KPN_00497 R→K at the position 49 (Fig. 3.8) and a signal peptide cleavage (1–19) was observed. The search results also indicated a signal peptide cleavage of the first 19 amino acids of the CsbD protein sequence previously unrecorded in the Uniprot database.

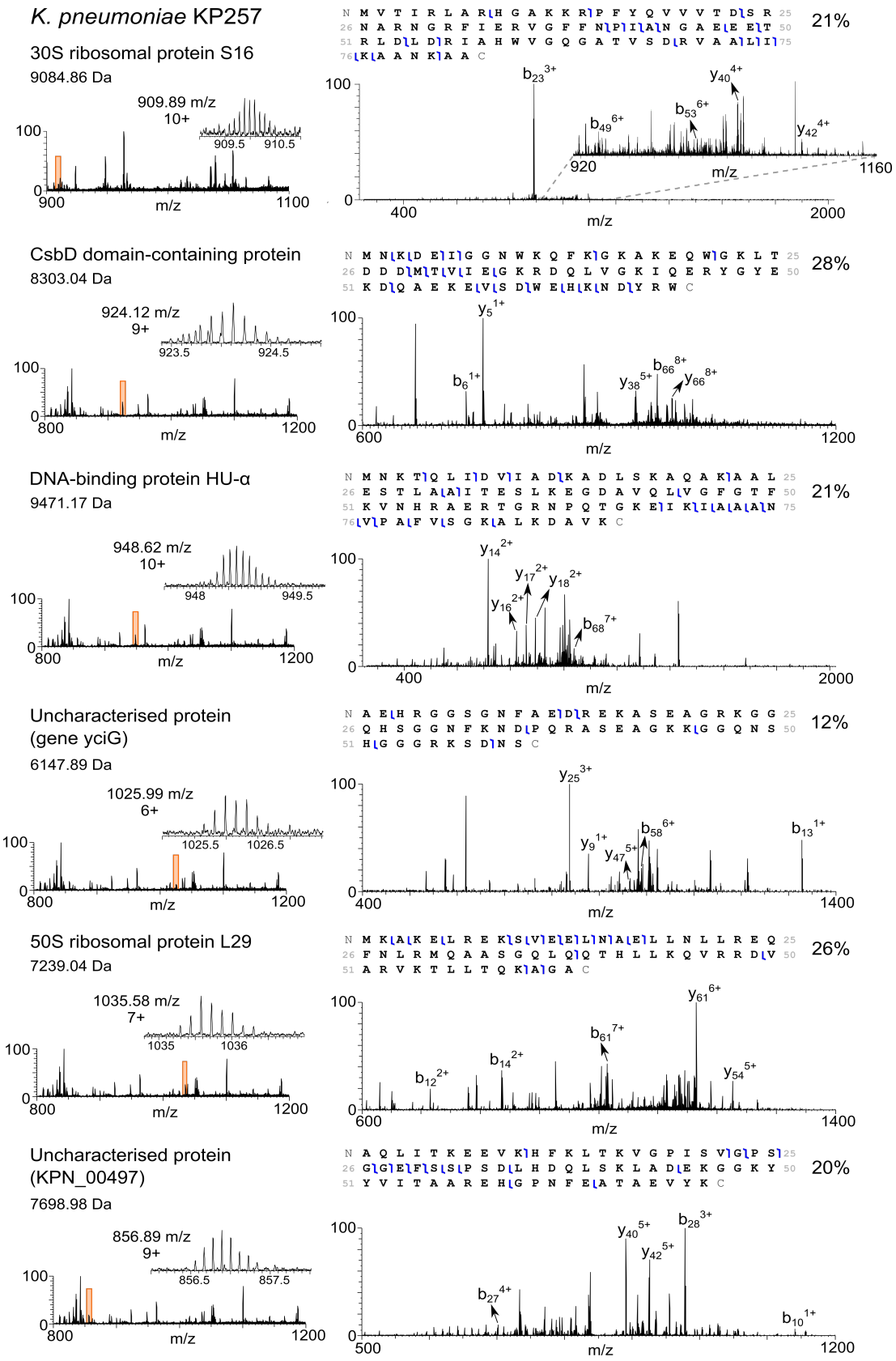


Figure 3.8: Representative MS/MS mass spectra of five newly-identified proteins for the *K. pneumoniae* KP257.

3.3.2.3 *Acinetobacter baumannii*

LESA MS analysis of the Gram-negative reference strain *A. baumannii* AYE yielded mass spectra with highly abundant protein peaks (Fig. 3.9, Fig. 3.10). The CID fragmentation of six intact precursors resulted in four protein identifications (Fig. 3.9) – three uncharacterised proteins (genes ABAYE1298, ABAYE2274 and ABAYE1876) and bacteriolytic lipoprotein entericidin B. All four proteins were observed in both biological replicates (Fig. 3.5). For ABAYE2274, cleavage of the initial methionine was detected from the MS/MS data (Fig. 3.9). The amino acid sequence of ABAYE1876 contains a signal peptide (1–14), information not yet recorded in the Uniprot database. Entericidin B is the first lipoprotein to be identified by LESA MS (Fig. 3.9). Lipoproteins are important for bacterial physiology, virulence and as activators of host innate immune response [174]. Bacterial lipoproteins are characterised by a conserved N-terminal lipid-modified cysteine residue. In this case, a mass shift of 813.72 Da was detected, however the exact structure of the lipid group attached to the N-terminus remains unknown and would require further analysis. Despite the mass shift, there is high confidence in the protein assignment due to the high sequence coverage (82%) obtained. *A. baumannii* AYE was compared to the clinical strain *A. baumannii* AC02 (Fig. 3.10). Again, LESA MS resulted in detection of several highly abundant protein peaks (Fig. 3.10), however their identification proved challenging. Six MS/MS spectra were searched against AYE database, however no protein IDs were assigned, suggesting dissimilarity in the protein amino acid sequences of these strains and a requirement for a new database for the clinical strain. The dissimilarity can also be observed from the comparison of the AYE and AC02 mass spectra (Fig. 3.10).

3.3.2.4 *Enterobacter cloacae*

Mass spectra obtained following LESA of *E. cloacae* S11 contained many abundant peaks corresponding to proteins (Fig. 3.1c). LESA MS/MS analysis of six intact precursors resulted in identification of five proteins (Fig. 3.11) – 50S ribosomal protein L29, DNA-binding protein, CsbD family protein, UPF0391 membrane protein and DUF1471

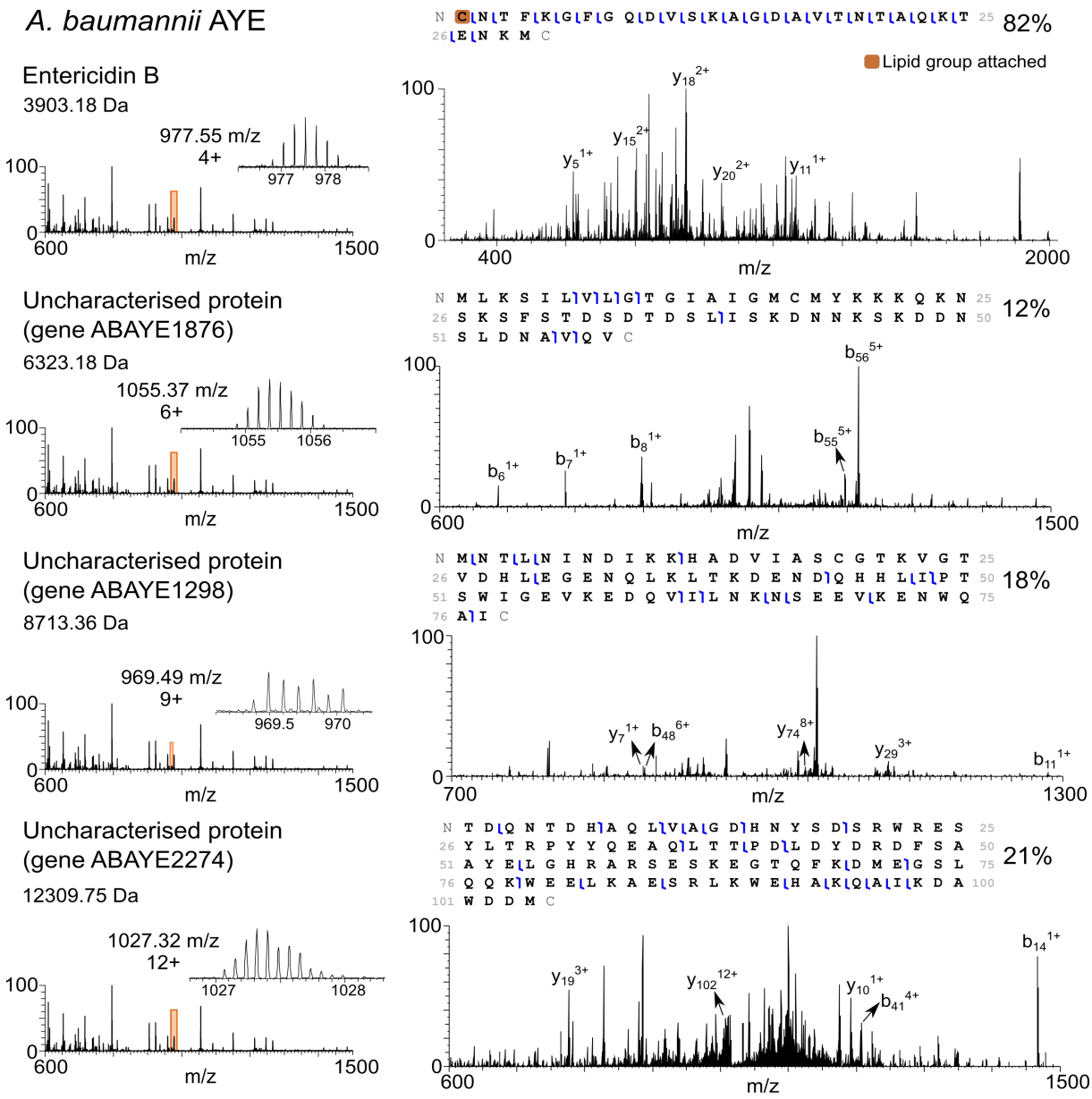


Figure 3.9: Representative MS/MS mass spectra of five newly-identified proteins for the *A. baumannii* AYE.

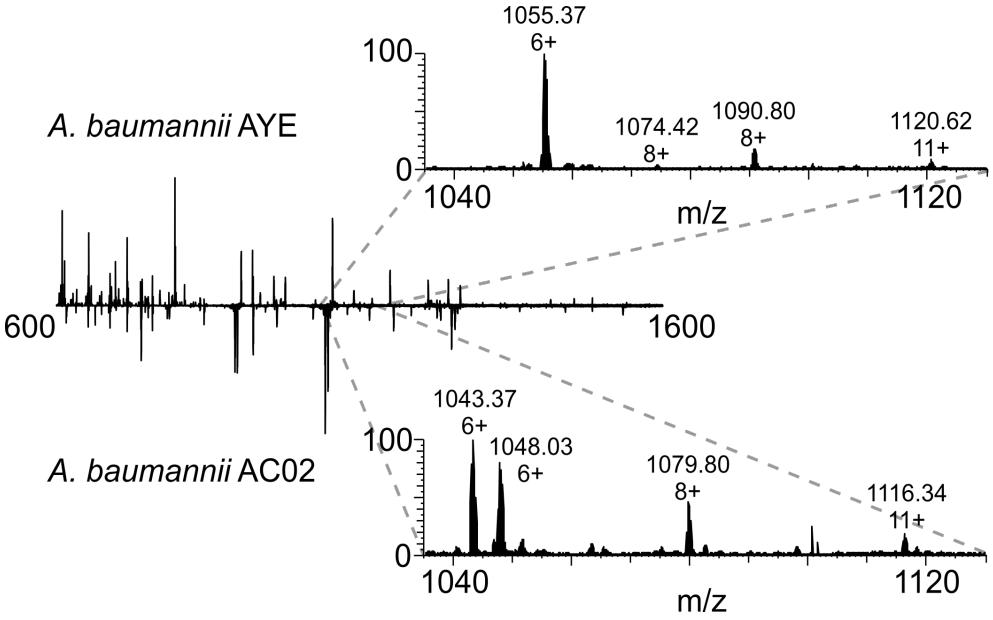


Figure 3.10: Comparison of the reference strain *A. baumannii* AYE to the clinical strain *A. baumannii* AC02.

domain-containing binding protein. Four out of five proteins were detected in both biological replicates. The representative mass spectra shown in Fig. 3.5 contain peaks corresponding to four (top) and three (bottom; the fourth protein was detected in a different mass spectrum) of those proteins. For the UPF0391 membrane protein a new PTM – formylation at the N-terminus, not yet reported in the Uniprot database, was revealed (Fig. 3.11). DUF1471 protein sequence contains a signal peptide (1–21) and a mutation (E→N) at one of two potential positions – either 37 or 45, however, no fragments were observed in this region to allow unambiguous localisation.

3.3.3 Identification of ESKAPE pathogens from multiple protein databases

Initially, the goal was to investigate ESKAPE pathogens by TD LESA MS combined with searching of individual species databases. If LESA MS is to find use as a diagnostic tool, however, correct identification of proteins (and species) from multiple databases is necessary. To evaluate that, a data analysis workflow was constructed in the ProSightPC software in which each MS/MS mass spectrum was searched against all six individual

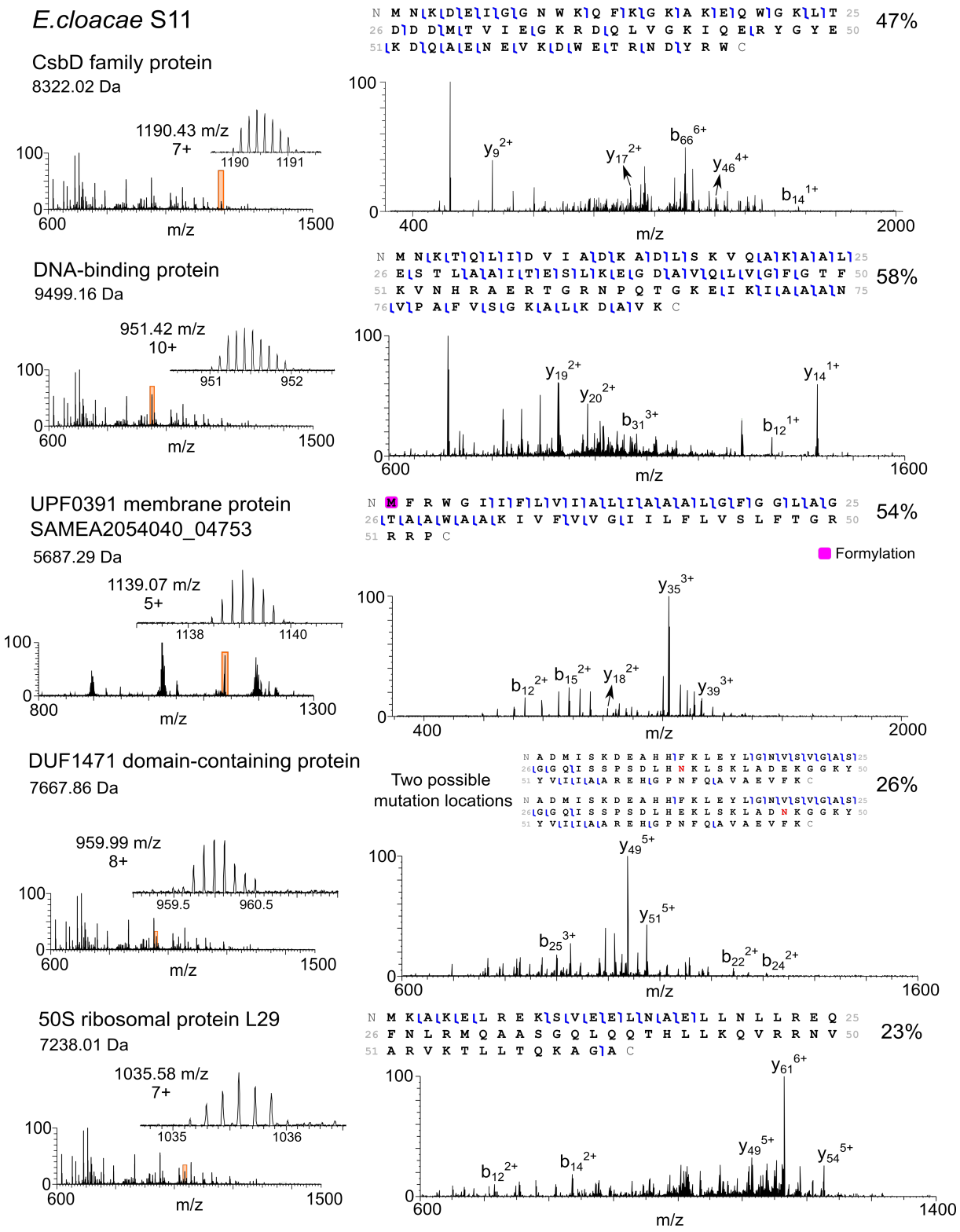


Figure 3.11: Representative MS/MS mass spectra of five newly-identified proteins for the *E. cloacae* S11.

ESKAPE protein databases (including *S. aureus* and *P. aeruginosa*), i.e. an automated concurrent search of individual databases, using the absolute mass search function. The 24 MS/MS spectra corresponding to the newly-identified proteins described above were used for the searches. The lowest e-score (expectation value) was used as the indicator of protein assignment and that assignment was compared with the known protein ID (assigned after searches against individual databases). E-score expresses how many hits are expected to have similar quality compared to those observed only by chance. The lower the e-score, the better the match and lower probability of reporting a false positive result [175]. The results are summarised in Table 3.1.

In total, 19/24 proteins were correctly assigned, both in terms of protein ID and bacterial species. For two MS/MS spectra (corresponding to CscD domain containing protein (*K. pneumoniae* KP257), and uncharacterised protein (gene ABAYE1876) (*A. baumannii* AYE)), no protein assignments were returned. Both CscD domain-containing protein, and uncharacterised protein (gene ABAYE1876) were identified above using the biomarker search function in the ProSightPC software. That function is designed for identification of truncated proteins, therefore the absolute mass search used in this automated data analysis workflow did not find any matching sequences.

Three proteins were misassigned, all from *E. cloacae* S11 – DUF1471 domain-containing protein (*E. cloacae* S11) was assigned as an uncharacterised protein EF_2117 (*E. faecalis*), 50S ribosomal protein L29 was identified as the same protein but from *K. pneumoniae* and UPF0391 membrane protein SAMEA2054040_04753 was assigned to UPF0391 membrane protein KPN_04833, again from *K. pneumoniae*. The DUF1471 domain-containing protein was identified above by use of a biomarker search within the ProSightPC software, and the misassignment appears to be the result of the absolute mass search. Both *Enterobacter* and *Klebsiella* belong to the family of Enterobacteriaceae, therefore similarities in protein amino acid sequences between these species might be observed.

The overall success rate was 79% when both Gram-positive and Gram-negative species

Table 3.1: Protein ID assignments after searches against all six individual ESKAPE databases.

Species	Protein ID	Lowest e-score	Database assignment	Correct?
E. faecium E745	Uncharacterised protein (gene HMPREF0351_11703)	3.02×10^{-34}	Uncharacterised protein (gene HMPREF0351_11703)	yes
	Uncharacterised protein (gene HMPREF0351_11270)	4.20×10^{-7}	Uncharacterised protein (gene HMPREF0351_11270)	yes
	50S ribosomal protein L29	2.83×10^{-10}	50S ribosomal protein L29	yes
	30S ribosomal protein S20	4.40×10^{-28}	30S ribosomal protein S20	yes
	Uncharacterised protein (gene HMPREF0351_12038)	4.90×10^{-25}	Uncharacterised protein (gene HMPREF0351_12038)	yes
E. faecalis V583	Uncharacterised protein (gene EF_0665)	1.10×10^{-39}	Uncharacterised protein (gene EF_0665)	yes
	50S ribosomal protein L29	3.70×10^{-39}	50S ribosomal protein L29	yes
	UPF0337 protein EF_1180	4.90×10^{-48}	UPF0337 protein EF_1180	yes
	DNA-binding protein HU	1.40×10^{-4}	DNA-binding protein HU	yes
K. pneumoniae KP257	CscD domain-containing protein	N/A	–	no
	30S ribosomal protein S16	6.50×10^{-12}	30S ribosomal protein S16	yes
	DNA-binding protein HU-α	7.63×10^{-15}	DNA-binding protein HU-α	yes
	Uncharacterised protein (gene yciG)	3.50×10^{-8}	Uncharacterised protein (gene yciG)	yes
	50S ribosomal protein L29	9.45×10^{-8}	50S ribosomal protein L29	yes
	Uncharacterised protein KPN_00497	2.54×10^{-18}	Uncharacterised protein KPN_00497	yes
A. baumannii AYE	Uncharacterised protein (gene ABAYE1298)	1.36×10^{-6}	Uncharacterised protein (gene ABAYE1298)	yes
	Entericidin B	1.70×10^{-20}	Entericidin B	yes
	Uncharacterised protein (gene ABAYE2274)	1.95×10^{-11}	Uncharacterised protein (gene ABAYE2274)	yes
	Uncharacterised protein (gene ABAYE1876)	N/A	–	no
E. cloacae S11	DNA-binding protein HU	5.83×10^{-57}	DNA-binding protein HU	yes
	CsbD family protein gene YjbJ	6.20×10^{-1}	CsbD family protein gene YjbJ	yes
	50S ribosomal protein L29	2.10×10^{-15}	50S ribosomal protein L29 (*K. pneumoniae*)	no
	UPF0391 membrane protein SAMEA2054040_04753	1.69×10^{-38}	UPF0391 membrane protein KPN_04833 (*K. pneumoniae*)	no
	DUF1471 domain-containing protein	1.80×10^{1}	Uncharacterised protein EF_2117 (*E. faecalis*)	no

are considered. Generally, the accuracy of assignment of Gram-negative species is higher than for Gram-positives, which is also observed in MALDI TOF MS [176]. Identification success rates for MALDI TOF MS at the species level vary between 84.1–94.9% [177,178] for aerobic bacteria and routine isolates and 81.8% for both aerobes and anaerobes [176]. Suggested improvements for LESA MS include the use of multiple databases at the genus and strain level, which might increase the success rate of protein identification in the future.

To further address the question of correct species identification, a BLAST search of the 24 identified proteins was performed to determine the specificity of the protein sequences. All of the protein sequences identified for *E. faecium, E. faecalis, K. pneumoniae* and *A. baumannii* were unique to their species. Two out of the five *E. cloacae* proteins were species-specific, whereas DNA-binding protein HU shared a 100% sequence homology with *Cedecea davisae*, CsbD family protein shared a 100% sequence homology with *Enterobacter hormaechei*, and UPF0391 membrane protein shared 100% sequence homology with *Lelliotia amnigena*.

3.3.4 LESA MS analysis of *P. aeruginosa* PS1054 and *C. glabrata* biofilm model

The optimal surface provides an ideal environment for culturing and attachment conditions for the microbes and biofilm development [179]. A correct choice of culturing plate, providing the surface for cell attachment, was essential prior to the experiment. Typical tissue culture treated (i.e. hydrophilic, negatively charged surface optimised for cell attachment) dishes of size 35 mm were selected – these allowed simple sample transfer between laboratories, fit into the Nanomate setup and required lower volumes of media and smaller microbial samples for biofilm growth. Next, an an optimal liquid supplying nutrients for the attached cells needed to be selected. Yeast extract peptone dextrose (YPD) broth was initially considered, however no cell growth of *P. aeruginosa* 1054 was observed in this medium. Unlike YPD, LB broth (LBB) provided good culturing

conditions for both *P. aeruginosa* 1054 and *C. glabrata*. Biofilms formed at the bottom of the tissue culture treated dish, where cells were attached to the plate surface, and liquid LBB medium was regularly changed for optimal culture growth. It was observed that a static 48 hour incubation was necessary to form a thick layer of the biofilm containing *P. aeruginosa* PS1054 and *C. glabrata* (Fig. 3.12). The layer of *C. glabrata* cells needed to be checked with an optical microscope because the biofilm could barely be seen with a naked eye (Fig. 3.12). Biofilms also contained a slimy substance corresponding to EPS induced by *P. aeruginosa* [180, 181], confirming its presence in the sample.

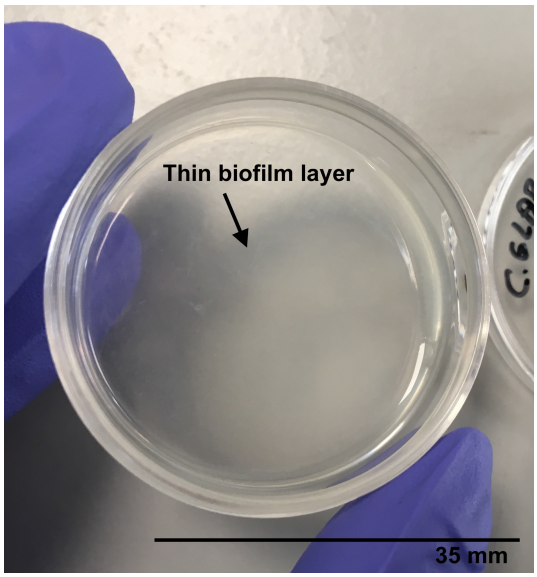


Figure 3.12: Photograph of the *C. glabrata* and *P. aeruginosa* PS1054 biofilm formed in a tissue culture treated dish after 48 hours of static incubation (LBB medium is removed).

The LBB culturing medium was carefully removed with a pipette from all plates prior to MS experiments so that LESA extraction could take place (see Fig. 3.12). First, background noise from the control culturing plate (containing no microbes) was investigated by LESA MS. LESA extraction was carried out directly from the plate surface after LBB removal (Fig. 3.13). Background signal is always present when analysing microbes growing on solid media such as LB agar (LBA) by LESA MS (Fig. 3.13). LBA is comprised of LBB and agar added to solidify the mixture. The goal was to compare the background signals from both solid and liquid media. The comparison showed that LESA mass spectra of the tissue culture dishes (after LBB removal) and LBA plates were

practically identical (Fig. 3.13).

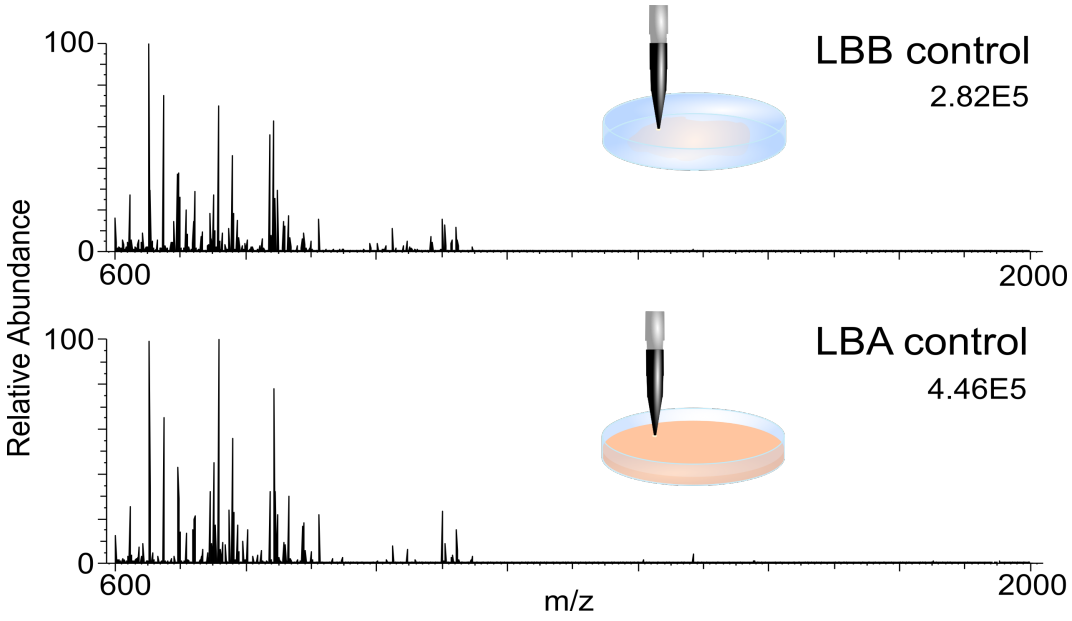


Figure 3.13: Comparison of control mass spectra of liquid LBB (LBB removed prior to LESA MS) and solid LBA. Traces of LBB are present in the plate.

Next, single-species cultures of *P. aeruginosa* PS1054 and *C. glabrata* grown in LBB were subjected to LESA MS. Empirical observations suggest that the signal intensities after LESA analysis of colonies grown on solid media (e.g., as in section 3.3.2 above) are generally higher than those from microbes cultured in the liquid media. In the single-species mass spectra, one protein for each microbe was detected and identified after MS/MS analysis (Fig. 3.14). Again, the observed signal intensities for tandem mass spectra of microbes cultured in the liquid media were lower than those of LESA MS/MS mass spectra of bacterial colonies grown on the solid substrate.

In the mass spectra of single-species *P. aeruginosa* PS1054 culture (Fig. 3.14), one peak corresponding to protein was observed. This was identified by tandem MS as BON domain-containing protein (gene PA4739) (Fig. 3.14, also see below Fig. 3.16). The BON domain-containing protein undergoes signal peptide cleavage (1–32), previously observed by LESA MS [32, 33].

LESA MS analysis of single-species *C. glabrata* culture resulted in detection of one protein (Fig. 3.14), identified as an uncharacterised protein (gene HSP12), also known as

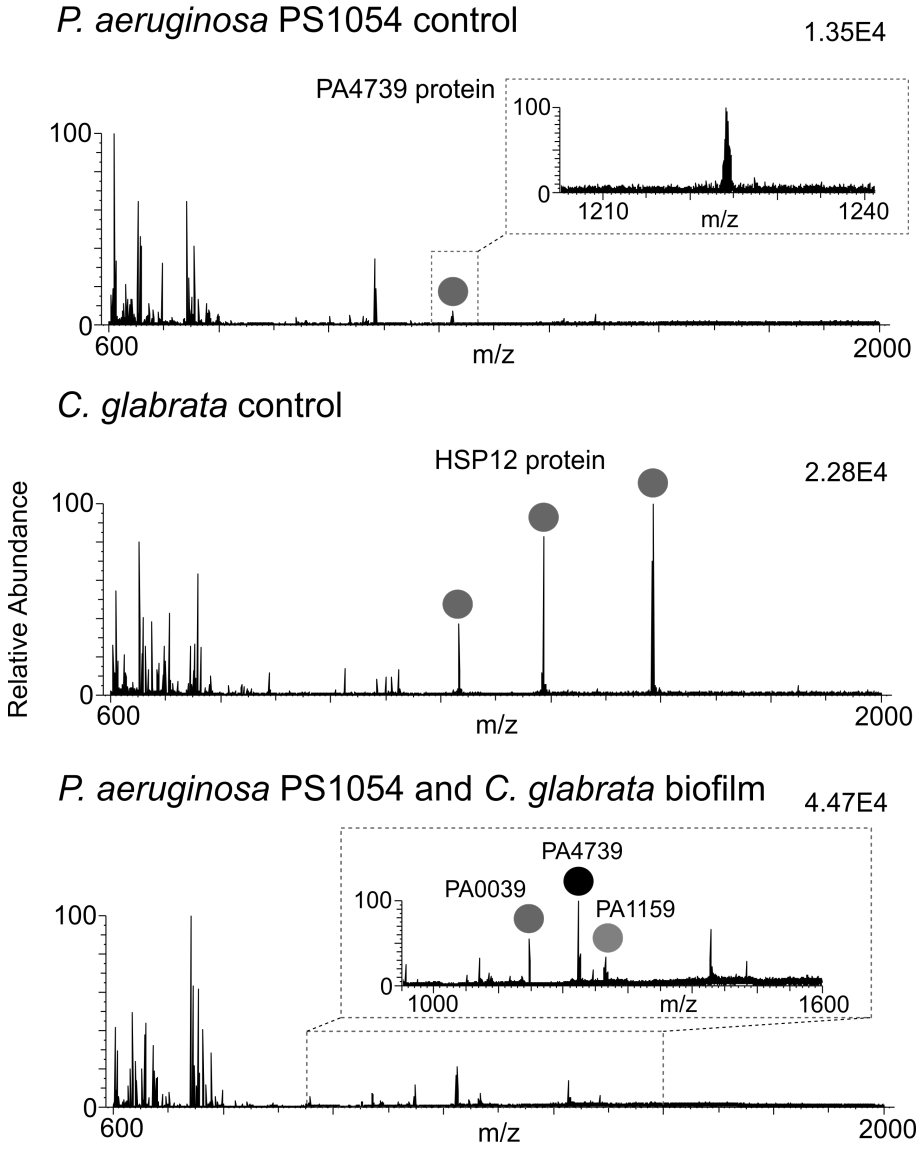


Figure 3.14: A comparison of mass spectra of controls (single-species cultures) *C. glabrata* and *P. aeruginosa* PS1054 with *P. aeruginosa* PS1054 and *C. glabrata* biofilm.

12 kDa heat shock protein (Fig. 3.15). This protein was previously observed in the LESA mass spectra of electroporated *C. glabrata* colonies [44]. Without the electroporation process, only one putative protein was observed in *C. glabrata* mass spectra after LESA extraction [44]. Here, the presence of highly abundant HSP12 may be explained by stress induced to the yeast – sudden LBB medium removal, change in the culturing conditions such as air humidity and drop of temperature. Use of the optimised LESA solvent system, probably contributing to the cell lysis, resulted in extraction of the HSP12 protein from the *C. glabrata* cells.

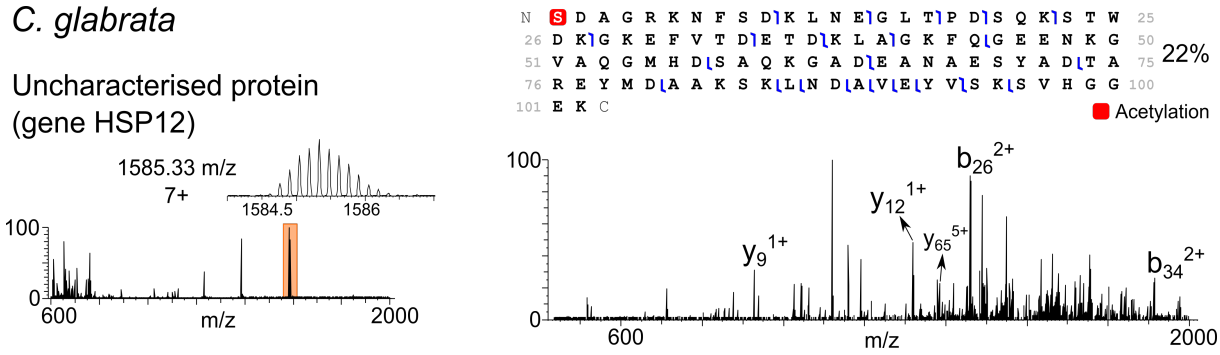


Figure 3.15: Tandem mass spectrum of the fragmented protein precursor of *C. glabrata*. This protein was identified from single-species *C. glabrata* culture.

The sample containing mixed-species biofilm of *P. aeruginosa* PS1054 and *C. glabrata* contained multiple peaks corresponding to proteins. Four peaks were selected for MS/MS analysis: 753.62 *m/z* (4+), 1148.01 *m/z* (5+), 1223.94 *m/z* (7+) and 1266.49 *m/z* (6+). With the exception of 753.62 *m/z*, the proteins were identified as uncharacterised protein (gene PA0039), BON domain-containing protein (gene PA4739) and probable cold-shock protein (gene PA1159), respectively (Fig. 3.16). No peaks corresponding to the uncharacterised protein HSP12 were detected (see Fig. 3.14), nor were any other *C. glabrata* proteins identified. This finding could potentially be explained by the antagonistic relationship of the *P. aeruginosa* and *Candida* species. Growth of *P. aeruginosa* PS1054 could be enhanced in the presence of *C. glabrata*, enabling detection and identification of three *P. aeruginosa* proteins but no *C. glabrata* proteins. SEM images of mixed-species biofilms of *P. aeruginosa* and *C. glabrata* show that 48 hour incubation reduces the biofilm density and changes the morphology of the yeast cells [169]. In this experiment, mixed-species biofilm was checked with the light microscope only, where *C. glabrata* cells were clearly visible, however no detailed morphology could be observed. It is known that *P. aeruginosa* is cytotoxic to the filamentous form of yeast such as *C. albicans* but it is unable to kill the yeast cells [170]. *C. glabrata*, in contrast with *C. albicans*, lacks formation of true hyphae. Bandara *et al.* studied the effect of lipopolysaccharide (LPS), a bacterial endotoxin present on the outer cell wall of Gram negative bacteria, on biofilm formation of *Candida* species [182]. The results showed that

Pseudomonas LPS inhibits the metabolic activity of *C. glabrata* in the first 90 minutes of the biofilm formation, while the metabolic activity of *C. albicans* was increased [182]. Interestingly, for both *C. glabrata* and *C. albicans* biofilm maturation was unaffected by *Pseudomonas* LPS after 48 hours of incubation [182]. The study presented here followed a protocol originally developed for co-culturing *P. aeruginosa* and *C. albicans* [171]. Based on the currently available information, differences between the *C. glabrata* and *C. albicans* mixed-species biofilms with *P. aeruginosa* must be considered, therefore changes will need to be made to the biofilm development and more time points will need to be investigated by LESA MS.

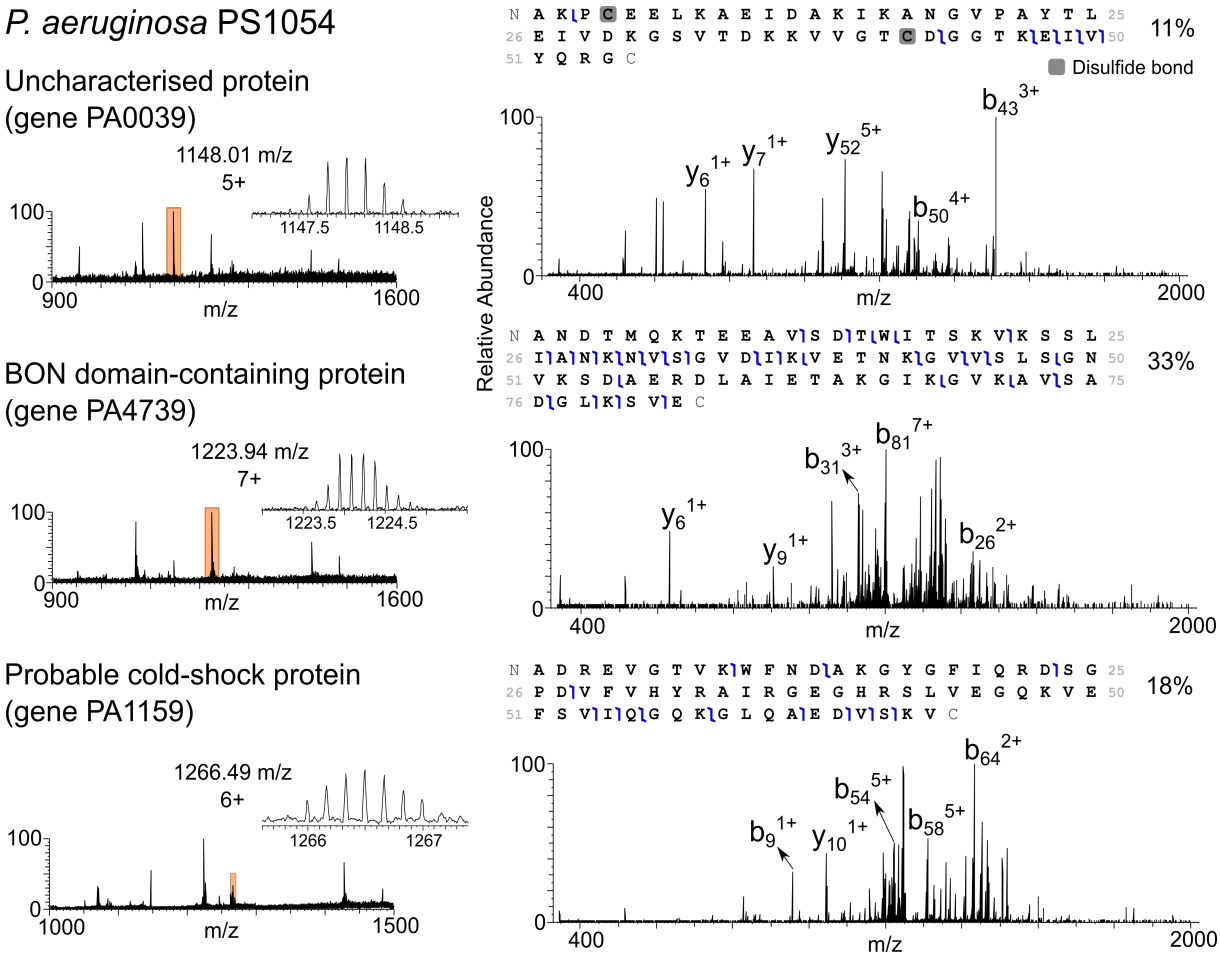


Figure 3.16: Tandem mass spectra of the fragmented protein precursors of *P. aeruginosa* PS1054 detected in the mixed-species biofilm sample. Only BON domain-containing protein was detected, fragmented and identified in the single-species *P. aeruginosa* PS1054 culture.

3.4 Conclusion

The results show that a LESA MS sampling solvent system comprising 60:35:5 acetonitrile:water:formic acid is capable of extracting proteins from both Gram-positive and Gram-negative ESKAPE pathogens. Four out of six ESKAPE microbes *E. faecium, K. pneumoniae, A. baumannii* and *E. cloacae*, including the clinically important strain *E. faecalis* and clinical isolate *A. baumannii* were investigated by TD LESA MS/MS method. The MS/MS mass spectra searches resulted in identification of 24 proteins. Proteins were also identified following searching of MS/MS spectra against databases from multiple species. In this work, the overall identification success rate was determined to be 79%, suggesting a potential of LESA MS as a useful tool in clinical microbiology.

Preliminary experiments suggest LESA MS is capable of extraction and identification of proteins from mixed-species biofilms. Three proteins were identified from the co-culture of *P. aeruginosa* PS1054 and *C. glabrata* and one protein was identified for each single-species culture of *P. aeruginosa* PS1054 and *C. glabrata*.

Chapter 4

Top-down LESA MS as a tool for direct analysis of proteins from complex substrates

Part of the work presented in this chapter was published in Havlikova *et al.* [33] (see Appendix A). Portions of the text were re-used in agreement with the open access publishing policy.

4.1 Background

Trauma is currently one of leading causes of death of people under the age of 49 [183]. Complications due to infection are the primary cause of death in patients who survive the first few days after a traumatic injury. In recent military conflicts, over half of the total injuries sustained were the result of trauma to the extremities [184]. Over 25% of those patients suffered complications due to infection (rising to 50% for those requiring intensive care), either in the bones (osteomyelitis) or deep-wound infection [184, 185]. Similar rates are associated with civilian trauma [184]. Infectious complications result in significantly higher amputation rates, which are also of greater severity.

Pathogens closely involved with traumatic wound infections are bacteria and fungi. The most commonly found bacteria in traumatic injuries are the ESKAPE pathogens (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa* and *Enterobacter cloacae*) [156], while the most

common fungal species found in wound infections are *Candida spp., Aspergillus spp., Fusarium spp., Mucor spp., Scedosporium spp.* and *Rhizopus spp.* [186, 187]. Rapid diagnosis and treatment of wound infection is of utmost importance. Current diagnosis involves visual inspection of the wound for signs of infection (inflammation), followed by collection of swabs or tissues, microbial culture and identification in a clinical laboratory, which takes hours or even days. The ideal solution would present a point-of-care diagnostic in which the microbes are identified directly from the wound, thus reducing the time from patient to result and enabling rapid deployment of the appropriate narrow spectrum antibiotic or antifungal treatment. In Chapter 3, LESA MS was applied to analysis of bacteria growing on solid media. The unique features of LESA MS with respect to its potential as a point-of-care diagnostic for wound infection are:

1. LESA sampling can be applied to any surface. To date, LESA sampling has been performed on substrates placed within the sampling platform; however, there is no inherent restriction on the nature of the substrate, e.g., a wound on a patient.

2. LESA MS enables the characterisation of proteins (as can matrix-assisted laser desorption ionisation time-of-flight (MALDI TOF) MS) and can be performed on living bacteria (unlike MALDI TOF MS).

3. LESA MS has the potential to identify not just microbial proteins but also proteins from the patient, thus providing an indication of host response.

To confirm the potential of LESA MS for direct identification of microbial wound infections, it was essential to perform LESA MS analysis of bacteria growing on more complex substrates mimicking the wound/real skin conditions. Therefore, LESA MS was first applied to the analysis of examples of all the ESKAPE species cultured on a more complex medium, i.e., blood agar containing horse blood. Protein peaks corresponding to both bacteria and substrate were detected in the mass spectra. Next, LESA MS was used for investigation of bacteria and yeast growing in wounded three-dimensional *in vitro* living skin equivalents ("Labskin"). Labskin comprises a dermal layer, consisting

of primary fibroblasts embedded in a fibrin matrix, and an epidermal layer created by seeding keratinocytes on the dermis, and the ability to culture microbes on the model is proven [188]. Clench and co-workers have applied matrix-assisted laser desorption ionisation (MALDI) MS and MALDI MS imaging of transverse sections of Labskin for the analysis of lipids and small molecules [189–194]. In this work, the analysis and identification of three ESKAPE pathogens *S. aureus*, *K. pneumoniae*, *P. aeruginosa* and one yeast species *C. glabrata* directly from wounded and infected Labskin by LESA MS was demonstrated. For *S. aureus*, two strains were considered, an MRSA reference strain and an MSSA clinical isolate. Bacterial, yeast and human proteins were detected and identified. Five of the identified human skin proteins are known to have antimicrobial activity. The protein δ-hemolysin was identified from both strains of *S. aureus*, however, the sequence of δ-hemolysin differs between the two strains, with an associated mass difference which is easily detected by MS. Detection of these proteins therefore allows the differentiation of these species. For the detection of *C. glabrata*, a process called electroporation was applied to the inoculated samples. Electroporation applies high voltages to cells to allow their permeation or lysis and it was recently shown that this technique supports the extraction and identification of proteins from yeast colonies which are particularly challenging for LESA MS [44]. One yeast protein was detected in the LESA mass spectra of an electroporated wounded and infected skin sample.

4.2 Materials and methods

Materials are listed in section 2.1. Preparation of culturing media and microbial samples is described in section 2.2. Preparation of *in vitro* 3D skin models (Labskin) is described in section 2.2.2.

The custom-built electroporation device developed in house by Dr. Klaudia Kocurek (described in detail in [44]) was used for lysis of *C. glabrata* cells grown in Labskin wounds. The electroporator device uses planar stainless steel electrodes, set to generate a field

strength of 15 kV/cm, to deliver pulses into the microbial sample (see below in the text) [44]. The height of electroporator electrodes was adjusted such that the electrodes were lowered into the infected wounded sample and the delivery of pulses was controlled through a custom software. Various electroporator settings were tested for LESA MS analysis of Labskin samples inoculated with yeast. The voltage applied varied between 1 kV – 3 kV, the number of pulses (set 1 s apart with pulse length of 20 µs) was tested in the range 3 – 60. The covered samples were stored on the laboratory bench at room temperature after electroporation and prior to MS analysis up to 2 hours.

Various extraction solvent systems were utilised for LESA MS analysis of microbes on complex substrates: 60% acetonitrile, 35% water and 5% formic acid for colonies grown on blood agar, acetonitrile-based solvent (50% acetonitrile, 45% water and 5% formic acid) and ethanol-based solvent (60% ethanol, 35%water and 5% formic acid) for Labskin samples inoculated with bacteria and 40% acetonitrile, 55%water and 5% formic acid for Labskin samples inoculated with yeast.

The LESA MS analysis of colonies cultured on blood agar followed workflow described in section 2.2.4. Labskin samples were placed in 60 mm Petri dishes adjacent to the 96-well microtiter plate. The robotic pipette of the Triversa Nanomate aspirated 3 µL of the extraction solvent from the microtiter plate. Next, the robotic arm was relocated to a position above the sample, and lowered towards the Labskin surface to enable the formation of a liquid microjunction and 2 µL of the solvent system was dispensed. After the sampling, 2.5 µL was re-aspirated back into the pipette tip and introduced into the mass spectrometer *via* chip-based nanoESI. The rest of the LESA and MS settings for the Orbitrap Elite were used as described in section 2.2.4. Part of the data was acquired on Q Exactive HF mass spectrometer. The mass spectra were acquired for at least 3 min in full scan mode, positive ion mode, in mass range 600 – 2000 *m/z*, at resolution 120 000 at 400 *m/z*. Top-down MS/MS analysis of proteins was performed by using HCD at varying collision energies in range 15 – 60 eV depending on the protein charge state. When fragmentation was carried out, each scan comprised 5 co-added microscans. The data

analysis and protein identification workflow is presented in section 2.2.5.

4.3 Results and discussion

4.3.1 LESA MS analysis of bacterial colonies growing on a complex substrate

To explore the potential of LESA MS as a diagnostic tool, bacteria were cultured on blood agar. This type of medium contains horse blood added in the preparation process, therefore it is considered as a more complex substrate than LBA or BHI typically used for growing ESKAPE pathogens. After LESA MS analysis of the blood agar only, additional peaks were observed in the mass spectra in the higher mass range (Fig. 4.1) which were not seen in the mass spectra of BHI agar nor LBA (see Fig. 4.1). The extra peaks were subjected to MS/MS analysis and a database search against the horse proteome confirmed the presence of haemoglobin α and β subunits (see Appendix B.2). Subsequently, LESA MS analysis of ESKAPE pathogens grown on blood agar was performed. For all of the ESKAPE species, both bacterial and substrate protein peaks were observed in the mass spectra (Fig. 4.2). This was an important step in the LESA MS analysis of complex substrates, suggesting that LESA MS is capable of extracting microbial and substrate proteins that were both observable in the same mass spectra and that the signal abundance of these species allows for their further tandem MS analysis and protein ID assignment (see examples of two MS/MS identifications in Appendix B.2).

4.3.2 Bacterial infection of the Labskin models and LESA MS analysis of Labskin samples

After the successful analysis of bacteria growing on a more complex blood agar substrate, LESA MS was applied to the analysis of Labskin models. The workflow is summarised in Fig. 4.3. The Labskin samples were wounded with a scalpel blade and inoculated with

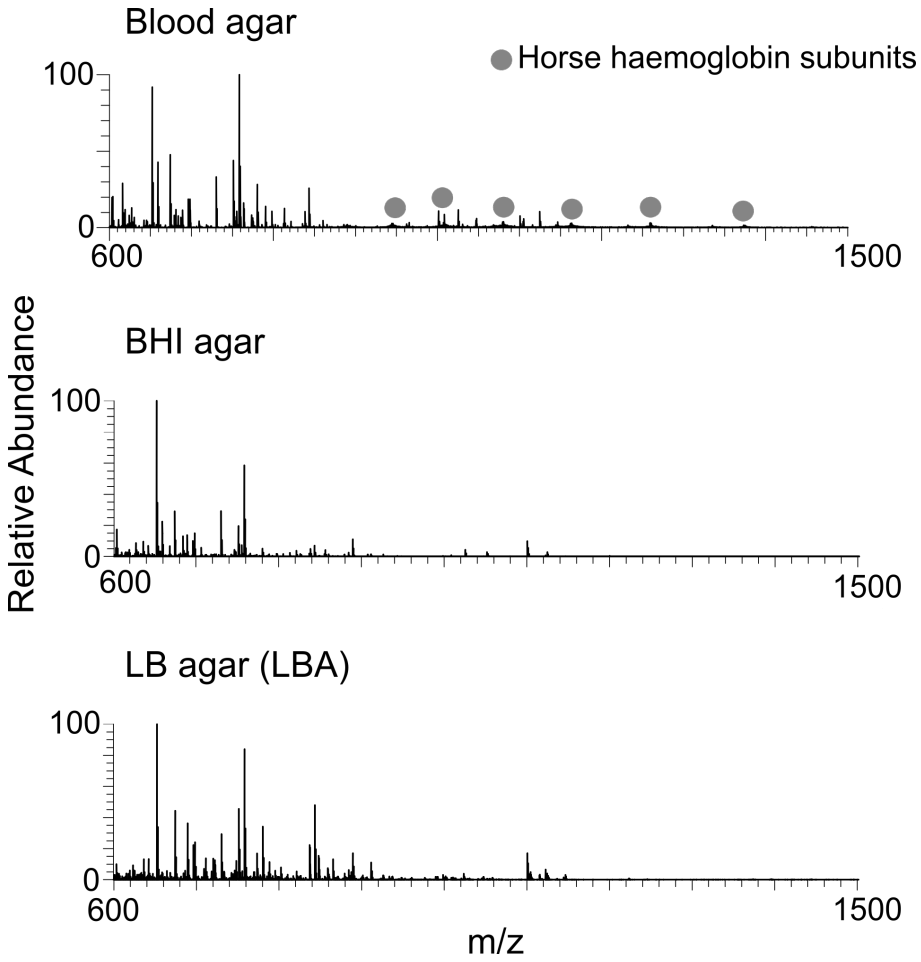


Figure 4.1: Comparison of LESA mass spectra of the more complex substrate blood agar with LBA and BHI agars (no microbial colonies present). Additional peaks observed for the blood agar were identified as horse haemoglobin α and β subunits.

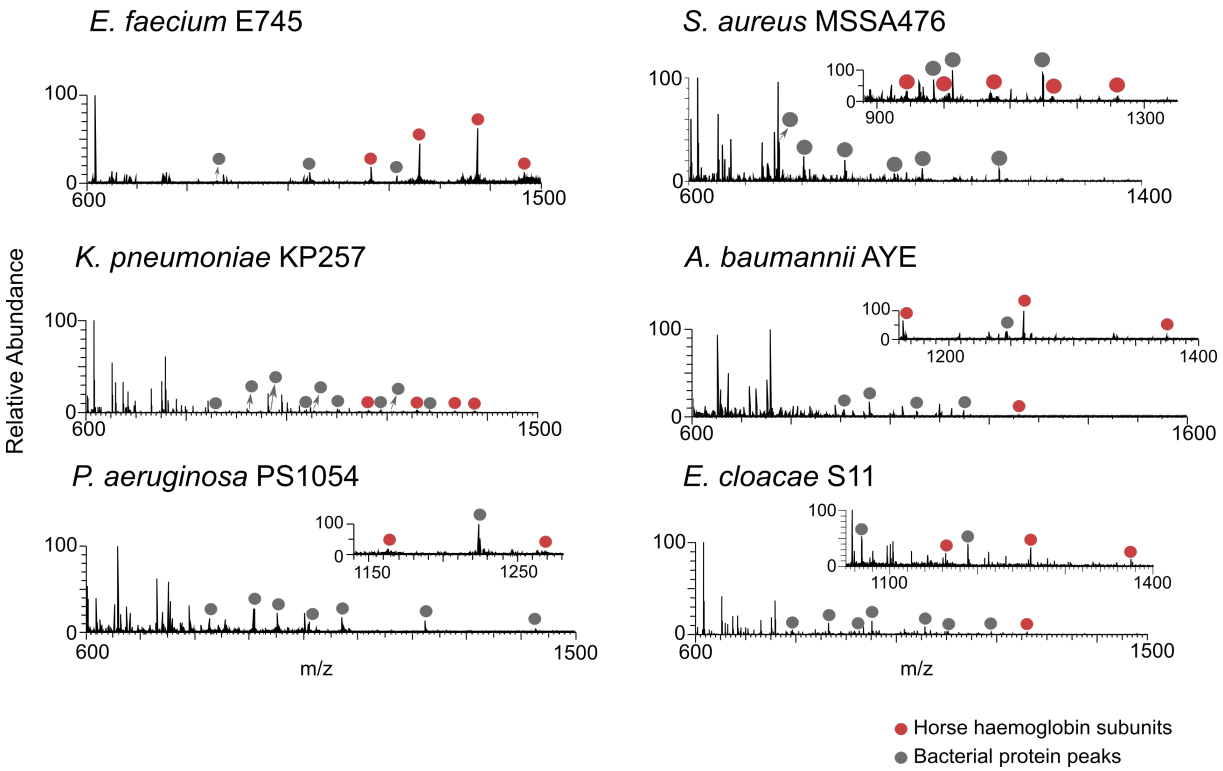


Figure 4.2: LESA mass spectra of ESKAPE bacterial colonies cultured on blood agar. Both bacterial and substrate protein peaks can be observed together in the mass spectra.

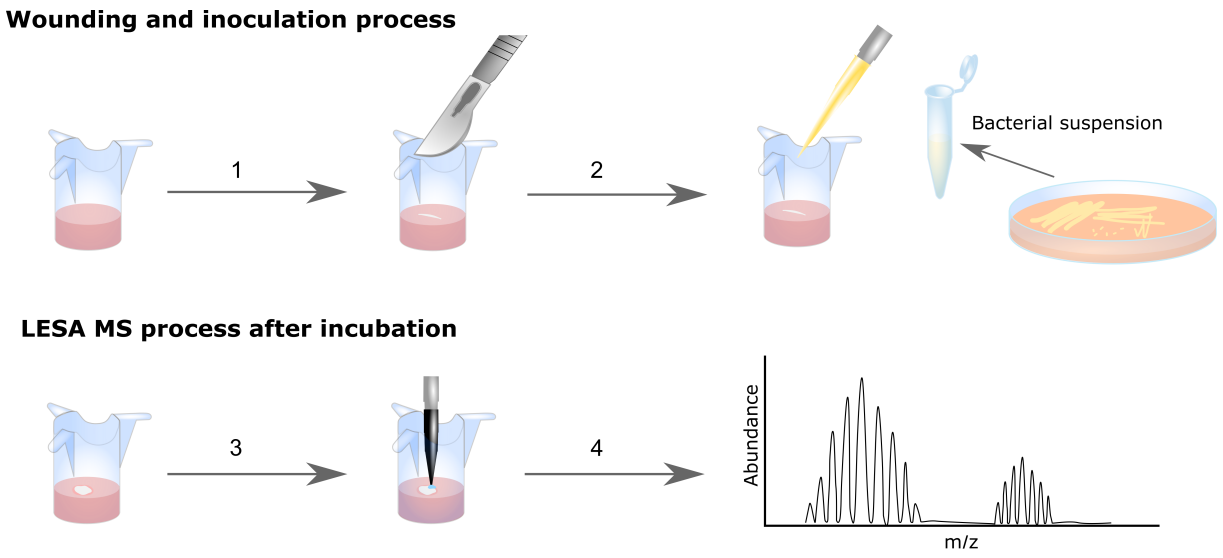


Figure 4.3: Labskin wounding, inoculation and analysis workflow. The Labskin sample (in the cell insert) is wounded with a scalpel blade (1) and inoculated with bacteria (2). After incubation (3), the infected Labskin sample is analysed by LESA MS (4).

bacterial suspensions of *S. aureus* NCTC13435, *S. aureus* MSSA476, *K. pneumoniae* KP257 and *P. aeruginosa* PS1054. The infectious dose in this experiment represents the minimum number of colony forming units (CFUs) necessary for development of infection. The required infectious dose was determined based on information provided by Labskin supplier Innovenn (*S. aureus*), literature values (*P. aeruginosa*) [195] or inferred experimentally by inoculating Labskin samples with three different bacterial concentrations (*K. pneumoniae*). The infectious doses are also in agreement with the typical concentration range (1–100 CFU/mm^2) of nosocomial pathogens on surfaces in hospital settings [196]. To ensure the correct infectious doses were administered, colonies of NCTC13435, MSSA476, PS1054 and KP257 were counted. CFU/mL in the suspension used for inoculation was back-calculated, serving as an estimate of the infectious dose administered (see Table 4.1).

The sampling solvent was optimised for extraction of intact proteins from control Labskin models and determined to be ethanol, water and formic acid (Fig. 4.4). Initially, two extraction solvent systems were tested, an acetonitrile-based (acetonitrile:water:formic acid, 50:45:5) and ethanol-based (ethanol:water:formic acid, 60:35:5). Both solvent systems were capable of extracting both human and bacterial proteins; however, a higher S/N was observed with the ethanol-based solvent system. The ethanol-based solvent system was therefore used for further experiments. An important consideration in translation of this approach to a point-of-care diagnostic is patient-friendly extraction such as that presented by an ethanol-based solvent system. The optimised sampling solvent was subsequently validated for bacterial protein detection in LESA MS of microbial colonies growing on agar (Fig. 4.5).

Four incubation time points were investigated: 24, 48, 72 and 96 h. After 24 h of incubation, there were no visible signs of colony growth inside the wound. The LESA mass spectra did not contain any peak corresponding to bacterial proteins; however, several peaks corresponding to human skin proteins were observed (see below for details of protein identification).

Table 4.1: Infectious doses for each bacterium. The ratio in brackets corresponds to dilution of the solution with OD 0.2.

Bacterial strain	**Required infectious dose (CFU)**	**CFU/mL at OD 0.2[a]**	**Volume used for infection [μL]**	**Back-calculated values of CFU/mL**
S. aureus NCTC13435	100	3.4×10^8	4 (1:13600)	6.4×10^7
S. aureus MSSA476	120	3.0×10^8	4 (1:10000)	5.7×10^7
K. pneumoniae KP257	80	8.6×10^7	10 (1:100000)	4.4×10^7
P. aeruginosa PS1054	15	6.7×10^8	3 (1:100000)	3.7×10^8

[a]As determined in previous colony counting experiments.

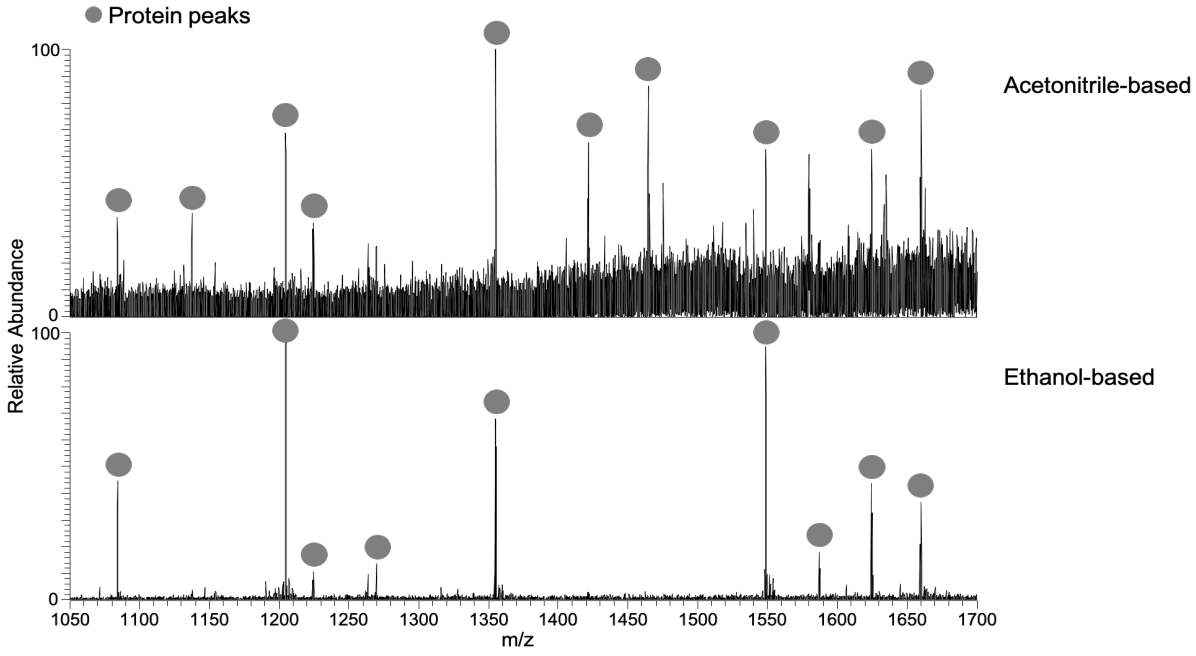


Figure 4.4: Comparison of mass spectra when (top) acetonitrile-based solvent system and (bottom) ethanol-based solvent system were used for protein extraction. The grey dots represent peaks corresponding to proteins. Both mass spectra were recorded for 3 min.

After 48 h of incubation, colonies had formed in the wound for NCTC13435, MSSA476 and KP257 (Fig. 4.6). For PS1054, significant changes in the Labskin structure were

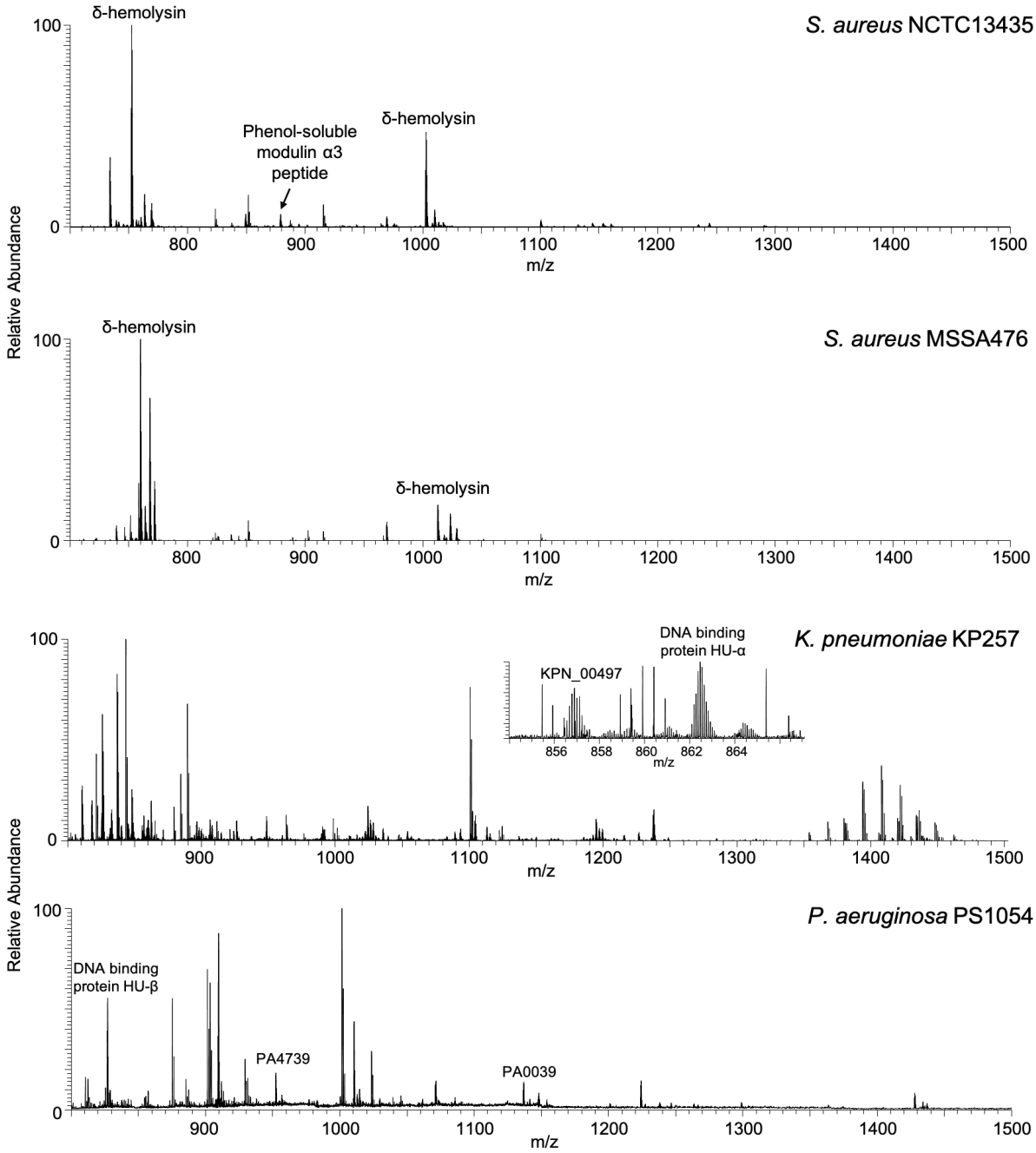


Figure 4.5: Example of mass spectra of bacterial colonies sampled with ethanol-based solvent (60:35:5, ethanol:water:formic acid). All of the mass spectra were acquired for 3 min. The LESA extraction was performed from bacterial colonies growing on LB agar. The sampling conditions were as described in section 2.2.

observed, together with the presence of the typical subtle green colour associated with secretion of pyoverdine and pyocyanin in *P. aeruginosa* [195]. Pyocyanin has a molecular weight of 210.231 Da which falls below the m/z range of this experiment. The molecular weight of pyoverdine (1365.424 Da) is within the m/z range of the experiment; however, no corresponding peaks were detected. Peaks corresponding to bacterial proteins were present in the mass spectra for all of the infected and wounded Labskin samples, in addition to peaks corresponding to human skin proteins (Fig. 4.6). Non-wounded Labskin samples inoculated with *S. aureus, K pneumoniae* and *P. aeruginosa* did not exhibit colony growth or changes in their structure during the experiment (Fig. 4.7). This result is unsurprising as colonisation of intact skin by opportunistic pathogens does not result in an infection. Comparison of the LESA mass spectra obtained from the control samples (both intact and wounded) with the infected samples (Fig. 4.6) reveals that no peaks corresponding to bacterial proteins were identified in the samples that were not inoculated, confirming that no bacterial cross-contamination occurred.

After 72 h incubation, two further proteins were detected in the mass spectra of the wounded samples infected with PS1054 (Fig. 4.8), suggesting further bacterial growth in the wound. In some LESA mass spectra obtained from the samples infected with PS1054, no skin proteins were detected, likely due to the rapid progress of the bacterial colonisation in the infected wound. After 96 h, no additional proteins, either human or bacterial, were observed.

4.3.3 Bacterial and human skin proteins identified following LESA MS analysis

The human and bacterial proteins identified following LESA MS are summarised in Table 4.3 (details of protein assignments are given in Appendix B.2). Five of the human protein-like species (β-defensin 4A, elafin, S100A7, S100A8 and S100A9) were identified as antimicrobial peptides (AMPs) and four (S100A6, S100A7, S100A8 and S100A9) belong to the low molecular weight S100 family of calcium-binding proteins.

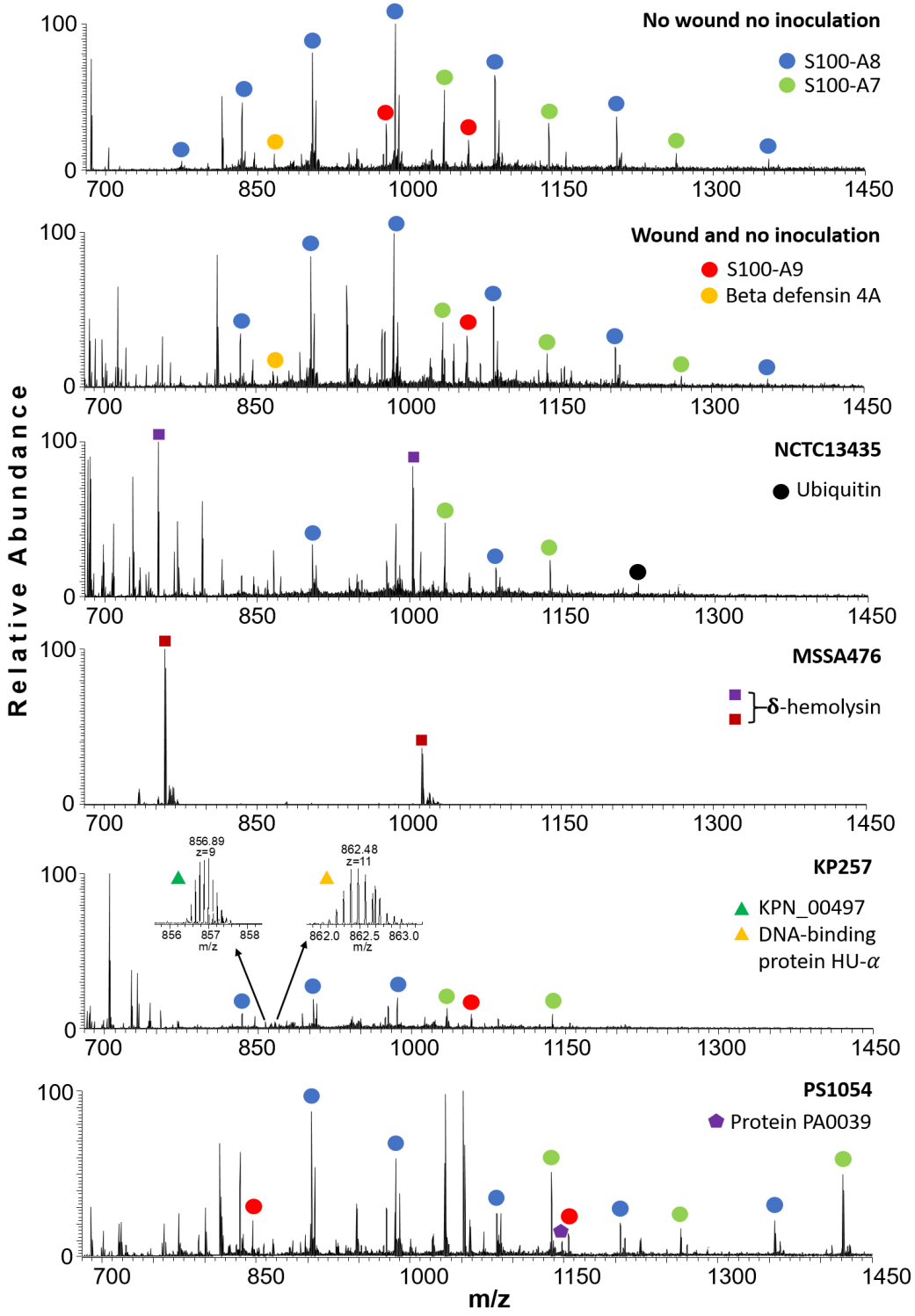


Figure 4.6: LESA mass spectra obtained from the intact control, wounded control, and Labskin samples wounded and infected with *S. aureus* NCTC13435, *S. aureus* MSSA476, *K. pneumoniae* KP257 and *P. aeruginosa* PS1054 after 48 h of incubation. Identified proteins are labelled. No bacterial proteins were detected in the control samples.

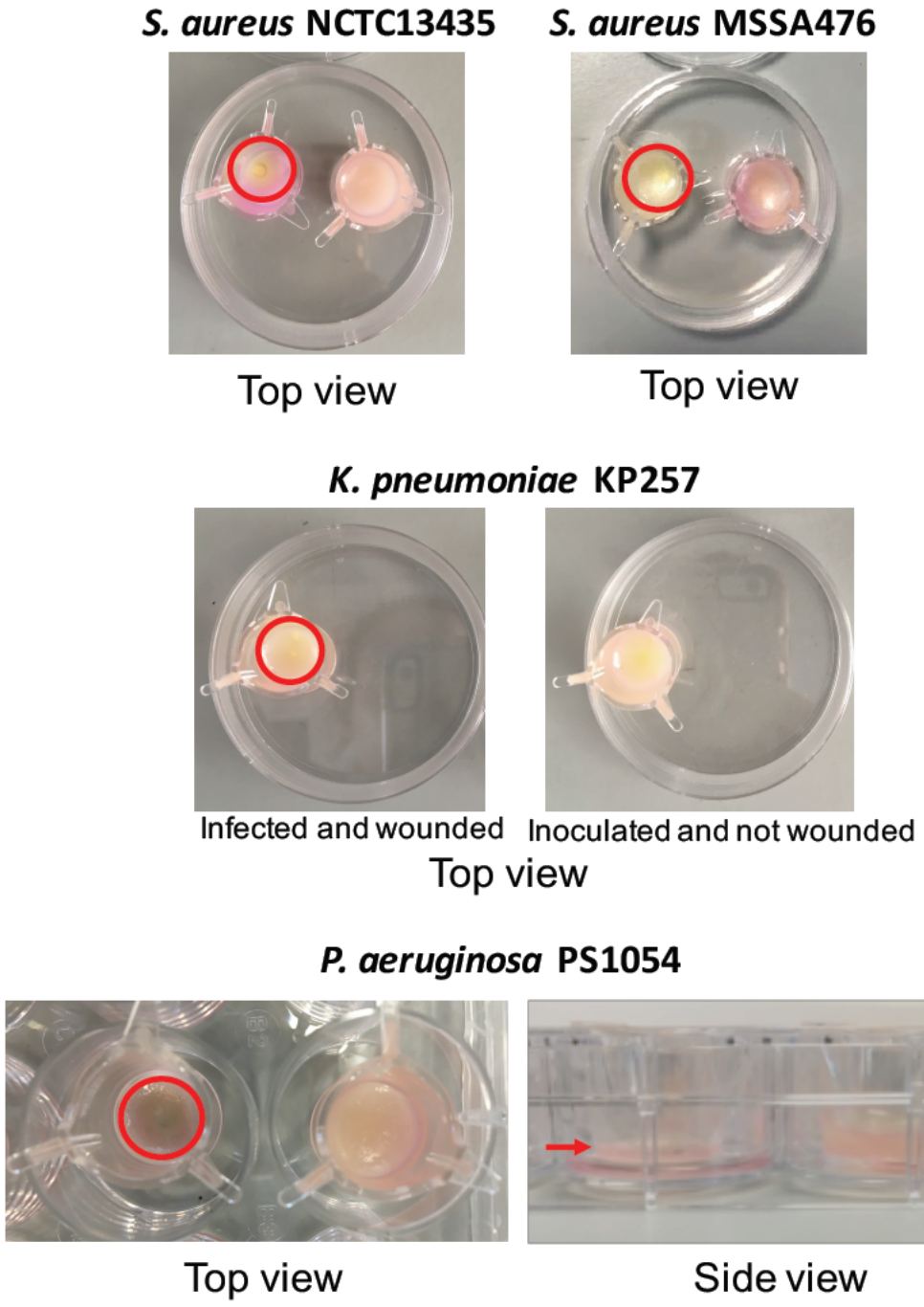


Figure 4.7: Visual comparison of infected/wounded (left) and inoculated/not wounded (right) Labskin samples for each of the bacterium studied. Red circles indicate colonies formed inside the wound. In the case of *P. aeruginosa*, the infection development is different and leads to thinning of the Labskin model (see side view, thinning is indicated by the red arrow).

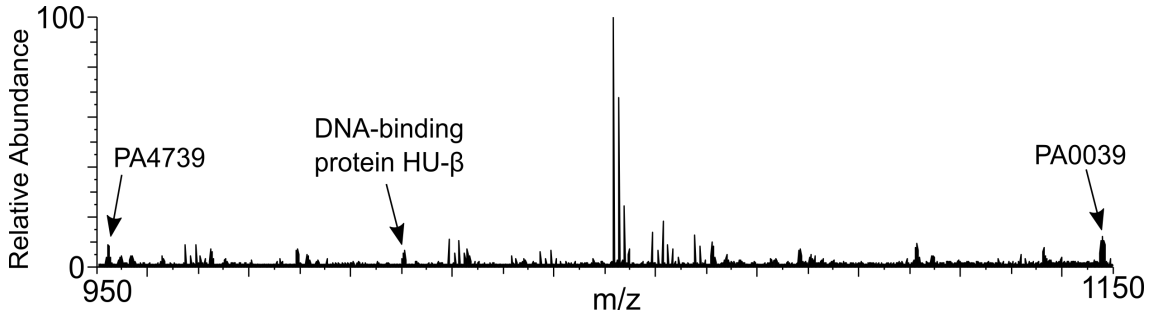


Figure 4.8: LESA mass spectrum obtained from *P. aeruginosa* PS1054 after 72 hours of incubation. Proteins PA4739 and DNA-binding protein HU-β were identified.

S100 proteins are present in the human body in both intracellular and extracellular forms. In its native state, S100A6 (also known as calcyclin) exists as a homodimer binding two Ca^{2+} ions. The protein was detected as a monomer in these experiments as a result of the denaturing extraction solvent system. S100A6 is known to be overexpressed in skin melanomas [197]. It is also known that S100A6 is upregulated by epidermal growth factor (EGF) and foetal calf serum [198, 199], both of which are used in the construction of the skin model [188] and may explain the presence of S100A6 in the mass spectra. S100A7 (also known as psoriasin) is a calcium and zinc-binding protein expressed by normal cultured (and malignant) keratinocytes [200], and therefore the presence on the skin models might be expected. Studies have shown that this protein has antimicrobial activity against *E. coli*, whilst also targeting *S. aureus, S. epidermidis* and *P. aeruginosa*, albeit less effectively [201]. S100A8 (calgranulin A) and S100A9 (calgranulin B) typically exist as the heterodimer calprotectin, which exhibits antimicrobial activity, but both are also known to act separately [202, 203]. Both proteins were detected (as monomers) in the mass spectra. S100A9 was identified in two different forms. One form was truncated at the N-terminus and the sequence was identified with serine acetylation at the (new) N-terminus. The second form was the full length protein, with S-nitrosylated cysteine at position 3, as indicated by UniProt. Elevated expression of calprotectin may be induced by skin dehydration, i.e., transepithelial water loss (TEWL) [204]. TEWL is known to be greater for skin models when compared to real skin [188], and this may explain the high abundance of S100A8 and S100A9 in the mass spectra. The function of S100A8/A9 is

also believed to be important in wound healing tissue [205], however S100A8 was detected in all of the Labskin samples, including non-wounded controls.

Proteins identified outside of the S100 family included β-defensin 4A, elafin and ubiquitin. Antimicrobial β-defensin 4A (or β-defensin 2) protects the skin from Gram-positive and Gram-negative bacteria [206]. This protein contains three disulfide bonds, specific to the family of defensins [207]. The high number of disulfide bonds and relatively short amino acid sequence resulted in poor sequence coverage (7.5%) but with a high confidence in the manual fragment assignment (see Appendix B.2). To obtain higher sequence coverage, it might be possible to reduce the disulfide bonds *in situ* prior to protein analysis. Peaks corresponding to β-defensin 4A were detected in all mass spectra regardless of the presence of wound or infection. The expression of β-defensin 4A has been shown to be upregulated in wounded and infected skin [208]. In the experiment here, the abundance of β-defensin 4A ions was higher in the control samples (both intact and wounded) when compared to the infected and wounded skin models. The higher abundance suggests that this protein may be a potential biomarker for infected tissue, however, validation experiments would be required. Elafin (skin antileucoprotease – SKALP, elastase specific inhibitor) is a peptide whose functions include antimicrobial activity and inhibition of proteases [209]. Its expression has been shown to be upregulated during inflammation, and it is a biomarker for graft-versus-host disease [210]. In this work, elafin was only detected in the mass spectra obtained from non-wounded skin samples which had been inoculated with *P. aeruginosa*. Finally, ubiquitin was observed in the majority of the acquired mass spectra. Ubiquitin is also commonly observed in the LESA mass spectra of tissue sections [94].

In terms of bacterial proteins, both extracellular and intracellular proteins were identified, in agreement with our previous findings for *S. aureus* MSSA476 and *P. aeruginosa* PS1054 grown on agar [32]. For both *S. aureus* strains, peaks were detected which correspond to two extracellular toxins – phenol-soluble modulins α3 and δ-hemolysin. These highly cytolytic peptides belong to the group of phenol-soluble

modulins specific to *S. aureus* species [211]. Top-down LESA MS revealed that both had formylated methionine at their N-termini. The accumulation of formylated δ-hemolysin is known to occur during the post-exponential growth phase [212], reflecting previous findings [32] that the abundance of this protein increases with incubation time.

In the wounded skin samples infected with *K. pneumoniae* KP257, two proteins were identified. Firstly, an uncharacterised protein predicted on the basis of gene KPN_00497 was identified. The protein was detected with a cleaved signal peptide (1–19) and a R→K substitution at position 49. Secondly, DNA-binding protein HU-α, which is involved in stabilisation of DNA under extreme environmental conditions, was identified.

Three bacterial proteins were identified from the samples infected with *P. aeruginosa* PS1054: two uncharacterised proteins and DNA-binding protein HU-β. All have been identified previously from colonies of PS1054 growing on agar [32]. The first uncharacterised protein is predicted on the basis of gene PA0039 with signal peptide (1–21) cleaved. PA0039 was the only protein detected in the mass spectra after 48 h that was sufficiently abundant to perform MS/MS analysis. The second uncharacterised protein is predicted on the basis of gene PA4739. Information about this protein available from UniProt suggests that the signal peptide is cleaved after amino acid 25, however the results presented here confirm that the signal peptide cleaves after amino acid 32, as observed previously [32]. PA4739 and DNA-binding protein HU-β were both detected after 72 h of incubation.

4.3.4 Mass difference between the δ-hemolysins of *S. aureus* strains

The sequence of δ-hemolysin differs between the two *S. aureus* strains NCTC13435 and MSSA476 as a result of a glycine to serine substitution at position 10 (G10S). This is an allelic variant related to a mutation in the hld gene and is characteristic of certain ST1 and ST59 strains of *S. aureus*, including ST1 strain MSSA476 [213]. The G→S substitution results in a mass difference, $\Delta m = +30.0105$ Da, which can be used to differentiate

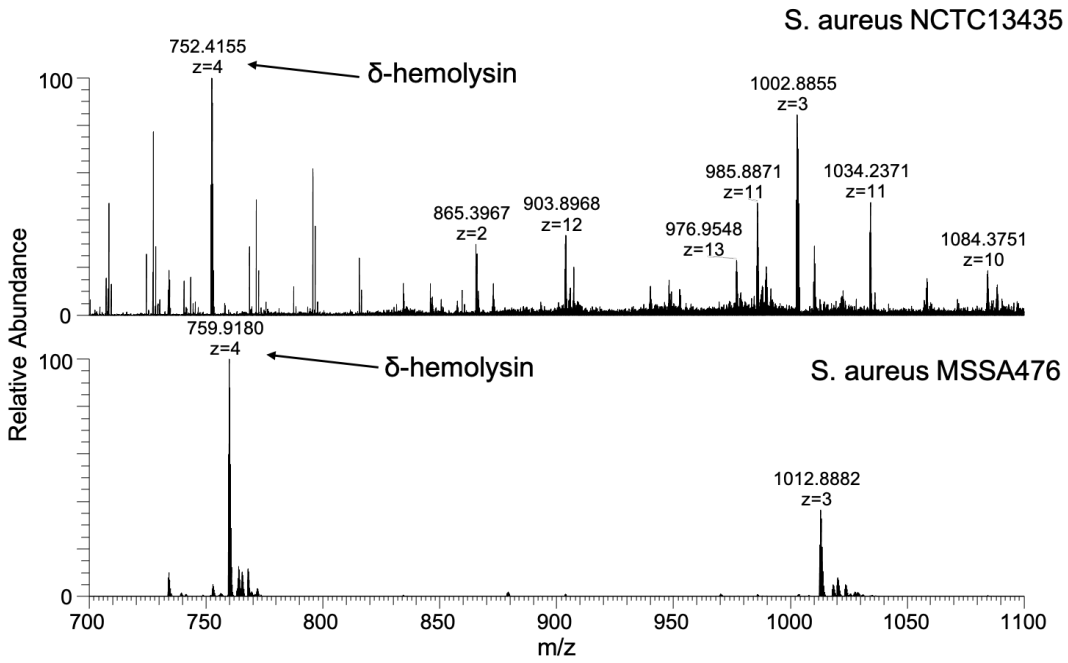


Figure 4.9: LESA mass spectra obtained from *S. aureus* NCTC13435 and *S. aureus* MSSA476. A mass shift of 30.01 Da is detected between the variants of δ-hemolysin from the two strains.

between the two strains by LESA MS. The site of substitution can be confirmed by LESA MS/MS, see Fig. 4.9. That is, LESA MS presents a rapid tool for direct differentiation between allelic variants in strains of *S. aureus*.

4.3.5 LESA MS analysis of Labskin models inoculated with yeast *C. glabrata*

The Labskin samples were inoculated with *C. glabrata* suspensions and incubated. The infectious dose for *C. glabrata* was not found to be available in the literature, therefore the infectious dose of 10^6 available for the closest well-studied species *C. albicans* was used [214]. The *C. albicans* infectious dose was, however, established to develop an *in vivo* endocarditis infection [214], therefore further optimisation of this value for inoculation of skin samples was necessary. Four different infectious doses of *C. glabrata* were tested (see Table 4.2). One out of four inoculated skin model samples developed a visible infection inside the wound and was suitable for electroporation and subsequent LESA MS analysis

throughout the whole duration of the experiment (Fig. 4.10).

Table 4.2: Infectious doses for yeast. The ratio in brackets corresponds to dilution of the solution with OD 0.2.

Yeast strain	**Infectious dose tested (CFU)**	**CFU/mL at OD 0.2[b]**	**Volume used for infection [μL]**	**Back-calculated values of CFU/mL**	**Infection developed**
C. glabrata	8100	2.7×10^6	3 (no dilution)	2.6×10^6	no
	2700		10 (1:10)		no
	1350		5 (1:10)		no
	810		3 (1:10)		yes

[b]As determined in previous yeast colony counting experiments.

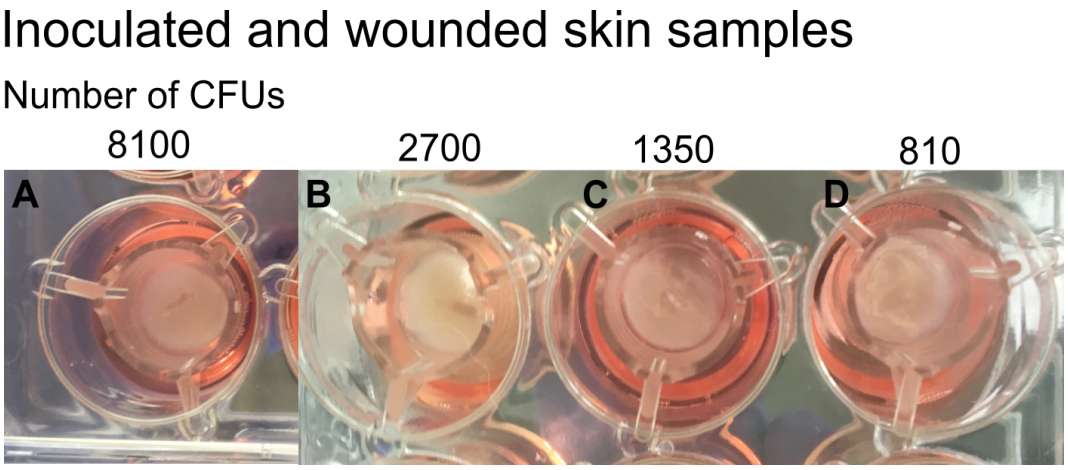


Figure 4.10: Comparison of infection development in four inoculated and wounded skin models. Infection did not progress well in samples A and B, while the sample C detached from the wall of the insert. Sample D shows signs of infection and remained suitable for electroporation and LESA MS analysis.

Inoculated samples were subjected to electroporation prior to MS analysis to aid the extraction of proteins from the skin wound (Fig. 4.11). Parameters such as voltage and number of pulses needed to be optimised for skin samples. Voltages in range 1 kV – 3 kV were tested in combination with different numbers of pulses. While arcing was observed only for the first one to three pulses when applying high voltages of 3 kV to

yeast colonies [44], it was present in most pulses above 1 kV when electroporating Labskin samples regardless of number of pulses applied (3 – 60). Arcing should be avoided as it can potentially damage the apparatus and when it occurs, the pulses are not delivered into the sample [44]. Presence of salts also contributes to the higher risk of arcing, however this cannot be completely avoided due to the fact that Labskin samples need to be incubated in a cell culture medium which is rich in salts and nutrients.

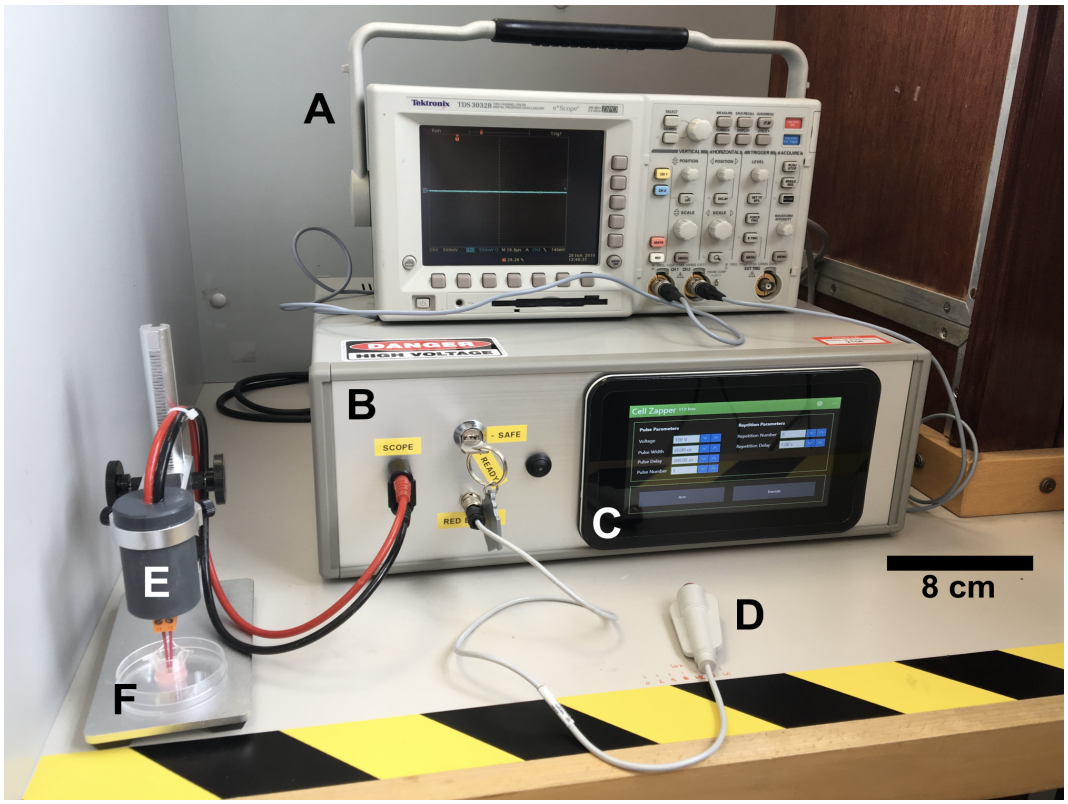


Figure 4.11: Home-built electroporator device. A – oscilloscope; B – final electroporator system encased in an aluminium box; C – device with a custom software operating the electroporator; D – a capacitor charge button (delivering pulses when pressed only); E – a set of electrodes delivering the pulses; F – Labskin sample in a Petri dish.

The summary of human skin and microbial proteins identified is in Table 4.3. Due to arcing, only 1 kV pulses could be delivered to control Labskin samples (Fig. 4.12). On day 5 of incubation, 20 pulses of 3 kV were successfully delivered into one skin sample with signs of developed infection. Protein peaks corresponding to *C. glabrata* were observed in the LESA mass spectra, suggesting the presence of intact and truncated HSP12 protein previously detected in the mass spectra of *C. glabrata* colonies (Fig.

4.13). Unfortunately, only one such sample was available for electroporation, therefore a comparison to a mass spectrum of wounded and infected sample with no voltage applied cannot be provided. Further MS/MS analysis of the putative HSP12 peak confirmed the presence of a truncated HSP12 form (Fig. 4.14), also previously observed in electroporated *C. glabrata* colonies [44]. It is likely that 3 kV pulses are necessary to be delivered to skin samples to detect any *C. glabrata* proteins, however further experiments will need to be carried out to confirm these findings.

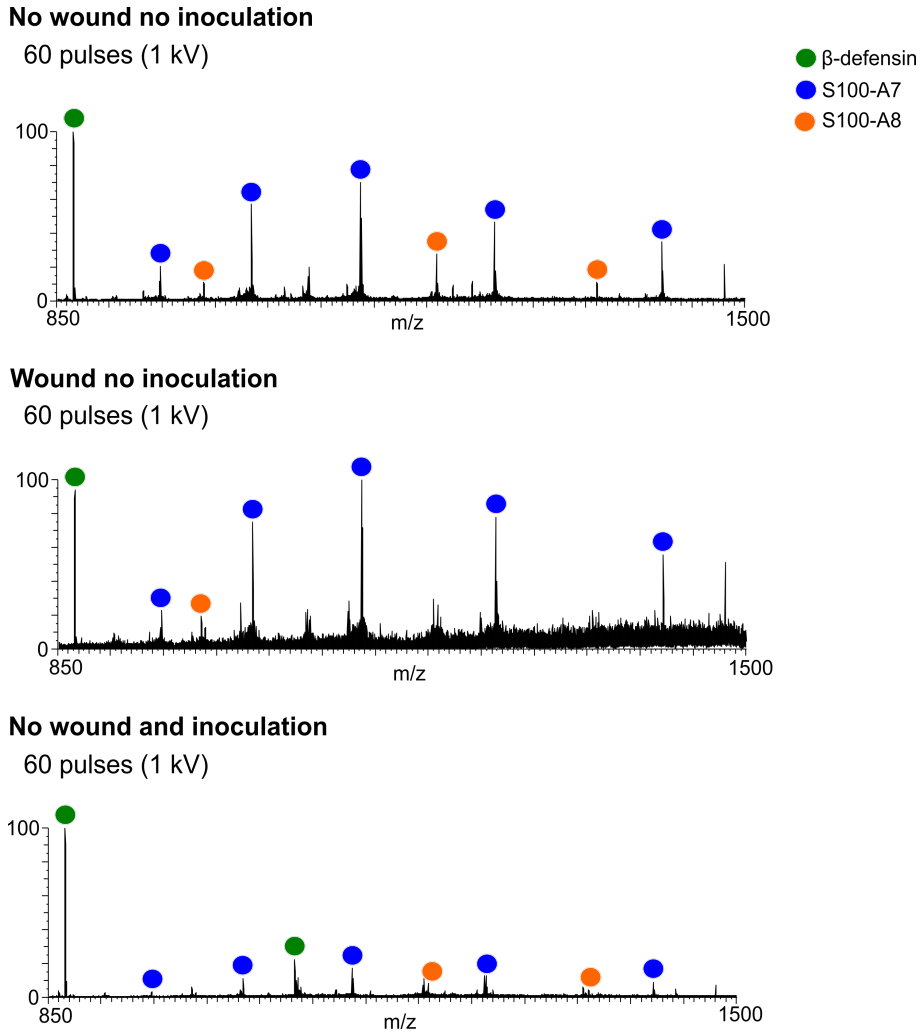


Figure 4.12: LESA mass spectra of intact control, wounded control and intact inoculated control samples subjected to electroporation. The mass spectra were acquired after 5 days of incubation. Human skin proteins are labelled.

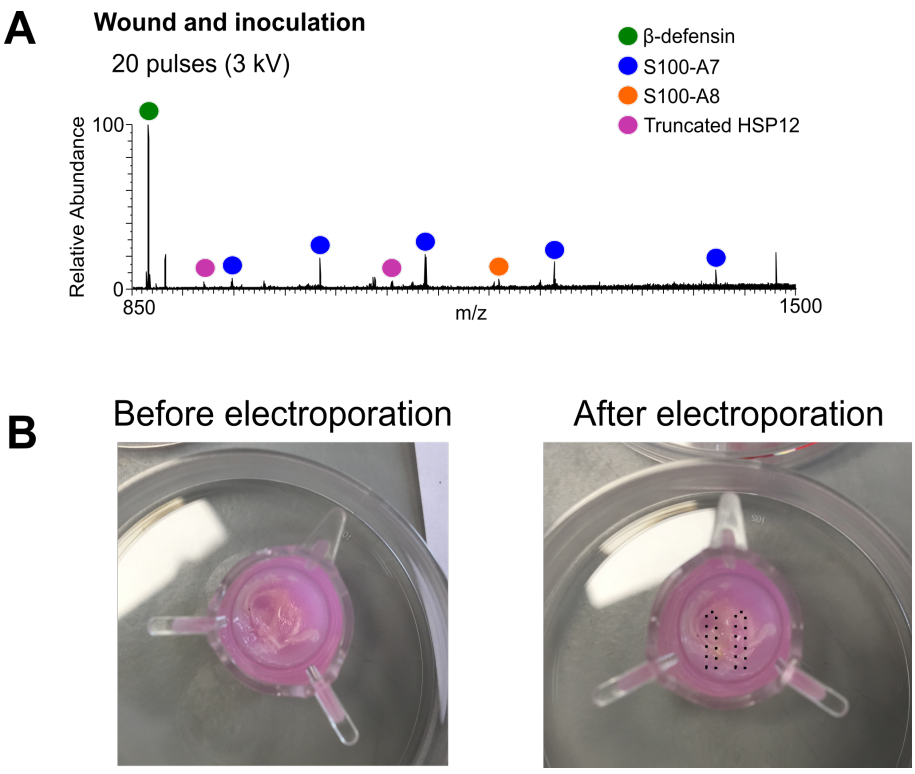


Figure 4.13: (A) LESA mass spectrum of a wounded and infected Labskin sample subjected to electroporation. Human skin proteins and the truncated HSP12 peaks are labelled. (B) Photographs of the wounded and infected sample before and after electroporation. The dashed lines indicate the area where the electrodes were placed.

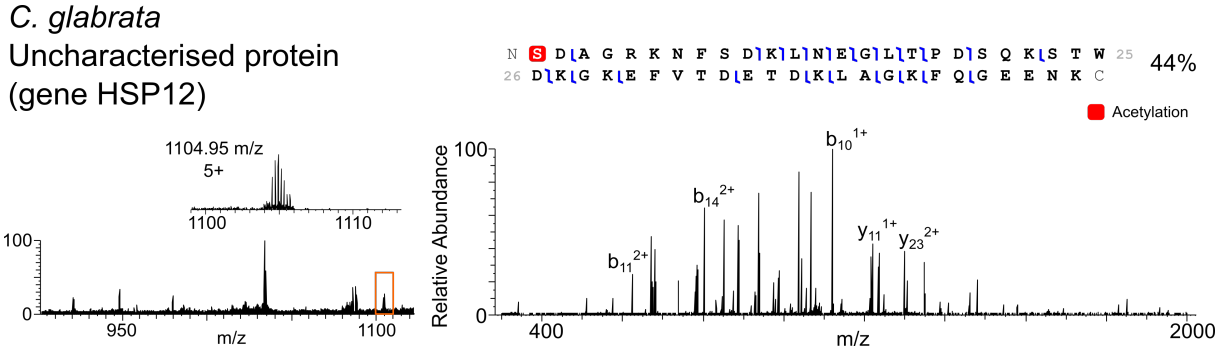


Figure 4.14: MS/MS mass spectrum of the truncated HSP12 protein of *C. glabrata*.

Table 4.3: Summary of human skin, bacterial and yeast proteins identified from Labskin samples.

Observed monoisotopic mass [Da]	**Theoretical monoisotopic mass [Da]**	**Mass difference [ppm]**	**Protein name**	**Uniprot accession number**	**Modification**	**Sequence coverage [%]**	**Sample type (I inoculated, C control)**
Human skin proteins							
4325.1296	4325.1371	–1.7387	β-defensin 4A	O15263	-Signal peptide (1–23), disulfide bonds (31–60, 38–53, 43–61)	12	I/C
5994.8239	5994.8202	0.6205	Elafin	P19957	Disulfide bonds (76–105, 83–109, 92–104, 98–113)	9	I
8559.6029	8559.6167	–1.6146	Ubiquitin	P62987	–	16	I/C
10027.3088	10027.2911	1.7632	S100-A6	P06703	-Met, acetylation of N-terminus (A)	8	I/C
10827.6373	10827.6492	–1.1046	S100-A8	P05109	–	53	I/C
11360.4787	11360.5189	–3.5403	S100-A7	P31151	-Met, acetylation of N-terminus (S), E27D substitution	26	I/C
12682.2879	12681.2806	–0.4274	S100-A9	P06702	Acetylation of N-terminus (S) (truncated form), S-nitrosylation (C3) (full length form)	17	I/C
Observed monoisotopic mass [Da]	**Theoretical monoisotopic mass [Da]**	**Mass difference [ppm]**	**Protein name**	**Uniprot accession number**	**Modification**	**Sequence coverage [%]**	**Hours of incubation**
S. aureus NCTC13435							
2633.4082	2633.4080	0.0573	Phenol-soluble modulin α3	P0C805	fMet	95	48
3004.6109	3004.6302	–0.1664	δ-hemolysin	Q2FWM8	fMet	60	48
S. aureus MSSA476							
2633.4094	2633.4080	0.5240	Phenol-soluble modulin α3	P0C805	fMet	95	48
3034.6413	3034.6413	–0.0066	δ-hemolysin	Q6G7S2	fMet, G10S substitution	88	48
K. pneumoniae KP257							
7698.9926	7698.9638	3.7382	KPN_00497	A6T5S6	-Signal peptide (1–19), R49K substitution	19	48
9471.1664	9471.1468	2.0673	DNA-binding protein HU-α	A6TGQ7	–	30	48
P. aeruginosa PS1054							
5731.9826	5731.9816	0.1814	PA0039	Q9I793	-Signal peptide (1–21)	13	48
8557.5098	8557.5077	0.2431	PA4739	Q9HV60	-Signal peptide (1–32)	44	72
9081.0514	9081.0445	0.7576	DNA-binding protein HU-β	P05384	–	26	72
C. glabrata							
5516.6756	5516.6608	2.6792	HSP12	Q6FPF6	-Met, acetylation of N-terminus (S) (truncated form)	44	120

4.4 Conclusion

In this Chapter, LESA MS analysis of microbes growing on complex substrates was presented. LESA MS analysis of the ESKAPE pathogens cultured on the more complex medium blood agar revealed that both bacterial and substrate protein peaks can be observed in the mass spectra and subsequently identified. The ESKAPE pathogens including different *S. aureus* strains were used for inoculation of Labskin models and were rapidly identified from infected wounds of 3D skin models. LESA MS of *in vitro* 3D skin models utilised an ethanol-based extraction solvent system. Intact proteins from both human skin and the infecting bacteria were identified with high confidence. Bacterial proteins could be identified within minutes once visible signs of infection were apparent (i.e., 48 h after inoculation of the skin model with bacteria). Detection of allelic variants of the protein δ-hemolysin enabled differentiation between the two strains of *S. aureus*. The skin models were also inoculated with the yeast *C. glabrata*. Electroporation of inoculated skin samples was performed to aid extraction of yeast proteins from the sample. One *C. glabrata* protein HSP12, also previously observed in electroporated yeast colonies, was identified after 5 days of incubation.

Chapter 5

Top-down LESA MS analysis of *ex vivo* human skin

5.1 Background

In the work presented in Chapter 4, *in vitro* 3D skin models were used as a model for development of infection in a skin wound. The 3D skin models consist of a dermal equivalent containing human fibroblasts and an epidermis formed with human keratinocytes [188,215]. The *in vitro* skin models represent a great environment mimicking the natural skin, but in fact, human skin is composed of three layers: epidermis, dermis and subcutaneous tissue containing a layer of fat cells [216]. Skin is a complex system also comprising many different types of structures such as cells synthesizing pigment, mechanoreceptors, cells involved in immune system responses, sweat glands, nerves, hair follicles and nails [216,217]. Utilising *ex vivo* human skin grafts would therefore represent a more realistic model for investigating infected skin wounds.

MS proteomics approaches have already been employed in studies of *ex vivo* human skin samples and bacterial models associated with skin. The most common approaches involve liquid chromatography MS (LC-MS) and MALDI TOF MS (including MALDI TOF MS imaging) [218]. Sample preparation prior to LC-MS analysis, however, includes a series of steps necessary to extract proteins from the sample followed by digestion of extracted proteins into peptides, while MALDI TOF MS typically uses histological cryosections of frozen human skin tissue. Both LC-MS and MALDI TOF MS are suitable for studying bacterial-skin interaction or identification of proteins and, additionally,

MALDI TOF MSI offers unique information about spatial distribution of proteins across the investigated tissue. An approach suitable for diagnostic purposes would, however, ideally be performed directly from the skin sample (or skin directly) without any (or only minimal) requirement for sample pre-treatment and it was observed that LESA MS is capable of such analysis (Chapter 4).

The work in this Chapter extends the TD LESA MS approach from the Labskin models to *ex vivo* human skin grafts. The sample preparation process is described, followed by optimisation of the extraction solvent system and the remainder of the workflow. The human skin samples were also wounded and inoculated with *S. aureus*. In total, two human skin proteins were identified by TD LESA MS, both of which were observed in non-inoculated and inoculated samples, while no bacterial proteins were identified so far by the developed workflow.

5.2 Materials and methods

Materials are listed in section 2.1. Specifications of the *ex vivo* human skin sample preparation prior to incubation are described in section 2.2.3.

A range of methanol-based and acetonitrile-based extraction solvent systems was tested for LESA MS analysis. The human skin samples sutured onto cell strainers were placed in 60 mm Petri dishes adjacent to half of the 96-well microtiter plate. The robotic pipette of the Triversa Nanomate aspirated 3 µL of the extraction solvent system from the microtiter plate and the robotic arm was subsequently relocated to a position above the sample. Next, the robotic pipette descended above the human skin sample to allow formation of the liquid microjunction and 2 µL of the solvent system was dispensed. After the sampling process, 2.5 µL of the sample was re-aspirated back into the pipette tip and introduced into the mass spectrometer *via* chip-based nanoESI. The remainder of the LESA MS settings was used as described in 2.2.4. Data analysis and protein identification followed workflow in section 2.2.5.

For a subset of investigated samples, the Triversa Nanomate was coupled to a miniaturised ultra-FAIMS device (Owlstone, Cambridge, UK), which was coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). FAIMS separation was performed in positive ion mode using the microchip device [219]. The chip temperature was set to 100°C and FAIMS was operated in a static mode. Ranges of values for dispersion field (DF) 180–250 Td (step size 10 Td) and compensation field (CF) 1.5–2.5 Td (step size 0.1 Td) were tested.

5.3 Results and discussion

5.3.1 General overview of human skin sample preparation prior to LESA MS analysis

A summary of sample preparation of the *ex vivo* human skin grafts and LESA MS analysis is presented in Fig. 5.1A (the detailed sample preparation is described in section 2.2.3). Briefly, skin samples were collected from Human Biomaterials Resource Centre (HBRC) of the University of Birmingham and transferred in a low glucose DMEM cell culturing medium supplemented with 1% P/S on ice to the laboratory. The skin samples were disinfected with ethanol and wiped to remove any traces of blood. Next, skin grafts were cut by scalpel and scissors into smaller pieces of size 1.5 × 1.5 cm and sutured onto cell strainers (Fig. 5.1B) as described by Xu *et al.* [220]. Samples were either left intact (Fig. 5.1B, sample 1) or wounded with a scalpel blade (Fig. 5.1B, sample 2). Suturing of the skin samples to cell strainers was necessary to keep the skin's natural tension. Skin grafts were also incubated in cell inserts only (Fig. 5.1C), leaving out the suturing step. These conditions were, however, not suitable for the subsequent LESA MS analysis (data not shown). The skin grafts were soaked with culturing medium and the skin surface was less even than surface of samples sutured to the cell strainer (Fig. 5.1C). Furthermore, manipulation of the skin sample, e.g. wounding, would be more challenging than wounding

of the sutured sample (see Fig. 5.1C).

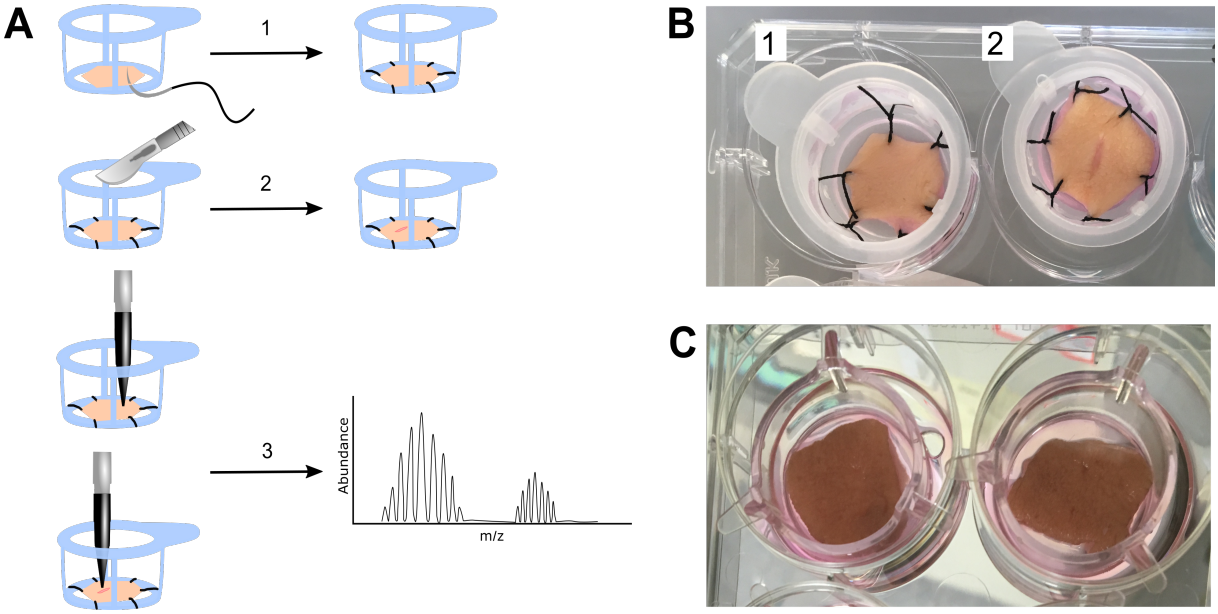


Figure 5.1: The general workflow of sample preparation and photographs of *ex vivo* human skin grafts. A general workflow is presented in (A): small pieces of skin grafts were sutured onto cell strainers (A1), wounded with a scalpel blade (A2), incubated and subjected to LESA MS at different time points (A3). A close view of the intact (B1) and wounded (B2) sutured human skin samples. The non-sutured samples are presented in (C).

5.3.2 Optimisation of LESA MS analysis of *ex vivo* human skin samples

The human skin samples were analysed at four different time points – after 24, 48, 72 and 96 hours of incubation. The first step in the LESA MS analysis was to find the most suitable solvent system for LESA extraction of proteins. Three extraction solvent systems were considered: acetonitrile-based, ethanol-based and methanol-based. For all three types of the solvent systems, various ratios of the organic solvent, water and formic acid were mixed and tested. The combinations of 50:49:1 acetonitrile:water:formic acid, 50:45:5 methanol:water:formic acid and 60:35:5 ethanol:water:formic acid resulted in detection of multiply charged species or putative protein peaks in the mass spectra (Fig. 5.2). The signal intensities were comparable in all analysed mass spectra (see Fig. 5.2). Despite the fact that all three solvent systems were capable of protein extraction, it was decided to

continue using the acetonitrile-based and methanol-based solvents for further optimisation because it was established in Chapter 4 that the acetonitrile-based solvent is suitable for extraction of skin proteins from *in vitro* skin models [33] while the previous work on LESA MS of thin tissue sections utilised methanol-based solvents [93]. Once the workflow is optimised for the analysis of *ex vivo* samples, it would be suggested to move towards the patient-friendly ethanol-based extraction solvent system.

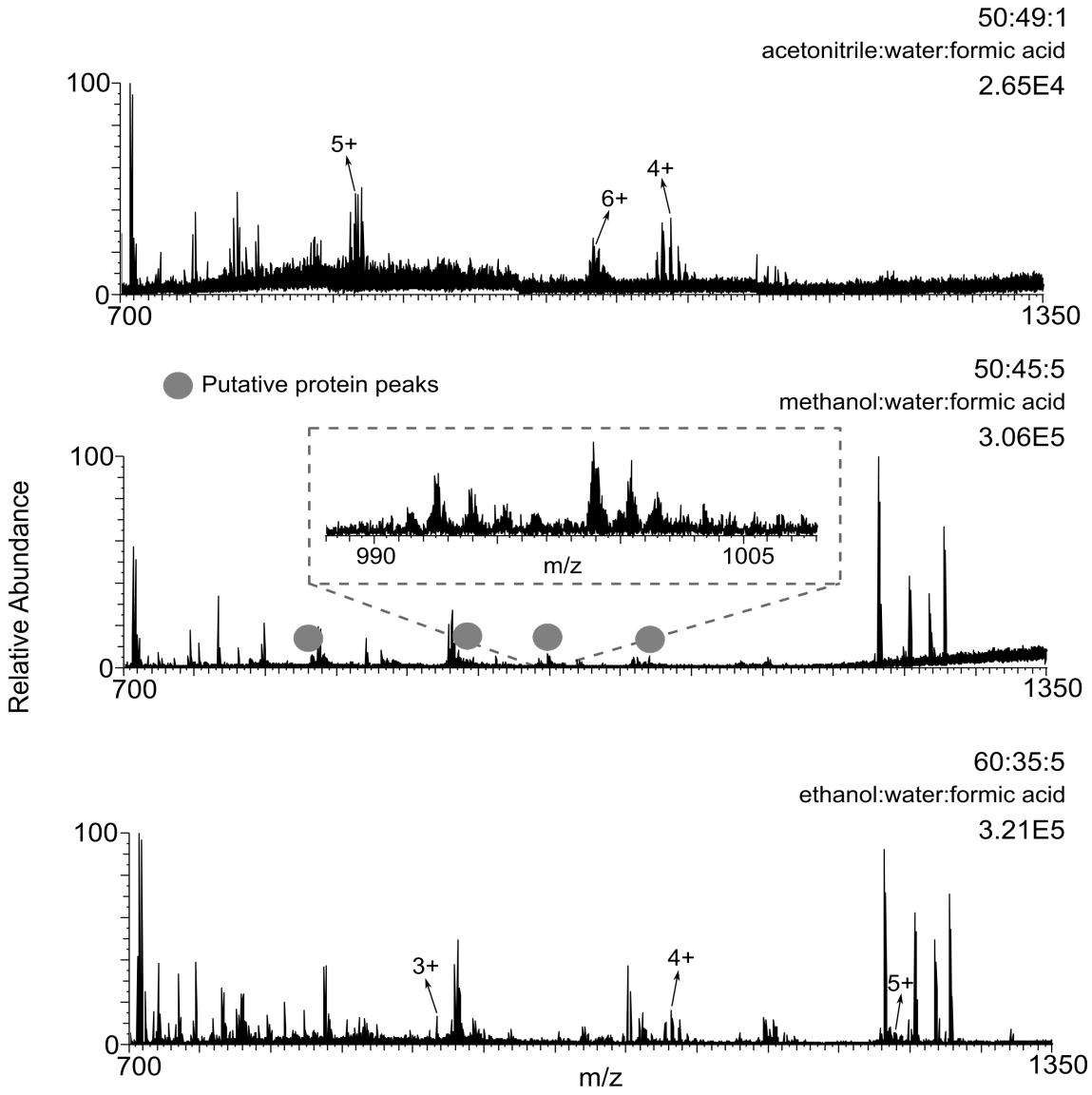


Figure 5.2: Comparison of three different solvent systems used for LESA extraction of proteins from *ex vivo* human skin samples. Multiply charged species and putative protein peaks were observed in the mass spectra.

A major challenge posed during the *ex vivo* human skin analysis was the presence of

salts in the samples. The skin samples were incubated in a cell culturing medium rich in salts and nutrients and the medium was found to be present in the epidermal layer – the top of the sample. A washing step was therefore included prior to MS analysis: either the whole surface of the sample was washed with water or the sampling spot was washed directly by the robotic pipette, i.e., the robotic arm dispensed 3 μL of water on the sampling spot, aspirated the water droplet, discarded the volume into an empty well and performed the sampling step from the same spot as described in section 5.2. Unfortunately, the washing step did not successfully remove all the remaining salts from the surface (Fig. 5.3). Fig. 5.3 shows no difference between the washed and unwashed human skin, observed in majority of the investigated samples, suggesting that washing did not help with efficient salt removal (and may also reduce signal corresponding to proteins). Interestingly, for some of the samples tested, it was still possible to detect proteins (Fig. 5.4) even without the washing step. Mass spectra presented in Fig. 5.4 were acquired as described in section 5.2 (with no washing step included). Differences between the mass spectra in Fig. 5.3 and 5.4 suggest presence of high sample-to-sample variability (or spot-to-spot variability within one sample) resulting in protein detection only in few LESA extractions, thus leading to low reproducibility. Based on the presented findings it also appears that the salts (or medium) is probably not distributed equally across the surface of the sample and possibly a different method, rather than sample washing, might be more suitable to reduce the background noise (see text below).

The mass spectra of *ex vivo* human skin grafts containing protein peaks were compared to the mass spectra of *in vitro* skin models (Labskin) and peaks of two different proteins were found to match (Fig. 5.4). In total, seven protein peaks from the *ex vivo* human skin samples mass spectra were subjected to MS/MS analysis – 832.43 *m/z* (4+), 860.88 *m/z* (4+), 865.80 *m/z* (5+), 913.69 *m/z* (5+), 1033.69 *m/z* (11+), 1221.77 *m/z* (7+), 1318.88 *m/z* (6+) and 1420.82 *m/z* (8+). Only two human skin proteins were successfully identified: S100-A7 (psoriasin) and ubiquitin (see Fig. 5.5). Ubiquitin was identified in a truncated form, missing the last six C-terminal amino acid residues (Fig. 5.5).

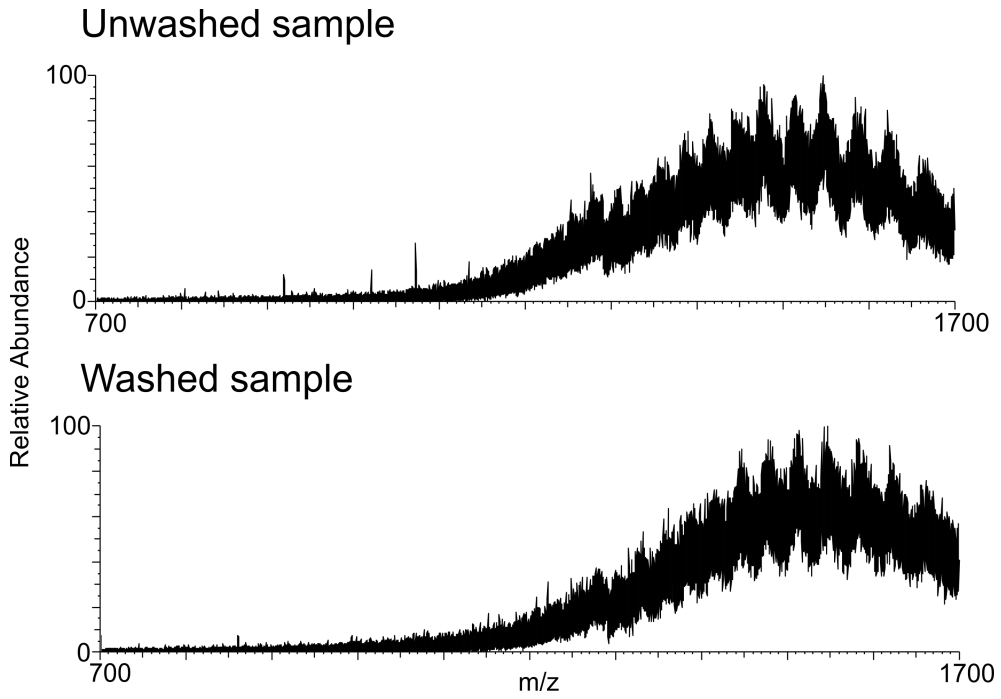


Figure 5.3: Comparison of LESA mass spectra of *ex vivo* human skin samples before and after washing step. The extraction solvent system 50:49:1 acetonitrile:water:formic acid was used.

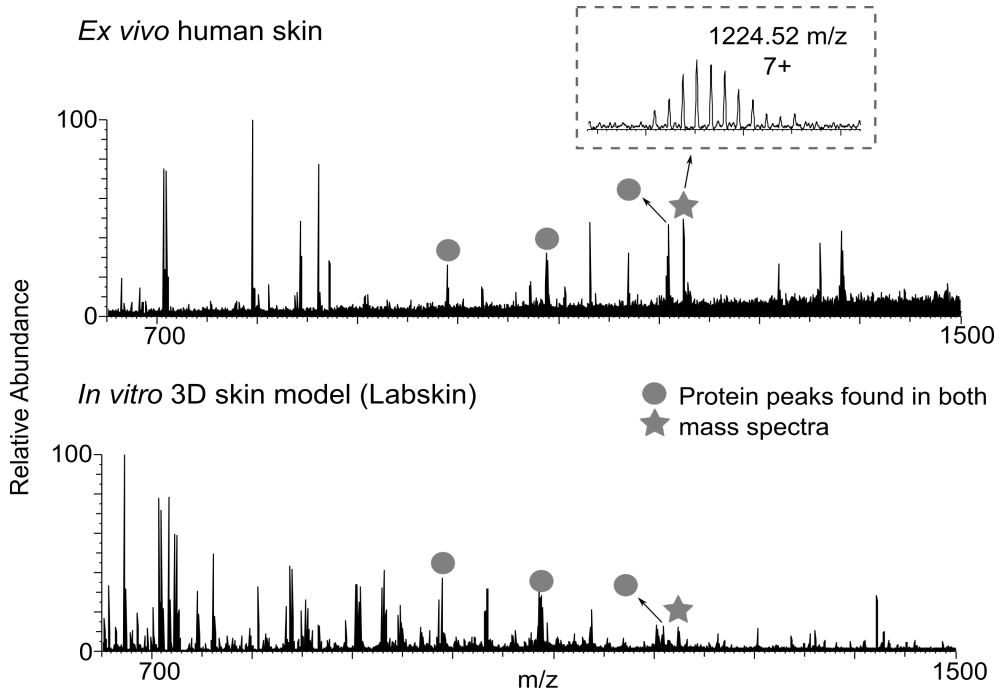


Figure 5.4: Comparison of the LESA mass spectra of *ex vivo* human skin (unwashed) and *in vitro* 3D skin model samples. Three protein peaks corresponding to different proteins were observed in both mass spectra.

Both proteins were observed in non-inoculated and inoculated samples (see below). The sequence coverage of these proteins is lower than those of proteins identified from Labskin samples (see Chapter 4), which is likely caused by the lower signal abundance due to presence of salts interfering with the protein signal (see the mass spectra of protein precursors in Fig. 5.5).

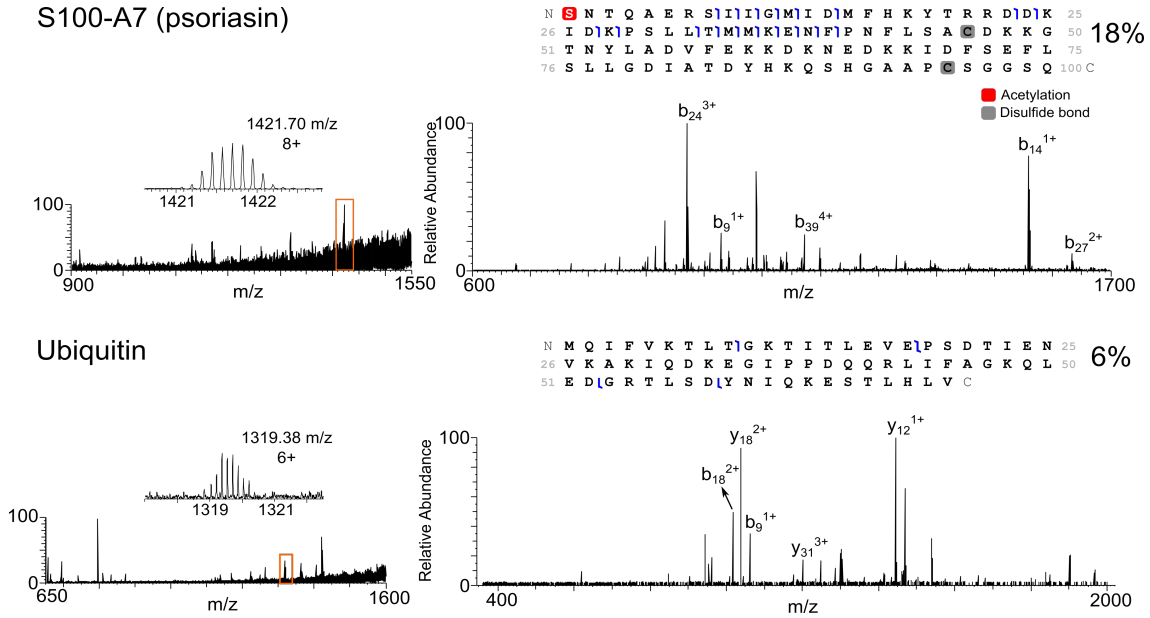


Figure 5.5: MS/MS mass spectra of human skin proteins identified from *ex vivo* human skin grafts. Ubiquitin was identified as a truncated protein, where the last six C-terminal amino acid residues are missing.

A different workflow involving chip-based high-field asymmetric waveform ion mobility spectrometry (planar FAIMS or p-FAIMS) device was employed to aid the removal of interfering signal from salts. It was shown previously, that FAIMS device coupled to LESA MS reduces the background noise and the signal-to-noise ratio (S/N) increases, while higher numbers of proteins were detected [94, 221]. LESA FAIMS MS was used here to investigate a subset of *ex vivo* human skin samples. A range of CF and DF values were applied, while only CF = 1.8 Td and DF = 200 Td resulted in detection of three different proteins (Fig. 5.6). Unfortunately, the IDs of the observed proteins could not be confirmed by the MS/MS analysis due to the low signal abundance of the protein precursor. A drop of signal intensity is commonly observed in FAIMS mass spectra [221]. The results show

that FAIMS could be a useful tool for analysis of proteins from human skin samples, while the application of the new cylindrical FAIMS Pro device may improve the S/N and allow for protein identification after MS/MS analysis. FAIMS Pro uses non-homogenous electric fields that gives rise to ion focusing and hence maintaining signal intensity [85, 222], while p-FAIMS devices use homogenous fields giving rise to narrow peak widths and optimal peak separation that are accompanied by dramatic loss of signal intensity [85, 223] (see also section 1.2.1).

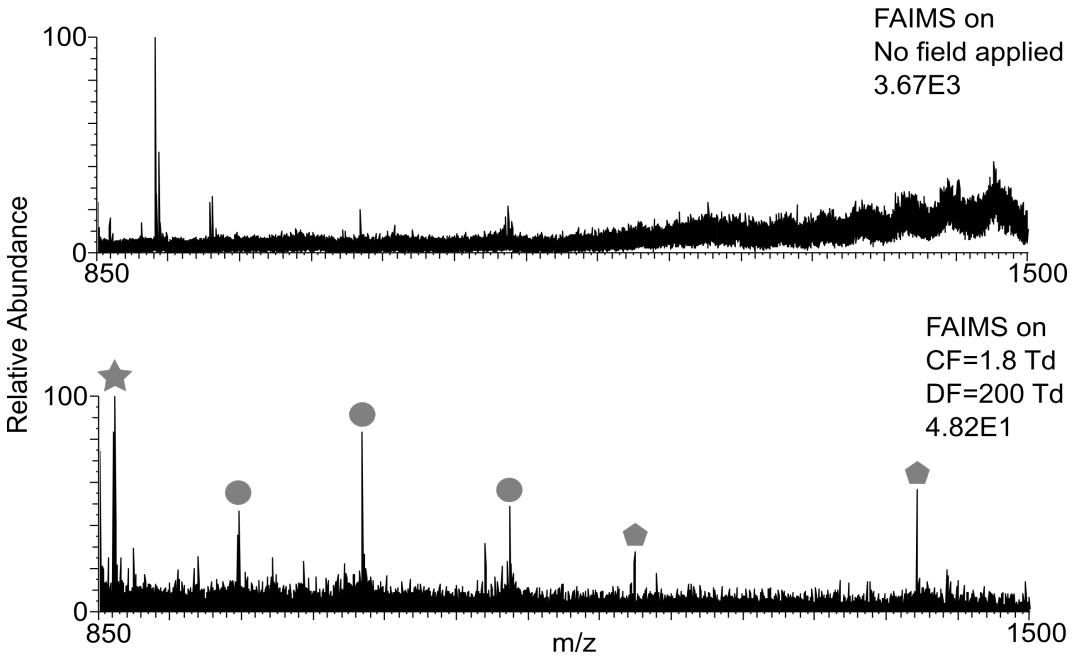


Figure 5.6: LESA FAIMS mass spectra of *ex vivo* human skin. Top – LESA mass spectrum with no FAIMS voltages applied, bottom – LESA mass spectrum acquired with the FAIMS field on (CF = 1.8 Td and DF = 200 Td).

5.3.3 LESA MS analysis of wounded *ex vivo* human skin samples and wounded samples after bacterial inoculation

The next step in the analysis of *ex vivo* human skin samples was their wounding and inoculation with bacterial species followed by LESA MS analysis. The human skin samples were wounded with a scalpel blade as presented in Fig. 5.1. First, only wounded samples with no inoculation were investigated by LESA MS. Unfortunately, the same issue as with the non-wounded human skin samples regarding presence of salts was observed. Difficulties were also experienced when sampling the wound itself. The whole volume of the dispensed extraction solvent system entered the wound and could not be re-aspirated back into the pipette tip. Moreover, the sample aspirated from the wounded site contained a lot of salts from the culturing medium which did not allow detection of proteins, or the S/N was very low (Fig. 5.7). A washing step was also considered, but if this method was to be used as a diagnostic approach for bacterial identification, microorganisms would be washed away from the sample and their identification would be prevented.

Second, inoculation with the Gram-positive bacterium *S. aureus* MSSA476 was carried out. The same infectious dose of 120 CFUs was used based on the inoculation protocol of Labskin samples (for more detail see Chapter 4). The samples were analysed after 24, 48, 72 and 96 hours of incubation. No protein peaks corresponding to bacteria were observed in the LESA mass spectra at any of the investigated time points, only protein peaks of psoriasin and ubiquitin were isolated and subjected to MS/MS (see Fig. 5.5). An example of a LESA mass spectrum of wounded and inoculated sample is presented in Fig. 5.7. The *ex vivo* human skin graft is a much more complex system than the 3D skin model. Real skin samples also contain cells which are involved in immune responses, therefore it is possible that the infectious dose was not high enough or longer incubation time would be necessary to develop a skin infection, however more experiments will need to be carried out to optimise the LESA extraction, infectious dose and incubation time of the inoculated samples.

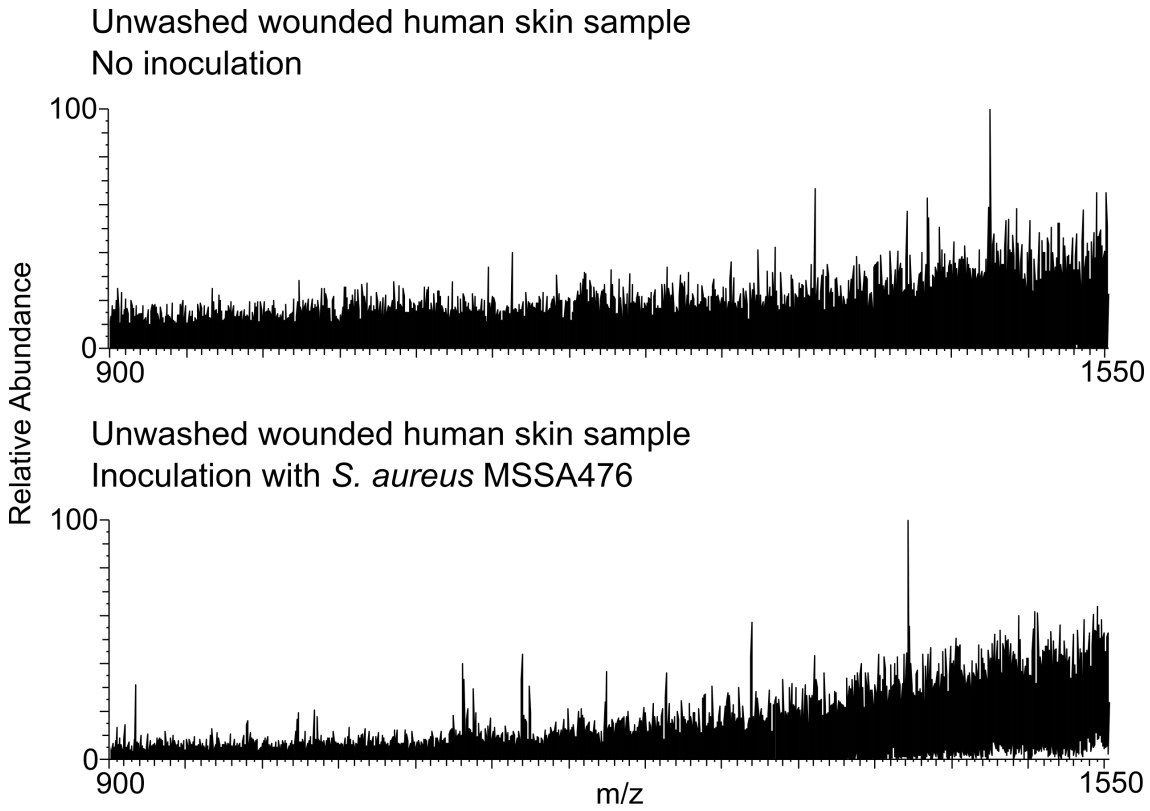


Figure 5.7: LESA mass spectra of unwashed human skin samples with no inoculation and inoculation with *S. aureus* MSSA476. No peaks corresponding to bacterial proteins were observed in the bottom mass spectrum.

5.4 Conclusion

In this Chapter, sample preparation and LESA MS analysis workflows for non-wounded, wounded and wounded and inoculated *ex vivo* human skin grafts were developed. Three solvent systems were tested and different washing steps were included to optimise the LESA extraction of proteins. Multiple peaks corresponding to proteins were detected in the mass spectra and seven of them were subjected to MS/MS analysis, resulting in identification of two human skin proteins psoriasin and ubiquitin, also previously observed in the LESA mass spectra of Labskin samples. This finding suggests that Labskin is mimicking real skin conditions and that the same proteins can be observed in both Labskin and human skin mass spectra. LESA MS was coupled to FAIMS to reduce the background noise. LESA FAIMS mass spectra allowed detection of three different proteins and improvement of the S/N, however the reduced signal intensity precluded their MS/MS analysis. Analysis of the wounded and inoculated human skin samples resulted in successful detection of human skin but no bacterial proteins in the mass spectra, suggesting that further optimisation of the culturing, inoculation and LESA sampling processes will be necessary.

Chapter 6

Top-down LESA MS analysis of microbes: Towards high-throughput protein identification

6.1 Background

In previous work, TD LESA MS was applied for direct identification of proteins from living bacterial colonies of ESKAPE pathogens (see also Chapter 3) [32, 33, 43]. The goal of this Chapter is to build on this approach and develop an automated method for rapid identification of microorganisms by TD LESA MS with use of state-of-art instrumentation. Numerous challenges had been observed while analysing bacterial cultures by LESA MS including interference from agar background and small molecules, presence of chimeric protein peaks and signal suppression due to a dominant protein peak. P-FAIMS was used in combination with LESA MS to remove (or at least reduce) the undesired signal in order to detect new proteins from various sample types [35, 92–94, 96, 221] and was also utilised for analysis of bacterial species *E. coli, P. aeruginosa* and *S. aureus* [93, 95]. The new workflow aims to significantly enhance protein signal intensity (by using new mass spectrometer), improve S/N and decrease false identification rate (FDR) value (both by using new c-FAIMS device). FDR is a measure of acceptance how many incorrect protein matches may be present in the dataset [131, 132], very important for the consistency of reported results, and has been implemented in TD proteomics searches [101, 133].

Until now, LESA MS protein identification was always performed manually, i.e.,

selection of the protein precursor, and the whole data acquisition process and protein ID assignment, were controlled by the person conducting the experiment. The manual identification process is successful, however it is laborious and may sometimes be subjective [118, 133]. If the future aim is to translate bacterial identification by LESA FAIMS MS into a diagnostic tool, manual protein identification must be replaced by an automated approach, therefore development of workflows offering higher throughput is of high importance.

In this Chapter, TD LESA MS was coupled to the new cylindrical FAIMS Pro device and the new state-of-art tribrid orbitrap Eclipse mass spectrometer. A high-throughput workflow, which employed data-dependent analysis for TD MS/MS, was optimised and applied to the analysis of bacterial proteins from two Gram-positive (*S. aureus, E. faecium*) and three Gram-negative (*E. coli, K. pneumoniae, P. aeruginosa*) species. The tandem mass spectra were subsequently searched against respective databases in the ProSightPD software. ProSightPC has been widely used in the TD proteomics community and is capable of high-throughput searches with the scoring system included [120, 128, 130, 224, 225], however with no FDR. ProSightPC was recently incorporated into the search nodes of the Proteome Discoverer software, named ProSightPD, with the facility for FDR determination included in the search. In addition, ProSightPD nodes are also suitable for high-throughput FAIMS analysis of proteins and identification of proteoforms [226].

The number of identified proteins from two workflows (with and without the FAIMS device coupled to the mass spectrometer) were compared. Analysis of FAIMS mass spectra showed improvements in protein p-scores and higher numbers of identified proteins and proteoforms, particularly for Gram-positive species. Furthermore, when less negative FAIMS CV voltages were applied, higher molecular weight (MW) proteins were observed and identified. Previous work described that the mobility of large proteins (mass >30 kDa) increases with increasing electric field [73, 227], agreeing with expected dipole alignment [228]. The higher MW proteins are the largest proteins identified directly from bacterial

colonies by denaturing TD LESA FAIMS MS to date.

6.2 Materials and methods

Materials are listed in section 2.1. Preparation of bacterial samples is described in section 2.2.

6.2.1 LESA MS analysis

LESA MS analysis was carried out by use of the Advion Triversa Nanomate (Advion, Ithaca, NY) coupled to a Thermo Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific, San Jose, CA). LESA extraction was performed as described in section 2.2.4. The mass spectra were acquired from different colonies for "no FAIMS" and "FAIMS" data, i.e., different locations of multiple colonies across the plate were sampled for each dataset. The sample was introduced into the mass spectrometer via chip-based nanoESI at gas pressure 0.3 psi and a tip voltage 1.65 kV. Mass spectra were acquired for 3 min for each species in full scan mode, positive ion mode, in normal mass range 600 – 2000 m/z at resolution 240000 at 200 m/z. The temperature of the ion transfer tube was set to 250 °C, injection time to 100 ms, RF lens value to 100%, AGC target to 200% (corresponding to 8×10^5 charges) and 1 scan comprised 1 microscan. Top-down MS/MS analysis was performed by using CID in the ion trap at 35% normalised collision energy. For manual MS/MS data collection, the mass spectra were acquired for 5 min and each scan comprised 30 co-added microscans.

An automated instrument acquisition "top 10" method was set up for data-dependent analysis. The instrument method was designed to collect data for 10 min in the intact protein mode at standard pressure (0.008 Torr) with the advanced peak determination option on. The instrument method comprised two steps: first, mass spectra were acquired for 0.5 min in the full scan mode and normal mass range 600 – 2000 m/z at resolution 120000 at 200 m/z. Maximum injection time was set to 500 ms, normalised AGC target

value to 200% (corresponding to 8×10^5 charges) and RF lens value to 100%. One scan comprised 3 co-added microscans. Second, survey scans were performed and followed by MS/MS of the 10 most abundant protein precursor peaks. The survey scans were acquired under the same conditions as the control mass spectra, except for the number of co-added microscans which was increased to 18. Quadrupole isolation of width 2 m/z was applied and CID was carried out in the ion trap by using helium gas at fixed 35% normalised collision energy (activation time 10 ms, Q = 0.25). MS/MS data were collected in the high mass range at 120000 resolution at 200 m/z, maximum injection time was set to 800 ms and normalised AGC target value to 250% (corresponding to 1.25×10^5 charges). Only precursor ions with positive charge states in range 3 – 25 and above intensity threshold 1×10^3 were subjected to MS/MS analysis. Dynamic exclusion was applied for 10 min such that only one charge state per precursor within 10 ppm mass tolerance was fragmented with exclusion of isotopes.

6.2.2 FAIMS analysis

The FAIMS Pro device (Thermo Fisher Scientific, San Jose, CA) was coupled to the Thermo Orbitrap Eclipse instrument. The dispersion voltage was -5 kV. Standard resolution setting for FAIMS was selected, the ion transfer tube temperature was set to 250°C with no additional gas flow. A FAIMS CV scan (also known as CV sweep) was performed for each bacterial species separately to determine optimum CV values. The CV range was -100 V to +30 V in step size 2 V per scan. Each scan comprised 1 microscan.

An automated instrument acquisition "top 10" method using the same settings as described above was set up for data-dependent analysis with FAIMS added to the workflow. CV voltages in the range -60 V to -10 V in 10 V steps were applied. For each CV voltage (six in total), a separate file was acquired hence creating 6 data files/microbe. FAIMS mass spectra recorded manually (for higher MW proteins) followed settings for manual protein identification described above (section 6.2.1), while two voltages -10 V and -20 V were applied.

6.2.3 Data processing

Protein databases for each bacterial species (PSCW format) were created from XML files downloaded from the UniProtKB database (www.uniprot.org) as described in section 2.2.5.

Data were analysed in Proteome Discoverer 2.4 software with the addition of ProSightPD 3.0 nodes for data processing. The processing workflow "Comprehensive Discovery Proteomics with FDR for HI HI data" was selected together with the consensus workflow "1% FDR Consensus". Here, only changes to the processing workflow template are listed, the remainder of the settings were left as default and no changes were made to the consensus workflow.

Parameters for the ProSightPD high/high cRAWler node were changed to following: fragmentation Xtract parameter S/N Threshold was set to 2 and Highest m/z to 3000 Da. For the ProSightPD Annotated Proteoform node, precursor and fragment tolerances were changed to 15 ppm and ΔM mode to true. In the ProSightPD Subsequence search node the subsequence precursor and fragment tolerances were both set to 15 ppm and ΔM mode to true. Fragment tolerance was set to 15 ppm for the ProSightPD Annotated Proteoform wider search node.

For protein ID searches of the higher MW proteins, the data processing workflow described in section 2.2.5 was followed.

6.3 Results and discussion

6.3.1 General overview of the TD LESA FAIMS MS workflow development

In the previous work, performed on an Orbitrap Elite, the manual selection of protein precursors for MS/MS analysis was followed by their manual identification in the ProSightPC search software. Analysis of *E. coli* K12 on the Orbitrap Eclipse revealed that

there were many previously undetected protein peaks present in the mass spectra (Fig. 6.1) when compared with data acquired on the Elite mass spectrometer. It was quickly apparent that the "traditional" manual precursor selection and protein ID assignment would be very labour intensive and inefficient. Moreover, with the application of FAIMS, the number of detected proteins was expected to be the same or increased. It was therefore required to develop a workflow for automated protein precursor selection, its subsequent MS/MS analysis and protein ID assignment.

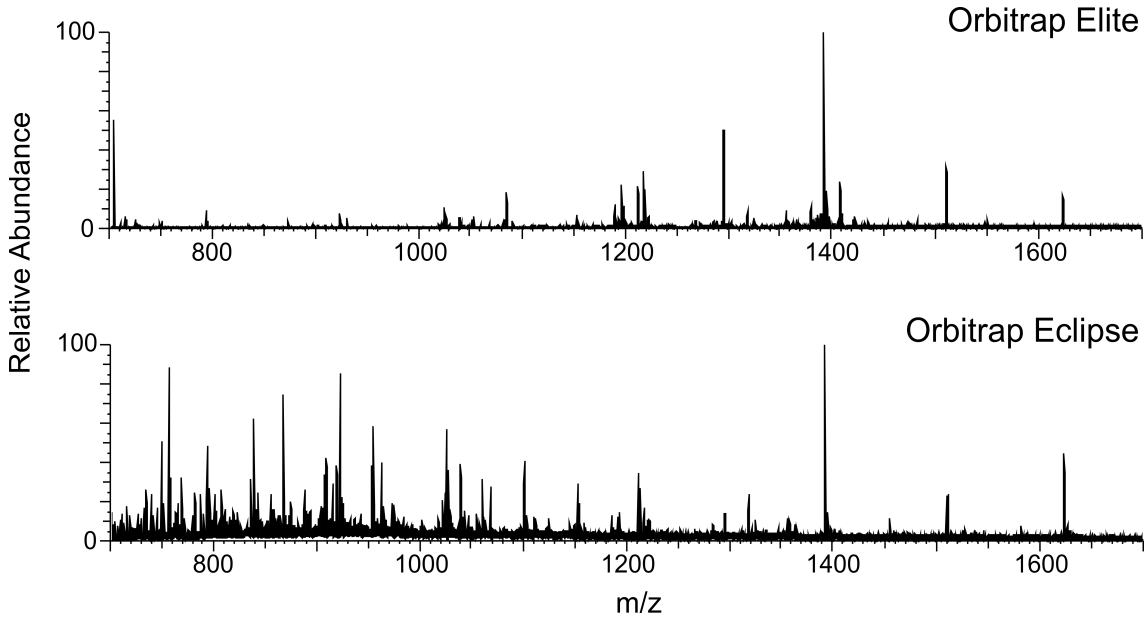


Figure 6.1: A comparison of LESA mass spectra of *E. coli* K12 acquired on the Orbitrap Elite (top) and Orbitrap Eclipse (bottom) mass spectrometers. Many abundant protein peaks emerged in the lower m/z range of the bottom spectrum.

A general overview of the workflow is presented in Fig. 6.2. The aim was to identify as many proteins as possible in a relatively short time using one or few (in the case of FAIMS analysis) LESA extractions from bacterial colonies. The new workflow consisted of two steps: first, a control MS1 mass spectrum was recorded and, second, data-dependent analysis (DDA) was performed where CID was used as a fragmentation technique. A "top 10" DDA method was created, i.e., the ten most abundant peaks from the mass spectrum were selected for MS/MS analysis. Once fragmented, precursor m/z values were added to an exclusion list, so that another set of protein ions could be selected for MS/MS

analysis. The method duration was 10 minutes; longer acquisition times were not suitable mostly due to the bacterial sample clogging the Nanomate chip nozzles resulting in a sudden nanoESI ion current drop. Both LESA and LESA FAIMS data were acquired, with FAIMS mass spectra recorded for each applied CV separately. Lastly, the MS data were searched against proteome databases by use of the ProSightPD software.

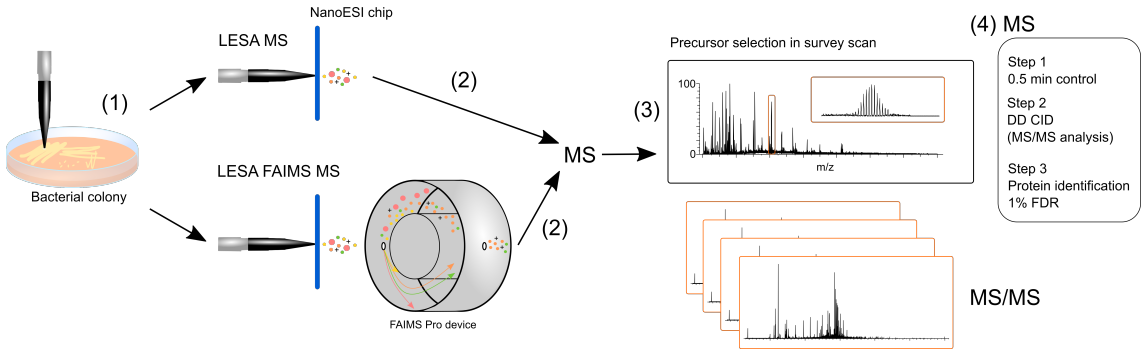


Figure 6.2: General workflow of the high-throughput LESA and LESA FAIMS MS method. Proteins are extracted from bacterial colonies by LESA and the sample is introduced into MS *via* chip-based nanoESI (1) with and without the FAIMS device attached (2). In both methods, protein precursor is selected for tandem MS (3) and MS/MS data is acquired and analysed in the three steps described (4).

Initial experiments were performed on the Gram-negative *E. coli* K12, followed by the Gram-positive *S. aureus* MSSA476. First, the bacterial colonies were subjected to LESA MS/MS with no FAIMS device attached to the mass spectrometer (see Fig. 6.3 and Fig. 6.4). Next, the colonies were analysed by LESA FAIMS MS/MS. A scan of CVs (CV sweep) was performed to determine the optimal CV values for detection of proteins. The CV scan was performed each time a new microbial species was introduced for FAIMS analysis. Highly abundant ions were present for all bacteria investigated between CV values -60 V and 0 V (Fig. 6.5A), thus the optimal range was selected to be from -60 V to -10 V in 10 V steps. Positive CV values were not investigated as the signal, even though seemingly abundant, did not contain any resolved peaks corresponding to proteins (see an example mass spectrum in Fig. 6.5B). For LESA FAIMS MS/MS analysis, a separate datafile was acquired for each CV, creating six files in total which were subsequently combined for database searches and protein ID assignments (Fig. 6.3 and Fig. 6.4). Most

of the high intensity signal was observed at CVs -40 V and -30 V, while the less negative CV voltages -20 V and -10 V were found to be suitable for identification of rarely observed higher MW proteins (see text below).

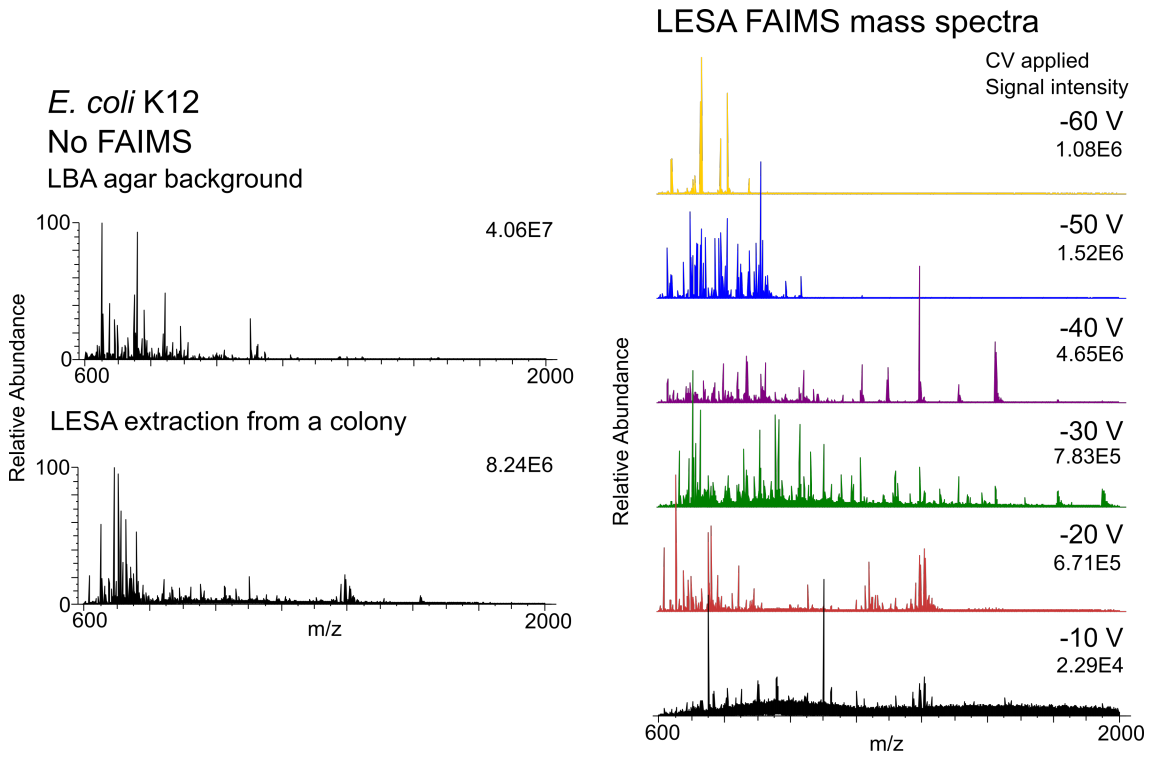


Figure 6.3: Left (top): LESA mass spectrum of LBA agar background. Left (below): LESA mass spectrum of *E. coli* K12 colony with no FAIMS attached to the mass spectrometer. Right: LESA FAIMS mass spectra across the optimised range of CV voltages -60 and -10 V with corresponding signal intensities.

The protein identification results for *E. coli* K12 and *S. aureus* MSSA476 are presented in Fig. 6.6 and in Appendix available online at DOI: 10.25500/edata.bham.00000755. Numbers of protein IDs obtained by manually selected precursors from LESA mass spectra, high-throughput LESA and high-throughput LESA FAIMS workflows were compared for both species (Fig. 6.6A). The protein IDs corresponding to the manual identification for *E. coli* K12 and *S. aureus* MSSA476 were published previously in [32, 43, 95] and all of the mass spectra were acquired on the Orbitrap Elite. Here, the results show that the highest numbers of identified proteins – 48 (*E. coli*) and 29 (*S. aureus*) – were achieved with the high-throughput LESA FAIMS method while the data were acquired on the Orbitrap Eclipse. A similar trend was observed for the number

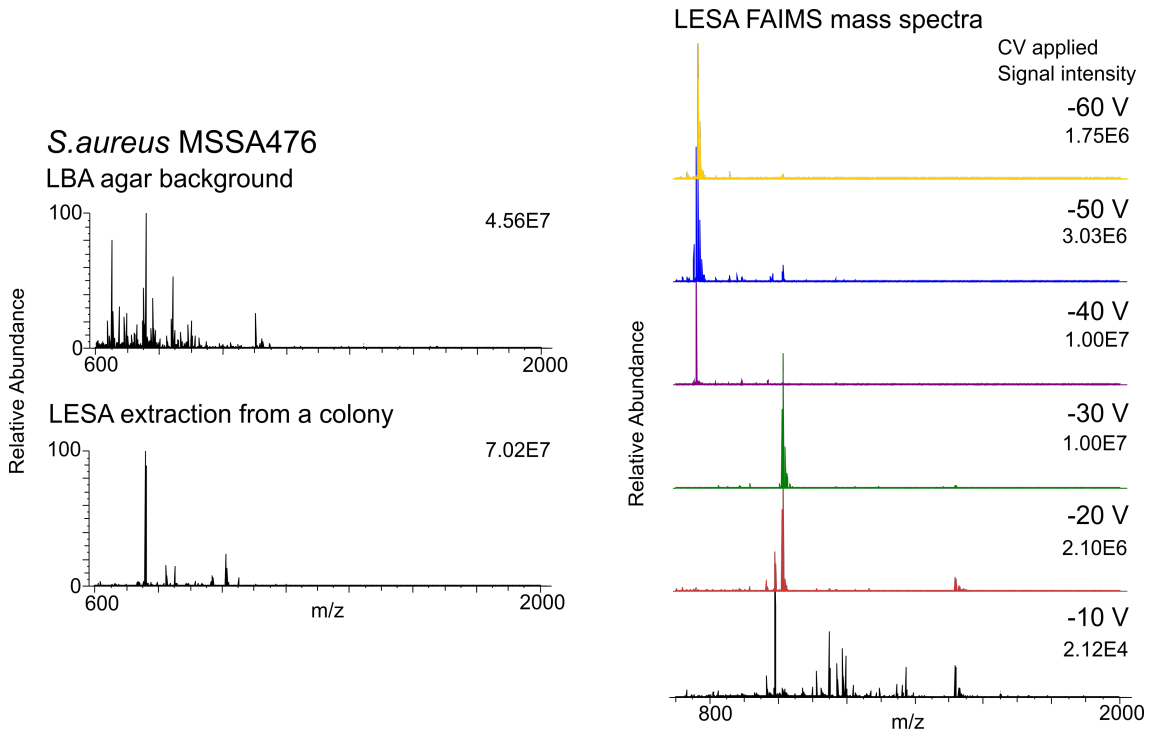


Figure 6.4: Left (top): LESA mass spectrum of LBA agar background. Left (below): LESA mass spectrum of *S. aureus* MSSA476 colony with no FAIMS attached to the mass spectrometer. Right: LESA FAIMS mass spectra across the optimised range of CV voltages -60 and -10 V with corresponding signal intensities.

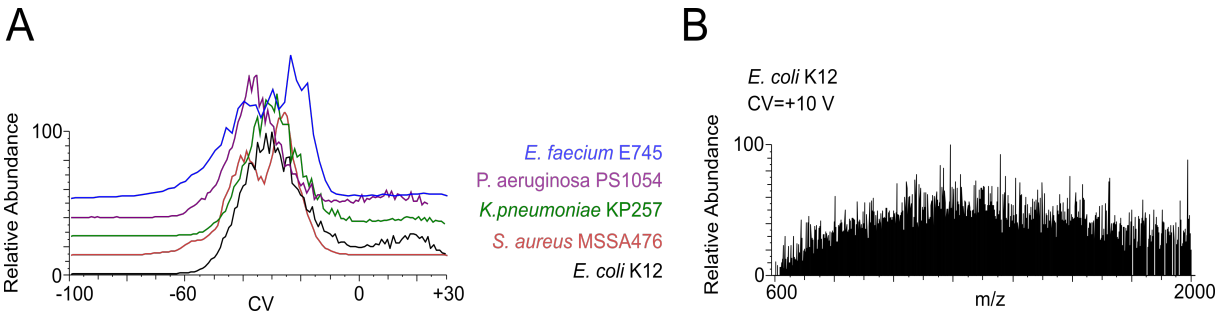


Figure 6.5: (A) CV traces of all investigated bacterial species. The most abundant signal corresponding to proteins was observed between -60 V and 0 V. (B) An example of a mass spectrum of *E. coli* K12 acquired at CV +10 V; no resolved protein peaks were observed at positive CV values.

of identified proteoforms, where only high-throughput data were compared – information about the number of proteoforms is unavailable for the manual protein ID assignment. ID search of *E. coli* K12 resulted in 85 proteoform IDs and 53 proteoforms were assigned in the *S. aureus* MSSA476 search (Fig. 6.6B). The study of Dupre *et al.* [229] used LC-MS/MS (without the FAIMS device incorporated in the workflow) for TD analysis of bacterial

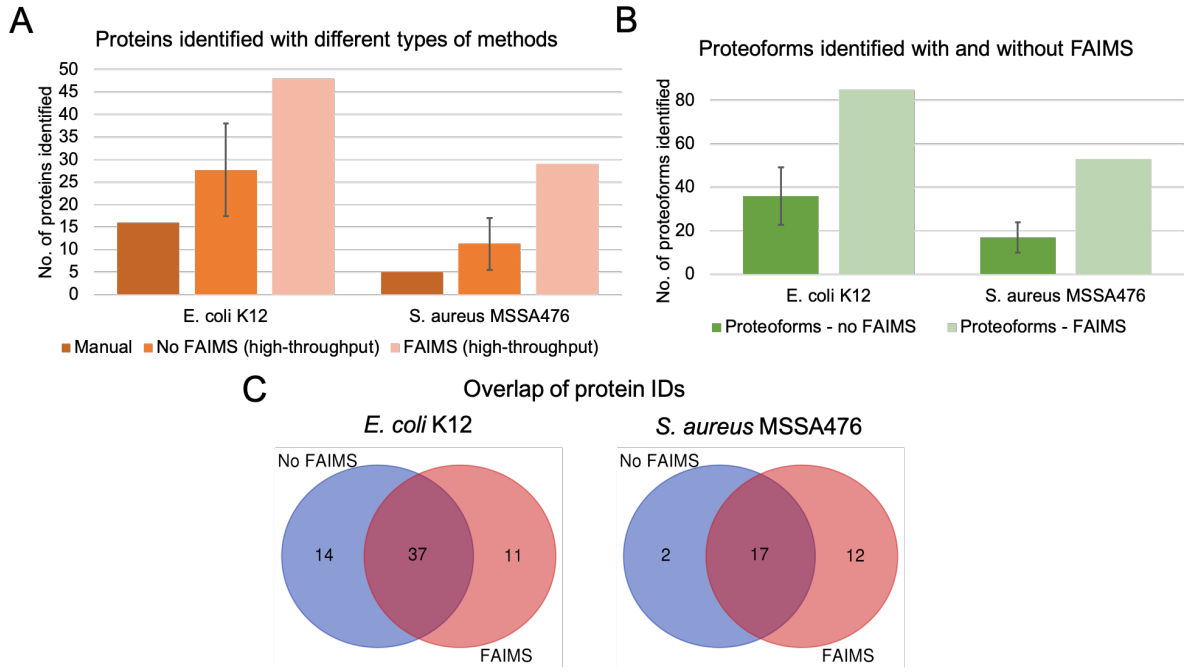


Figure 6.6: (A) Comparison of number of identified proteins from manual protein precursor selection, high-throughput LESA and high-throughput LESA FAIMS methods. *E. coli* K12: 16 (n=1), 28 ± 10 (n=3), 48 (n=1) and *S. aureus* MSSA476: 5 (n=1), 11 ± 6 (n=3) and 29 (n=1), respectively. The error bars represent standard deviation. (B) Comparison of the number of proteoforms identified without and with the FAIMS device attached. *E. coli* K12: 36 ± 13 (n=3) and 85 (n=1), *S. aureus* MSSA476: 17 ± 7 (n=3) and 53 (n=1), respectively. Only high-throughput methods were compared. (C) Overlap between the protein assignments for "no FAIMS" and FAIMS data. The number of proteins identified with "no FAIMS" was obtained from an ID search of all three "no FAIMS" replicates.

proteins extracted from *E. coli* K12 with eight different lysis buffers. That analysis resulted in identification of 242 – 328 proteins and 720 – 1386 proteoforms, depending on the extraction buffer used. This number may be even higher if the FAIMS is employed – it should be however emphasised that LESA FAIMS MS identifies proteins from a single colony extraction, while the LC-MS/MS approach analysed lysates containing approximately 10^9 cells/mL.

The overlap between protein assignments from "no FAIMS" and FAIMS approaches was investigated (Fig. 6.6C). In the case of "no FAIMS" data, all three files (replicates acquired in a single day) were searched together to obtain an overall number of identified proteins. Due to the nature of sampling, the protein IDs differ to some extent between the

replicates, therefore the combined search yields higher number of protein IDs compared to each individual replicate search. The results show that the numbers of unique proteins for *E. coli* K12 "no FAIMS" and FAIMS are similar – 14 and 11, respectively (Fig. 6.6C).

The FAIMS analysis of *S. aureus* MSSA476 resulted in identification of 12 unique proteins compared to only two with no FAIMS attached (Fig. 6.6C). LESA mass spectra of *S. aureus* MSSA476 always contain a very dominant peak – δ-hemolysin, also visible across mass spectra in Fig. 6.4. Even though the signal for the dominant peptide was present throughout the CV sweep, application of FAIMS CV voltages allowed "new" proteins to be identified from the mass spectra. The fact that FAIMS reduces, but does not remove δ-hemolysin was observed in the previous study [95]. It was suggested that the efficiency of cell lysis may limit protein identification [32, 95]; however, the mass spectrometry instrumentation itself (Orbitrap Elite) presented a limiting factor. Here, 29 proteins (including δ-hemolysin) were identified from the LESA mass spectra acquired with the new instrumentation (Orbitrap Eclipse) compared to five proteins from manual assignments [32], representing a major improvement in the LESA FAIMS analysis which is in contrast with the previous FAIMS experiment where no previously undetected proteins were identified [95].

6.3.2 TD LESA FAIMS MS of Gram-negative *Klebsiella pneumoniae* KP257, *Pseudomonas aeruginosa* PS1054 and Gram-positive *Enterococcus feacium* E745

After optimisation of FAIMS parameters, Gram-negative *K. pneumoniae* KP257, *P. aeruginosa* PS1054 and a Gram-positive *E. faecium* E745 were subjected to analysis by the LESA MS and LESA FAIMS MS workflows. Representative mass spectra are shown in Fig. 6.7 – Fig. 6.9. Again, a comparison of the three approaches was carried out (Fig. 6.10A). Six, nineteen and five proteins from *K. pneumoniae* KP257, *P. aeruginosa* PS1054 and *E. faecium* E745, respectively, were identified manually in the previous work [33,164].

The number of protein IDs assigned to each species was dramatically improved by the high-throughput LESA MS approach: 21 ± 5 (*K. pneumoniae*), 22 ± 9 (*P. aeruginosa*) and 13 ± 5 (*E. faecium*) proteins were identified without and 34, 49 and 19 with the FAIMS device, respectively. In terms of proteoforms, higher numbers were identified from the FAIMS mass spectra (Fig. 6.10B). Comparison of the protein ID assignments in Fig. 6.10C confirms the previous observations – some identifications are unique to the "no FAIMS" and FAIMS analysis, while the majority of the identified proteins overlap. It was also noticed that the LESA FAIMS MS improved the search outcomes, specifically p-scores that are associated with the confidence of the protein match. A large number of the identified proteins (for both approaches) comprised ribosomal proteins and uncharacterised proteins, typical for LESA MS analysis. All files containing protein identifications are available in Appendix at DOI: 10.25500/edata.bham.00000755.

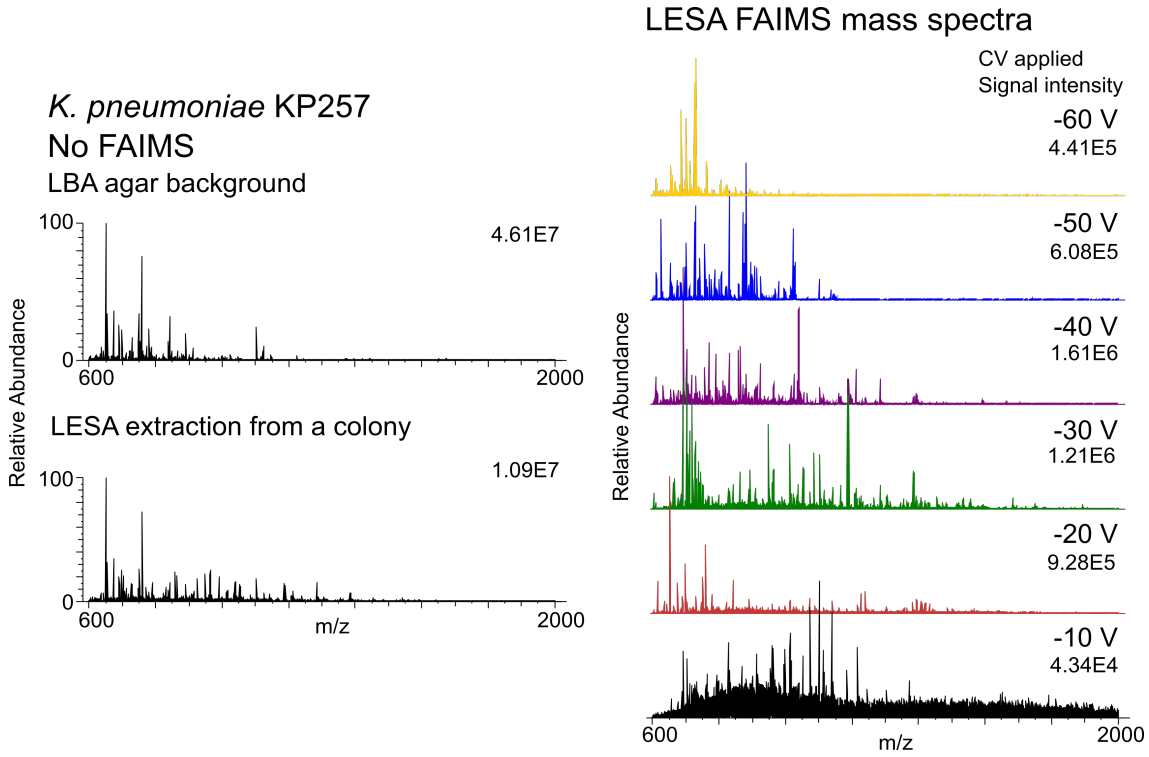


Figure 6.7: Left (top): LESA mass spectrum of LBA agar background. Left (below): LESA mass spectrum of *K. pneumoniae* KP257 colony with no FAIMS attached to the mass spectrometer. Right: LESA FAIMS mass spectra across the optimised range of CV voltages -60 and -10 V with corresponding signal intensities.

Two biological replicates were investigated for *K. pneumoniae* KP257, *P. aeruginosa*

PS1054 and *E. faecium* E745 (Fig. 6.11). The protein IDs obtained from the biological replicates were also compared (see Fig. 6.12). The differences are relatively high, which may also be related to the inherent variability of LESA extractions and the natural variability of the biological samples such as colony thickness and height; however, more biological replicates would be required to investigate the inter-day variability.

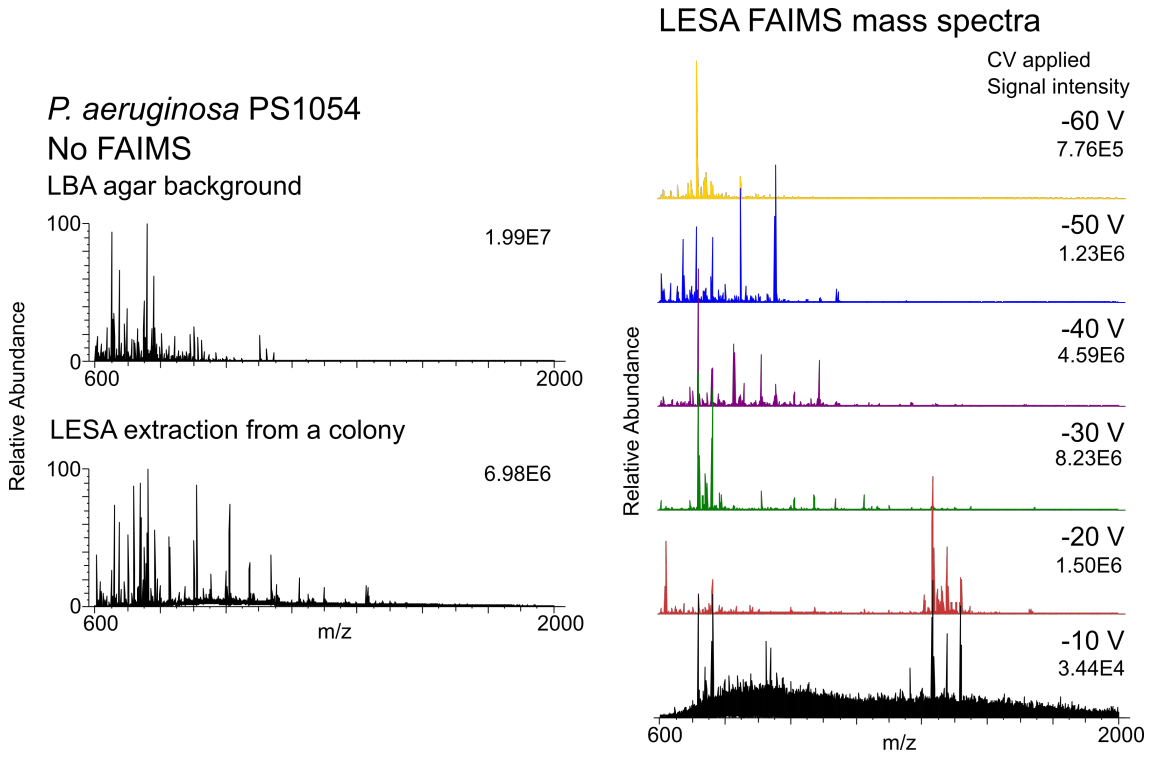


Figure 6.8: Left (top): LESA mass spectrum of LBA agar background. Left (below): LESA mass spectrum of *P. aeruginosa* PS1054 colony with no FAIMS attached to the mass spectrometer. Right: LESA FAIMS mass spectra across the optimised range of CV voltages -60 and -10 V with corresponding signal intensities.

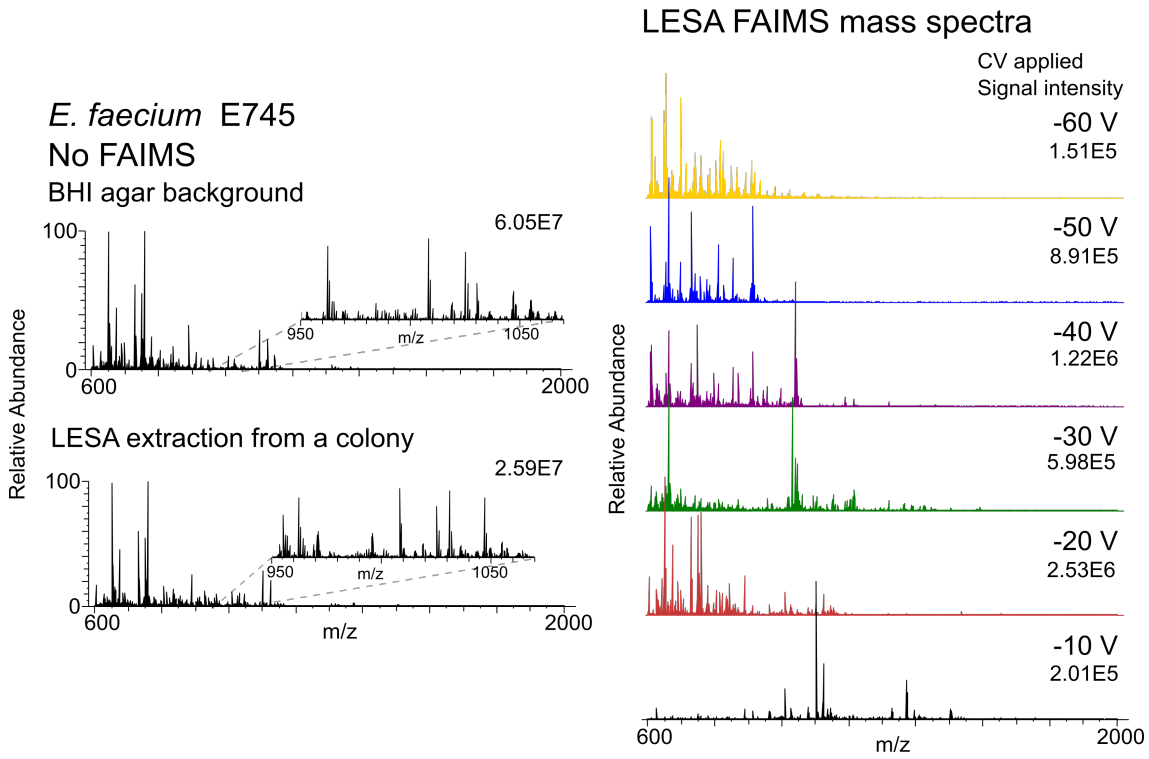


Figure 6.9: Left (top): LESA mass spectrum of BHI agar background. Left (below): LESA mass spectrum of *E. faecium* E745 colony with no FAIMS attached to the mass spectrometer. Right: LESA FAIMS mass spectra across the optimised range of CV voltages -60 and -10 V with corresponding signal intensities.

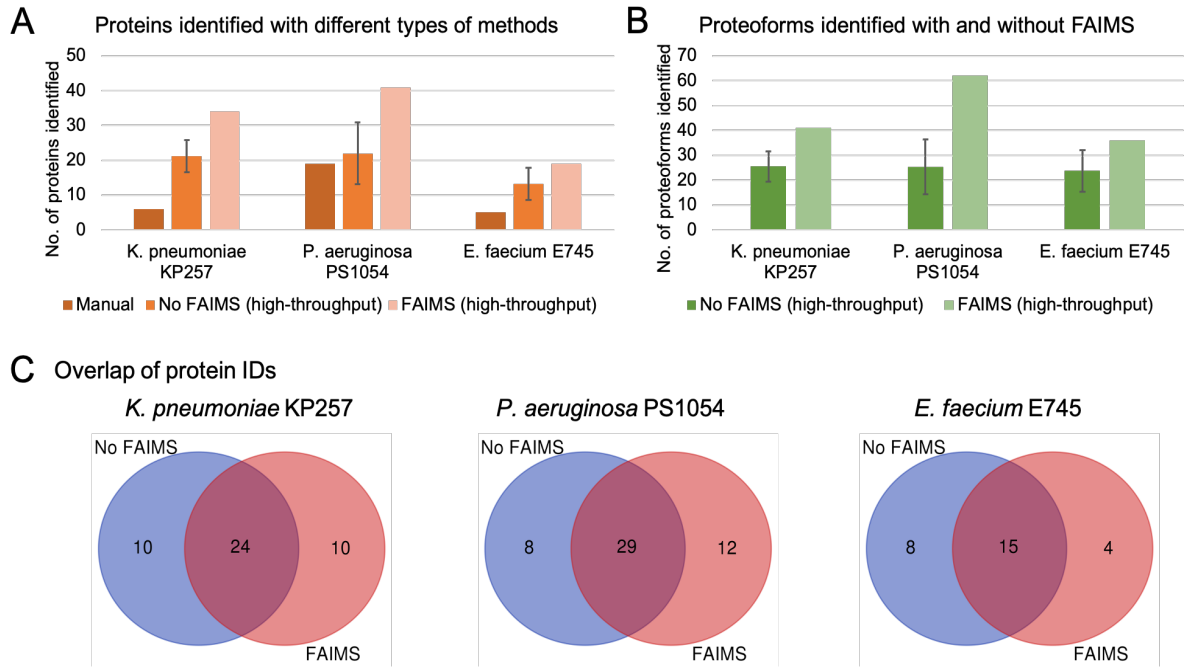


Figure 6.10: (A) Comparison of number of identified proteins from manual protein precursor selection, high-throughput LESA and high-throughput LESA FAIMS methods. *K. pneumoniae* KP257: 6 (n=1), 21 ± 5 (n=4), 34 (n=1), *P. aeruginosa* PS1054: 19 (n=1), 22 ± 9 (n=3) and 49 (n=1) and *E. faecium* E745: 5 (n=1), 13 ± 5 (n=4), 19 (n=1), respectively. The error bars represent one standard deviation. (B) Comparison of the number of proteoforms identified without and with the FAIMS device attached. *K. pneumoniae* KP257: 26 ± 6 (n=4) and 41 (n=1), *P. aeruginosa* PS1054: 25 ± 11 (n=3) and 62 (n=1) and *E. faecium* E745: 24 ± 8 (n=4) and 36 (n=1) proteins, respectively. Only high-throughput methods were compared. (C) Overlap between the protein assignments for "no FAIMS" and FAIMS data. The number of proteins identified with "no FAIMS" was obtained from an ID search of all "no FAIMS" replicates.

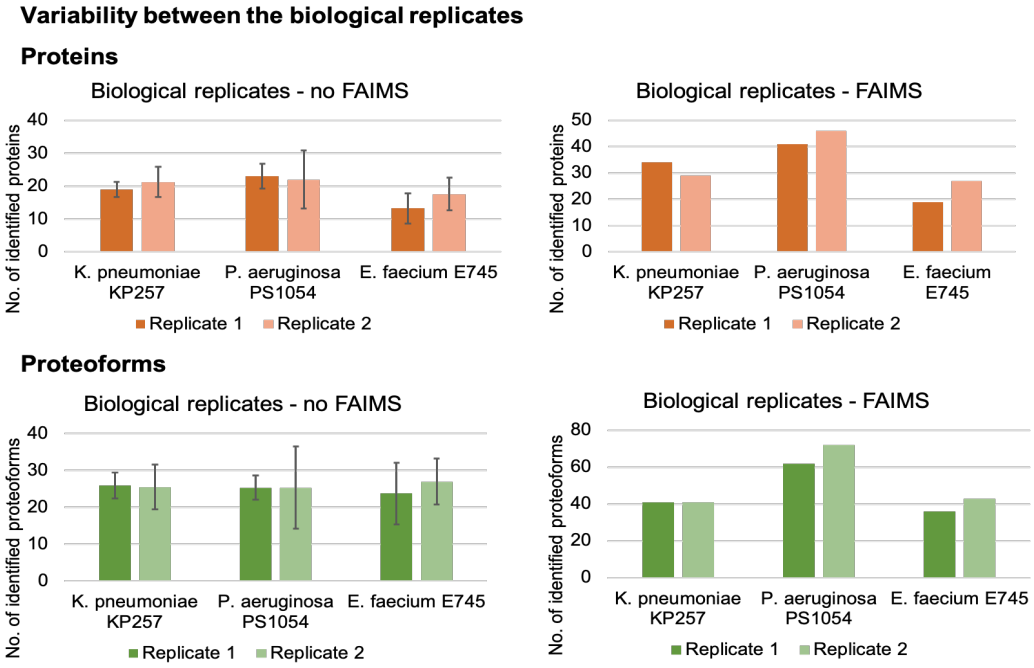


Figure 6.11: Comparison of variability between two biological replicates between proteins (top) and proteoforms (bottom) identified with and without the FAIMS device. Error bars represent one standard deviation. Proteins – "no FAIMS" (replicate 1, replicate 2) and FAIMS (replicate 1, replicate 2). *K. pneumoniae* KP257: 19 ± 2 (n=4), 21 ± 5 (n=4) and 34 (n=1), 29 (n=1); *P. aeruginosa*PS1054: 23 ± 4 (n=3), 22 ± 9 (n=3) and 41 (n=1), 46 (n=1); *E. faecium* E745: 13 ± 5 (n=4), 18 ± 5 (n=5) and 19 (n=1), 27 (n=1). Proteoforms – "no FAIMS" (replicate 1, replicate 2) and FAIMS (replicate 1, replicate 2). *K. pneumoniae* KP257: 26 ± 4 (n=4), 26 ± 6 (n=4) and 41 (n=1), 41 (n=1); *P. aeruginosa*PS1054: 25 ± 3 (n=3), 25 ± 11 (n=3) and 62 (n=1), 72 (n=1); *E. faecium* E745: 24 ± 8 (n=4), 27 ± 6 (n=5) and 36 (n=1), 43 (n=1).

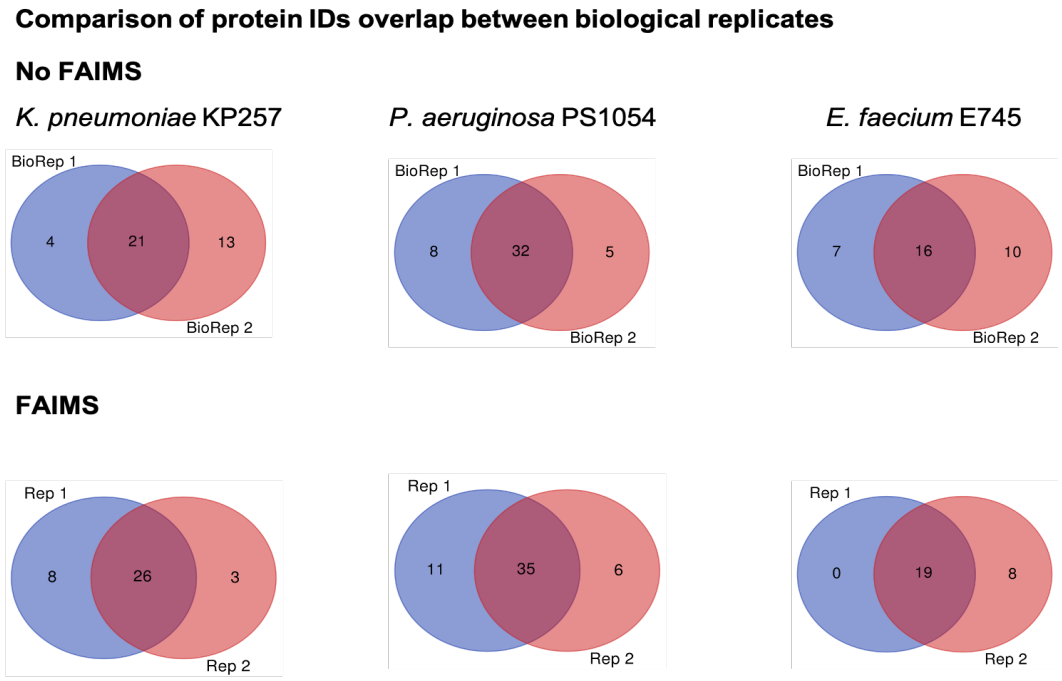


Figure 6.12: Overlap of the protein IDs between the biological replicates. The inter-day variability is observed between both "no FAIMS" and FAIMS data.

6.3.3 Identification of high-molecular weight proteins by TD LESA FAIMS MS

Application of less negative CV voltages resulted in detection of higher MW proteins in the LESA FAIMS mass spectra. Specifically, CV voltages -10 V and -20 V allowed detection of six proteins (Fig. 6.13 – 6.17, Table 6.1), not previously observed in the LESA and LESA FAIMS mass spectra, of which three were of MW above 30 kDa (see proteins in bold in Fig. 6.13 – 6.15 and Table 6.1). None of the six proteins were identified by the data-dependent "top 10" analysis workflow – the low abundance of the precursor ions during the survey scan was insufficient for selection for MS/MS; longer acquisition times (than those of survey scans) were necessary to observe peaks of higher MW proteins in the mass spectra. Generally, MS/MS data of the proteins identified below also needed to be collected for a longer amount of time (see section 6.2.1) than in the optimised top 10 method workflow for the successful ID assignment. All MS/MS mass spectra and ID assignments are listed in Appendix B.4.

The first protein of MW>30 kDa was identified from the *E. coli* K12 mass spectra as the outer membrane porin C (OmpC) (Table 6.1). According to the UniProt, this protein is capable of forming pores, hence allowing diffusion of small molecules across the membrane. OmpC exists as a homotrimer or a heterotrimer in its native conditions, but it was detected and identified as a monomer by the TD LESA FAIMS MS (Fig. 6.13). Outer membrane porin A (OmpA) was identified from *K. pneumoniae* KP257 (Table 6.1). OmpA is, according to the UniProt, a protein inferred from homology required for conjugation with F-type plasmids, while existing as a monomer or homodimer. Again, OmpA was detected in this experiment as a monomer. Both OmpC and OmpA were observed at FAIMS CV voltage -10 V. Unfortunately, the sequence coverage of OmpA is rather low (2%) (see Fig. 6.14), suggesting that further MS/MS analysis will be required. The last protein with MW above 30 kDa was detected in the mass spectra of *P. aeruginosa* PS1054 at CV = -20 V. This protein was identified as the probable binding

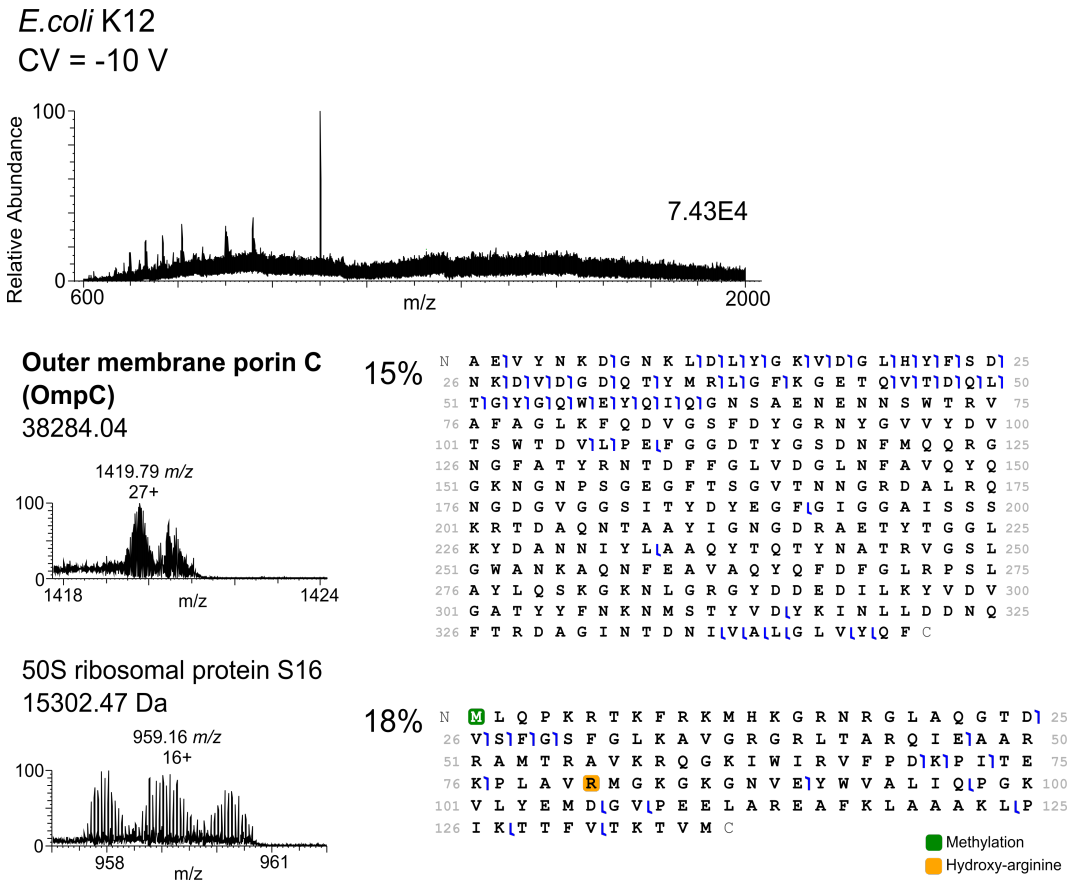


Figure 6.13: Proteins identified from *E. coli* K12 at FAIMS CV -10 V. Top: full scan LESA FAIMS mass spectrum. Left: isolated precursor peaks. Right: protein sequences, sequence coverages and detected PTMs.

protein component of ABC transporter or D-amino acid transporter (Table 6.1). Based on the UniProt database, the D-amino acid transporter is a predicted protein present in the extracellular space containing a signal peptide (1–23). The MS/MS analysis revealed presence of a disulfide bond, observed as a mass shift of 2.01 Da and missing fragments between the two cysteines (see Fig. 6.15), not yet reported on the UniProt website.

Other proteins with lower MWs, and previously undetected in the LESA mass spectra, nor identified by the top 10 method, were subjected to tandem MS analysis. The MWs of these proteins were found to be in the range 15–30 kDa and were detected in the LESA FAIMS mass spectra when CVs -10 V and -20 V were applied. For *E. coli* K12, two protein peaks were selected for MS/MS – 959.1000 (16+) and 1101.3500. The first precursor was identified as 50S ribosomal protein L16 (see Table 6.1 and Appendix B.4).

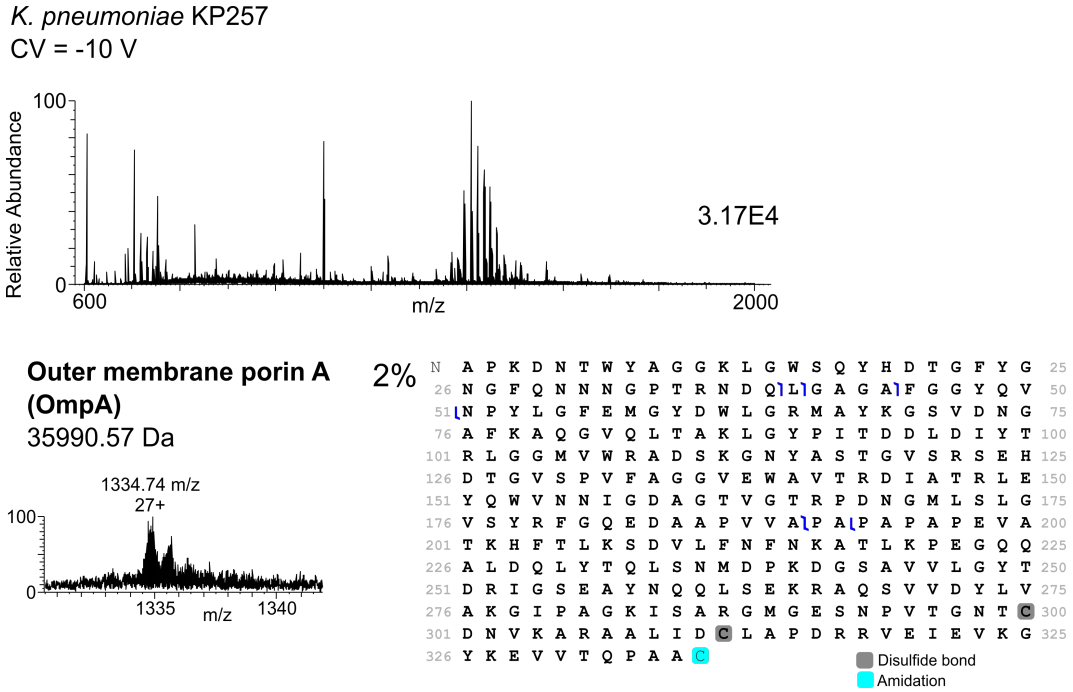


Figure 6.14: Protein identified from *K. pneumoniae* KP257 at FAIMS CV -10 V. Top: full scan LESA FAIMS mass spectrum. Left: isolated precursor peaks. Right: protein sequence, sequence coverage and detected PTMs.

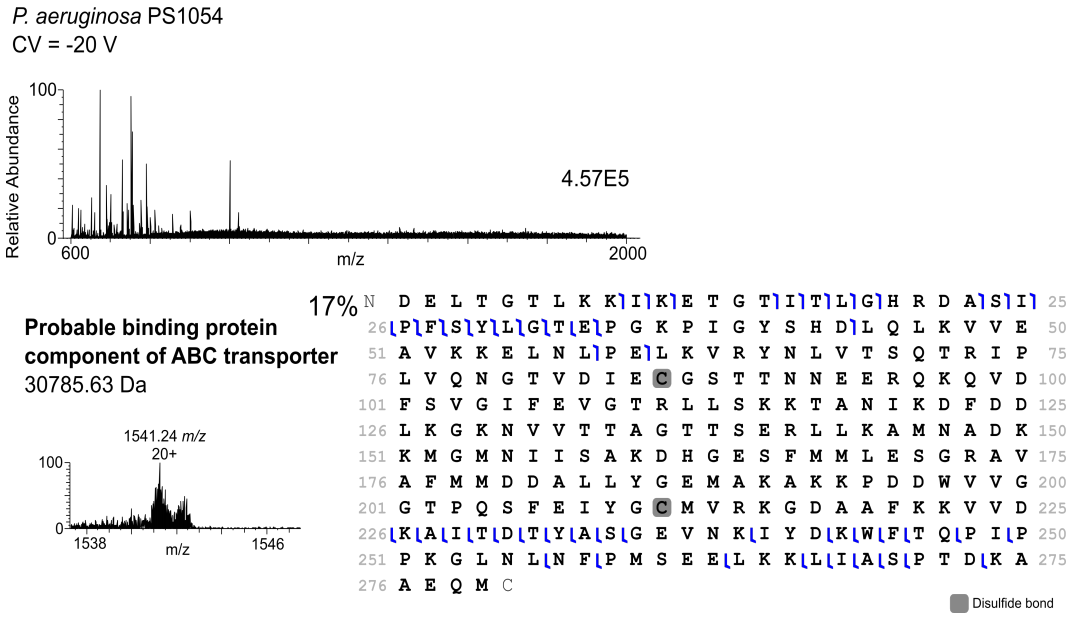


Figure 6.15: Protein identified from *P. aeruginosa* PS1054 at FAIMS CV -20 V. Top: full scan LESA FAIMS mass spectrum. Left: isolated precursor peak. Right: protein sequence, sequence coverage and detected PTMs.

This ribosomal protein was not observed previously in the LESA nor LESA FAIMS mass spectra. Interestingly, the second precursor (1101.3500) was detected with charge state

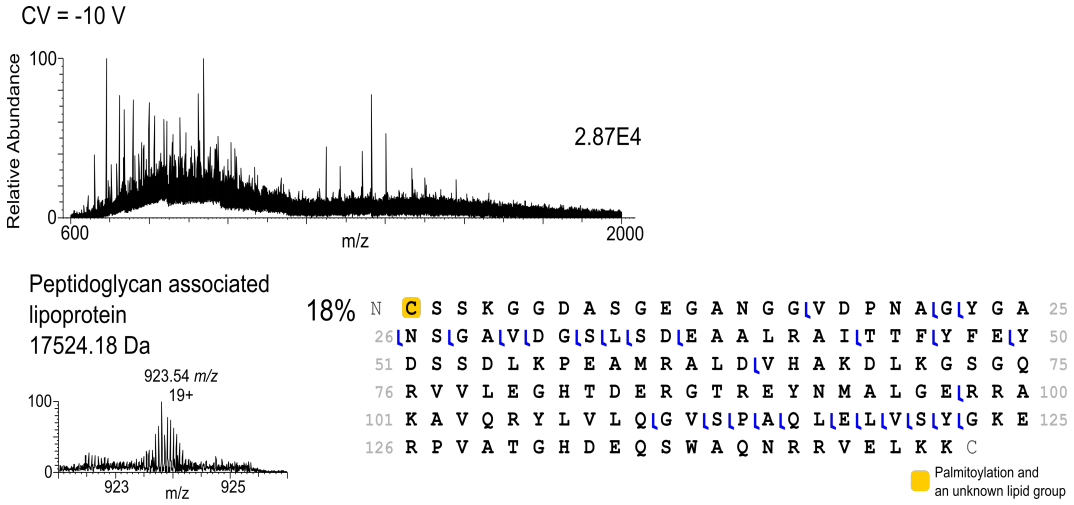


Figure 6.16: Protein identified from *P. aeruginosa* PS1054 at FAIMS CV -10 V. Top: full scan LESA FAIMS mass spectrum. Left: isolated precursor peak. Right: protein sequence, sequence coverage and detected PTM.

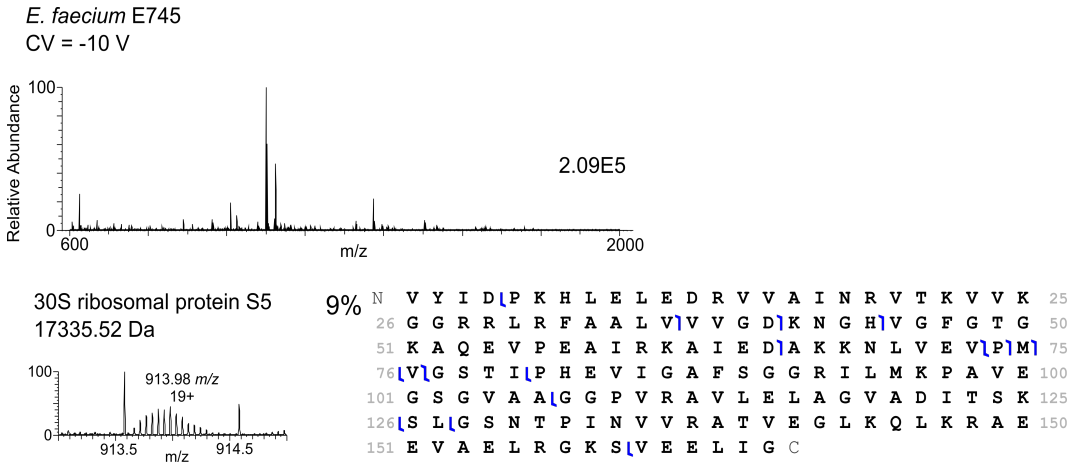


Figure 6.17: Protein identified from *E. faecium* E745 at FAIMS CV -10 V. Top: full scan LESA FAIMS mass spectrum. Left: isolated precursor peaks. Right: protein sequence and sequence coverage.

1+, but when subjected to MS/MS analysis, peaks with charge states 23+ emerged in the MS^2 mass spectra. MS^3 analysis of these peaks was performed, however the signal abundance of the precursors was very low (1×10^2) yielding only very few low abundant fragments, resulting in no ID assignment.

Another protein was identified from colonies of *P. aeruginosa* PS1054 when CV voltage -10 V was applied – peptidoglycan associated lipoprotein (Table 6.1). It is reported, that the N-terminal cysteine of peptidoglycan associated lipoprotein is modified by

attachment of N-palmitoyl and S-diacylglycerol groups. Despite the available information on PTMs, the mass shift detected was higher than the theoretical protein mass (with PTMs included). Further experiments will be necessary to determine the type of lipid group attached to the N-terminus. Lipoproteins were observed by LESA MS for the first time in Chapter 3 and, similarly, a mass shift was detected without information about the lipid group attached to the N-terminus. Nevertheless, the high number of fragments assigned with high confidence confirm the protein presence.

In the LESA FAIMS mass spectra of Gram-positive *E. faecium* E745, 30S ribosomal protein S5 was identified at CV -10 V (Table 6.1). This protein was not observed in the previous high-throughput ID assignments and it is, so far, the highest MW protein observed by denatured TD LESA FAIMS MS analysis of *E. faecium* E745.

Table 6.1: Summary of newly identified proteins at low FAIMS CVs -10 V and -20 V.

Protein name*	**Observed monoisotopic mass [Da]**	**Theoretical monoisotopic mass [Da]**	**Mass difference [Da]**	**UniProt accession no.**	**CV applied [V]**
E. coli K12					
Outer membrane porin C	38284.04	38284.12	-1.88	P06996	-10
50S ribosomal protein L16	15302.43	15302.47	-0.04	P0ADY7	-10
K. pneumoniae KP257					
Outer membrane porin A	35990.57	35990.77	-0.20	A6T751	-10
P. aeruginosa PS1054					
Probable binding protein component of ABC transporter (D-amino acid transporter)	30785.63	30785.81	-5.79	Q9I402	-20
Peptidoglycan associated lipoprotein	17524.1769	16001.7900	1522.39	Q9I4Z4	-10
E. faecium E745					
30S ribosomal protein S5	17335.52	17335.57	-0.05	Q3XYX2	-10

*Proteins of MW higher than 30 kDa are written in bold.

6.4 Conclusion

The work described in this chapter demonstrates a workflow for TD identification of proteins from bacteria which offers higher throughput than previous approaches. Two methods, one using LESA MS coupled to an Orbitrap Elite mass spectrometer and a second employing the new FAIMS Pro device coupled to the Orbitrap Eclipse, were compared. Initial testing and optimisation was performed on the well-studied bacterial species *E. coli* K12 and *S. aureus* MSSA476. The TD analysis showed that the highest number of proteins and proteoforms was identified with the high-throughput LESA FAIMS approach. FAIMS analysis significantly improved the number of IDs assigned for *S. aureus* MSSA476 by reducing the signal of the dominant δ-hemolysin peak.

Two Gram-negative and one Gram-positive species were further tested with the same, now optimised, instrument method. The results obtained were very similar to those of *E. coli* and *S. aureus*. Again, the high-throughput LESA FAIMS analysis yielded the highest number of protein and proteoform identifications. Two biological replicates were compared for each species and the inter-day variability was examined, suggesting that the differences may be caused by the natural variability between the bacterial colonies and LESA extraction process itself.

Lastly, three new proteins of MW above 30 kDa were identified at low FAIMS CVs -10 V and -20 V. Two of them were identified as membrane proteins. To date, these are the highest MW proteins identified by denaturing LESA FAIMS MS.

Chapter 7

Conclusion and future work

The overall aim of this thesis was to focus on exploration of LESA MS as a potential future diagnostic tool for infected skin wounds, where microbial classification is achieved by TD identification of proteins instead of typically used spectral matching. In this work, TD LESA MS was extended to optimisation and development of workflows for analysis of clinically important microbes growing on simple culturing media and complex substrates including blood agar, *in vitro* 3D skin models (Labskin) and *ex vivo* human skin grafts. LESA MS of ESKAPE pathogens grown on simple substrates allowed identification of 24 proteins and searches against multiple databases determined the protein identification rate to be 79% at protein and species level. Mixed-species biofilm of *P. aeruginosa* and *C. glabrata* was grown and its LESA MS analysis resulted in identification of three *P. aeruginosa* proteins. LESA MS of ESKAPE microbes cultured on blood agar showed that both substrate and bacterial protein peaks may be observed together in the mass spectra. Analysis of inoculated/infected Labskin samples yielded identification of seven human, nine bacterial and one yeast proteins. Application of the LESA MS workflow to *ex vivo* human skin grafts allowed identification of two human skin proteins from both non-inoculated and inoculated samples. Ultimately, high-throughput bacterial identification was investigated with aid of the state-of-art instrumentation, with a total of 179 proteins and 277 proteoforms were identified from colonies of *E. coli* and four ESKAPE species. In this concluding chapter, the outcomes from experiments in the individual Chapters are discussed and suggestions are provided for future work.

Chapter 3 described the optimisation of LESA solvent system for extracting proteins from both Gram-positive and Gram-negative ESKAPE pathogens. In previous work,

two different extraction solvent systems were used depending on the species investigated (Gram-positive or Gram-negative). Development of a universal extraction solvent system for LESA MS for all of the ESKAPE species was therefore one of the important goals achieved towards *in situ* LESA sampling of microbes. The applicability of the optimised extraction solvent will, however, need to be tested for microbes outside of the ESKAPE group in the future to confirm its universality. Next, four out of six ESKAPE microbes *E. faecium, K. pneumoniae, A. baumannii* and *E. cloacae*, including the clinically important strain *E. faecalis* and clinical isolate *A. baumannii* growing on simple substrates were investigated by TD LESA MS and the MS/MS spectra searches resulted in identification of 24 proteins. Interestingly, despite the presence of numerous proteins in the LESA mass spectra of the clinical isolate *A. baumannii* AC02, none of them were identified after tandem MS analysis. This finding highlights a drawback of protein-based TD MS microbial identification which is the low availability of well-characterised (annotated and reviewed) microbial protein databases. Many of the microbial proteome databases include only proteins inferred from homology with close-related species or non-reviewed proteins, while there are missing sequences and PTMs, all precluding correct protein ID assignment. An alternative approach to the typical database search (such as spectral clustering [230]) may also need to be considered in the future experiments.

To compare TD LESA MS to other techniques for bacterial characterisation, databases of MS/MS spectra from multiple species were searched and the overall identification success rate was determined to be 79% (including both Gram-positive and Gram-negative species). This number was compared to the identification success rates of MALDI TOF MS, but it is important to emphasise that the aim of this work is not to compete with MALDI TOF MS, but to explore suitability of LESA MS as a direct surface sampling method. *E. cloacae* presented the biggest challenge in terms of species identification based on results from both the multidatabase search and the BLAST search. Suggested improvements for LESA MS include the use of multiple databases at the genus and strain level, which might increase the success rate of protein identification in the future. Further

development of LESA MS as a diagnostic requires application to a broader range of species outside ESKAPE microbes and that will need to be implemented in future work.

Chapter 3 explored LESA MS analysis of mixed-species biofilms. A biofilm model of *P. aeruginosa* PS1054 and yeast *C. glabrata* was developed and subjected to LESA MS, leading to identification of three bacterial, but no yeast, proteins from the mixed-species sample. It is suggested that the detection of bacterial proteins only is a result of the antagonistic relationship between *P. aeruginosa* and *Candida* species. Most of the published studies focus on *P. aeruginosa* and *C. albicans* and the work presented in this thesis utilised protocol for biofilm formation originally developed for co-culture of *P. aeruginosa* and *C. albicans*. Nevertheless, the results show that LESA MS is capable of detecting proteins from a microbial co-culture. The future work suggests aiming at optimisation of the culturing process while LESA MS analysis should be carried out at more time points to investigate the progress of biofilm formation. A more in-depth analysis could potentially reveal how the presence of proteins changes with time during the biofilm development. It would be desirable to focus on analysis of mixed-species biofilms not only of *P. aeruginosa* and *C. albicans*, but also *P. aeruginosa* in combination with other ESKAPE species. Another aim will focus on improvement of number of proteins detected from the biofilm samples. Currently, three proteins would not be enough to correctly identify microbial species in clinical settings. The number of identified proteins could potentially be increased by using the new instrument (Orbitrap Eclipse) with the option to increase the S/N by incorporating FAIMS in the workflow.

Chapter 4 investigated the potential of LESA MS as an *in situ* diagnostic tool, i.e., LESA MS was used to analyse bacteria growing on more complex substrates including *in vitro* 3D skin models (Labskin). First, LESA MS analysis of ESKAPE species grown on blood agar was performed. This experiment showed that both bacterial and haemoglobin protein peaks (from the blood agar substrate) may be detected together in the LESA mass spectra which was one of the important steps towards sampling of microbes cultured on the complex substrates. Second, a wounded Labskin model was studied, where the wounded

samples were inoculated with two strains of *S. aureus*, *K. pneumoniae* and *P. aeruginosa*. Detection and identification of seven human skin and nine bacterial proteins after 48 hours of incubation showed that LESA MS offers the opportunity to analyse both host and microbial proteins in near real-time. Surprisingly, the ethanol-based extraction solvent system achieved better results (in terms of S/N) than the acetonitrile-based one. The use of a more patient-friendly extraction solvent represents a potential for *in situ* analysis from a patient in the future. Detection of allelic variants of the protein δ-hemolysin enabled differentiation between the two closely related strains of *S. aureus*. LESA MS may be used to provide dynamic information about the host-pathogen relationship at the wounded site. Future work will focus on identification of proteins from mixtures of pathogens in infected wounds. If this approach is to be used directly on patients, the proteins of human skin microbial commensals (e.g. *Staphylococcus epidermidis, Malassezia* spp. and *Propionibacterium* spp. [188]) need to be considered. LESA MS of commensal-pathogen co-culture should therefore be carried out. Furthermore, LESA FAIMS MS using the new FAIMS Pro device and the new mass spectrometer will likely be accompanied by improved numbers of proteins detected from the Labskin samples as well as various microbial species. The intact *in vitro* 3D skin models may also be attractive for native mass spectrometry analysis as some of the detected proteins are under native conditions present as a homodimer (calcyclin) or a heterodimer (calprotectin).

Labskin samples were also inoculated with yeast *C. glabrata* and these samples were subjected to electroporation. Subsequent LESA MS analysis resulted in identification of a truncated *C. glabrata* HSP12 protein from wounded and infected samples after 5 days of incubation. Only one sample that developed infection could be studied. The use of electroporator on Labskin samples proved challenging. The electrodes could not be properly inserted into the wound and the high salt content of culturing medium contributed to arcing between the electrodes. Future work will need to focus on modifications of the electroporation experimental setup, otherwise the pulses delivered into the samples will not be able to lyse the yeast cells. Another suggestion includes

testing of the optimised *C. glabrata* infectious dose and possible use of other types of clinically relevant yeast such as *C. albicans* or *C. neoformans*.

In Chapter 5, TD LESA MS was extended to analysis of *ex vivo* human skin grafts. Several challenges were experienced: the sample preparation and culturing conditions needed to be optimised together with extraction solvent system such that LESA MS could take place; skin grafts were sutured to cell strainers to maintain the skin tension, while the subsequent LESA MS analysis showed high presence of salts in the majority of the investigated samples resulting in low S/N and a high background noise. Nevertheless, proteins were detected and the comparison with Labskin mass spectra showed similarity in the observed protein peaks. Two proteins were identified by LESA MS/MS: psoriasin and ubiquitin. A chip-based FAIMS device was employed to aid the removal of background noise, and in fact, three proteins were detected, but their abundance was too low to perform tandem MS. LESA MS analysis of wounded and inoculated human skin samples resulted in identification of two human skin, but no bacterial proteins. The LESA sampling of the wound itself was very difficult to perform – the extraction solvent system dispensed in the wound could not be all re-aspirated, and the re-aspirated sample contained salts. This needs to be considered for future direct wound analysis. The re-aspirated sample may contain not only bacteria, but also other undesired analytes that need to be filtered out. Suggestions for future experiments would be to use the new cylindrical FAIMS device together with the state-of-art tribrid orbitrap mass spectrometer, potentially leading to more successful protein detection and identification, however the matrix effects need to be taken into account as they may suppress ionisation of the proteins of interest. Furthermore, the infectious dose of *S. aureus*, the same as used in previous experiments in Chapter 4, was probably not sufficient to cause an infection in the real human skin sample. *Ex vivo* human skin is a much more complex system and also contains cells involved in immune responses which may suppress the infection development, however, more studies will need to be conducted to confirm these findings and to understand the infection development in these samples.

Lastly, Chapter 6 described the development of a high-throughput LESA FAIMS MS workflow for analysis of *E. coli* K12 and ESKAPE microbes using state-of-art instrumentation – cylindrical FAIMS Pro device and orbitrap tribrid mass spectrometer. Two methods using LESA and LESA FAIMS coupled to the mass spectrometer were compared. Initially, the methods were optimised using the well-characterised species *E. coli* K12 and *S. aureus* MSSA476. It was shown that the high-throughput LESA FAIMS approach results in identification of the highest number of proteins and proteoforms. Significant improvements in protein identification were observed for FAIMS analysis of *S. aureus* MSSA476. The optimised method was extended to three ESKAPE species *E. faecium*, *K. pneumoniae* and *P. aeruginosa*, while the experiments confirmed that LESA FAIMS MS yields the highest number of protein and proteoform identifications. Two biological replicates were investigated for each species together with inter-day variability. Further work will be required, as in this Chapter only one replicate for each LESA FAIMS experiment is presented. If LESA MS is considered to be implemented in clinical practice in the future, LESA extraction reproducibility needs to be addressed. This could potentially be improved by sampling multiple locations/sample or by increasing dwell time (i.e., the contact between the droplet and the sampling surface, currently limited by the Advion software). For complete characterisation, the developed approach will need to be extended to all of the ESKAPE microbes and to other species outside this group. Investigation of Gram-positive species may be of high interest, particularly a comparison of closely related strains such as *S. aureus* MSSA476 and NCTC13435. Such comparison may potentially reveal other differences (other than the detected allelic variants) between these strains. In addition, the duration of LESA FAIMS analysis will need to be optimised. Acquisition of "top 10" data at six different CVs may take up to one day, that is (in terms of numbers of proteins identified) significantly shorter compared to the previous FAIMS analyses, but not suitable for rapid protein-based analysis of clinical samples. Future experiments may consider using a smaller CV range, where the highest numbers of proteins are detected. This step could significantly reduce the sample

analysis and protein identification time. Ultimately, three proteins of MW higher than 30 kDa were identified from *E. coli*, *K. pneumoniae* and *P. aeruginosa* at low FAIMS CVs -10 V and -20 V, which are to date proteins of highest MW identified by denaturing TD LESA FAIMS MS. This finding opens up an opportunity to study the ESKAPE microbes by native MS, as these proteins exist as a homotrimer or heterotrimer (OmpC) and a homodimer (OmpA).

In summary, this thesis describes the development of LESA MS as a tool capable of relatively fast protein extraction from microbes grown on simple and complex substrates with subsequent tandem MS analysis and identification. This work optimised the previously established workflow and was expanded to analysis of ESKAPE pathogens, mixed-species biofilms, intact and inoculated *in vitro* skin models and *ex vivo* human skin grafts. The progress in instrumentation development allowed high-throughput analysis and significantly improved identification of proteins and proteoforms in much shorter analysis time. The acquired knowledge can be further applied to analysis of other microbial species including yeasts cultured on various complex substrates and development of LESA MS-based tools capable of sampling surfaces with potential of direct *in situ* microbial identification.

References

[1] J. H. Gross, *Mass Spectrometry: A textbook*. Springer International Publishing, 2017.

[2] H.-F. Grützmacher, "Massenspektrometrie herausgegeben von Hermann Kienitz, verfaßt von Fritz Aulinger, Gerhard Franke, Karleugen Habfast, Hermann Kienitz und Gerhard Spiteller. Verlag Chemie, Weinheim/Bergstraße, 1968. 883 Seiten mit 342 Abbildungen und 52 Tabellen im Text u," *Organic Mass Spectrometry*, vol. 2, pp. 431–432, aug 1969.

[3] J. H. Gross, *Mass Spectrometry: A Textbook*. Berlin, Heidelberg: Springer, 2nd editio ed., 2011.

[4] J. J. Thomson, "LIV. Positive rays," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 16, pp. 657–691, oct 1908.

[5] J. J. Thomson, "LXXXIII. Rays of positive electricity," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 20, pp. 752–767, oct 1910.

[6] J. J. Thomson, "XIX. Further experiments on positive rays," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 24, pp. 209–253, aug 1912.

[7] R. Aebersold and M. Mann, "Mass spectrometry-based proteomics," *Nature*, vol. 422, no. 6928, pp. 198–207, 2003.

[8] E. de Hoffmann and V. Stroobant, *Mass Spectrometry Principles and Applications*. John Wiley & Sons, Ltd., 3rd ed., 2007.

[9] M. Karas and F. Hillenkamp, "Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10 000 Daltons," *Analytical Chemistry*, vol. 60, no. 20, pp. 2299–2301, 1988.

[10] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse, "Electrospray ionization for mass spectrometry of large biomolecules," *Science*, vol. 246, pp. 64 LP – 71, oct 1989.

[11] M. Mann, C. K. Meng, and J. B. Fenn, "Interpreting mass spectra of multiply charged ions," *Analytical Chemistry*, vol. 61, pp. 1702–1708, aug 1989.

[12] K. Tang, J. S. Page, R. T. Kelly, and I. Marginean, *Electrospray ionization in mass spectrometry*. Elsevier Ltd., 3 ed., 2017.

[13] G. Taylor, "Disintegration of water drops in an electric field," *Proceeding of The Royal Society A*, vol. 280, no. 1382, pp. 383–397, 1964.

[14] L. Konermann, E. Ahadi, A. D. Rodriguez, and S. Vahidi, "Unraveling the mechanism of electrospray ionization," *Analytical Chemistry*, vol. 85, no. 1, pp. 2–9, 2013.

[15] M. S. Wilm and M. Mann, "Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last?," *International Journal of Mass Spectrometry and Ion Processes*, vol. 136, no. 2-3, pp. 167–180, 1994.

[16] M. Wilm and M. Mann, "Analytical properties of the nanoelectrospray ion source," *Analytical Chemistry*, vol. 68, no. 1, pp. 1–8, 1996.

[17] H. Metwally, Q. Duez, and L. Konermann, "Chain Ejection Model for Electrospray Ionization of Unfolded Proteins: Evidence from Atomistic Simulations and Ion Mobility Spectrometry," *Analytical Chemistry*, vol. 90, no. 16, pp. 10069–10077, 2018.

[18] D. J. Pimlott and L. Konermann, "Using covalent modifications to distinguish protein electrospray mechanisms: Charged residue model (CRM) vs. chain ejection model (CEM)," *International Journal of Mass Spectrometry*, vol. 469, p. 116678, 2021.

[19] V. Kertesz, M. J. Ford, and G. J. Van Berkel, "Automation of a Surface Sampling Probe/Electrospray Mass Spectrometry System," *Analytical Chemistry*, vol. 77, no. 22, pp. 7183–7189, 2005.

[20] V. Kertesz and G. J. Van Berkel, "Fully automated liquid extraction-based surface sampling and ionization using a chip-based robotic nanoelectrospray platform," *Journal of Mass Spectrometry*, vol. 45, no. 3, pp. 252–260, 2010.

[21] Advion, "Advion ESI Chip." `https://www.advion.com/products/triversa-nanomate/`, 2020. Accessed: 2020-07-20.

[22] D. Eikel, M. Vavrek, S. Smith, C. Bason, S. Yeh, W. A. Korfmacher, and J. D. Henion, "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*, vol. 25, pp. 3587–3596, 2011.

[23] W. B. Parson, S. L. Koeniger, R. W. Johnson, J. Erickson, Y. Tian, C. Stedman, A. Schwartz, E. Tarcsa, R. Cole, and G. J. Van Berkel, "Analysis of chloroquine and metabolites directly from whole-body animal tissue sections by liquid extraction surface analysis (LESA) and tandem mass spectrometry," *Journal of Mass Spectrometry*, vol. 47, no. 11, pp. 1420–1428, 2012.

[24] J. G. Swales, J. W. Tucker, M. J. Spreadborough, S. L. Iverson, M. R. Clench, P. J. H. Webborn, and R. J. A. Goodwin, "Mapping Drug Distribution in Brain Tissue Using Liquid Extraction Surface Analysis Mass Spectrometry Imaging," *Analytical Chemistry*, vol. 87, no. 19, pp. 10146–10152, 2015.

[25] R. Almeida, Z. Berzina, E. C. Arnspang, J. Baumgart, J. Vogt, R. Nitsch, and C. S. Ejsing, "Quantitative Spatial Analysis of the Mouse Brain Lipidome by Pressurized Liquid Extraction Surface Analysis," *Analytical Chemistry*, vol. 87, no. 3, pp. 1749–1756, 2015.

[26] Z. Hall, Y. Chu, and J. L. Griffin, "Liquid Extraction Surface Analysis Mass Spectrometry Method for Identifying the Presence and Severity of Nonalcoholic Fatty Liver Disease," *Analytical Chemistry*, vol. 89, pp. 5161–5170, 2017.

[27] A. N. Conner, J. R. Jarvis, L. J. Alderwick, and R. L. Griffiths, "Direct liquid extraction surface analysis mass spectrometry of cell wall lipids from mycobacteria: Salt additives for decreased spectral complexity," *Rapid Communications in Mass Spectrometry*, no. February, pp. 1–9, 2019.

[28] A. Koulman, S. Furse, M. Baumert, G. Goldberg, and L. Bluck, "Rapid profiling of triglycerides in human breast milk using liquid extraction surface analysis Fourier transform mass spectrometry reveals new very long chain fatty acids and differences within individuals," *Rapid Communications in Mass Spectrometry*, vol. 33, no. 15, pp. 1267–1276, 2019.

[29] J. R. Haler, E. K. Sisley, Y. L. Cintron-Diaz, S. N. Meitei, H. J. Cooper, and F. Fernandez-Lima, "Workflow for fast lipid tissue screening using LESA-FT-ICR-MS," *Analytical Methods*, vol. 11, no. 18, pp. 2385–2395, 2019.

[30] R. L. Edwards, P. Griffiths, J. Bunch, and H. J. Cooper, "Top-Down Proteomics and Direct Surface Sampling of Neonatal Dried Blood Spots: Diagnosis of Unknown Hemoglobin Variants," *Journal of The American Society for Mass Spectrometry*, vol. 23, no. 11, pp. 1921–1930, 2012.

[31] J. Sarsby, N. J. Martin, P. F. Lalor, J. Bunch, and H. J. Cooper, "Top-down and bottom-up identification of proteins by liquid extraction surface analysis mass spectrometry of healthy and diseased human liver tissue," *Journal of the American Society for Mass Spectrometry*, vol. 25, no. 11, pp. 1953–1961, 2014.

[32] 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," *Journal of The American Society for Mass Spectrometry*, vol. 28, pp. 2066–2077, oct 2017.

[33] J. Havlikova, R. C. May, I. B. Styles, and H. J. Cooper, "Direct identification of bacterial and human proteins from infected wounds in living 3D skin models," *Scientific Reports*, vol. 10, no. 1, pp. 1–9, 2020.

[34] N. J. Martin, R. L. Griffiths, R. L. Edwards, and H. J. Cooper, "Native Liquid Extraction Surface Analysis Mass Spectrometry: Analysis of Noncovalent Protein Complexes Directly from Dried Substrates," *Journal of The American Society for Mass Spectrometry*, vol. 26, no. 8, pp. 1320–1327, 2015.

[35] R. L. Griffiths, E. K. Sisley, A. F. Lopez-Clavijo, A. L. Simmonds, I. B. Styles, and H. J. Cooper, "Native mass spectrometry imaging of intact proteins and protein complexes in thin tissue sections," *International Journal of Mass Spectrometry*, 2017.

[36] E. Illes-Toth and H. J. Cooper, "Probing the Fundamentals of Native Liquid Extraction Surface Analysis Mass Spectrometry of Proteins: Can Proteins Refold during Extraction?," *Analytical Chemistry*, vol. 91, no. 19, pp. 12246–12254, 2019.

[37] O. J. Hale, E. Illes-Toth, T. H. Mize, and H. J. Cooper, "High-Field Asymmetric Waveform Ion Mobility Spectrometry and Native Mass Spectrometry: Analysis of Intact Protein Assemblies and Protein Complexes," *Analytical Chemistry*, vol. 92, no. 10, pp. 6811–6816, 2020.

[38] O. J. Hale and H. J. Cooper, "Native Mass Spectrometry Imaging and In Situ Top-Down Identification of Intact Proteins Directly from Tissue," *Journal of the American Society for Mass Spectrometry*, 2020.

[39] J. G. Swales, N. Strittmatter, J. W. Tucker, M. R. Clench, P. J. H. Webborn, and R. J. A. Goodwin, "Spatial Quantitation of Drugs in tissues using Liquid Extraction Surface Analysis Mass Spectrometry Imaging," *Scientific Reports*, vol. 6, no. 37648, pp. 1–9, 2016.

[40] J. Havlikova, E. Randall, R. Griffiths, J. Swales, R. Goodwin, J. Bunch, I. Styles, and H. Cooper, "Quantitative Imaging of Proteins in Tissue by Stable Isotope Labeled Mimetic Liquid Extraction Surface Analysis Mass Spectrometry," *Analytical Chemistry*, vol. 91, pp. 14198–14202, 2019.

[41] R. L. Edwards, A. J. Creese, M. Baumert, P. Griffiths, J. Bunch, and H. J. Cooper, "Hemoglobin Variant Analysis via Direct Surface Sampling of Dried Blood Spots Coupled with High-Resolution Mass Spectrometry," *Analytical Chemistry*, vol. 83, pp. 2265–2270, mar 2011.

[42] R. L. Edwards, P. Griffiths, J. Bunch, and H. J. Cooper, "Compound heterozygotes and beta-thalassemia: Top-down mass spectrometry for detection of hemoglobinopathies," *PROTEOMICS*, vol. 14, pp. 1232–1238, jan 2014.

[43] E. C. Randall, J. Bunch, and H. J. Cooper, "Direct analysis of intact proteins from Escherichia coli colonies by liquid extraction surface analysis mass spectrometry," *Analytical Chemistry*, vol. 86, no. 21, pp. 10504–10510, 2014.

[44] K. Kocurek, J. Havlikova, E. Buchan, A. Tanner, R. C. May, and H. J. Cooper, "Electroporation and Mass Spectrometry: a New Paradigm for in Situ Analysis of Intact Proteins Direct From Living Yeast Colonies," *Analytical Chemistry*, vol. 92, no. 3, pp. 2605–2611, 2020.

[45] O. J. Hale and H. J. Cooper, "In situ mass spectrometry analysis of intact proteins and protein complexes from biological substrates," *Biochemical Society Transactions*, vol. 48, pp. 317–326, 2020.

[46] D. J. Douglas, "Linear quadrupoles in mass spectrometry," *Mass Spectrometry Reviews*, vol. 28, no. 6, pp. 937–960, 2009.

[47] J. C. Schwartz, M. W. Senko, and J. E. Syka, "A two-dimensional quadrupole ion trap mass spectrometer," *Journal of the American Society for Mass Spectrometry*, vol. 13, no. 6, pp. 659–669, 2002.

[48] A. A. Makarov, "Mass spectrometer." US Patent, 00692883, 1999.

[49] A. Makarov, "Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis," *Analytical Chemistry*, vol. 72, pp. 1156–1162, mar 2000.

[50] E. S. Hecht, M. Scigelova, S. Eliuk, and A. Makarov, *Fundamentals and Advances of Orbitrap Mass Spectrometry*. John Wiley & Sons, Ltd., 2019.

[51] R. A. Zubarev and A. Makarov, "Orbitrap Mass Spectrometry," *Analytical Chemistry*, vol. 85, pp. 5288–5296, jun 2013.

[52] S. Eliuk and A. Makarov, "Evolution of Orbitrap Mass Spectrometry Instrumentation," *Annual Review of Analytical Chemistry*, vol. 8, pp. 61–80, 2015.

[53] J. V. Olsen, J. C. Schwartz, J. Griep-Raming, M. L. Nielsen, E. Damoc, E. Denisov, O. Lange, P. Remes, D. Taylor, M. Splendore, E. R. Wouters, M. Senko, A. Makarov, M. Mann, and S. Horning, "A dual pressure linear ion trap orbitrap instrument with very high sequencing speed," *Molecular and Cellular Proteomics*, vol. 8, no. 12, pp. 2759–2769, 2009.

[54] A. Michalski, E. Damoc, O. Lange, E. Denisov, D. Nolting, M. Müller, R. Viner, J. Schwartz, P. Remes, M. Belford, J. J. Dunyach, J. Cox, S. Horning, M. Mann, and A. Makarov, "Ultra high resolution linear ion trap orbitrap mass spectrometer (orbitrap elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes," *Molecular and Cellular Proteomics*, vol. 11, no. 3, pp. 1–11, 2012.

[55] A. Michalski, E. Damoc, J. P. Hauschild, O. Lange, A. Wieghaus, A. Makarov, N. Nagaraj, J. Cox, M. Mann, and S. Horning, "Mass spectrometry-based proteomics using Q exactive, a high-performance benchtop quadrupole orbitrap mass spectrometer," *Molecular and Cellular Proteomics*, vol. 10, no. 9, pp. 1–11, 2011.

[56] Thermo Fisher Scientific, "Orbitrap Elite." `https://assets.thermofisher.com/TFS-Assets/CMD/Specification-Sheets/PS-30229-MS-Orbitrap-Elite-PS30229-EN.pdf`, 2016. Accessed: 2021-10-29.

[57] Thermo Fisher Scientific, "Orbitrap Q Exactive HF Mass Spectrometer." Specification Sheet: Orbitrap Q Exactive HF Orbitrap LC-MS/MS System - Higher-Quality Data, Faster Than Ever, 2017. Accessed: 2021-11-01.

[58] M. W. Senko, P. M. Remes, J. D. Canterbury, R. Mathur, Q. Song, S. M. Eliuk, C. Mullen, L. Earley, M. Hardman, J. D. Blethrow, H. Bui, A. Specht, O. Lange,

E. Denisov, A. Makarov, S. Horning, and V. Zabrouskov, "Novel parallelized quadrupole/linear ion trap/orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates," *Analytical Chemistry*, vol. 85, no. 24, pp. 11710–11714, 2013.

[59] A. M. Brunner, P. Lössl, F. Liu, R. Huguet, C. Mullen, M. Yamashita, V. Zabrouskov, A. Makarov, A. F. Altelaar, and A. J. Heck, "Benchmarking multiple fragmentation methods on an orbitrap fusion for top-down phospho-proteoform characterization," *Analytical Chemistry*, vol. 87, no. 8, pp. 4152–4158, 2015.

[60] Thermo Fisher Scientific, "Orbitrap Eclipse Tribrid Mass Spectrometer." `https://assets.thermofisher.com/TFS-Assets/CMD/Specification-Sheets/ps-65451-orbitrap-eclipse-tribrid-ms-ps65451-en.pdf`, 2019. Accessed: 2021-10-30.

[61] L. Sleno and D. A. Volmer, "Ion activation methods for tandem mass spectrometry," *Journal of Mass Spectrometry*, vol. 39, no. 10, pp. 1091–1112, 2004.

[62] L. A. Macias, I. C. Santos, and J. S. Brodbelt, "Ion Activation Methods for Peptides and Proteins," *Analytical Chemistry*, vol. 92, pp. 227–251, 2019.

[63] R. Craig and R. C. Beavis, "TANDEM: Matching proteins with tandem mass spectra," *Bioinformatics*, vol. 20, no. 9, pp. 1466–1467, 2004.

[64] V. H. Wysocki, G. Tsaprailis, L. L. Smith, and L. A. Breci, "Mobile and localized protons: A framework for understanding peptide dissociation," *Journal of Mass Spectrometry*, vol. 35, no. 12, pp. 1399–1406, 2000.

[65] R. Boyd and Á. Somogyi, "The mobile proton hypothesis in fragmentation of protonated peptides: A perspective," *Journal of the American Society for Mass Spectrometry*, vol. 21, no. 8, pp. 1275–1278, 2010.

[66] W. Cui, H. Rohrs, and M. Gross, "Top-down mass spectrometry: recent developments, applications and perspectives," *Analyst*, vol. 136, pp. 3854–3864, 2011.

[67] J. S. Brodbelt, "Ion Activation Methods for Peptides and Proteins," *Analytical Chemistry*, vol. 88, no. 1, pp. 30–51, 2016.

[68] J. V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, and M. Mann, "Higher-energy C-trap dissociation for peptide modification analysis," *Nature Methods*, vol. 4, no. 9, pp. 709–712, 2007.

[69] D. B. Wilburn, A. L. Richards, D. L. Swaney, and B. C. Searle, "CIDer: A Statistical Framework for Interpreting Differences in CID and HCD Fragmentation," *Journal of Proteome Research*, vol. 20, no. 4, pp. 1951–1965, 2021.

[70] J. E. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz, and D. F. Hunt, "Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 26, pp. 9528–9533, 2004.

[71] L. M. Mikesh, B. Ueberheide, A. Chi, J. J. Coon, J. E. Syka, J. Shabanowitz, and D. F. Hunt, "The Utility of ETD Mass Spectrometry in Proteomic Analysis," *Biochim Biophys Acta*, vol. 1764, no. 12, pp. 1811–1822, 2006.

[72] J. J. Coon, J. E. Syka, J. Shabanowitz, and D. F. Hunt, "Tandem mass spectrometry for peptide and protein sequence analysis," *BioTechniques*, vol. 38, no. 4, pp. 519–523, 2005.

[73] H. J. Cooper, "To What Extent is FAIMS Beneficial in the Analysis of Proteins?," *Journal of the American Society for Mass Spectrometry*, vol. 27, no. 4, pp. 566–577, 2016.

[74] V. Gabelica, A. A. Shvartsburg, C. Afonso, P. Barran, J. L. Benesch, C. Bleiholder, M. T. Bowers, A. Bilbao, M. F. Bush, J. L. Campbell, I. D. Campuzano, T. Causon, B. H. Clowers, C. S. Creaser, E. De Pauw, J. Far, F. Fernandez-Lima, J. C. Fjeldsted, K. Giles, M. Groessl, C. J. Hogan, S. Hann, H. I. Kim, R. T. Kurulugama, J. C. May, J. A. McLean, K. Pagel, K. Richardson, M. E. Ridgeway, F. Rosu, F. Sobott, K. Thalassinos, S. J. Valentine, and T. Wyttenbach, "Recommendations for reporting ion mobility Mass Spectrometry measurements," *Mass Spectrometry Reviews*, pp. 1–30, 2019.

[75] R. Cumeras, E. Figueras, C. E. Davis, J. I. Baumbach, and I. Gràcia, "Review on Ion Mobility Spectrometry. Part 1: Current instrumentation," *Analyst*, vol. 140, no. 5, pp. 1376–1390, 2015.

[76] J. N. Dodds and E. S. Baker, "Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead," *Journal of the American Society for Mass Spectrometry*, vol. 30, no. 11, pp. 2185–2195, 2019.

[77] V. Gabelica and E. Marklund, "Fundamentals of ion mobility spectrometry," *Current Opinion in Chemical Biology*, vol. 42, pp. 51–59, 2018.

[78] E. A. Mason and E. W. McDaniel, "Transport properties of ions in gases," *NASA STI/Recon Technical Report A*, vol. 89, p. 15174, jan 1988.

[79] C. Ieritano, A. Lee, J. Crouse, Z. Bowman, N. Mashmoushi, P. M. Crossley, B. P. Friebe, J. L. Campbell, and W. S. Hopkins, "Determining collision cross sections from differential ion mobility spectrometry," *Analytical Chemistry*, 2021.

[80] L. J. Brown, D. E. Toutoungi, N. A. Devenport, J. C. Reynolds, G. Kaur-Atwal, P. Boyle, and C. S. Creaser, "Miniaturized ultra high field asymmetric waveform ion mobility spectrometry combined with mass spectrometry for peptide analysis," *Analytical Chemistry*, vol. 82, no. 23, pp. 9827–9834, 2010.

[81] R. Guevremont, "High-field asymmetric waveform ion mobility spectrometry: A new tool for mass spectrometry," *Journal of Chromatography A*, vol. 1058, no. 1-2, pp. 3–19, 2004.

[82] A. A. Shvartsburg, F. Li, K. Tang, and R. D. Smith, "High-resolution field asymmetric waveform ion mobility spectrometry using new planar geometry analyzers," *Analytical Chemistry*, vol. 78, no. 11, pp. 3706–3714, 2006.

[83] A. A. Shvartsburg, K. Tang, R. D. Smith, M. Holden, M. Rush, A. Thompson, and D. Toutoungi, "Ultrafast differential ion mobility spectrometry at extreme electric fields coupled to mass spectrometry," *Analytical Chemistry*, vol. 81, no. 19, pp. 8048–8053, 2009.

[84] K. E. Swearingen, J. M. Winget, M. R. Hoopmann, U. Kusebauch, and R. L. Moritz, "Decreased gap width in a cylindrical FAIMS device improves protein discovery," *Analytical Chemistry*, vol. 87, no. 24, pp. 12230–12237, 2016.

[85] R. W. Purves, S. Prasad, M. Belford, A. Vandenberg, and J. J. Dunyach, "Optimization of a New Aerodynamic Cylindrical FAIMS Device for Small Molecule Analysis," *Journal of the American Society for Mass Spectrometry*, vol. 28, no. 3, pp. 525–538, 2017.

[86] R. W. Purves and R. Guevremont, "Electrospray Ionization High-Field Asymmetric Waveform Ion Mobility Spectrometry - Mass Spectrometry," *Analytical Chemistry*, vol. 71, no. 13, pp. 2346–2357, 1999.

[87] R. W. Purves, D. A. Barnett, and R. Guevremont, "Separation of protein conformers using electro spray-high field asymmetric waveform ion mobility spectrometry-mass spectrometry," *International Journal of Mass Spectrometry*, vol. 197, no. 1-3, pp. 163–177, 2000.

[88] R. Guevremont, D. A. Barnett, R. W. Purves, and J. Vandermey, "Analysis of a tryptic digest of pig hemoglobin using ESI-FAIMS-MS," *Analytical Chemistry*, vol. 72, no. 19, pp. 4577–4584, 2000.

[89] M. Sans, C. L. Feider, and L. S. Eberlin, "Advances in mass spectrometry imaging coupled to ion mobility spectrometry for enhanced imaging of biological tissues," *Current Opinion in Chemical Biology*, vol. 42, no. February 2018, pp. 138–146, 2018.

[90] K. Y. Garza, C. L. Feider, D. R. Klein, J. A. Rosenberg, J. S. Brodbelt, and L. S. Eberlin, "Desorption Electrospray Ionization Mass Spectrometry Imaging of Proteins Directly from Biological Tissue Sections," *Analytical Chemistry*, vol. 90, no. 13, pp. 7785–7789, 2018.

[91] C. L. Feider, N. Elizondo, and L. S. Eberlin, "Ambient ionization and FAIMS mass spectrometry for enhanced imaging of multiply charged molecular ions in biological tissues," *Analytical Chemistry*, vol. 88, no. 23, pp. 11533–11541, 2016.

[92] R. L. Griffiths, A. Dexter, A. J. Creese, and H. J. Cooper, "Liquid extraction surface analysis field asymmetric waveform ion mobility spectrometry mass spectrometry for the analysis of dried blood spots," *Analyst*, vol. 140, no. 20, pp. 6879–6885, 2015.

[93] J. Sarsby, R. L. Griffiths, A. M. Race, J. Bunch, E. C. Randall, A. J. Creese, and H. J. Cooper, "Liquid Extraction Surface Analysis Mass Spectrometry Coupled with Field Asymmetric Waveform Ion Mobility Spectrometry for Analysis of Intact Proteins from Biological Substrates," *Analytical Chemistry*, vol. 87, pp. 6794–6800, jul 2015.

[94] R. L. Griffiths, A. J. Creese, A. M. Race, J. Bunch, and H. J. Cooper, "LESA FAIMS Mass Spectrometry for the Spatial Profiling of Proteins from Tissue," *Analytical Chemistry*, vol. 88, no. 13, pp. 6758–6766, 2016.

[95] K. I. Kocurek, R. C. May, and H. J. Cooper, "Application of High-Field Asymmetric Waveform Ion Mobility Separation to LESA Mass Spectrometry of Bacteria," *Analytical Chemistry*, vol. 91, no. 7, pp. 4755–4761, 2019.

[96] R. L. Griffiths, J. W. Hughes, S. E. Abbatiello, M. W. Belford, I. B. Styles, and H. J. Cooper, "Comprehensive LESA Mass Spectrometry Imaging of Intact Proteins by Integration of Cylindrical FAIMS," *Analytical Chemistry*, 2020.

[97] H. Chen, G. Gamez, and R. Zenobi, "What Can We Learn from Ambient Ionization Techniques?," *Journal of the American Society for Mass Spectrometry*, vol. 20, no. 11, pp. 1947–1963, 2009.

[98] T. K. Toby, L. Fornelli, and N. L. Kelleher, "Progress in Top-Down Proteomics and the Analysis of Proteoforms," *Annual Review of Analytical Chemistry*, vol. 9, no. 1, pp. 499–519, 2016.

[99] L. M. Smith and N. L. Kelleher, "Proteoform: A single term describing protein complexity," *Nature Methods*, vol. 10, no. 3, pp. 186–187, 2013.

[100] L. M. Smith and N. L. Kelleher, "Proteoforms as the next proteomics currency," *Science*, vol. 359, no. 6380, pp. 1106–1108, 2018.

[101] L. V. Schaffer, R. J. Millikin, R. M. Miller, L. C. Anderson, R. T. Fellers, Y. Ge, N. L. Kelleher, R. D. LeDuc, X. Liu, S. H. Payne, L. Sun, P. M. Thomas, T. Tucholski, Z. Wang, S. Wu, Z. Wu, D. Yu, M. R. Shortreed, and L. M. Smith, "Identification and Quantification of Proteoforms by Mass Spectrometry," *Proteomics*, vol. 19, no. 10, pp. 1–15, 2019.

[102] L. M. Smith, P. M. Thomas, M. R. Shortreed, L. V. Schaffer, R. T. Fellers, R. D. LeDuc, T. Tucholski, Y. Ge, J. N. Agar, L. C. Anderson, J. Chamot-Rooke, J. Gault, J. A. Loo, L. Paša-Tolić, C. V. Robinson, H. Schlüter, Y. O. Tsybin, M. Vilaseca, J. A. Vizcaíno, P. O. Danis, and N. L. Kelleher, "A five-level classification system for proteoform identifications," *Nature Methods*, vol. 16, no. 10, pp. 939–940, 2019.

[103] N. L. Kelleher, H. Y. Lin, G. A. Valaskovic, D. J. Aaserud, E. K. Fridriksson, and F. W. McLafferty, "Top down versus bottom up protein characterization by tandem high- resolution mass spectrometry," *Journal of the American Chemical Society*, vol. 121, no. 4, pp. 806–812, 1999.

[104] G. E. Reid and S. A. McLuckey, "'Top down' protein characterization via tandem mass spectrometry," *Journal of Mass Spectrometry*, vol. 37, no. 7, pp. 663–675, 2002.

[105] K. I. Kocurek, R. L. Griffiths, and H. J. Cooper, "Ambient ionisation mass spectrometry for in situ analysis of intact proteins," *Journal of Mass Spectrometry*, vol. 53, pp. 565–578, apr 2018.

[106] T. Tucholski and Y. Ge, "Fourier-transform ion cyclotron resonance mass spectrometry for characterizing proteoforms," *Mass Spectrometry Reviews*, pp. 1–20, 2020.

[107] V. H. Wysocki, K. A. Resing, Q. Zhang, and G. Cheng, "Mass spectrometry of peptides and proteins," *Methods*, vol. 35, no. 3 SPEC.ISS., pp. 211–222, 2005.

[108] R. A. Zubarev, D. M. Horn, E. K. Fridriksson, N. L. Kelleher, N. A. Kruger, M. A. Lewis, B. K. Carpenter, and F. W. McLafferty, "Electron capture dissociation for structural characterization of multiply charged protein cations," *Analytical Chemistry*, vol. 72, no. 3, pp. 563–573, 2000.

[109] J. B. Shaw, W. Li, D. D. Holden, Y. Zhang, J. Griep-Raming, R. T. Fellers, B. P. Early, P. M. Thomas, N. L. Kelleher, and J. S. Brodbelt, "Complete protein characterization using top-down mass spectrometry and ultraviolet photodissociation," *Journal of the American Chemical Society*, vol. 135, no. 34, pp. 12646–12651, 2013.

[110] Y. M. Fung, C. M. Adams, and R. A. Zubarev, "Electron ionization dissociation of singly and multiply charged peptides," *Journal of the American Chemical Society*, vol. 131, no. 29, pp. 9977–9985, 2009.

[111] M. A. Zenaidee, C. Lantz, T. Perkins, W. Jung, R. R. Loo, and J. A. Loo, "Internal Fragments Generated by Electron Ionization Dissociation Enhance Protein Top-Down Mass Spectrometry," *Journal of the American Society for Mass Spectrometry*, vol. 31, no. 9, pp. 1896–1902, 2020.

[112] A. L. McCormack, J. L. Jones, and V. H. Wysocki, "Surface-Induced Dissociation of Multiply Protonated Peptides," *Journal of the American Society for Mass Spectrometry*, vol. 3, no. 8, pp. 859–862, 1992.

[113] A. D. Catherman, O. S. Skinner, and N. L. Kelleher, "Top Down proteomics: Facts and perspectives," *Biochemical and Biophysical Research Communications*, vol. 445, no. 4, pp. 683–693, 2014.

[114] D. P. Donnelly, C. M. Rawlins, C. J. DeHart, L. Fornelli, L. F. Schachner, Z. Lin, J. L. Lippens, K. C. Aluri, R. Sarin, B. Chen, C. Lantz, W. Jung, K. R. Johnson, A. Koller, J. J. Wolff, I. D. Campuzano, J. R. Auclair, A. R. Ivanov, J. P. Whitelegge, L. Paša-Tolić, J. Chamot-Rooke, P. O. Danis, L. M. Smith, Y. O. Tsybin, J. A. Loo, Y. Ge, N. L. Kelleher, and J. N. Agar, "Best practices and benchmarks for intact protein analysis for top-down mass spectrometry," *Nature Methods*, vol. 16, no. 7, pp. 587–594, 2019.

[115] R. Aebersold, J. N. Agar, I. J. Amster, M. S. Baker, C. R. Bertozzi, E. S. Boja, C. E. Costello, B. F. Cravatt, C. Fenselau, B. A. Garcia, Y. Ge, J. Gunawardena, R. C. Hendrickson, P. J. Hergenrother, C. G. Huber, A. R. Ivanov, O. N. Jensen, M. C. Jewett, N. L. Kelleher, L. L. Kiessling, N. J. Krogan, M. R. Larsen, J. A. Loo, R. R. Ogorzalek Loo, E. Lundberg, M. J. Maccoss, P. Mallick, V. K. Mootha, M. Mrksich, T. W. Muir, S. M. Patrie, J. J. Pesavento, S. J. Pitteri, H. Rodriguez,

A. Saghatelian, W. Sandoval, H. Schlüter, S. Sechi, S. A. Slavoff, L. M. Smith, M. P. Snyder, P. M. Thomas, M. Uhlén, J. E. Van Eyk, M. Vidal, D. R. Walt, F. M. White, E. R. Williams, T. Wohlschlager, V. H. Wysocki, N. A. Yates, N. L. Young, and B. Zhang, "How many human proteoforms are there?," *Nature Chemical Biology*, vol. 14, no. 3, pp. 206–214, 2018.

[116] T. K. Toby, L. Fornelli, K. Srzentić, C. J. DeHart, J. Levitsky, J. Friedewald, and N. L. Kelleher, "A comprehensive pipeline for translational top-down proteomics from a single blood draw," *Nature Protocols*, vol. 14, no. 1, pp. 119–152, 2018.

[117] L. M. Smith, J. N. Agar, J. Chamot-rooke, P. O. Danis, Y. Ge, and J. A. Loo, "The Human Proteoform Project: Defining the Human Proteome," *Science Advances*, vol. 7, p. eabk0734, 2021.

[118] J. C. Tran, L. Zamdborg, D. R. Ahlf, J. E. Lee, A. D. Catherman, K. R. Durbin, J. D. Tipton, A. Vellaichamy, J. F. Kellie, M. Li, C. Wu, S. M. Sweet, B. P. Early, N. Siuti, R. D. Leduc, P. D. Compton, P. M. Thomas, and N. L. Kelleher, "Mapping intact protein isoforms in discovery mode using top-down proteomics," *Nature*, vol. 480, no. 7376, pp. 254–258, 2011.

[119] J. A. Melby, D. S. Roberts, E. J. Larson, K. A. Brown, E. F. Bayne, S. Jin, and Y. Ge, "Novel Strategies to Address the Challenges in Top-Down Proteomics," *J. Am. Soc. Mass Spectrom*, vol. 32, pp. 1278–1294, 2021.

[120] R. D. LeDuc, G. K. Taylor, Y. B. Kim, T. E. Januszyk, L. H. Bynum, J. V. Sola, J. S. Garavelli, and N. L. Kelleher, "ProSight PTM: An integrated environment for protein identification and characterization by top-down mass spectrometry," *Nucleic Acids Research*, vol. 32, no. WEB SERVER ISS., pp. 340–345, 2004.

[121] N. M. Karabacak, L. Li, A. Tiwari, L. J. Hayward, P. Hong, M. L. Easterling, and J. N. Agar, "Sensitive and specific identification of wild type and variant proteins from 8 to 669 kDa using top-down mass spectrometry," *Molecular and Cellular Proteomics*, vol. 8, no. 4, pp. 846–856, 2009.

[122] X. Liu, Y. Sirotkin, Y. Shen, G. Anderson, Y. S. Tsai, Y. S. Ting, D. R. Goodlett, R. D. Smith, V. Bafna, and P. A. Pevzner, "Protein identification using top-down," *Molecular and Cellular Proteomics*, vol. 11, no. 6, p. M111.008524, 2012.

[123] Q. Kou, L. Xun, and X. Liu, "TopPIC: A software tool for top-down mass spectrometry-based proteoform identification and characterization," *Bioinformatics*, vol. 32, no. 22, pp. 3495–3497, 2016.

[124] R. X. Sun, L. Luo, L. Wu, R. M. Wang, W. F. Zeng, H. Chi, C. Liu, and S. M. He, "pTop 1.0: A High-Accuracy and High-Efficiency Search Engine for Intact Protein Identification," *Analytical Chemistry*, vol. 88, no. 6, pp. 3082–3090, 2016.

[125] J. Park, P. D. Piehowski, C. Wilkins, M. Zhou, J. Mendoza, G. M. Fujimoto, B. C. Gibbons, J. B. Shaw, Y. Shen, A. K. Shukla, R. J. Moore, T. Liu, V. A. Petyuk, N. Tolić, L. Paša-Tolić, R. D. Smith, S. H. Payne, and S. Kim,

"Informed-Proteomics: Open-source software package for top-down proteomics," *Nature Methods*, vol. 14, no. 9, pp. 909–914, 2017.

[126] A. J. Cesnik, M. R. Shortreed, L. V. Schaffer, R. A. Knoener, B. L. Frey, M. Scalf, S. K. Solntsev, Y. Dai, A. P. Gasch, and L. M. Smith, "Proteoform Suite: Software for Constructing, Quantifying, and Visualizing Proteoform Families," *Journal of Proteome Research*, vol. 17, no. 1, pp. 568–578, 2018.

[127] C. Lantz, M. A. Zenaidee, B. Wei, Z. Hemminger, R. R. Ogorzalek Loo, and J. A. Loo, "ClipsMS: An Algorithm for Analyzing Internal Fragments Resulting from Top-Down Mass Spectrometry," *Journal of Proteome Research*, vol. 20, no. 4, pp. 1928–1935, 2021.

[128] F. Meng, B. J. Cargile, L. M. Miller, A. J. Forbes, J. R. Johnson, and N. L. Kelleher, "Informatics and multiplexing of intact protein identification in bacteria and the archaea," *Nature Biotechnology*, vol. 19, no. 10, pp. 952–957, 2001.

[129] L. Zamdborg, R. D. LeDuc, K. J. Glowacz, Y. B. Kim, V. Viswanathan, I. T. Spaulding, B. P. Early, E. J. Bluhm, S. Babai, and N. L. Kelleher, "ProSight PTM 2.0: Improved protein identification and characterization for top down mass spectrometry," *Nucleic Acids Research*, vol. 35, no. SUPPL.2, pp. 701–706, 2007.

[130] R. D. Leduc, R. T. Fellers, B. P. Early, J. B. Greer, P. M. Thomas, and N. L. Kelleher, "The C-Score: A bayesian framework to sharply improve proteoform scoring in high-throughput top down proteomics," *Journal of Proteome Research*, vol. 13, no. 7, pp. 3231–3240, 2014.

[131] L. Käll, J. D. Storey, M. J. MacCoss, and W. S. Noble, "Posterior error probabilities and false discovery rates: Two sides of the same coin," *Journal of Proteome Research*, vol. 7, no. 1, pp. 40–44, 2008.

[132] S. Aggarwal and A. K. Yadav, *False Discovery Rate Estimation in Proteomics*, pp. 119–128. New York, NY: Springer New York, 2016.

[133] R. D. LeDuc, R. T. Fellers, B. P. Early, J. B. Greer, D. P. Shams, P. M. Thomas, and N. L. Kelleher, "Accurate estimation of context-dependent false discovery rates in top-down proteomics," *Molecular and Cellular Proteomics*, vol. 18, no. 4, pp. 796–805, 2019.

[134] J. Zhang, J. Rector, J. Q. Lin, J. H. Young, M. Sans, N. Katta, N. Giese, W. Yu, C. Nagi, J. Suliburk, J. Liu, A. Bensussan, R. J. DeHoog, K. Y. Garza, B. Ludolph, A. G. Sorace, A. Syed, A. Zahedivash, T. E. Milner, and L. S. Eberlin, "Nondestructive tissue analysis for ex vivo and in vivo cancer diagnosis using a handheld mass spectrometry system," *Science Translational Medicine*, vol. 9, sep 2017.

[135] P. G. Simmonds, "Whole Microorganisms Studied by Pyrolisis-Gas Chromatography-Mass Spectrometry: Significance for Extraterrestrial Life Detection Experiments," *Applied Microbiology*, vol. 20, no. 4, pp. 567–572, 1970.

[136] H. L. Meuzelaar and P. G. Kistemaker, "A Technique for Fast and Reproducible Fingerprinting of Bacteria by Pyrolysis Mass Spectrometry," *Analytical Chemistry*, vol. 45, no. 3, pp. 587–590, 1973.

[137] J. P. Anhalt and C. Fenselau, "Identification of Bacteria Using Mass Spectrometry," *Analytical Chemistry*, vol. 47, no. 2, pp. 219–225, 1975.

[138] M. A. Claydon, S. N. Davey, V. Edwards-Jones, and D. B. Gordon, "The Rapid Identification of Intact Microorganisms Using Mass Spectrometry," *Nature Biotechnology*, vol. 14, no. 11, pp. 1584–1586, 1996.

[139] T. Krishnamurthy and P. L. Ross, "Rapid identification of bacteria by direct matrix-assisted laser desorption/ionization mass spectrometric analysis of whole cells," *Rapid Communications in Mass Spectrometry*, vol. 10, no. 15, pp. 1992–1996, 1996.

[140] R. D. Holland, J. G. Wilkes, F. Rafii, J. B. Sutherland, C. C. Persons, K. J. Voorhees, and J. O. Lay, "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*, vol. 10, no. 10, pp. 1227–1232, 1996.

[141] S. Lévesque, P. J. Dufresne, H. Soualhine, M. C. Domingo, S. Bekal, B. Lefebvre, and C. Tremblay, "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*, vol. 10, no. 12, pp. 1–21, 2015.

[142] N. Strittmatter, M. Rebec, E. A. Jones, O. Golf, A. Abdolrasouli, J. Balog, V. Behrends, K. A. Veselkov, and Z. Takats, "Characterization and Identification of Clinically Relevant Microorganisms Using Rapid Evaporative Ionization Mass Spectrometry," *Analytical Chemistry*, vol. 86, pp. 6555–6562, jul 2014.

[143] N. Strittmatter, E. A. Jones, K. A. Veselkov, M. Rebec, J. G. Bundy, and Z. Takats, "Analysis of intact bacteria using rapid evaporative ionisation mass spectrometry," *Chemical Communications*, vol. 49, no. 55, p. 6188, 2013.

[144] F. Bolt, S. J. Cameron, T. Karancsi, D. Simon, R. Schaffer, T. Rickards, K. Hardiman, A. Burke, Z. Bodai, A. Perdones-Montero, M. Rebec, J. Balog, and Z. Takats, "Automated High-Throughput Identification and Characterization of Clinically Important Bacteria and Fungi using Rapid Evaporative Ionization Mass Spectrometry," *Analytical Chemistry*, vol. 88, no. 19, pp. 9419–9426, 2016.

[145] P. Pruski, D. A. MacIntyre, H. V. Lewis, P. Inglese, G. D. S. Correia, T. T. Hansel, P. R. Bennett, E. Holmes, and Z. Takats, "Medical Swab Analysis Using Desorption Electrospray Ionization Mass Spectrometry: A Noninvasive Approach for Mucosal Diagnostics," *Analytical Chemistry*, vol. 89, pp. 1540–1550, feb 2017.

[146] P. Pruski, G. D. Correia, H. V. Lewis, K. Capuccini, P. Inglese, D. Chan, R. G. Brown, L. Kindinger, Y. S. Lee, A. Smith, J. Marchesi, J. A. McDonald, S. Cameron, K. Alexander-Hardiman, A. L. David, S. J. Stock, J. E. Norman, V. Terzidou, T. G.

Teoh, L. Sykes, P. R. Bennett, Z. Takats, and D. A. MacIntyre, "Direct on-swab metabolic profiling of vaginal microbiome host interactions during pregnancy and preterm birth," *Nature Communications*, vol. 12, no. 1, pp. 1–14, 2021.

[147] V. P. Sica, H. A. Raja, T. El-Elimat, and N. H. Oberlies, "Mass spectrometry imaging of secondary metabolites directly on fungal cultures," *RSC Advances*, vol. 4, no. 108, pp. 63221–63227, 2014.

[148] A. Tata, C. J. Perez, M. O. Ore, D. Lostun, A. Passas, S. Morin, and D. R. Ifa, "Evaluation of imprint DESI-MS substrates for the analysis of fungal metabolites," *RSC Advances*, vol. 5, no. 92, pp. 75458–75464, 2015.

[149] A. Tata, C. Perez, M. L. Campos, M. A. Bayfield, M. N. Eberlin, and D. R. Ifa, "Imprint Desorption Electrospray Ionization Mass Spectrometry Imaging for Monitoring Secondary Metabolites Production during Antagonistic Interaction of Fungi," *Analytical Chemistry*, vol. 87, no. 24, pp. 12298–12305, 2015.

[150] H. Hu, S. Smith, X. Li, Z. Qian, Y. Su, M. Lin, J. Tu, and Y.-M. Liu, "Fast quantification of free amino acids in food by microfluidic voltage–assisted liquid desorption electrospray ionization–tandem mass spectrometry," *Analytical and Bioanalytical Chemistry*, vol. 412, no. 8, pp. 1947–1954, 2020.

[151] I. Lanekoff, O. Geydebrekht, G. E. Pinchuk, A. E. Konopka, and J. Laskin, "Spatially resolved analysis of glycolipids and metabolites in living Synechococcus sp. PCC 7002 using nanospray desorption electrospray ionization," *Analyst*, vol. 138, no. 7, pp. 1971–1978, 2013.

[152] J. Watrous, P. Roach, B. Heath, T. Alexandrov, J. Laskin, and P. C. Dorrestein, "Metabolic Profiling Directly from the Petri Dish Using Nanospray Desorption Electrospray Ionization Imaging Mass Spectrometry," *Analytical Chemistry*, vol. 85, pp. 10385–10391, nov 2013.

[153] C.-C. Hsu, M. S. ElNaggar, Y. Peng, J. Fang, L. M. Sanchez, S. J. Mascuch, K. A. Møller, E. K. Alazzeh, J. Pikula, R. A. Quinn, Y. Zeng, B. E. Wolfe, R. J. Dutton, L. Gerwick, L. Zhang, X. Liu, M. Månsson, and P. C. Dorrestein, "Real-Time Metabolomics on Living Microorganisms Using Ambient Electrospray Ionization Flow-Probe," *Analytical Chemistry*, vol. 85, pp. 7014–7018, aug 2013.

[154] C. F. F. Angolini, P. H. Vendramini, F. D. S. Araújo, W. L. Araújo, R. Augusti, M. N. Eberlin, and L. G. de Oliveira, "Direct Protocol for Ambient Mass Spectrometry Imaging on Agar Culture," *Analytical Chemistry*, vol. 87, pp. 6925–6930, jul 2015.

[155] S. Mukhopadhyay, A. S. Bharath Prasad, C. H. Mehta, and U. Y. Nayak, "Antimicrobial peptide polymers: no escape to ESKAPE pathogens—a review," *World Journal of Microbiology and Biotechnology*, vol. 36, no. 9, pp. 1–14, 2020.

[156] S. Santajit and N. Indrawattana, "Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens," *BioMed Research International*, vol. 2016, p. 2475067, may 2016.

[157] M. S. Mulani, E. E. Kamble, S. N. Kumkar, M. S. Tawre, and K. R. Pardesi, "Emerging strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: A review," *Frontiers in Microbiology*, vol. 10, no. APR, 2019.

[158] M. Navidinia, "The Clinical Importance of Emerging Eskape Pathogens in Nosocomial Infections," *Journal of Paramedical Sciences*, vol. 7, no. 3, pp. 43–57, 2016.

[159] L. B. Rice, "Federal funding for the study of antimicrobial resistance in nosocomial pathogens: No ESKAPE," *Journal of Infectious Diseases*, vol. 197, no. 8, pp. 1079–1081, 2008.

[160] G. Suleyman and G. J. Alangaden, "Nosocomial Fungal Infections: Epidemiology, Infection Control, and Prevention," *Infectious Disease Clinics of North America*, vol. 30, no. 4, pp. 1023–1052, 2016.

[161] E. D. Ashley, R. Drew, M. Johnson, R. Danna, D. Dabrowski, V. Walker, M. Prasad, B. Alexander, G. Papadopoulos, and J. Perfect, "Cost of invasive fungal infections in the era of new diagnostics and expanded treatment options," *Pharmacotherapy*, vol. 32, no. 10, pp. 890–901, 2012.

[162] T. R. Kozel and B. Wickes, "Fungal diagnostics," *Cold Spring Harbor Perspectives in Medicine*, vol. 4, p. a019299, 2014.

[163] M. Backx, P. L. White, and R. A. Barnes, "New fungal diagnostics," *British Journal of Hospital Medicine*, vol. 75, no. 5, pp. 271–276, 2014.

[164] J. Havlikova, R. C. May, I. B. Styles, and H. J. Cooper, "Liquid Extraction Surface Analysis Mass Spectrometry of ESKAPE Pathogens," *Journal of the American Society for Mass Spectrometry*, 2021.

[165] S. L. Percival, K. E. Hill, D. W. Williams, S. J. Hooper, D. W. Thomas, and J. W. Costerton, "A review of the scientific evidence for biofilms in wounds," *Wound Repair and Regeneration*, vol. 20, no. 5, pp. 647–657, 2012.

[166] C. J. Nobile and A. D. Johnson, "Candida albicans Biofilms and Human Disease," *Annual Review Microbiology*, vol. 69, pp. 71–92, 2015.

[167] M. B. Lohse, M. Gulati, A. D. Johnson, and C. J. Nobile, "Development and regulation of single-and multi-species Candida albicans biofilms," *Nature Reviews Microbiology*, vol. 16, no. 1, pp. 19–31, 2018.

[168] B. P. Singh, S. Ghosh, and A. Chauhan, "Development, dynamics and control of antimicrobial-resistant bacterial biofilms: a review," *Environmental Chemistry Letters*, 2021.

[169] H. M. Bandara, J. Y. Yau, R. M. Watt, L. J. Jin, and L. P. Samaranayake, "Pseudomonas aeruginosa inhibits in-vitro Candida biofilm development.," *BMC microbiology*, vol. 10, p. 125, 2010.

[170] A. Trejo-Hernández, A. Andrade-Domínguez, M. Hernández, and S. Encarnación, "Interspecies competition triggers virulence and mutability in Candida albicans-Pseudomonas aeruginosa mixed biofilms," *ISME Journal*, vol. 8, no. 10, pp. 1974–1988, 2014.

[171] F. Alam, D. Catlow, A. Di Maio, J. M. Blair, and R. A. Hall, "Candida albicans enhances meropenem tolerance of Pseudomonas aeruginosa in a dual-species biofilm," *Journal of Antimicrobial Chemotherapy*, vol. 75, no. 4, pp. 925–935, 2020.

[172] M. Beganovic, M. K. Luther, L. B. Rice, C. A. Arias, M. J. Rybak, and K. L. Laplante, "A review of combination antimicrobial therapy for enterococcus faecalis bloodstream infections and infective endocarditis," *Clinical Infectious Diseases*, vol. 67, no. 2, pp. 303–309, 2018.

[173] S. Ranotkar, P. Kumar, S. Zutshi, K. S. Prashanth, B. Bezbaruah, J. Anand, and M. Lahkar, "Vancomycin-resistant enterococci: Troublemaker of the 21st century," *Journal of Global Antimicrobial Resistance*, vol. 2, no. 4, pp. 205–212, 2014.

[174] H. Nakayama, K. Kurokawa, and B. L. Lee, "Lipoproteins in bacteria: Structures and biosynthetic pathways," *FEBS Journal*, vol. 279, no. 23, pp. 4247–4268, 2012.

[175] Thermo Fisher Scientific, "ProSightPC 4.0 User Guide." `https://assets.thermofisher.com/TFS-Assets/CMD/manuals/Man-XCALI-97801-ProSightPC-User-ManXCALI97801-EN.pdf`, 2016. Accessed: 2022-07-05.

[176] J. F. Lin, M. C. Ge, T. P. Liu, S. C. Chang, and J. J. Lu, "A simple method for rapid microbial identification from positive monomicrobial blood culture bottles through matrix-assisted laser desorption ionization time-of-flight mass spectrometry," *Journal of Microbiology, Immunology and Infection*, vol. 51, no. 5, pp. 659–665, 2018.

[177] T. C. Dingle and S. M. Butler-Wu, "MALDI-TOF mass spectrometry for microorganism identification," *Clinics in Laboratory Medicine*, vol. 33, no. 3, pp. 589–609, 2013.

[178] M. C. Ge, A. J. Kuo, K. L. Liu, Y. H. Wen, J. H. Chia, P. Y. Chang, M. H. Lee, T. L. Wu, S. C. Chang, and J. J. Lu, "Routine identification of microorganisms by matrix-assisted laser desorption ionization time-of-flight mass spectrometry: Success rate, economic analysis, and clinical outcome," *Journal of Microbiology, Immunology and Infection*, vol. 50, no. 5, pp. 662–668, 2017.

[179] R. M. Donlan, "Biofilms: Microbial Life on Surfaces," *Emerging Infectious Diseases*, vol. 8, no. 9, pp. 881–890, 2002.

[180] M. Harmsen, L. Yang, S. J. Pamp, and T. Tolker-Nielsen, "An update on Pseudomonas aeruginosa biofilm formation, tolerance, and dispersal," *FEMS Immunology and Medical Microbiology*, vol. 59, no. 3, pp. 253–268, 2010.

[181] M. T. T. Thi, D. Wibowo, and B. H. Rehm, "Pseudomonas aeruginosa biofilms," *International Journal of Molecular Sciences*, vol. 21, no. 22, pp. 1–25, 2020.

[182] H. M. Bandara, O. L. Lam, R. M. Watt, L. J. Jin, and L. P. Samaranayake, "Bacterial lipopolysaccharides variably modulate in vitro biofilm formation of Candida species," *Journal of Medical Microbiology*, vol. 59, no. 10, pp. 1225–1234, 2010.

[183] "Injuries and Violence: The Facts." `http://www.who.int/violence{\_}injury{\_}prevention/key{\_}facts/VIP{\_}key{\_}facts.pdf?ua=1`. Accessed: 2021-04-25.

[184] H. C. Yun, C. K. Murray, K. J. Nelson, and M. J. Bosse, "Infection After Orthopaedic Trauma: Prevention and Treatment," *Journal of Orthopaedic Trauma*, vol. 30, 2016.

[185] D. R. Tribble, N. G. Conger, S. Fraser, T. D. Gleeson, K. Wilkins, T. Antonille, A. Weintrob, A. Ganesan, L. J. Gaskins, P. Li, G. Grandits, M. L. Landrum, D. R. Hospenthal, E. V. Millar, L. H. Blackbourne, J. R. Dunne, D. Craft, K. Mende, G. W. Wortmann, R. Herlihy, J. McDonald, and C. K. Murray, "Infection-associated clinical outcomes in hospitalized medical evacuees after traumatic injury: Trauma infectious disease outcome study," *Journal of Trauma - Injury, Infection and Critical Care*, vol. 71, no. SUPPL. 1, 2011.

[186] D. R. Tribble and C. J. Rodriguez, "Combat-Related Invasive Fungal Wound Infections," *Current Fungal Infection Reports*, vol. 8, pp. 277–286, dec 2014.

[187] A. C. Weintrob, A. B. Weisbrod, J. R. Dunne, C. J. Rodriguez, D. Malone, B. A. Lloyd, T. E. Warkentien, W. J, C. K. Murray, W. Bradley, F. Shaikh, J. Shah, D. Aggarwal, M. L. Carson, and D. R. Tribble, "Combat trauma-associated invasive fungal wound infections: epidemiology and clinical classification," *Epidemiology and Infection*, vol. 143, no. 1, pp. 214–224, 2015.

[188] D. B. Holland, R. A. Bojar, A. H. T. Jeremy, E. Ingham, and K. T. Holland, "Microbial colonization of an in vitro model of a tissue engineered human skin equivalent – a novel approach," *FEMS Microbiology Letters*, vol. 279, pp. 110–115, dec 2007.

[189] C. A. Mitchell, H. Long, M. Donaldson, S. Francese, and M. R. Clench, "Lipid changes within the epidermis of living skin equivalents observed across a time-course by MALDI-MS imaging and profiling," *Lipids in Health and Disease*, vol. 14, p. 84, aug 2015.

[190] A. Harvey, L. M. Cole, R. Day, M. Bartlett, J. Warwick, R. A. Bojar, D. Smith, N. Cross, and M. R. Clench, "MALDI-MSI for the analysis of a 3D tissue-engineered psoriatic skin model," *PROTEOMICS*, vol. 16, pp. 1718–1725, may 2016.

[191] C. A. Mitchell, M. Donaldson, S. Francese, and M. R. Clench, "MALDI MSI analysis of lipid changes in living skin equivalents in response to emollient creams containing palmitoylethanolamide," *Methods*, vol. 104, pp. 93–100, 2016.

[192] E. E. L. Lewis, M. R. T. Barrett, L. Freeman-Parry, R. A. Bojar, and M. R. Clench, "Examination of the skin barrier repair/wound healing process using a living skin equivalent model and matrix-assisted laser desorption-ionization-mass spectrometry imaging," *International Journal of Cosmetic Science*, vol. 40, pp. 148–156, jan 2018.

[193] C. Russo, E. E. L. Lewis, L. Flint, and M. R. Clench, "Mass Spectrometry Imaging of 3D Tissue Models," *PROTEOMICS*, vol. 18, p. 1700462, apr 2018.

[194] C. Russo, N. Brickelbank, C. Duckett, S. Mellor, S. Rumbelow, and M. R. Clench, "Quantitative Investigation of Terbinafine Hydrochloride Absorption into a Living Skin Equivalent Model by MALDI-MSI," *Analytical Chemistry*, vol. 90, no. 16, pp. 10031–10038, 2018.

[195] J. Cohen, W. G. Powderly, and S. M. Opal, *Infectious Diseases E-Book*. Elsevier Health Sciences, 4th ed., 2017.

[196] G. Borkow, *Use of biocidal surfaces for reduction of healthcare acquired infections*, vol. 9783319080. Springer, Cham, 2014.

[197] A. Halawi, O. Abbas, and M. Mahalingam, "S100 proteins and the skin: A review," *Journal of the European Academy of Dermatology and Venereology*, vol. 28, no. 4, pp. 405–414, 2014.

[198] F. Ghezzo, E. Lauret, S. Ferrari, and R. Baserga, "Growth factor regulation of the promoter for calcyclin, a growth-regulated gene," *Journal of Biological Chemistry*, vol. 263, no. 10, pp. 4758–4763, 1988.

[199] R. Donato, G. Sorci, and I. Giambanco, "S100A6 protein: functional roles," *Cellular and Molecular Life Sciences*, vol. 74, no. 15, pp. 2749–2760, 2017.

[200] P. H. Watson, E. R. Leygue, and L. C. Murphy, "Psoriasin (S100A7)," *International Journal of Biochemistry and Cell Biology*, vol. 30, no. 5, pp. 567–571, 1998.

[201] R. Gläser, J. Harder, H. Lange, J. Bartels, E. Christophers, and J.-M. Schröder, "Antimicrobial psoriasin (S100A7) protects human skin from Escherichia coli infection," *Nature Immunology*, vol. 6, p. 57, nov 2004.

[202] K. Hsu, C. Champaiboon, B. Guenther, B. Sorenson, A. Khammanivong, K. Ross, C. Geczy, and M. Herzberg, "Anti-Infective Protective Properties of S100 Calgranulins," *Anti-Inflammatory & Anti-Allergy Agents in Medicinal Chemistry*, vol. 8, no. 4, pp. 290–305, 2009.

[203] W. Leśniak and A. Graczyk-Jarzynka, "The S100 proteins in epidermis: Topology and function," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1850, no. 12, pp. 2563–2572, 2015.

[204] A. Zhong, W. Xu, J. Zhao, P. Xie, S. Jia, J. Sun, R. D. Galiano, T. A. Mustoe, and S. J. Hong, "S100A8 and S100A9 Are Induced by Decreased Hydration in the Epidermis and Promote Fibroblast Activation and Fibrosis in the Dermis," *American Journal of Pathology*, vol. 186, no. 1, pp. 109–122, 2016.

[205] C. Kerkhoff, A. Voss, T. E. Scholzen, M. M. Averill, K. S. Zänker, and K. E. Bornfeldt, "Novel Insights into the Role of S100A8/A9 in Skin Biology," *Experimental dermatology*, vol. 21, pp. 822–826, nov 2012.

[206] J. M. Schröder and J. Harder, "Antimicrobial skin peptides and proteins," *Cellular and Molecular Life Sciences*, vol. 63, no. 4, pp. 469–486, 2006.

[207] M. Schäfer-Korting and J. Rolff, "Skin delivery of antimicrobial peptides," in *Emerging Nanotechnologies in Immunology: The Design, Applications and Toxicology of Nanopharmaceuticals and Nanovaccines*, pp. 23–45, Elsevier Inc., 2018.

[208] B. Wang, B. J. McHugh, A. Qureshi, D. J. Campopiano, D. J. Clarke, J. R. Fitzgerald, J. R. Dorin, R. Weller, and D. J. Davidson, "IL-1β–Induced Protection of Keratinocytes against Staphylococcus aureus-Secreted Proteases Is Mediated by Human β-Defensin 2," *Journal of Investigative Dermatology*, vol. 137, no. 1, pp. 95–105, 2017.

[209] O. Wiedow, J. M. Schroder, H. Gregory, J. A. Young, and E. Christophers, "Elafin: An elastase-specific inhibitor of human skin. Purification, characterization, and complete amino acid sequence," *Journal of Biological Chemistry*, vol. 265, no. 25, pp. 14791–14795, 1990.

[210] S. Paczesny, T. M. Braun, J. E. Levine, J. Hogan, J. Crawford, B. Coffing, S. Olsen, S. W. Choi, H. Wang, V. Faca, S. Pitteri, Q. Zhang, A. Chin, C. Kitko, S. Mineishi, G. Yanik, E. Peres, D. Hanauer, Y. Wang, P. Reddy, S. Hanash, and J. L. Ferrara, "Elafin is a biomarker of graft-versus-host disease of the skin," *Science Translational Medicine*, vol. 2, no. 13, 2010.

[211] A. Peschel and M. Otto, "Phenol-soluble modulins and staphylococcal infection," *Nature Reviews Microbiology*, vol. 11, no. 11, p. 814, 2013.

[212] G. A. Somerville, A. Cockayne, M. Dürr, A. Peschel, M. Otto, and J. M. Musser, "Synthesis and Deformylation of Staphylococcus aureus δ-Toxin Are Linked to Tricarboxylic Acid Cycle Activity," *Journal of Bacteriology*, vol. 185, no. 22, pp. 6686–6694, 2003.

[213] G. Y. Cheung, A. J. Yeh, D. Kretschmer, A. C. Duong, K. Tuffuor, C. L. Fu, H. S. Joo, B. A. Diep, M. Li, Y. Nakamura, G. Nunez, A. Peschel, and M. Otto, "Functional characteristics of the Staphylococcus aureus δ-toxin allelic variant G10S," *Scientific Reports*, vol. 5, pp. 1–12, 2015.

[214] W. M. Scheld, R. A. Calderone, J. P. Brodeur, and M. A. Sande, "Influence of performed antibody on the pathogenesis of experimental Candida albicans endocarditis," *Infection and Immunity*, vol. 40, no. 3, pp. 950–955, 1983.

[215] R. A. Bojar, "Studying the Human Skin Microbiome Using 3D In Vitro Skin Models," *Applied In Vitro Toxicology*, vol. 1, pp. 165–171, jun 2015.

[216] P. A. J. Kolarsick, M. A. Kolarsick, and C. Goodwin, "Anatomy and Physiology of the Skin," *Journal of the Dermatology Nurses' Association*, vol. 3, no. 4, 2011.

[217] M. W. Findlay and G. C. Gurtner, "Chapter 35 - Engineering Niches for Skin and Wound Healing," in *Biology and Engineering of Stem Cell Niches* (A. Vishwakarma, J. M. B. T. B. Karp, and E. of Stem Cell Niches, eds.), pp. 559–579, Boston: Academic Press, 2017.

[218] D. Taverna, L. B. Nanney, A. C. Pollins, G. Sindona, and R. Caprioli, "Spatial mapping by imaging mass spectrometry offers advancements for rapid definition of human skin proteomic signatures," *Experimental Dermatology*, vol. 20, no. 8, pp. 642–647, 2011.

[219] A. A. Shvartsburg, R. D. Smith, A. Wilks, A. Koehl, D. Ruiz-Alonso, and B. Boyle, "Ultrafast Differential Ion Mobility Spectrometry at Extreme Electric Fields in Multichannel Microchips," *Analytical Chemistry*, vol. 81, no. 15, pp. 6489–6495, 2009.

[220] W. Xu, S. Jong Hong, S. Jia, Y. Zhao, R. D. Galiano, and T. A. Mustoe, "Application of a partial-thickness human ex vivo skin culture model in cutaneous wound healing study," *Laboratory Investigation*, vol. 92, p. 584, jan 2012.

[221] E. K. Sisley, E. Illes-Toth, and H. J. Cooper, "In situ analysis of intact proteins by ion mobility mass spectrometry," *TrAC - Trends in Analytical Chemistry*, vol. 124, p. 115534, 2020.

[222] R. Guevremont and R. W. Purves, "Atmospheric pressure ion focusing in a high-field asymmetric waveform ion mobility spectrometer," *Review of Scientific Instruments*, vol. 70, no. 2, pp. 1370–1383, 1999.

[223] B. B. Schneider, E. G. Nazarov, F. Londry, P. Vouros, and T. R. Covey, "Differential mobility spectrometry/mass spectrometry history, theory, design optimisation, simulations, and applications," *Mass Spectrometry Reviews*, vol. 35, pp. 687–737, 2016.

[224] G. E. Reid, H. Shang, J. M. Hogan, G. U. Lee, and S. A. McLuckey, "Gas-phase concentration, purification and identification of whole proteins from complex mixtures," *Proceedings 50th ASMS Conference on Mass Spectrometry and Allied Topics*, no. 8, pp. 47–48, 2002.

[225] I. Ntai, T. K. Toby, R. D. LeDuc, and N. L. Kelleher, *A Method for Label-Free, Differential Top-Down Proteomics*, pp. 121–133. New York, NY: Springer New York, 2016.

[226] M. Greer, V. R. Gerbasi, J. Greer, R. Fellers, D. Horn, R. Huguett, S. Abbatiello, M. Belford, P. Compton, K. R. Durbin, N. L. Kelleher, R. D. Leduc, S. Peterman, and P. Thomas, "High-throughput Top-down FAIMS Data Analysis with ProSightPD Nodes in Proteome Discoverer Software," 2019.

[227] A. A. Shvartsburg and R. D. Smith, "Protein analyses using differential ion mobility microchips with mass spectrometry," *Analytical Chemistry*, vol. 84, no. 17, pp. 7297–7300, 2012.

[228] R. D. Melani, K. Srzentić, V. R. Gerbasi, J. P. McGee, R. Huguet, L. Fornelli, and N. L. Kelleher, "Direct measurement of light and heavy antibody chains using ion mobility and middle-down mass spectrometry," *mAbs*, vol. 11, no. 8, pp. 1351–1357, 2019.

[229] M. Dupré, M. Duchateau, C. Malosse, D. Borges-Lima, V. Calvaresi, I. Podglajen, D. Clermont, M. Rey, and J. Chamot-Rooke, "Optimization of a Top-Down Proteomics Platform for Closely Related Pathogenic Bacterial Discrimination," *Journal of Proteome Research*, 2020.

[230] D. Borges Lima, M. Dupré, M. D. Mariano Santos, P. C. Carvalho, and J. Chamot-Rooke, "DiagnoTop: A Computational Pipeline for Discriminating Bacterial Pathogens without Database Search," *Journal of the American Society for Mass Spectrometry*, vol. 32, no. 6, pp. 1295–1299, 2021.

Appendix A
Publications

1. Havlikova, J.; May, R. C.; Styles, I. B. and Cooper, H. J. Direct identification of bacterial and human proteins from infected wounds in living 3D skin models. *Scientific Reports* 2020, 10, 11900.

2. Havlikova, J.; May, R. C.; Styles, I. B. and Cooper, H. J. Liquid extraction surface analysis mass spectrometry of ESKAPE pathogens. *Journal of the American Society for Mass Spectrometry.* 2021, 32, 1345–1351

SCIENTIFIC REPORTS
nature research

Check for updates

OPEN

Direct identification of bacterial and human proteins from infected wounds in living 3D skin models

Jana Havlikova[1,2], Robin C. May[2,3], Iain B. Styles[1,4,5,6] & Helen J. Cooper[2]✉

Trauma is one of the leading causes of death in people under the age of 49 and complications due to wound infection are the primary cause of death in the first few days after injury. The ESKAPE pathogens are a group of bacteria that are a leading cause of hospital-acquired infections and a major concern in terms of antibiotic resistance. Here, we demonstrate a novel and highly accurate approach for the rapid identification of ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp.) directly from infected wounds in 3D in vitro skin models. Wounded skin models were inoculated with bacteria and left to incubate. Bacterial proteins were identified within minutes, directly from the wound, by liquid extraction surface analysis mass spectrometry. This approach was able to distinguish closely related strains and, unlike genomic approaches, can be modified to provide dynamic information about pathogen behaviour at the wound site. In addition, since human skin proteins were also identified, this method offers the opportunity to analyse both host and pathogen biomarkers during wound infection in near real-time.

Trauma is one of the leading causes of death in people under the age of 49[1]. Complications due to infection are the primary cause of death in patients who survive the first few days after a traumatic injury. In recent military conflicts, over half of the total injuries sustained were the result of trauma to the extremities[2]. Over 25% of those patients suffered complications due to infection (rising to 50% for those requiring intensive care), either in the bones (osteomyelitis) or deep-wound infection[2,3]. Similar rates of infection are associated with civilian trauma[2]. Infectious complications result in significantly higher amputation rates, which are also of greater severity. The ESKAPE pathogens (***Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa***, and ***Enterobacter*** spp.) are a group of opportunistic pathogens which account for most hospital-acquired infections, and their antibiotic resistance is rising[4,5]. Rapid diagnosis and treatment of wound infection is therefore of utmost importance. Current diagnosis involves visual inspection of the wound for signs of infection (inflammation), followed by collection of swabs or tissues, microbial culture and identification in a clinical laboratory, which can take hours or even days.

Mass spectrometry (MS) has emerged as a powerful technique for identification and classification of microorganisms. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS has been in development for that purpose since the late 1990s[6,7] and is now well-established in the clinic[8]. In the analysis, patient-derived samples are cultured and individual colonies are smeared onto a MALDI target plate. A matrix is applied and the sample is subjected to MALDI-TOF MS. Identification is achieved by spectral fingerprinting, that is, matching protein mass spectra acquired from unknown samples against databases filled with mass spectra of previously identified microorganisms. An often-quoted advantage of MALDI-TOF is its speed. It is certainly true that the analysis time is short, especially in comparison with biochemical techniques, but this belies the time taken for culturing of the sample on solid media. Recent research has focused on the use of blood cultures to reduce the time of the sample preparation, but even with this approach it takes approximately 2 days to obtain the final

[1]EPSRC Centre for Doctoral Training in Physical Sciences for Health, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. [2]School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. [3]Institute of Microbiology and Infection, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. [4]School of Computer Science, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. [5]Centre of Membrane Proteins and Receptors, The Universities of Birmingham and Nottingham, The Midlands, Birmingham, UK. [6]Alan Turing Institute, 96 Euston Road, London NW1 2DB, UK. ✉email

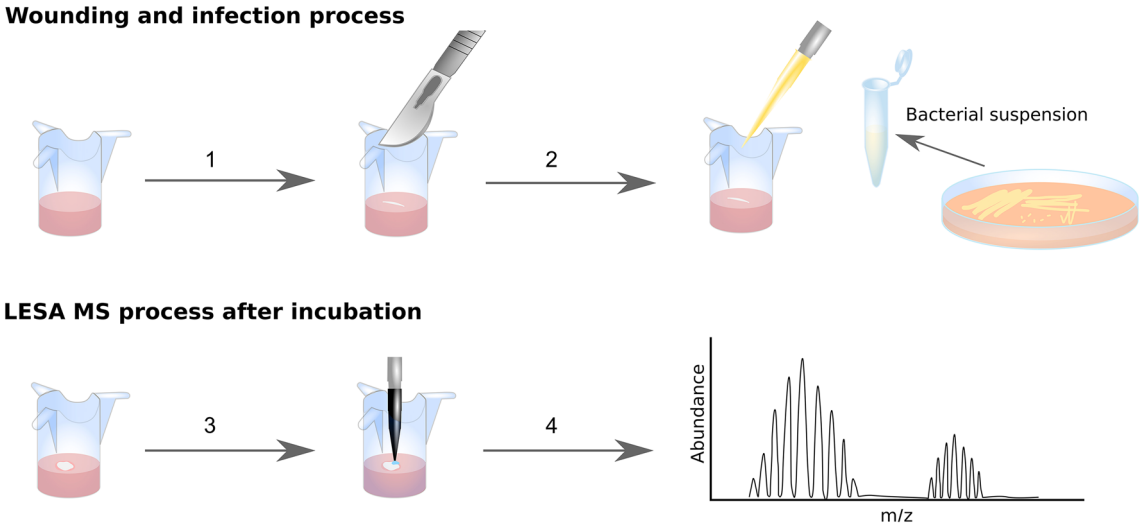


Figure 1. Labskin wounding, inoculation and analysis workflow. The Labskin sample (in the cell insert) is wounded with a scalpel blade (1) and inoculated with bacteria (2). After incubation (3), the infected Labskin sample is analysed by LESA MS (4).

result[9]. Other disadvantages of MALDI-TOF MS, particularly for the analysis of biofilms, arise because analysis takes place under vacuum. Problems include poor matrix saturation and sample flaking[10].

The ideal solution would involve a point-of-care diagnostic in which the microorganisms are identified directly from the wound, thus reducing the time from patient to result and enabling rapid deployment of the appropriate narrow spectrum antibiotic. Previous work in our laboratory has focused on the development of liquid extraction surface analysis (LESA) MS for direct analysis of bacteria growing on solid media[11–13]. LESA MS[14] is an ambient mass spectrometry technique, i.e., is conducted under atmospheric conditions in the open laboratory. Other ambient mass spectrometry techniques that have been applied to the analysis of bacteria include desorption electrospray ionisation (DESI) MS[15–23], rapid evaporative ionisation mass spectrometry (REIMS)[24–27], paper spray (PS) MS[28–30], nano-DESI[31–33], and Flowprobe[34]. These techniques have so far been limited to small molecules (lipids and metabolites), whereas crucially LESA MS is capable of analysing intact proteins in bacteria. The unique features of LESA MS with respect to its potential as a point-of-care diagnostic for wound infection are: (1) LESA sampling can be applied to any surface. To date, LESA sampling has been performed on substrates placed within the sampling platform; however, there is no inherent restriction on the nature of the substrate, e.g., a wound on a patient. (2) LESA MS enables the characterisation of proteins (like MALDI-TOF MS) and can be performed on living bacteria (unlike MALDI-TOF MS); (3) LESA MS has the potential to identify not just microbial proteins but also proteins from the patient, thus providing an indication of host response.

To confirm the potential of LESA MS for direct identification of bacterial wound infections, we have applied LESA MS to the analysis of bacteria growing in wounded three-dimensional in vitro living skin equivalents ("Labskin"). Labskin comprises a dermal layer, consisting of primary fibroblasts embedded in a fibrin matrix, and an epidermal layer created by seeding keratinocytes on the dermis, and the ability to culture microbes on the model is proven[35]. Clench and co-workers have applied MALDI MS and MALDI MS imaging of transverse sections of Labskin for the analysis of lipids and small molecules[36–41]. Here, we demonstrate the analysis and identification of three ESKAPE pathogens, *S. aureus*, *K. pneumoniae,* and *P. aeruginosa* directly from wounded and infected Labskin by LESA MS. For *S. aureus*, two strains were considered, an MRSA reference strain and an MSSA clinical isolate. Both bacterial and human proteins were detected and identified. Protein identification was achieved by top-down mass spectrometry[42] in which intact protein ions are fragmented to provide sequence information. The mass-to-charge ratios of the resulting sequence fragments are searched against protein databases by use of dedicated algorithms and putative identifications and associated scores are returned. Seven human proteins and nine bacterial proteins were identified in total. Five of the identified human skin proteins are known to have antimicrobial activity. The protein δ-hemolysin was identified from both strains of *S. aureus*; however, the sequence of δ-hemolysin differs between the two, with an associated mass difference which is easily detected by mass spectrometry. Detection of these proteins therefore allows the differentiation of these species.

Results and discussion

Infection of the Labskin models and LESA MS analysis of Labskin samples. The workflow is summarised in Fig. 1. Briefly, the Labskin samples were wounded with a scalpel blade and inoculated with bacterial suspensions of *S. aureus* NCTC13435, *S. aureus* MSSA476, *K. pneumoniae* KP257, and *P. aeruginosa* PS1054 (SI, Table S1). The sampling solvent was optimised for extraction of intact proteins from control Labskin models and determined to be ethanol, water and formic acid (60:35:5) (SI, Fig. S1). An important consideration in translation of this approach to a point-of-care diagnostic is patient-friendly extraction such as that presented by an ethanol-based solvent system. The optimised sampling solvent was subsequently validated for bacterial protein detection in LESA MS of microbial colonies growing on agar (SI, Fig. S2).

Four incubation time points were investigated: 24, 48, 72, and 96 h. After 24 h of incubation, there were no visible signs of colony growth inside the wound. The LESA mass spectra did not contain any peaks corresponding to bacterial proteins; however, several peaks corresponding to human skin proteins were observed (see below for details of protein identification).

After 48 h of incubation, colonies had formed in the wound for NCTC13435, MSSA476, and KP257 (Fig. S3). For PS1054, significant changes in the Labskin structure were observed, together with the presence of the typical subtle green colour associated with secretion of pyoverdine and pyocyanin in *P. aeruginosa*[43]. Pyocyanin has molecular weight of 210.231 Da which falls below the *m*/*z* range of this experiment. The molecular weight of pyoverdine (1,365.424 Da) is within the *m*/*z* range of the experiment; however, no corresponding peaks were detected. Peaks corresponding to bacterial proteins were present in the mass spectra for all of the infected and wounded Labskin samples, in addition to peaks corresponding to human skin proteins (Fig. 2). Non-wounded Labskin samples inoculated with *S. aureus*, *K. pneumoniae,* and *P. aeruginosa* did not exhibit colony growth or changes in their structure during the experiment (Fig. S3). This result is unsurprising as colonisation of intact skin by opportunistic pathogens does not result in an infection. Comparison of the LESA mass spectra obtained from the control samples (both intact and wounded) with the infected samples (Fig. 2) reveals that no peaks corresponding to bacterial proteins were identified in the samples that were not inoculated, confirming that no bacterial cross-contamination occurred.

After 72 h incubation, two further proteins were detected in the mass spectra of the wounded samples infected with PS1054 (Fig. S4), suggesting further bacterial growth in the wound. In some LESA mass spectra obtained from the samples infected with PS1054, no skin proteins were detected, likely due to the rapid progress of the bacterial colonisation in the infected wound. After 96 h, no additional proteins, either human or bacterial, were observed.

Bacterial and human skin proteins identified following LESA MS analysis. The human and bacterial proteins identified following LESA MS are summarized in Table 1 (details of protein assignments are given in the SI, Fig. S6–S22 and Tables S2–S18). Five of the human proteins identified (β-defensin 4A, elafin, S100A7, S100A8, and S100A9) are antimicrobial peptides (AMPs) and four (S100A6, S100A7, S100A8, and S100A9) belong to the low molecular weight S100 family of calcium-binding proteins.

S100 proteins are present in the human body in both intracellular and extracellular forms. In its native state, S100A6 (also known as calcyclin) exists as a homodimer binding two Ca^{2+} ions. The protein was detected as a monomer in our experiments as a result of the denaturing extraction solvent system. S100A6 is known to be overexpressed in skin melanomas[44]. It is also known that S100A6 is upregulated by epidermal growth factor (EGF) and foetal calf serum[45,46], both of which are used in the construction of the skin model[35] and may explain the presence of S100A6 in the mass spectra. S100A7 (also known as psoriasin) is a calcium and zinc-binding protein expressed by normal cultured (and malignant) keratinocytes[47], and therefore the presence on the skin models might be expected. Studies have shown that this protein has antimicrobial activity against *Escherichia coli*, whilst also targeting *S. aureus*, *S. epidermidis,* and *P. aeruginosa,* albeit less effectively[48]. S100A8 (calgranulin A) and S100A9 (calgranulin B) typically exist as the heterodimer calprotectin, which exhibits antimicrobial activity, but both are also known to act separately[49,50]. Both proteins were detected (as monomers) in the mass spectra. S100A9 was identified in two different forms. One form was truncated at the N-terminus and the sequence was identified with serine acetylation at the (new) N-terminus. The second form was the full length protein, with S-nitrosylated cysteine at position 3, as indicated by UniProt. Elevated expression of calprotectin may be induced by skin dehydration, i.e., transepithelial water loss (TEWL)[51]. TEWL is known to be greater for skin models when compared to real skin[35], and this may explain the high abundance of S100A8 and S100A9 in the mass spectra. The function of S100A8/A9 is also believed to be important in wound healing tissue[52], however S100A8 was detected in all of the Labskin samples, including non-wounded controls.

Proteins identified outside of the S100 family included β-defensin 4A, elafin, and ubiquitin. Antimicrobial β-defensin 4A (or β-defensin 2) protects the skin from gram-positive and gram-negative bacteria[53]. This protein contains three disulphide bonds, specific to the family of defensins[54]. The high number of disulphide bonds and relatively short amino acid sequence resulted in poor sequence coverage (7.5%) but with a high confidence in the manual fragment assignment (see SI, Fig. S6, Table S2). Peaks corresponding to β-defensin 4A were detected in all mass spectra regardless of the presence of wound or infection. The expression of β-defensin 4A has been shown to be upregulated in wounded and infected skin[55]. In our experiment, the abundance of β-defensin 4A ions was higher in the control samples (both intact and wounded) when compared to the infected and wounded skin models. The higher abundance suggests that this protein may be a potential biomarker for infected tissue; however, validation experiments would be required. Elafin (skin antileucoprotease—SKALP, elastase specific inhibitor) is a peptide whose functions include antimicrobial activity and inhibition of proteases[56]. Its expression has been shown to be upregulated during inflammation, and it is a biomarker for graft-versus-host disease[57]. In our experiments, elafin was only detected in the mass spectra obtained from non-wounded skin samples which had been inoculated with *P. aeruginosa*. Finally, ubiquitin was observed in the majority of the acquired mass spectra. Ubiquitin is also commonly observed in the LESA mass spectra of tissue sections[58].

In terms of bacterial proteins, both extracellular and intracellular proteins were identified, in agreement with our previous findings for *S. aureus* MSSA476 and *P. aeruginosa* PS1054 grown on agar[12]. For both *S. aureus* strains, peaks were detected which correspond to two extracellular toxins – phenol-soluble modulin α3 and δ-hemolysin. These highly cytolytic peptides belong to the group of phenol-soluble modulins specific to *S. aureus* species[59]. Top-down LESA MS revealed that both had formylated methionine at their N-termini. The accumulation of formylated δ-hemolysin is known to occur during the post-exponential growth phase[60], reflecting our findings that the abundance of this protein increases with incubation time.

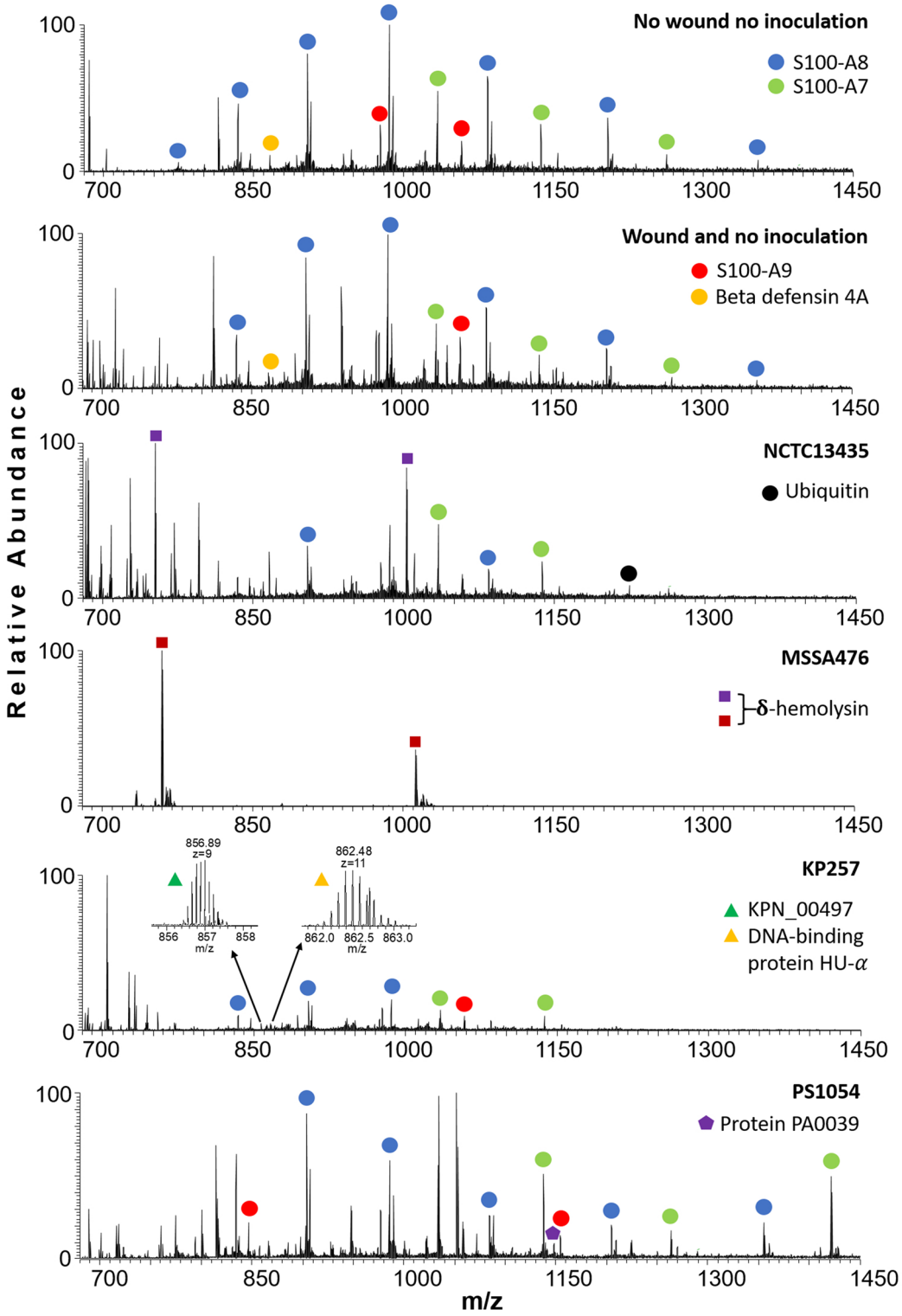


Figure 2. LESA mass spectra obtained from the intact control, wounded control, and Labskin samples wounded and infected with *S. aureus* NCTC13435, *S. aureus* MSSA476, *K. pneumoniae* KP257 and *P. aeruginosa* PS1054 after 48 h of incubation. Identified proteins are labelled. No bacterial proteins were detected in the control samples.

In the wounded skin samples infected with *K. pneumoniae* KP257, two proteins were identified. Firstly, an uncharacterized protein predicted on the basis of gene KPN_00497 was identified. The protein was detected with

Observed monoisotopic mass (Da)	Theoretical monoisotopic mass (Da)	Mass difference (ppm)	Protein name	Uniprot accession number	Modification	Sequence coverage (%)	Sample type (*I* inoculated, *C* control)
Human skin proteins							
4,325.1296	4,325.1371	− 1.7387	β-defensin 4A	O15263	-Signal peptide (1–23), disulfide bonds (31–60, 38–53, 43–61)	12	I/C
5,994.8239	5,994.8202	0.6205	Elafin	P19957	Disulfide bonds (76–105, 83–109, 92–104, 98–113)	9	I
8,559.6029	8,559.6167	− 1.6146	Ubiquitin	P62987	–	16	I/C
10,027.3088	10,027.2911	1.7632	S100-A6	P06703	-Met; acetylation of N-terminus (A)	8	I/C
10,827.6373	10,827.6492	− 1.1046	S100-A8	P05109	–	53	I/C
11,360.4787	11,360.5189	− 3.5403	S100-A7	P31151	-Met; acetylation of N-terminus (S), E27D substitution	26	I/C
12,682.2879	12,681.2806	− 0.4274	S100-A9	P06702	Acetylation of N-terminus (S) (truncated form) S-nitrosylation (C3) (full length form))	17	I/C
Observed monoisotopic mass (Da)	**Theoretical monoisotopic mass (Da)**	**Mass difference (ppm)**	**Protein name**	**Uniprot accession number**	**Modification**	**Sequence coverage (%)**	**Hours of incubation**
S. aureus NCTC13435							
2,633.4082	2,633.4080	0.0573	Phenol-soluble modulin α3	P0C805	fMet	95	48
3,004.6109	3,004.6302	− 0.1664	δ-hemolysin	Q2FWM8	fMet	60	48
S. aureus MSSA476							
2,633.4094	2,633.4080	0.5240	Phenol-soluble modulin α3	P0C805	fMet	95	48
3,034.6413	3,034.6413	− 0.0066	δ-Hemolysin	Q6G7S2	fMet, G10S substitution	88	48
K. pneumoniae KP257							
7,698.9926	7,698.9638	3.7382	KPN_00497	A6T5S6	-Signal peptide (1–19), R49K substitution	19	48
9,471.1664	9,471.1468	2.0673	DNA-binding protein HU-α	A6TGQ7	–	30	48
P. aeruginosa PS1054							
5,731.9826	5,731.9816	0.1814	PA0039	Q9I793	-Signal peptide (1–21)	13	48
8,557.5098	8,557.5077	0.2431	PA4739	Q9HV60	-Signal peptide (1–32)	44	72
9,081.0514	9,081.0445	0.7576	DNA-binding protein HU-β	P05384	–	26	72

Table 1. Summary of human skin and bacterial proteins identified from Labskin samples.

a cleaved signal peptide (1–19) and a R → K substitution at position 49. Secondly, DNA-binding protein HU-α, which is involved in stabilization of DNA under extreme environmental conditions, was identified.

Three bacterial proteins were identified from the samples infected with *P. aeruginosa* PS1054: two uncharacterized proteins and DNA-binding protein HU-β. All have been identified previously from colonies of PS1054 growing on agar[12]. The first uncharacterized protein is predicted on the basis of gene PA0039 with signal peptide (1–21) cleaved. PA0039 was the only protein detected in the mass spectra after 48 h that was sufficiently abundant to perform top-down tandem mass spectrometry (MS/MS) analysis. The second uncharacterized protein is predicted on the basis of gene PA4739. Information about this protein available from UniProt suggests that the signal peptide is cleaved after amino acid 25, however our results confirm that the signal peptide cleaves after amino acid 32, as observed previously[12]. PA4739 and DNA-binding protein HU-β were both detected after 72 h of incubation.

Mass difference between the δ-hemolysins of *S. aureus* strains. The sequence of δ-hemolysin differs between the two *S. aureus* strains NCTC13435 and MSSA476 as a result of a glycine to serine substitution at position 10 (G10S). This is an allelic variant related to a mutation in the hld gene and is characteristic of certain ST1 and ST59 strains of *S. aureus*, including ST1 strain MSSA476[61]. The G → S substitution results in a mass difference, Δm = + 30.0105 Da, which can be used to differentiate between the two strains by LESA MS. (The site of

substitution can be confirmed by LESA tandem mass spectrometry (MS/MS), see Fig. S5, SI). That is, LESA MS presents a rapid tool for direct differentiation between allelic variants in strains of *S. aureus*.

Conclusion

We have demonstrated the rapid identification of ESKAPE pathogens, including different strains of *S. aureus*, directly from infected wounds in 3D in vitro living skin equivalents. Our approach involves LESA mass spectrometry, an ambient technique which here makes use of ethanol-based solvents. Intact proteins from both human skin and the infecting bacteria were identified with high confidence. Bacterial proteins could be identified within minutes once visible signs of infection were apparent (i.e., 48 h after inoculation of the skin model with bacteria). Detection of allelic variants of the protein δ-hemolysin enabled differentiation between the two strains of *S. aureus*. Future work will focus on identification of proteins from mixtures of pathogens in infected wounds. Furthermore, improvements in mass spectrometry technology will likely be accompanied by improved numbers of proteins detected from the various microbial species.

Materials and methods

Materials. Analytical grade acetonitrile, ethanol, water and formic acid were purchased from Fisher Scientific (Loughborough, UK). LB broth [yeast extract (VWR, Lutterworth, UK), peptone (Sigma Aldrich, Gillingham, UK) and sodium chloride (Fisher Scientific, Loughborough, UK)] and LB agar [LB broth with added agar (Appleton Woods, Birmingham, UK)] were used for bacterial culture. In vitro 3D skin models "Labskin" were purchased from Innovenn (Sand Hutton, UK). *S. aureus* NCTC13435 was obtained from Public Health England (Porton Down, UK) via Innovenn. *S. aureus* MSSA476, *P. aeruginosa* PS1054 were obtained from Mark Webber (Institute of Microbiology and Infection, University of Birmingham, UK) and *K. pneumoniae* KP257 was obtained from Willem van Schaik (Institute of Microbiology and Infection, University of Birmingham, UK).

Skin sample preparation. On arrival, Labskin samples were transferred into a new 12-well plate with fresh proprietary Labskin medium (Innovenn, Sand Hutton, UK) and incubated overnight at 37 °C, 5% CO_2 and > 95% relative humidity. On day 2, the medium was replaced with the fresh Labskin medium, and the samples were wounded with a scalpel blade (Fig. 1). The wounded samples were infected with *S. aureus* NCTC13435 (n = 9), *S. aureus* MSSA476 (n = 3), *K. pneumoniae* KP257 (n = 3) and *P. aeruginosa* PS1054 (n = 8). Full details of the infection process and calculation of infectious dose is given in *Supplementary Information*. Three control samples were included: intact (n = 5), wounded but not inoculated (n = 5) and inoculated but not wounded for *S. aureus* NCTC13435 (n = 4), *S. aureus* MSSA476 (n = 2), *K. pneumoniae* KP257 (n = 2) and *P. aeruginosa* PS1054 (n = 7). All samples (control and infected/wounded) were incubated for 24, 48, 72 and 96 h after the infection at 37 °C, 5% CO_2 and > 95% relative humidity prior to MS analysis.

LESA mass spectrometry. LESA MS was performed using an Advion Triversa Nanomate (Advion, Ithaca, NY, USA) coupled to a Thermo Orbitrap Elite mass spectrometer or a Q Exactive HF mass spectrometer (both Thermo Fisher Scientific, Bremen, Germany) as described previously[11,12]. The solvent system used for LESA MS comprised ethanol (Fisher Scientific, Loughborough, UK), water (Fisher Scientific, Loughborough, UK) and formic acid (Fisher Scientific, Loughborough, UK) (60:35:5). Labskin samples were placed in the 60 mm Petri dishes adjacent to the half of the 96-well microtiter plate. The robotic arm of the Triversa Nanomate aspirated 3 µL of the extraction solvent system from the microtiter plate. Next, the robotic arm was relocated to a position above the sample, descended above the Labskin surface such that the formation of the liquid microjunction was allowed and dispensed 2 µL of the solvent system. After the sampling process, 2.5 µL of the solvent system with extracted analytes was re-aspirated back into the pipette tip and introduced into the mass spectrometer via chip-based nano-electrospray system at gas pressure 0.3 psi and a tip voltage 1.75 kV. The Triversa Nanomate was controlled with the advanced user interface (AUI) in the Chipsoft software 8.3.1. (Advion, Ithaca, NY, USA). The mass spectra were recorded for at least 3 min in positive ion mode in full scan mode, 600–2000 *m/z,* at resolution 120,000 at 400 *m/z*. Top-down MS/MS analysis of bacterial proteins was carried out using both collision induced dissociation (CID) in the ion trap using helium gas at a normalized collision energy 35% (Orbitrap Elite) and higher energy collision dissociation (HCD) at a varying collision energy between 15–60 eV depending on the protein charge state (Q Exactive HF). When performing fragmentation, each scan comprised 30 co-added microscans (Orbitrap Elite) or 5 co-added microscans (Q Exactive HF) and data were recorded for at least 5 min.

Data processing and protein identification. Top-down identification of proteins was performed by use of ProSight 4.1 software (Thermo Fisher Scientific, Bremen, Germany) as described previously[12]. Pre-built ProSight top-down protein databases for each model organism were downloaded from the Database Warehouse of Proteinaceous website (https://www.proteinaceous.net), which also includes link to Uniprot proteomes with the Proteome ID listed. Databases included *Homo sapiens* (UP000005640, 71,599 entries), *Staphylococcus aureus* NCTC 8325 (UP000008816, 2,889 entries), *Klebsiella pneumoniae* ATCC 700721 (UP000000265, 5,126 entries), and *Pseudomonas aeruginosa* ATCC 15692/PA01 (UP000002438, 5,563 entries). MS/MS spectra were imported into ProSightPC in profile mode and deconvoluted by the THRASH algorithm at default settings and S/N 3. Search parameters in absolute mass search mode accounted for all post-translational modifications and delta-mass (Δm) mode on for locating the unknown modifications or possible mutations. The search window width was 1,000 Da, with initial fragment mass tolerance ± 15 ppm and minimum matching fragments number set to 4. All identified protein sequences were subsequently checked with the Sequence Gazer function of ProSight software followed by manual peak assignment with the fragment tolerance narrowed to ± 5 ppm.

Data availability

Supplementary data supporting this research is openly available from the University of Birmingham data archive at https://doi.org/10.25500/edata.bham.00000489.

Received: 11 December 2019; Accepted: 19 June 2020
Published online: 17 July 2020

References

1. WHO. *Injuries and Violence: The Facts* (World Health Organization, Geneva, 2014).
2. Yun, H. C., Murray, C. K., Nelson, K. J. & Bosse, M. J. Infection after orthopaedic trauma: prevention and treatment. *J. Orthop. Trauma* **30**, S21–S26. https://doi.org/10.1097/bot.0000000000000667 (2016).
3. Tribble, D. R. *et al.* Infection-associated clinical outcomes in hospitalized medical evacuees after traumatic injury: trauma infectious disease outcome study. *J. Trauma* **71**, S33–S42. https://doi.org/10.1097/TA.0b013e318221162e (2011).
4. Santajit, S. & Indrawattana, N. Mechanisms of antimicrobial resistance in ESKAPE pathogens. *Biomed. Res. Int.* **2016**, 2475067. https://doi.org/10.1155/2016/2475067 (2016).
5. Mende, K. *et al.* Microbiology of combat-related extremity wounds: trauma infectious disease outcomes study. *Diagn. Microbiol. Infect. Dis.* **94**, 173–179. https://doi.org/10.1016/j.diagmicrobio.2018.12.008 (2019).
6. Holland, R. D. *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 Commun. Mass Spectrom.* **10**, 1227–1232. https://doi.org/10.1002/(sici)1097-0231(19960731)10:10<1227::Aid-rcm659>3.0.Co;2-6 (1996).
7. Claydon, M. A., Davey, S. N., Edwards-Jones, V. & Gordon, D. B. The rapid identification of intact microorganisms using mass spectrometry. *Nat. Biotechnol.* **14**, 1584–1586. https://doi.org/10.1038/nbt1196-1584 (1996).
8. Hou, T.-Y., Chiang-Ni, C. & Teng, S.-H. Current status of MALDI-TOF mass spectrometry in clinical microbiology. *J. Food Drug Anal.* **27**, 404–414. https://doi.org/10.1016/j.jfda.2019.01.001 (2019).
9. Lin, J. F., Ge, M. C., Liu, T. P., Chang, S. C. & Lu, J. J. A simple method for rapid microbial identification from positive monomicrobial blood culture bottles through matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *J. Microbiol. Immunol. Infect.* **51**, 659–665. https://doi.org/10.1016/j.jmii.2017.03.005 (2018).
10. Yang, J. Y. *et al.* Primer on agar-based microbial imaging mass spectrometry. *J. Bacteriol.* **194**, 6023–6028. https://doi.org/10.1128/Jb.00823-12 (2012).
11. 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).
12. Kocurek, K. I., Stones, L., Bunch, J., May, R. C. & Cooper, H. J. Top-down LESA mass spectrometry protein analysis of gram-positive and gram-negative bacteria. *J. Am. Soc. Mass Spectrom.* **28**, 2066–2077. https://doi.org/10.1007/s13361-017-1718-8 (2017).
13. Kocurek, K. I., May, R. C. & Cooper, H. J. Application of high-field asymmetric waveform ion mobility separation to LESA mass spectrometry of bacteria. *Anal. Chem.* **91**, 4755–4761 (2019).
14. Kertesz, V., Ford, M. J. & Van Berkel, G. J. Automation of a surface sampling probe/electrospray mass spectrometry system. *Anal. Chem.* **77**, 7183–7189. https://doi.org/10.1021/ac0510742 (2005).
15. Meetani, M. A., Shin, Y.-S., Zhang, S., Mayer, R. & Basile, F. Desorption electrospray ionization mass spectrometry of intact bacteria. *J. Mass Spectrom.* **42**, 1186–1193. https://doi.org/10.1002/jms.1250 (2007).
16. Song, Y., Talaty, N., Tao, W. A., Pan, Z. & Cooks, R. G. Rapid ambient mass spectrometric profiling of intact, untreated bacteria using desorption electrospray ionization. *Chem. Commun.* https://doi.org/10.1039/B615724F (2007).
17. Jackson, A. U. *et al.* Targeted metabolomic analysis of *Escherichia coli* by desorption electrospray ionization and extractive electrospray ionization mass spectrometry. *Anal. Biochem.* **375**, 272–281. https://doi.org/10.1016/j.ab.2008.01.011 (2008).
18. 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. https://doi.org/10.1039/b900069k (2009).
19. Zhang, J. I. *et al.* Rapid direct lipid profiling of bacteria using desorption electrospray ionization mass spectrometry. *Int. J. Mass Spectrom.* **301**, 37–44. https://doi.org/10.1016/j.ijms.2010.06.014 (2011).
20. Rodrigues, J. P., Prova, S. S., Moraes, L. A. B. & Ifa, D. R. Characterization and mapping of secondary metabolites of *Streptomyces* sp. from caatinga by desorption electrospray ionization mass spectrometry (DESI-MS). *Anal. Bioanal. Chem.* **410**, 7135–7144. https://doi.org/10.1007/s00216-018-1315-0 (2018).
21. Araújo, F. D. S. *et al.* Desorption electrospray ionization mass spectrometry imaging reveals chemical defense of *Burkholderia seminalis* against cacao pathogens. *RSC Adv.* **7**, 29953–29958. https://doi.org/10.1039/C7RA03895J (2017).
22. Angolini, C. F. F. *et al.* Direct protocol for ambient mass spectrometry imaging on agar culture. *Anal. Chem.* **87**, 6925–6930. https://doi.org/10.1021/acs.analchem.5b01538 (2015).
23. Pruski, P. *et al.* Medical swab analysis using desorption electrospray ionization mass spectrometry: a noninvasive approach for mucosal diagnostics. *Anal. Chem.* **89**, 1540–1550. https://doi.org/10.1021/acs.analchem.6b03405 (2017).
24. Schäfer, K.-C. *et al.* In vivo, in situ tissue analysis using rapid evaporative ionization mass spectrometry. *Angew. Chem. Int. Ed.* **48**, 8240–8242. https://doi.org/10.1002/anie.200902546 (2009).
25. Balog, J. *et al.* Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. *Sci. Transl. Med.* **5**, 194ra193. https://doi.org/10.1126/scitranslmed.3005623 (2013).
26. Strittmatter, N. *et al.* Analysis of intact bacteria using rapid evaporative ionisation mass spectrometry. *Chem. Commun.* **49**, 6188–6190. https://doi.org/10.1039/C3CC42015A (2013).
27. Strittmatter, N. *et al.* Characterization and identification of clinically relevant microorganisms using rapid evaporative ionization mass spectrometry. *Anal. Chem.* **86**, 6555–6562. https://doi.org/10.1021/ac501075f (2014).
28. Hamid, A. M. *et al.* Rapid discrimination of bacteria by paper spray mass spectrometry. *Anal. Chem.* **86**, 7500–7507. https://doi.org/10.1021/ac501254b (2014).
29. Wei, P. *et al.* Analysis of bacteria using zero volt paper spray. *Anal. Methods* **8**, 1770–1773. https://doi.org/10.1039/C6AY00063K (2016).
30. Chamberlain, C. A., Rubio, V. Y. & Garrett, T. J. Strain-level differentiation of bacteria by paper spray ionization mass spectrometry. *Anal. Chem.* **91**, 4964–4968. https://doi.org/10.1021/acs.analchem.9b00330 (2019).
31. Lanekoff, I., Geydebrekht, O., Pinchuk, G. E., Konopka, A. E. & Laskin, J. Spatially resolved analysis of glycolipids and metabolites in living *Synechococcus* sp. PCC 7002 using nanospray desorption electrospray ionization. *Analyst* **138**, 1971–1978. https://doi.org/10.1039/c3an36716a (2013).
32. Watrous, J. *et al.* Metabolic profiling directly from the Petri dish using nanospray desorption electrospray ionization imaging mass spectrometry. *Anal. Chem.* **85**, 10385–10391. https://doi.org/10.1021/ac4023154 (2013).
33. Uwakweh, A. O., Mwangi, J. N., Todd, D., Jia, Z. & Chiu, N. H. L. Nanospray desorption electrospray ionization mass spectrometry of untreated and treated probiotic *Lactobacillus reuteri* cells. *Anal. Bioanal. Chem.* **410**, 4237–4245. https://doi.org/10.1007/s00216-018-1071-1 (2018).

34. Hsu, C.-C. *et al.* Real-time metabolomics on living microorganisms using ambient electrospray ionization flow-probe. *Anal. Chem.* **85**, 7014–7018. https://doi.org/10.1021/ac401613x (2013).
35. Holland, D. B., Bojar, R. A., Jeremy, A. H., Ingham, E. & Holland, K. T. Microbial colonization of an in vitro model of a tissue engineered human skin equivalent—a novel approach. *FEMS Microbiol. Lett.* **279**, 110–115. https://doi.org/10.1111/j.1574-6968.2007.01021.x (2008).
36. Harvey, A. *et al.* MALDI-MSI for the analysis of a 3D tissue-engineered psoriatic skin model. *Proteomics* **16**, 1718–1725. https://doi.org/10.1002/pmic.201600036 (2016).
37. Lewis, E. E. L., Barrett, M. R. T., Freeman-Parry, L., Bojar, R. A. & Clench, M. R. Examination of the skin barrier repair/wound healing process using a living skin equivalent model and matrix-assisted laser desorption-ionization-mass spectrometry imaging. *Int. J. Cosmet. Sci.* **40**, 148–156. https://doi.org/10.1111/ics.12446 (2018).
38. Mitchell, C. A., Donaldson, M., Francese, S. & Clench, M. R. MALDI MSI analysis of lipid changes in living skin equivalents in response to emollient creams containing palmitoylethanolamide. *Methods* **104**, 93–100. https://doi.org/10.1016/j.ymeth.2016.02.001 (2016).
39. Mitchell, C. A., Long, H., Donaldson, M., Francese, S. & Clench, M. R. Lipid changes w thin the epidermis of living skin equivalents observed across a time-course by MALDI-MS imaging and profiling. *Lipids Health Dis.* **14**, 84. https://doi.org/10.1186/s12944-015-0089-z (2015).
40. Russo, C., Lewis, E. E. L., Flint, L. & Clench, M. R. Mass spectrometry imaging of 3D tissue models. *Proteomics* **18**, e1700462. https://doi.org/10.1002/pmic.201700462 (2018).
41. Russo, C. *et al.* Quant tative investigation of terbinafine hydrochloride absorption into a living skin equivalent model by MALDI-MSI. *Anal. Chem.* **90**, 10031–10038. https://doi.org/10.1021/acs.analchem.8b02648 (2018).
42. Catherman, A. D., Skinner, O. S. & Kelleher, N. L. Top down proteomics: facts and perspectives. *Biochem. Biophys. Res. Commun.* **445**, 683–693. https://doi.org/10.1016/j.bbrc.2014.02.041 (2014).
43. Wisplinghoff, H. In *Infectious Diseases* 4th edn (eds Cohen, J., Powderly, W. G. *et al.*) 1579–1599 (Elsevier, Amsterdam, 2017).
44. Halawi, A., Abbas, O. & Mahalingam, M. S100 proteins and the skin: a review. *J. Eur. Acad. Dermatol. Venereol.* **28**, 405–414. https://doi.org/10.1111/jdv.12237 (2014).
45. Donato, R., Sorci, G. & Giambanco, I. S100A6 protein: functional roles. *Cell. Mol. Life Sci.* **74**, 2749–2760. https://doi.org/10.1007/s00018-017-2526-9 (2017).
46. Ghezzo, F., Lauret, E., Ferrari, S. & Baserga, R. Growth factor regulation of the promoter for calcyclin, a growth-regulated gene. *J. Biol. Chem.* **263**, 4758–4763 (1988).
47. Watson, P. H., Leygue, E. R. & Murphy, L. C. Psoriasin (S100A7). *Int. J. Biochem. Cell Biol.* **30**, 567–571. https://doi.org/10.1016/s1357-2725(97)00066-6 (1998).
48. Glaser, R. *et al.* Antimicrobial psoriasin (S100A7) protects human skin from *Escherichia coli* infection. *Nat. Immunol.* **6**, 57–64. https://doi.org/10.1038/ni1142 (2005).
49. Kenneth, H. *et al.* Anti-infective protective properties of S100 calgranulins. *Antiinflamm. Antiallergy Agents Med. Chem.* **8**, 290–305. https://doi.org/10.2174/187152309789838975 (2009).
50. Lesniak, W. & Graczyk-Jarzynka, A. The S100 proteins in epidermis: topology and function. *Biochim. Biophys. Acta* **2563–2572**, 2015. https://doi.org/10.1016/j.bbagen.2015.09.015 (1850).
51. Zhong, A. *et al.* S100A8 and S100A9 are induced by decreased hydration in the epidermis and promote fibroblast activation and fibrosis in the dermis. *Am. J. Pathol.* **186**, 109–122. https://doi.org/10.1016/j.ajpath.2015.09.005 (2016).
52. Kerkhoff, C. *et al.* Novel insights into the role of S100A8/A9 in skin biology. *Exp. Dermatol.* **21**, 822–826. https://doi.org/10.1111/j.1600-0625.2012.01571.x (2012).
53. Schroder, J. M. & Harder, J. Antimicrobial skin peptides and proteins. *Cell. Mol. Life Sci.* **63**, 469–486. https://doi.org/10.1007/s00018-005-5364-0 (2006).
54. Schäfer-Korting, M. & Rolff, J. In *Emerging Nanotechnologies in Immunology* (eds Shegokar, R. & Souto, E. B.) 23–45 (Elsevier, Amsterdam, 2018).
55. Wang, B. *et al.* IL-1β-induced protection of keratinocytes against *Staphylococcus aureus*-secreted proteases Is mediated by human β-defensin 2. *J. Investig. Dermatol.* **137**, 95–105. https://doi.org/10.1016/j.jid.2016.08.025 (2017).
56. Wiedow, O., Schroder, J. M., Gregory, H., Young, J. A. & Christophers, E. Elafin: an elastase-specific inhibitor of human skin. Purification, characterization, and complete amino acid sequence. *J. Biol. Chem.* **265**, 14791–14795 (1990).
57. Paczesny, S. *et al.* Elafin is a biomarker of graft-versus-host disease of the skin. *Sci. Transl. Med.* **2**, 13ra12. https://doi.org/10.1126/scitranslmed.3000406 (2010).
58. Griffiths, R. L., Creese, A. J., Race, A. M., Bunch, J. & Cooper, H. J. LESA FAIMS mass spectrometry for the spatial profiling of proteins from tissue. *Anal. Chem.* **88**, 6758–6766. https://doi.org/10.1021/acs.analchem.6b01060 (2016).
59. Peschel, A. & Otto, M. Phenol-soluble modulins and staphylococcal infection. *Nat. Rev. Microbiol.* **11**, 667–673. https://doi.org/10.1038/nrmicro3110 (2013).
60. Somerville, G. A. *et al.* Synthesis and deformylation of *Staphylococcus aureus* delta-toxin are linked to tricarboxylic acid cycle activity. *J. Bacteriol.* **185**, 6686–6694. https://doi.org/10.1128/JB.185.22.6686-6694.2003 (2003).
61. Cheung, G. Y. C. *et al.* Functional characteristics of the *Staphylococcus aureus* δ-toxin allelic variant G10S. *Sci. Rep.* **5**, 18023. https://doi.org/10.1038/srep18023 (2015).

Acknowledgements

The authors thank Emma Keen for technical help. H.J.C. is an EPSRC Established Career Fellow (EP/L023490/1 and EP/S002979/1). J.H. is funded by the EPSRC Physical Sciences for Health Doctoral Training Centre (EP/L016346/1). 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.

Author contributions

H.J.C. led the research and designed the experiment. J.H. performed the experiment and analysed the data. J.H., H.J.C., R.C.M. and I.B.S wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-68233-6.

Correspondence and requests for materials should be addressed to H.J.C.

ACS Partner Journal

Journal of the American Society for Mass Spectrometry

pubs.acs.org/jasms

Research Article

Liquid Extraction Surface Analysis Mass Spectrometry of ESKAPE Pathogens

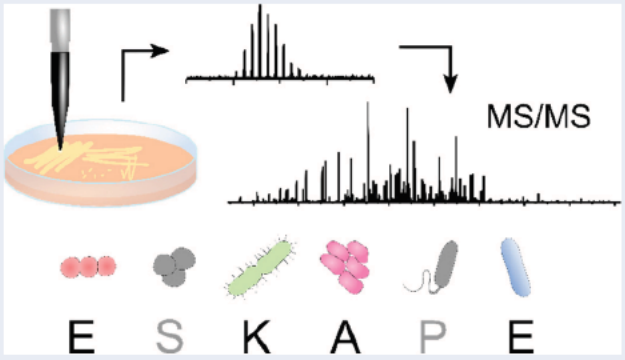
Jana Havlikova, Robin C. May, Iain B. Styles, and Helen J. Cooper*

Cite This: *J. Am. Soc. Mass Spectrom.* 2021, 32, 1345−1351

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The ESKAPE pathogens (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa,* and *Enterobacter cloacae*) represent clinically important bacterial species that are responsible for most hospital acquired drug resistant infections; hence, the need for rapid identification is of high importance. Previous work has demonstrated the suitability of liquid extraction surface analysis mass spectrometry (LESA MS) for the direct analysis of colonies of two of the ESKAPE pathogens (*Staphylococcus aureus* and *Pseudomonas aeruginosa*) growing on agar. Here, we apply LESA MS to the remaining four ESKAPE species (*E. faecium* E745, *K. pneumoniae* KP257, *A. baumannii* AYE, and *E. cloacae* S11) as well as *E. faecalis* V583 (a close relative of *E. faecium*) and a clinical isolate of *A. baumannii* AC02 using an optimized solvent sampling system. In each case, top down LESA MS/MS was employed for protein identification. In total, 24 proteins were identified from 37 MS/MS spectra by searching against protein databases for the individual species. The MS/MS spectra for the identified proteins were subsequently searched against multiple databases from multiple species in an automated data analysis workflow with a view to determining the accuracy of identification of unknowns. Out of 24 proteins, 19 were correctly assigned at the protein and species level, corresponding to an identification success rate of 79%.



INTRODUCTION

The ESKAPE pathogens are six clinically relevant bacterial species *Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa,* and *Enterobacter* spp., of which two are Gram positive (*E. faecium, S. aureus*) and the remainder are Gram negative.[1] The ESKAPE microbes are responsible for most hospital acquired (nosocomial) infections,[2,3] and their antibiotic resistance is rising.[4] In fact, WHO reports at least 700 000 deaths annually due to infections by the drug resistant strains,[5] affecting both developed and developing countries.[6] Development of improved tools for rapid and accurate identification of these bacteria and hence tailored treatment of patients is therefore of high importance.

The use of mass spectrometry (MS) for the analysis and identification of bacteria is well established. Currently, the gold standard, FDA approved mass spectrometry (MS) approach for identification of microbes is matrix assisted laser desorption ionization time of flight (MALDI TOF) MS together with dedicated software for spectral fingerprinting.[7] Nevertheless, MALDI TOF MS has some drawbacks for bacterial analysis including sample preparation requirements and the fact that analysis takes place under vacuum conditions, precluding analysis of live colonies. Ambient ionization MS techniques overcome these limitations.[8−16] One such ambient technique is liquid extraction surface analysis (LESA) MS, a liquid microjunction sampling tool, based on diffusion of analytes into a droplet of solvent.[17] The droplet is subsequently introduced into the mass spectrometer via chip based nano electrospray (nanoESI).

We have been developing LESA MS as a tool for the top down analysis of proteins directly from bacteria growing on solid substrates. This approach is different to the MALDI TOF MS diagnostic approach, because it focuses on identification of bacterial proteins rather than spectral matching. Another promising approach, which focuses on analysis of intact proteins, is liquid chromatography top down proteomics (LC TDP). Recent work by Chamot Rooke and co workers demonstrated identification of >200 proteins and >500 proteoforms by LC TDP.[18]

Our pilot LESA MS study focused on *Escherichia coli* K12,[19] while later work identified 39 proteins from multiple species including 2 of the ESKAPE pathogens *Staphylococcus aureus* and *Pseudomonas aeruginosa*.[20] More recent work focused on

Special Issue: Focus: Top Down Proteomics: Technology Advances and Biomedical Applications

Received: December 16, 2020
Revised: February 22, 2021
Accepted: February 22, 2021
Published: March 1, 2021

the analysis of bacteria growing on three dimensional (3D) living skin equivalents, with the aim of simultaneous identification of bacterial and human skin proteins.[21] *In vitro* skin models were inoculated with three ESKAPE pathogens, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and two different strains of *Staphylococcus aureus*. In each case, LESA MS analysis showed that one to two bacterial proteins could be detected after 48 h and that both bacterial and human skin proteins were observed in the same mass spectra.

Here, we extend the top down LESA MS approach to the remaining three ESKAPE pathogens—*Enterococcus faecium* E745 (and its close relative *Enterococcus faecalis* V583, commonly found in the hospital environment), *Acinetobacter baumannii* (the reference strain AYE and a clinical isolate AC02), and *Enterobacter cloacae* S11—as well as expanding the study of *Klebsiella pneumoniae* KP257 (previously only considered in the context of *in vitro* skin models). A key consideration if LESA MS is to find use as a diagnostic tool is a universal sampling approach, i.e., a single solvent system suitable for all Gram positive and Gram negative species. We first optimized the solvent system to enable successful protein extractions from all of the ESKAPE pathogens, before applying MS/MS for top down protein identification. In total, 24 proteins were identified from 37 MS/MS spectra. Finally, we compare protein identification from searches against multiple databases and associated success of LESA diagnosis.

MATERIALS AND METHODS

Materials. Analytical grade acetonitrile, water, and formic acid were purchased from Fisher Scientific (Loughborough, UK). Bacteriological agar was purchased from Appleton Woods (Birmingham, UK). LB broth (yeast extract (VWR, Lutterworth, UK), peptone (Sigma Aldrich, Gillingham, UK) and sodium chloride (Fisher Scientific, Loughborough, UK)), BHI broth (dehydrated brain heart infusion (VWR, Lutter worth, UK)), LB agar (LB broth with added bacteriological agar), and BHI agar (BHI broth with added bacteriological agar) were prepared for culturing bacterial species. Bacterial strains *E. faecium* E745, *E. faecalis* V583, and *K. pneumoniae* KP257 were obtained from Willem van Schaik (Institute of Microbiology and Infection (IMI), University of Birmingham, UK), *S. aureus* MSSA476 and *P. aeruginosa* PS1054 were obtained from Mark Webber (Quadram Institute, Norwich, UK), *A. baumannii* AYE and AC02 were obtained from Jessica Blair (IMI, University of Birmingham, UK), and *E. cloacae* S11 was obtained from Allan McNally (IMI, University of Birmingham, UK).

Preparation of Bacterial Colonies. Liquid cultures of each species were prepared. Approximately 1 μL of bacteria from the frozen glycerol stock was resuspended in 5−10 mL of liquid broth. Liquid cultures were incubated up to 18 h at shaking conditions (200 rpm) at 37 °C and plated on the solid agar media in 60 mm Petri dishes. Agar plates were incubated at static conditions and 37 °C overnight. ESKAPE pathogens were cultured in LB broth and LBA except for the *Enterococci* strains, which required culturing in BHI liquid broth and BHI agar. Two biological replicates were prepared for each species.

LESA MS Analysis. LESA MS of proteins from bacterial colonies growing on agar plates was performed by use of the Advion Triversa Nanomate (Advion, Ithaca, NY) coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) as described previously.[19,20] For each biological replicate, between 1−11 technical replicates were performed. The optimized extraction solvent system (see Results and Discussion section) comprised acetonitrile, water, and formic acid (60:35:5). The agar plates were placed next to half of a 96 well microtiter plate in the Triversa Nanomate. The robotic pipet of the Triversa Nanomate aspirated 3 μL of the sampling solvent system and relocated to the new position above the sample. The descending pipet tip touched the colony surface and dispensed 2 μL of the solvent system. Subsequently, 2.5 μL of the solvent system containing analytes was reaspirated back to the pipet tip and introduced into the mass spectrometer via chip based nanoESI at a gas pressure of 0.3 psi and a tip voltage of 1.75 kV. The Triversa Nanomate platform was controlled with the Chipsoft software 8.3.1. (Advion, Ithaca, NY, USA).

The mass spectra were acquired for (at least) 3 min in full scan positive ion mode with a mass range of 600−2000 m/z at a resolution of 120 000 at 400 m/z in the Orbitrap mass analyzer. The automatic gain control (AGC) target was 1×10^6 charges. Each MS scan comprised a single microscan. Precursor ions were selected for fragmentation with an isolation window of 3 m/z. Collision induced dissociation (CID) was performed in the ion trap with use of helium gas at a normalized collision energy of 35%. The AGC target was between 5×10^4 and 1×10^5 charges. MS/MS mass spectra were recorded in the Orbitrap for (at least) 5 min at a resolution of 120 000 at 400 m/z, and each scan comprised 30 coadded microscans.

Data Analysis and Identification of Proteins. Top down identification of proteins from fragmentation mass spectra was achieved by use of the ProSight 4.1 software (Thermo Fisher Scientific, Bremen, Germany). The whole organism proteome databases were downloaded in XML format from the UniProt Web site (uniprot.org) for *E. faecium* ATCC BAA 472/TX0016/DO (UP000005269, 3059 entries, 15 638 proteoforms), *E. faecalis* ATCC 700802/V583 (UP000001415, 3240 entries, 17,469 proteoforms), *S. aureus* NCTC8325 (UP000008816, 2889 entries, 14 793 proteo forms), *K. pneumoniae* ATCC 700721 (UP000000265, 5126 entries, 26 531 proteoforms), *A. baumannii* AYE (UP000002446, 3652 entries, 18 616 proteoforms), *P. aeruginosa* ATCC 15692/PA01 (UP000002438, 5563 entries, 29 775 proteoforms), and *E. cloacae* S611 (UP000017834, 3989 entries, 20 832 proteoforms). Each database was constructed as a standard top down database, taking into account cleavage of initial methionines, N terminal acetylation, and N terminal formylation. Single nucleotide polymorphisms (SNPs) and post translational modifications (PTMs) were considered with a maximum of 13 features per sequence and maximum mass of 70 kDa. The MS/MS spectra were deconvoluted using the THRASH algorithm in the ProSight import profile window with default parameters and an S/N set to 2 or 3 depending on the quality of acquired MS/MS mass spectra. The absolute mass search included a delta mass (Δm) option for locating possible unknown post translational modifications and mutations. The search window width was set to 1000 Da with a fragment tolerance of ±15 ppm and the minimum matching fragments number set to 4. Identified protein sequences were checked with the Sequence Gazer function of ProSight software followed by manual peak assignment where the fragment tolerance was narrowed to ±5 ppm.

BLAST (Basic Local Alignment Search Tool) searches were performed as follows: For each protein identified, the protein

https://doi.org/10.1021/jasms.0c00466
J. Am. Soc. Mass Spectrom. 2021, 32, 1345−1351

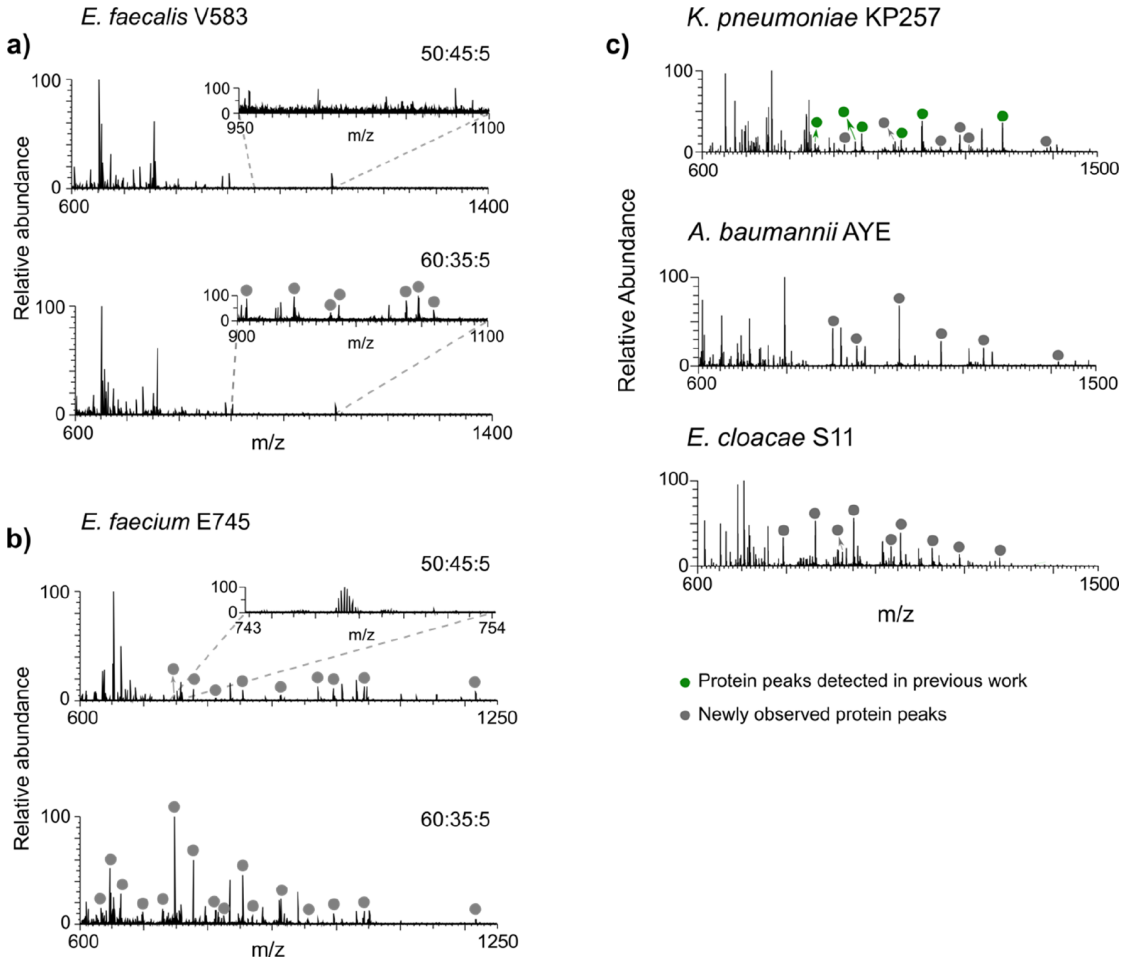


Figure 1. LESA mass spectra of bacterial colonies. Comparison of extraction solvent systems, acetonitrile/water/formic acid 50:45:5 and 60:35:5 for (a) *E. faecalis* V583 and (b) *E. faecium* E745. (c) LESA MS analysis of *K. pneumoniae* KP257, *A. baumannii* AYE, and *E. cloacae* S11 with the 60:35:5 extraction solvent system.

sequence was downloaded from Uniprot in FASTA format and searched against the nonredundant protein sequences (nr) database using the BLASTP algorithm with default parameters (blast.ncbi.nlm.nih.gov).

RESULTS AND DISCUSSION

Optimization of Extraction Solvent System for LESA MS Analysis. In our earlier work, two different solvent extraction systems comprising acetonitrile, water, and formic acid were employed for the LESA MS analysis of Gram positive (50:45:5) and Gram negative (40:59:1) bacteria.[20] The higher acetonitrile and formic acid content for the Gram positive species ensures the extraction of proteins, possibly due to cell lysis or increased ionization efficiency.[20] Utilization of two extraction solvents is, however, impractical when analyzing multiple and unknown bacterial species. In addition, initial LESA MS experiments with the 50:45:5 extraction solvent system of Gram positive *Enterococci* species resulted in the observation of no protein peaks for *E. faecalis* V583 and only a few, low abundance protein peaks for *E. faecium* E745 (Figure 1a,b). Subsequent experiments focused on *E. faecalis* V583 as a model organism for the optimization of extraction solvent composition. Solvents comprising various ratios of ethanol, acetonitrile, water, and formic acid were investigated (see Supporting Information, Figure S1). The best results, as determined by detection of the greatest number of protein peaks, were achieved with the 60:35:5 acetonitrile/water/formic acid solvent system (Figure 1a). (Interestingly, there appears to be a limit on the acetonitrile content: At 80% acetonitrile, no protein peaks were detected, likely a result of poor protein solubility). The optimized solvent system also resulted in detection of a greater number of proteins from *E. faecium* E745 (Figure 1b). The optimized LESA extraction solvent system was subsequently tested for the remaining three ESKAPE species that are the focus of this work (Figure 1c) as well as the previously studied *S. aureus* MSSA476 and *P. aeruginosa* PS1054 (Figure S2, Supporting Information). Representative LESA mass spectra from biological replicates are shown in Figure S3, Supporting Information. The results show that 60:35:5 acetonitrile/water/formic acid is suitable as a solvent for the extraction of proteins from both Gram positive and Gram negative ESKAPE microorganisms.

LESA MS Analysis of ESKAPE Pathogens. For each of the species studied, the most abundant protein peaks were selected for fragmentation by CID. The resulting MS/MS spectra were searched against the corresponding individual bacterial protein database. Top down LESA MS analysis of the ESKAPE species resulted in identification of 24 proteins from 37 MS/MS spectra. MS/MS spectra and fragment assignments for all proteins are shown in the Supporting Information, File

https://doi.org/10.1021/jasms.0c00466
J. Am. Soc. Mass Spectrom. 2021, 32, 1345−1351

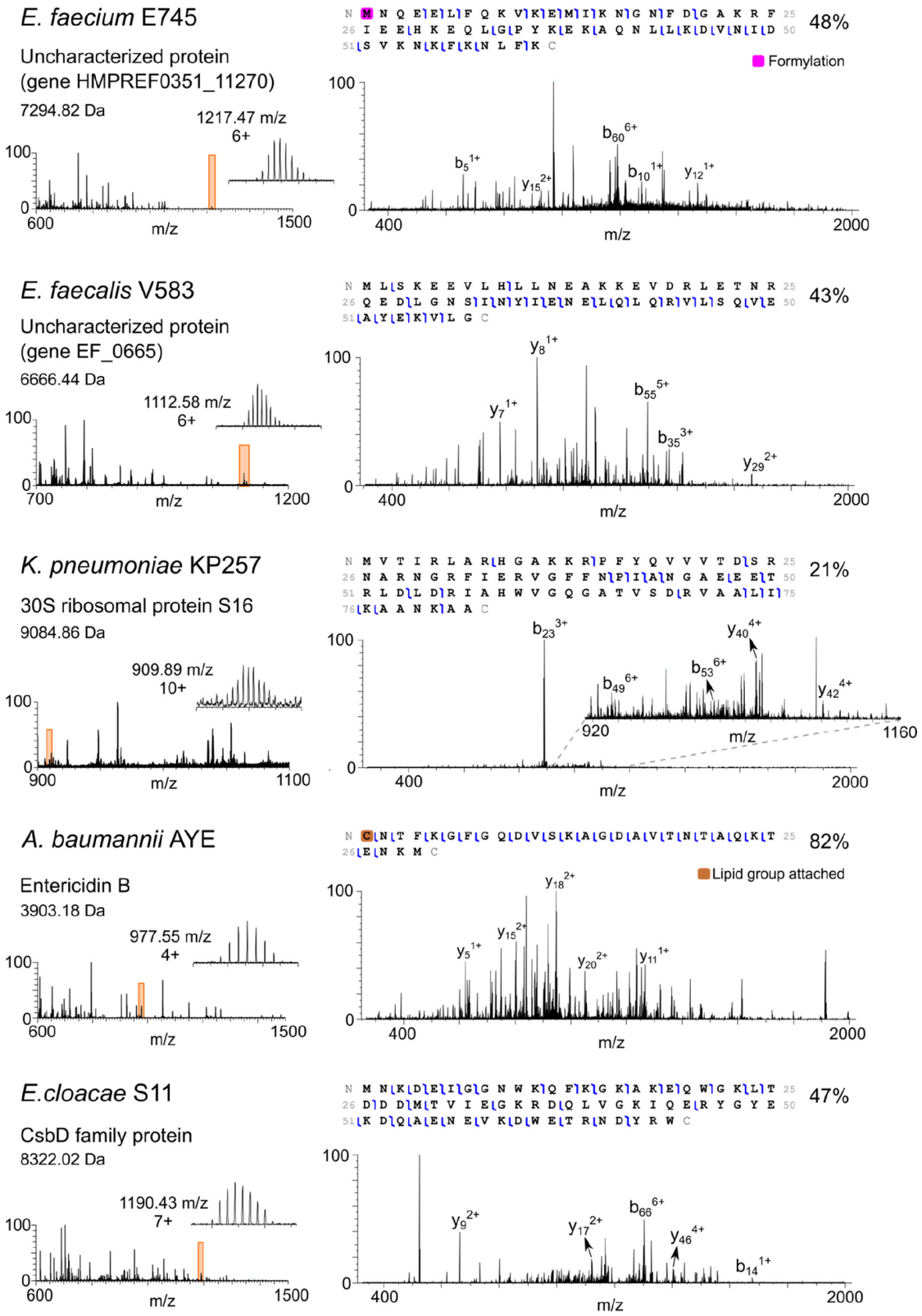


Figure 2. Representative MS/MS spectra corresponding to five newly identified proteins for the various ESKAPE species.

S1. Representative MS/MS spectra for the various species are shown in Figure 2.

Enterococci. For *E. faecium* E745, five proteins were identified after CID fragmentation of 11 MS/MS spectra (see File S1, Supporting Information): 50S ribosomal protein L29, 30S ribosomal protein S20, and three uncharacterized proteins HMPREF0351 11703, HMPREF0351 11270, and HMPREF0351 12038. Four out of five of the identified proteins were observed in both biological replicates. According to Uniprot, the predicted HMPREF0351 11703 protein

https://doi.org/10.1021/jasms.0c00466
J. Am. Soc. Mass Spectrom. 2021, 32, 1345−1351

sequence includes a signal peptide. Our data suggest no signal peptide cleavage and formylation of the N terminal methionine. The second uncharacterized protein HMPREF0351 11270 (see Figure 2) was identified with formylated methionine at the N terminus, a modification so far unrecorded in the Uniprot database. For the third uncharac terized protein HMPREF0351 12038, a mutation R → Q was detected at either position 38 or position 43; however, no fragments were observed in this region to allow unambiguous localization.

A close relative of *E. faecium* is *E. faecalis*, a clinically important species associated with infective endocarditis, biofilm formation, and antimicrobial resistance.[22] We chose to investigate this microbe due to its abundance in the hospital environment and increasing antibiotic resistance.[23] LESA MS/MS of five *E. faecalis* V583 intact precursor ions resulted in the identification of four proteins (see File S1, Supporting Information): DNA binding protein HU, 50S ribosomal protein L29, UPF0337 protein EF 1180, and uncharacterized protein (gene EF 0665). The UPF0337 protein EF 1180 is a protein inferred from homology and belongs to the bacterial general stress response protein (CsbD) family, while the uncharacterized protein (gene EF 0665) (see Figure 2) is a predicted protein. Comparison of the LESA mass spectra obtained from *E. faecalis* V583 and *E. faecium* E745 revealed that the 50S ribosomal protein L29 was observed for both species (Figure S4). The sequence of the protein from the two species differs by an N → K substitution at position 62 resulting in a mass difference Δm = 14.07 Da. This observation is potentially useful as a diagnostic for differentiation between these two microorganisms.

Klebsiella pneumoniae. Previous LESA MS experiments have only considered *Klebsiella pneumoniae* in the context of its growth on 3D living skin equivalents. In that work, two proteins were identified. A more in depth LESA MS/MS analysis of the bacteria cultured on agar plates is therefore warranted. LESA MS/MS of nine precursor proteins from *K. pneumoniae* KP257 resulted in identification of six proteins, two of which (DNA binding protein HU α and KPN 00497) were identified in the previous analysis of the *in vitro* 3D skin models.[21] All six proteins were observed in both biological replicates. The same protein mutation R → K at the position 49 and a signal peptide cleavage (1−19) was observed for the KPN 00497 protein. The four novel proteins (see File S1, Supporting Information) included two ribosomal proteins (50S ribosomal protein L29 and 30S ribosomal protein S16 (see Figure 2), one uncharacterized protein (gene yciG), and CsbD domain containing protein. The search results indicated a signal peptide cleavage of the first 19 amino acids of the CsbD protein sequence previously unrecorded in the Uniprot database.

Acinetobacter baumannii. LESA MS analysis of the Gram negative reference strain *A. baumannii* AYE yielded mass spectra with highly abundant protein peaks (Figures 1c and Figure S5). Subsequent CID fragmentation of six intact precursors resulted in four protein identifications—three uncharacterized proteins (genes ABAYE1298, ABAYE2274, and ABAYE1876) and bacteriolytic lipoprotein entericidin B (see File S1, Supporting Information). All four proteins were observed in both biological replicates. For ABAYE2274, cleavage of the initial methionine was detected from the MS/MS data. The amino acid sequence of ABAYE1876 contains a signal peptide (1−14), information not yet recorded in the Uniprot database. Entericidin B is the first lipoprotein to be identified by LESA MS (see Figure 2). Lipoproteins are important for bacterial physiology as well as virulence and as activators of the host innate immune response.[24] Bacterial lipoproteins are characterized by a conserved N terminal lipid modified cysteine residue. In this case, a mass shift of 813.72 Da was detected; however, the exact structure of the lipid group attached to the N terminus remains unknown and would require further analysis. Despite the mass shift, there is high confidence in the protein assignment due to the high sequence coverage (82%) obtained. *A. baumannii* AYE was compared to the clinical strain *A. baumannii* AC02 (Figure S5). Again, LESA MS resulted in detection of several highly abundant protein peaks (Figure S5); however, their identification proved challenging. Six MS/MS spectra were searched against the AYE database; however, no protein IDs were assigned, suggesting dissimilarity in the protein amino acid sequences of these strains and a requirement for a new database. The dissimilarity can also be observed from the comparison of the AYE and AC02 mass spectra (Figure S5).

Enterobacter cloacae. Mass spectra obtained following LESA of *E. cloacae* S11 contained many abundant peaks corresponding to proteins (Figure 1c). LESA MS/MS analysis of six intact precursors resulted in identification of five proteins (see File S1, Supporting Information)—50S ribosomal protein L29, DNA binding protein, CsbD family protein (see Figure 2), UPF0391 membrane protein, and DUF1471 domain containing binding protein. Four out of five proteins were detected in both biological replicates. The representative mass spectra shown in Figure S3, Supporting Information contain peaks corresponding to four and three of those proteins. For the UPF0391 membrane protein a new PTM—formylation at the N terminus, not yet reported in the Uniprot database—was revealed. The DUF1471 protein sequence contains a signal peptide (1−21) and a mutation (E → N) at one of two potential positions—either 37 or 45; however, no fragments were observed in this region to allow unambiguous local ization.

Identification of ESKAPE Pathogens from Multiple Protein Databases. Initially, our goal was to investigate ESKAPE pathogens by top down LESA MS combined with searching of individual species databases. If LESA MS is to find use as a diagnostic tool, however, correct identification of proteins (and species) from multiple databases is necessary. To evaluate that, a data analysis workflow was constructed in the ProSightPC software, in which each MS/MS mass spectrum was searched against all six individual ESKAPE protein databases (including *S. aureus* and *P. aeruginosa*) (i.e., an automated concurrent search of individual databases) using the absolute mass search function. The 24 MS/MS spectra corresponding to the newly identified proteins described above were used for the searches. The lowest e score value was used as the indicator of protein assignment, and that assignment was compared with the known protein ID. The results are summarized in Table S1, Supporting Information.

In total, 19/24 proteins were correctly assigned, both in terms of protein ID and bacterial species. For two MS/MS spectra (corresponding to CscD domain containing protein (*K. pneumoniae* KP257) and uncharacterized protein (gene ABAYE1876) (*A. baumannii* AYE)), no protein assignments were returned. Both CscD domain containing protein and uncharacterized protein (gene ABAYE1876) were identified above using the biomarker search function in the ProSightPC

https://doi.org/10.1021/jasms.0c00466
J. Am. Soc. Mass Spectrom. 2021, 32, 1345−1351

software. That function is designed for identification of truncated proteins; therefore, the absolute mass search used in this automated data analysis workflow did not find any matching sequences.

Three proteins were misassigned, all from *E. cloacae* S11: DUF1471 domain containing protein (*E. cloacae* S11) was assigned as an uncharacterized protein EF 2117 (*E. faecalis*), 50S ribosomal protein L29 was identified as the same protein but from *K. pneumoniae*, and UPF0391 membrane protein SAMEA2054040 04753 was assigned to UPF0391 membrane protein KPN 04833, again from *K. pneumoniae*. The DUF1471 domain containing protein was identified above by use of a biomarker search within the ProSightPC software, and the misassignment appears to be the result of the absolute mass search. Both *Enterobacter* and *Klebsiella* belong to the family of Enterobacteriaceae; therefore, similarities in protein amino acid sequences between these species might be observed.

The overall success rate was 79% when both Gram positive and Gram negative species are considered. Generally, the accuracy of assignment of Gram negative species is higher than for Gram positives, which is also observed in MALDI TOF MS.[25] Identification success rates for MALDI TOF MS at the species level vary between 84.1−94.9%[26,27] for aerobic bacteria and routine isolates and 81.8% for both aerobes and anaerobes.[25,27] Suggested improvements for LESA MS include the use of multiple databases at the genus and strain level, which might increase the success rate of protein identification in the future.

To further address the question of correct species identification, a BLAST search of the 24 identified proteins was performed to determine the specificity of the protein sequences. All of the protein sequences identified for *E. faecium*, *E. faecalis*, *K. pneumoniae*, and *A. baumannii* were unique to their species. Two out of the five *E. cloacae* proteins were species specific, whereas DNA binding protein HU shared a 100% sequence homology with *Cedecea davisae*, CsbD family protein shared a 100% sequence homology with *E. hormaechei*, and UPF0391 membrane protein shared 100% sequence homology with *Lelliotia amnigena*.

CONCLUSION

The results show that a LESA MS sampling solvent system comprising 60:35:5 acetonitrile/water/formic acid is capable of extracting proteins from both Gram positive and Gram negative ESKAPE pathogens. Four out of six ESKAPE microbes *E. faecium*, *K. pneumoniae*, *A. baumannii*, and *E. cloacae*, including the clinically important strain *E. faecalis* and the clinical isolate *A. baumannii*, were investigated by a top down LESA MS/MS method. The MS/MS mass spectra searches resulted in identifications of 24 proteins. The 50S ribosomal protein L29 was observed in three of the four ESKAPE species studied here. This protein has also previously been identified in the LESA mass spectra obtained from the ESKAPE pathogen *P. aeruginosa*.[20] This result suggests that this protein may potentially serve as a biomarker for these species. For top down LESA MS to be useful in bacterial identification, there is a requirement for proteins (and species) to be correctly identified following searching of MS/MS spectra against databases from multiple species. In this work, the overall identification success rate was determined to be 79%. *E. cloacae* presented the biggest challenge in terms of species identification based on results from both the multidatabase search and the BLAST search. Further development of LESA MS as a diagnostic requires application to a broader range of species. Nevertheless, the results presented suggest LESA MS has potential as a useful tool in clinical microbiology.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jasms.0c00466.

LESA extraction solvent system optimization (Figure S1); LESA MS of *S. aureus* and *P. aeruginosa* (Figure S2); LESA MS of biological replicates (Figure S3); comparison of LESA MS of enterococci 50S ribosomal protein L29 (Figure S4); comparison of LESA MS of *A. baumannii* strains (Figure S5); summary of results from multiple database search (Table S1); protein assignments (File S1) (PDF)

AUTHOR INFORMATION

Corresponding Author

Helen J. Cooper − *School of Biosciences, University of Birmingham, Birmingham B15 2TT, United Kingdom;* orcid.org/0000 0003 4590 9384; [redacted]

Authors

Jana Havlikova − *EPSRC Centre for Doctoral Training in Physical Sciences for Health and School of Biosciences, University of Birmingham, Birmingham B15 2TT, United Kingdom*

Robin C. May − *Institute of Microbiology and Infection, University of Birmingham, Birmingham B15 2TT, United Kingdom*

Iain B. Styles − *EPSRC Centre for Doctoral Training in Physical Sciences for Health and School of Computer Science, University of Birmingham, Birmingham B15 2TT, United Kingdom*

Complete contact information is available at: https://pubs.acs.org/10.1021/jasms.0c00466

Notes

The authors declare no competing financial interest.

Supplementary data supporting this research is openly available from the University of Birmingham data archive at DOI: 10.25500/edata.bham.00000615.

ACKNOWLEDGMENTS

H.J.C. is an EPSRC Established Career Fellow (EP/L023490/1 and EP/S002979/1). J.H. is funded by the EPSRC Physical Sciences for Health Doctoral Training Centre (EP/L016346/1). 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.

REFERENCES

(1) Mukhopadhyay, S.; Bharath Prasad, A. S.; Mehta, C. H.; Nayak, U. Y. Antimicrobial peptide polymers: no escape to ESKAPE pathogens a review. *World J. Microbiol. Biotechnol.* **2020**, *36* (9), 131.

(2) Santajit, S.; Indrawattana, N. Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens. *Biomed Res Int.* **2016**, *2016*, 2475067.
(3) Mulani, M. S.; Kamble, E. E.; Kumkar, S. N.; Tawre, M. S.; Pardesi, K. R. Emerging Strategies to Combat ESKAPE Pathogens in the Era of Antimicrobial Resistance: A Review. *Front. Microbiol.* **2019**, *10*, 539.
(4) Navidinia, M. The clinical importance of emerging ESKAPE pathogens in nosocomial infections. *Archives of Advances in Biosciences* **2016**, *7* (3), 43−57.
(5) IACG. *No Time to Wait: Securing the Future from Drug Resistant Infections*; World Health Organization, 2019.
(6) Rice, L. B. Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE. *J. Infect. Dis.* **2008**, *197* (8), 1079−1081.
(7) Lévesque, S.; Dufresne, P. J.; Soualhine, H.; Domingo, M. C.; Bekal, S.; Lefebvre, B.; Tremblay, C. 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), e0144878−e0144878.
(8) Hsu, C. C.; ElNaggar, M. S.; Peng, Y.; Fang, J.; Sanchez, L. M.; Mascuch, S. J.; Møller, K. A.; Alazzeh, E. K.; Pikula, J.; Quinn, R. A.; Zeng, Y.; Wolfe, B. E.; Dutton, R. J.; Gerwick, L.; Zhang, L.; Liu, X.; Månsson, M.; Dorrestein, P. C. Real Time Metabolomics on Living Microorganisms Using Ambient Electrospray Ionization Flow Probe. *Anal. Chem.* **2013**, *85* (15), 7014−7018.
(9) Angolini, C. F. F.; Vendramini, P. H.; Araújo, F. D. S.; Araújo, W. L.; Augusti, R.; Eberlin, M. N.; de Oliveira, L. G. Direct Protocol for Ambient Mass Spectrometry Imaging on Agar Culture. *Anal. Chem.* **2015**, *87* (13), 6925−6930.
(10) Meetani, M. A.; Shin, Y. S.; Zhang, S.; Mayer, R.; Basile, F. Desorption electrospray ionization mass spectrometry of intact bacteria. *J. Mass Spectrom.* **2007**, *42* (9), 1186−1193.
(11) Song, Y.; Talaty, N.; Tao, W. A.; Pan, Z.; Cooks, R. G. Rapid ambient mass spectrometric profiling of intact, untreated bacteria using desorption electrospray ionization. *Chem. Commun.* **2007**, No. 1, 61−63.
(12) 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.* **2013**, *85* (21), 10385−10391.
(13) 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.* **2013**, *49* (55), 6188−6190.
(14) 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.* **2014**, *86* (13), 6555−6562.
(15) 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.* **2014**, *86* (15), 7500−7507.
(16) Cameron, S. J. S.; Bodai, Z.; Temelkuran, B.; Perdones Montero, A.; Bolt, F.; Burke, A.; Alexander Hardiman, K.; Salzet, M.; Fournier, I.; Rebec, M.; Takáts, Z. Utilisation of Ambient Laser Desorption Ionisation Mass Spectrometry (ALDI MS) Improves Lipid Based Microbial Species Level Identification. *Sci. Rep.* **2019**, *9* (1), 3006.
(17) 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.* **2010**, *45* (3), 252−260.
(18) Dupré, M.; Duchateau, M.; Malosse, C.; Borges Lima, D.; Calvaresi, V.; Podglajen, I.; Clermont, D.; Rey, M.; Chamot Rooke, J. Optimization of a Top Down Proteomics Platform for Closely Related Pathogenic Bacterial Discrimination. *J. Proteome Res.* **2021**, *20* (1), 202−211.
(19) 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.* **2014**, *86* (21), 10504−10510.
(20) Kocurek, K. I.; Stones, L.; Bunch, J.; May, R. C.; Cooper, H. J. Top down LESA mass spectrometry protein analysis of gram positive and gram negative bacteria. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (10), 2066−2077.
(21) Havlikova, J.; May, R. C.; Styles, I. B.; Cooper, H. J. Direct identification of bacterial and human proteins from infected wounds in living 3D skin models. *Sci. Rep.* **2020**, *10* (1), 11900.
(22) Beganovic, M.; Luther, M. K.; Rice, L. B.; Arias, C. A.; Rybak, M. J.; LaPlante, K. L. A Review of Combination Antimicrobial Therapy for Enterococcus faecalis Bloodstream Infections and Infective Endocarditis. *Clin. Infect. Dis.* **2018**, *67* (2), 303−309.
(23) Ranotkar, S.; Kumar, P.; Zutshi, S.; Prashanth, K. S.; Bezbaruah, B.; Anand, J.; Lahkar, M. Vancomycin resistant enterococci: Trouble maker of the 21st century. *Journal of Global Antimicrobial Resistance* **2014**, *2* (4), 205−212.
(24) Nakayama, H.; Kurokawa, K.; Lee, B. L. Lipoproteins in bacteria: structures and biosynthetic pathways. *FEBS J.* **2012**, *279* (23), 4247−4268.
(25) Lin, J. F.; Ge, M. C.; Liu, T. P.; Chang, S. C.; Lu, J. J. A simple method for rapid microbial identification from positive monomicrobial blood culture bottles through matrix assisted laser desorption ionization time of flight mass spectrometry. *Journal of Microbiology, Immunology and Infection* **2018**, *51* (5), 659−665.
(26) Dingle, T. C.; Butler Wu, S. M. Maldi tof mass spectrometry for microorganism identification. *Clin Lab Med.* **2013**, *33* (3), 589−609.
(27) Ge, M. C.; Kuo, A. J.; Liu, K. L.; Wen, Y. H.; Chia, J. H.; Chang, P. Y.; Lee, M. H.; Wu, T. L.; Chang, S. C.; Lu, J. J. Routine identification of microorganisms by matrix assisted laser desorption ionization time of flight mass spectrometry: Success rate, economic analysis, and clinical outcome. *J. Microbiol Immunol Infect* **2017**, *50* (5), 662−668.

https://doi.org/10.1021/jasms.0c00466
J. Am. Soc. Mass Spectrom. 2021, 32, 1345−1351

Appendix B
MS/MS mass spectra and protein ID assignments

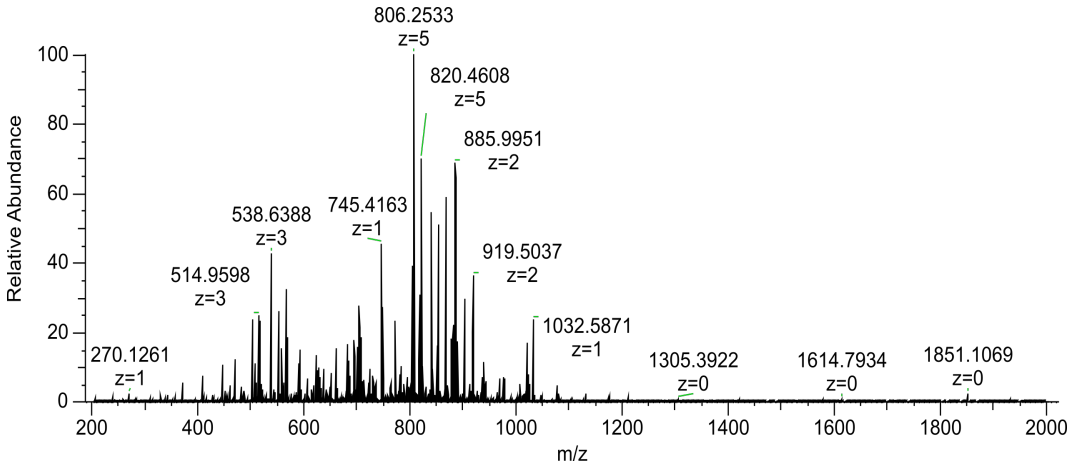
All masses listed are monoisotopic. PTMs: a – acetylation, f – formylation, m – methylation. Fragmentation methods: CID (collision-induced dissociation), HCD (higher-energy collision dissociation).

B.1 Chapter 3

Enterococcus faecium E745

Uncharacterised protein (gene HMPREF_11703)

Uniprot accession number: Q3XZ91
Charge state: 8+
Observed monoisotopic mass: 5967.3058 Da
Sequence:
fMKKFVSGILVGSLATAAAVAGLVASVKKTVIDPIDEKEAMIEENRKKAMRKRVSR

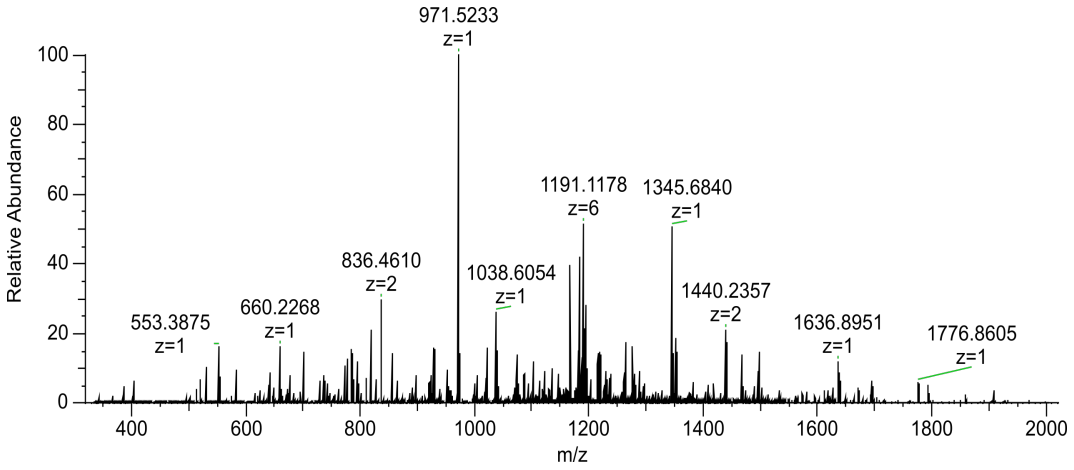


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
451.2499	451.2518	2	b8-H2O	-0.0038	-4.2105
460.2551	460.2571	2	b8	-0.0040	-4.3454
502.7996	502.8007	2	y8	-0.0022	-2.1877
507.7916	507.7939	2	b9-H2O	-0.0046	-4.5294
510.6454	510.6476	3	y12	-0.0066	-4.3083

516.7971	516.7992	2	b9	-0.0042	-4.0635
520.9632	520.9655	3	b15	-0.0069	-4.4149
544.6422	544.6445	3	b16	-0.0069	-4.2229
562.3174	562.3252	3	b17-H2O	-0.0078	-4.6237
566.3311	566.3334	2	b10	-0.0046	-4.0612
568.3211	568.3235	3	b17	-0.0072	-4.2229
592.0002	592.0026	3	b18	-0.0072	-4.0540
594.8416	594.8441	2	b11	-0.0050	-4.2028
618.589	618.5916	4	y20-H2O	-0.0104	-4.2031
623.0917	623.0942	4	y20	-0.0100	-4.0122
625.0228	625.0254	3	b19	-0.0078	-4.1598
638.3575	638.3601	2	b12	-0.0052	-4.0729
648.7015	648.7044	3	b20	-0.0087	-4.4704
662.3666	662.3694	1	b5	-0.0028	-4.2272
683.5507	683.5535	6	y36	-0.0168	-4.0962
694.8994	694.9021	2	b13	-0.0054	-3.8854
700.0619	700.0649	6	y37	-0.0180	-4.2853
703.6818	703.6817	7	y46-NH3	0.0007	0.1421
704.3823	704.3852	4	y23	-0.0116	-4.1171
711.9012	711.9044	6	y38	-0.0192	-4.4950
723.7411	723.7439	6	y39	-0.0168	-3.8688
730.4178	730.4207	2	b14	-0.0058	-3.9703
735.5798	735.5834	6	y40	-0.0216	-4.8941
761.2592	761.2625	6	y42-H2O	-0.0198	-4.3349
764.2616	764.2642	6	y42	-0.0156	-3.4020
780.2734	780.2738	6	y43-NH3	-0.0024	-0.5126
783.1093	783.1116	6	y43	-0.0138	-2.9370
805.8521	805.8553	5	y35	-0.0160	-3.9709
807.1172	807.1156	6	b46	0.0096	1.9824
816.4599	816.4631	2	b16	-0.0064	-3.9193
820.0597	820.0627	5	y36	-0.0150	-3.6583
823.6296	823.6319	6	y46	-0.0138	-2.7925
839.8731	839.8764	5	y37	-0.0165	-3.9291
851.9783	851.9817	2	b17	-0.0068	-3.9907
854.0803	854.0838	5	y38	-0.0175	-4.0980
868.288	868.2913	5	y39	-0.0165	-3.8006
882.4947	882.4987	5	y40	-0.0200	-4.5326
902.7045	902.7082	5	y41	-0.0185	-4.0988
913.5105	913.5103	5	y42-NH3	0.0010	0.2189
916.9123	916.9156	5	y42	-0.0165	-3.5990
937.0308	937.0344	2	b19	-0.0072	-3.8419
938.8411	938.8445	3	y23	-0.0102	-3.6215
968.3395	968.3372	5	b46	0.0115	2.3752
1007.0637	1007.0673	4	y35	-0.0144	-3.5747
1014.8777	1014.8815	3	y25	-0.0114	-3.7443
1024.8225	1024.8266	4	y36	-0.0164	-4.0007
1032.5871	1032.591	1	b9	-0.0039	-3.7769

Uncharacterised protein (gene HMPREF0351_11270)

Uniprot accession number: I3U1K6
Charge state: 6+
Observed monoisotopic mass: 7294.8163 Da
Sequence:
fMNQEELFQKVKEMIKNGNFDGAKRFIEEHKEQLGPYKEKAQNLLKDVNIDSVK
NKFKNLFK

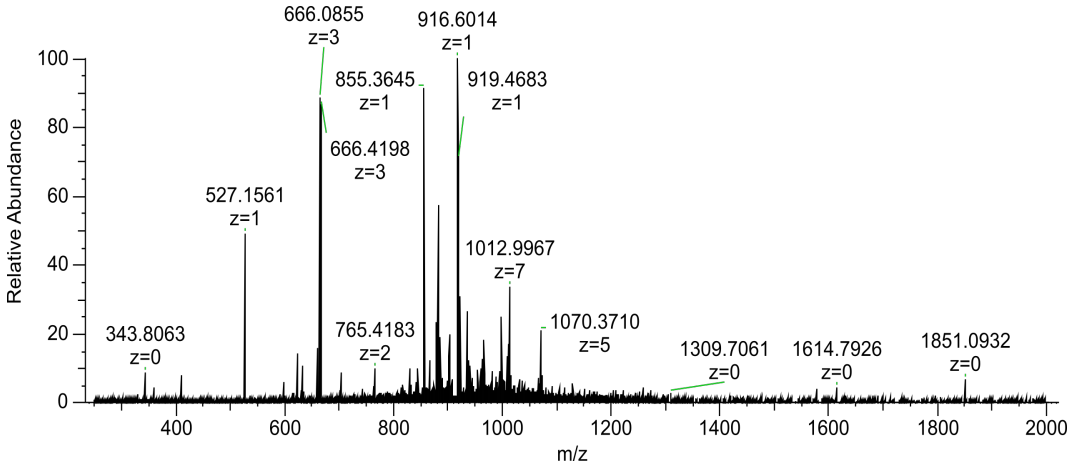


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
451.2499	451.2518	2	b8-H2O	-0.0038	-4.2105
460.2551	460.2571	2	b8	-0.0040	-4.3454
502.7996	502.8007	2	y8	-0.0022	-2.1877
507.7916	507.7939	2	b9-H2O	-0.0046	-4.5294
510.6454	510.6476	3	y12	-0.0066	-4.3083
516.7971	516.7992	2	b9	-0.0042	-4.0635
520.9632	520.9655	3	b15	-0.0069	-4.4149
544.6422	544.6445	3	b16	-0.0069	-4.2229
562.3174	562.32	3	b17-H2O	-0.0078	-4.6237
566.3311	566.3334	2	b10	-0.0046	-4.0612
568.3211	568.3235	3	b17	-0.0072	-4.2229
592.0002	592.0026	3	b18	-0.0072	-4.0540
594.8416	594.8441	2	b11	-0.0050	-4.2028
618.589	618.5916	4	y20-H2O	-0.0104	-4.2031
623.0917	623.0942	4	y20	-0.0100	-4.0122
625.0228	625.0254	3	b19	-0.0078	-4.1598
638.3575	638.3601	2	b12	-0.0052	-4.0729
648.7015	648.7044	3	b20	-0.0087	-4.4704
662.3666	662.3694	1	b5	-0.0028	-4.2272
683.5507	683.5535	6	y36	-0.0168	-4.0962

694.8994	694.9021	2	b13	-0.0054	-3.8854
700.0619	700.0649	6	y37	-0.0180	-4.2853
703.6818	703.6817	7	y46-NH3	0.0007	0.1421
704.3823	704.3852	4	y23	-0.0116	-4.1171
711.9012	711.9044	6	y38	-0.0192	-4.4950
723.7411	723.7439	6	y39	-0.0168	-3.8688
730.4178	730.4207	2	b14	-0.0058	-3.9703
735.5798	735.5834	6	y40	-0.0216	-4.8941
761.2592	761.2625	6	y42-H2O	-0.0198	-4.3349
764.2616	764.2642	6	y42	-0.0156	-3.4020
780.2734	780.2738	6	y43-NH3	-0.0024	-0.5126
783.1093	783.1116	6	y43	-0.0138	-2.9370
805.8521	805.8553	5	y35	-0.0160	-3.9709
807.1172	807.1156	6	b46	0.0096	1.9824
816.4599	816.4631	2	b16	-0.0064	-3.9193
820.0597	820.0627	5	y36	-0.0150	-3.6583
823.6296	823.6319	6	y46	-0.0138	-2.7925
839.8731	839.8764	5	y37	-0.0165	-3.9291
851.9783	851.9817	2	b17	-0.0068	-3.9907
854.0803	854.0838	5	y38	-0.0175	-4.0980
868.288	868.2913	5	y39	-0.0165	-3.8006
882.4947	882.4987	5	y40	-0.0200	-4.5326
902.7045	902.7082	5	y41	-0.0185	-4.0988
913.5105	913.5103	5	y42-NH3	0.0010	0.2189
916.9123	916.9156	5	y42	-0.0165	-3.5990
937.0308	937.0344	2	b19	-0.0072	-3.8419
938.8411	938.8445	3	y23	-0.0102	-3.6215
968.3395	968.3372	5	b46	0.0115	2.3752
1007.0637	1007.0673	4	y35	-0.0144	-3.5747
1014.8777	1014.8815	3	y25	-0.0114	-3.7443
1024.8225	1024.8266	4	y36	-0.0164	-4.0007
1032.5871	1032.591	1	b9	-0.0039	-3.7769

50S ribosomal protein L29

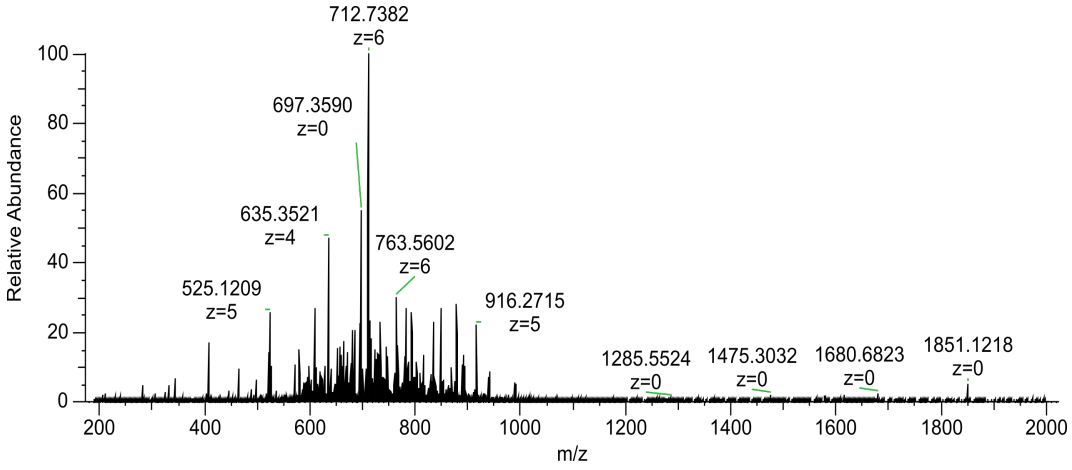
Uniprot accession number: Q3XYY1
Charge state: 8+
Observed monoisotopic mass: 7339.0618 Da
Sequence:
MKVKEIRELTTAEMLDQEKQLKEELFNLRFQLATGQLENTARIKEVRKSIARIKTV
LREQAK



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
666.0855	666.0882	3	y17	-0.0081	-4.0535
702.9274	702.9306	4	y24	-0.0128	-4.5524
765.4183	765.4214	2	b13	-0.0062	-4.0501
830.9382	830.9417	2	b14	-0.0070	-4.2121
965.955	965.9586	5	y41	-0.0180	-3.7269
1012.4237	1012.4272	7	y60	-0.0245	-3.4570
1069.7749	1069.7737	5	b45	0.0060	1.1217
1272.6768	1272.681	4	b43	-0.0168	-3.3001

30S ribosomal protein S20

Uniprot accession number: I3U1I4
Charge state: 13+
Observed monoisotopic mass: 9053.8009 Da
Sequence:
PNIESAIKRVRTSENANVKNSSQTSAMRTAIKKFEDAVASGADNVDALYKEAVKAI
DMAESKGLIHKNKANRDKSRLSKKIAK

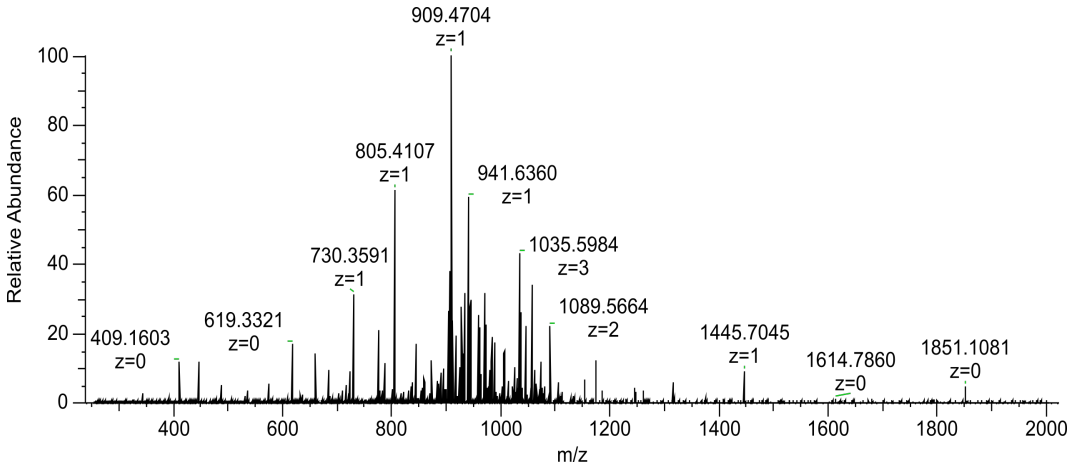


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
497.0555	497.0577	4	y17	-0.0088	-4.4260
524.9204	524.9227	5	y23	-0.0115	-4.3816
579.8851	579.8877	2	y10	-0.0052	-4.4836
602.8264	602.8294	6	b33	-0.0180	-4.9765
610.6323	610.6351	7	y38	-0.0196	-4.5854
624.7848	624.7878	7	y39	-0.0210	-4.8016
627.6666	627.6694	3	b17	-0.0084	-4.4609
638.5041	638.5067	7	y40-H2O	-0.0182	-4.0720
641.0764	641.0796	7	y40	-0.0224	-4.9916
654.3371	654.3402	7	b43	-0.0217	-4.7376
655.8988	655.9016	4	y23	-0.0112	-4.2689
662.3790	662.3818	6	y35	-0.0168	-4.2272
668.0163	668.0190	6	b36	-0.0162	-4.0418
670.6290	670.6321	7	b44	-0.0217	-4.6225
681.2257	681.2291	6	y36	-0.0204	-4.9910
684.7820	684.7847	7	b45	-0.0189	-3.9428
708.2063	708.2095	6	b39	-0.0192	-4.5184
709.2349	709.2380	6	y38-H2O	-0.0186	-4.3709
712.2369	712.2398	6	y38	-0.0174	-4.0717
722.7116	722.7148	6	b40	-0.0192	-4.4277
725.7465	725.7494	6	y39-H2O	-0.0174	-3.9959
728.7477	728.7512	6	y39	-0.0210	-4.8027
747.7556	747.7583	6	y40	-0.0162	-3.6108
760.3918	760.3913	6	b43-NH3	0.0030	0.6576
775.9303	775.9313	6	y42-NH3	-0.0060	-1.2888
782.2328	782.2362	6	b44	-0.0204	-4.3465
788.2688	788.2726	6	y43	-0.0228	-4.8207
835.4390	835.4425	5	b38	-0.0175	-4.1894
849.6463	849.6499	5	b39	-0.0180	-4.2370
867.0533	867.0563	5	b40	-0.0150	-3.4600
878.4569	878.4606	5	b41	-0.0185	-4.2119

892.6650	892.6680	5	b42	-0.0150	-3.3607
915.6697	915.6734	5	b43	-0.0185	-4.0407
938.4779	938.4820	5	b44	-0.0205	-4.3688
940.9963	941.0005	2	b17	-0.0084	-4.4633
990.5309	990.5347	2	b18	-0.0076	-3.8363

Uncharacterised protein (gene HMPREF0351_12038)

Uniprot accession number: Q3XY98
Charge state: 7+
Observed monoisotopic mass: 6596.5391 Da
Sequence:
MKWKEKVDEAAEKLYDLVKSEKYNIEVNIPKKGGKAVRVKSKRPTNHTKKWM
AKNR
This protein contains an R→Q mutation at two possible locations: 38 or 43.



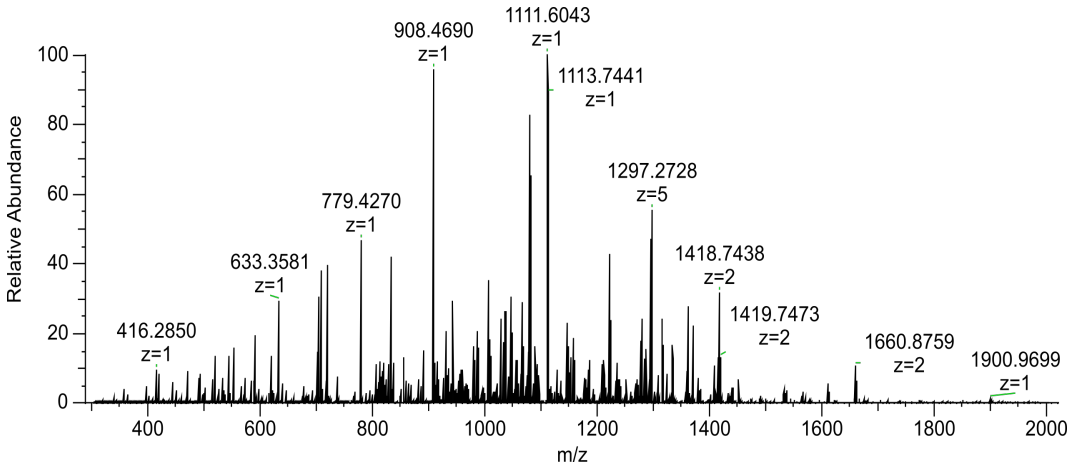
Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
1044.5397	1044.5426	b8	-0.0030	-2.8357
1173.5836	1173.5852	b9	-0.0016	-1.3906
1444.6946	1444.7020	b12	-0.0074	-5.1097
1444.6978	1444.7020	b12	-0.0042	-2.9362
1572.7949	1572.7970	b13	-0.0021	-1.3276
1685.8762	1685.8810	b14	-0.0049	-2.8804
1963.9565	1963.9713	b16	-0.0148	-7.5383
2176.117	2176.1238	b18	-0.0068	-3.1207
804.4053	804.4065	y6	-0.0012	-1.4719
2749.5197	2777.5823	y24	0.0400	14.3945
3102.7611	3130.8250	y27	0.0387	12.3514
3102.7655	3130.8250	y27	0.0431	13.7523

3428.9627	3457.0204	y30	0.0449	12.9768
3558.0035	3586.0630	y31	0.0430	12.0026
3948.1766	3976.2533	y34	0.0258	6.5001
4292.3505	4320.4229	y37	0.0302	6.9905
4292.3516	4320.4229	y37	0.0313	7.2393
4420.4481	4448.5179	y38	0.0328	7.3777
4632.5809	4660.6703	y40	0.0131	2.8198
4910.6742	4938.7606	y42	0.0162	3.2781
5422.9968	5451.0564	y47	0.0430	7.8851
6337.4108	6365.5062	y54	0.0072	1.1272

Enterococcus faecalis V583

Uncharacterised protein (gene EF_0665)

Uniprot accession number: Q838A7
Charge state: 6+
Observed monoisotopic mass: 6666.4375 Da
Sequence:
MLSKEEVLHLLNEAKKEVDRLETNRQEDLGNSINYIENELQLQRVLSQVEAYEK
VLG

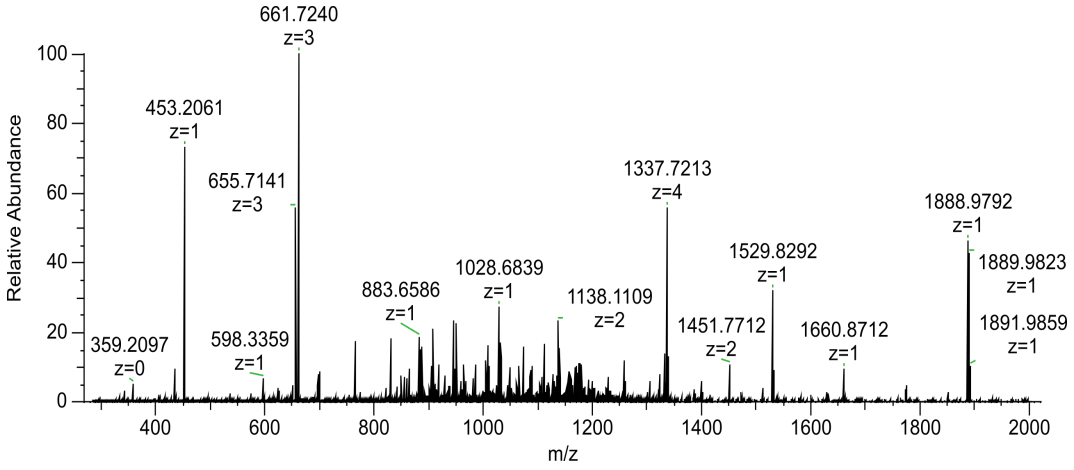


Observed mass (monoisotopic) *m/z*	Theoretical mass (monoisotopic) *m/z*	Charge	Fragment	Mass error [Da]	Mass error [ppm]
534.2792	534.2813	2	b9	-0.0042	-2.3396
545.3271	545.3293	1	y5	-0.0022	-4.5844
708.3899	708.3927	1	y6	-0.0028	-3.5291
779.4268	779.4298	1	y7	-0.0030	-3.2075
837.9329	837.9359	4	b28	-0.0120	-0.7459
908.4690	908.4724	1	y8	-0.0034	-2.7519

958.9978	959.0020	4	b33	-0.0168	-0.6517
980.5431	980.5467	2	y17	-0.0072	-1.2748
987.5086	987.5127	4	b34	-0.0164	-0.6329
1007.5375	1007.5408	1	y9	-0.0033	-2.4813
1028.2739	1028.2786	4	b35	-0.0188	-0.6078
1037.0851	1037.0888	2	y18	-0.0074	-1.2053
1047.5396	1047.5440	5	b44	-0.0220	-0.4773
1056.5448	1056.5496	4	b36	-0.0192	-0.5915
1064.2175	1064.2199	6	b54	-0.0144	-0.3915
1067.5527	1067.5554	1	b9	-0.0027	-2.3418
1080.7272	1080.7313	6	b55	-0.0246	-0.3855
1088.8067	1088.8102	4	b37	-0.0140	-0.5740
1149.5765	1149.5816	4	b39	-0.0204	-0.5437
1152.8023	1152.8063	5	b49	-0.0200	-0.4337
1158.6272	1158.6315	2	y20	-0.0086	-1.0789
1209.8627	1209.8673	4	b41	-0.0184	-0.5166
1223.1485	1223.1528	2	y21	-0.0086	-1.0219
1240.6361	1240.6389	3	b32	-0.0084	-0.6717
1251.2394	1251.2434	5	b53	-0.0200	-0.3996
1270.1511	1270.1529	4	b43	-0.0072	-0.4921
1285.4702	1285.4744	5	y55	-0.0210	-0.3890
1296.6714	1296.6761	5	b55	-0.0235	-0.3856
1309.1732	1309.1782	4	b44	-0.0200	-0.4774
1316.3424	1316.3479	3	b34	-0.0165	-0.6331
1333.9399	1333.9453	4	b45	-0.0216	-0.4685
1362.2160	1362.2163	4	b46	-0.0012	-0.4588
1370.6972	1370.7023	3	b35	-0.0153	-0.6080
1408.3870	1408.3970	3	b36	-0.0300	-0.5917
1418.2423	1418.2480	2	y24	-0.0114	-0.8814
1440.7527	1440.7561	4	b49	-0.0136	-0.4338
1451.4045	1451.4112	3	b37	-0.0201	-0.5742
1532.4346	1532.4397	3	b39	-0.0153	-0.5438
1660.3752	1660.3803	2	y29	-0.0102	-0.7528

50S ribosomal protein L29

Uniprot accession number: Q839F6
Charge state: 7+
Observed monoisotopic mass: 7324.9912 Da
Sequence:
MKVKEIRELTTAEMLDKEKQLKEELFNLRFQLATGQLENTARIKEVRQSIARIKT
VLREQAN

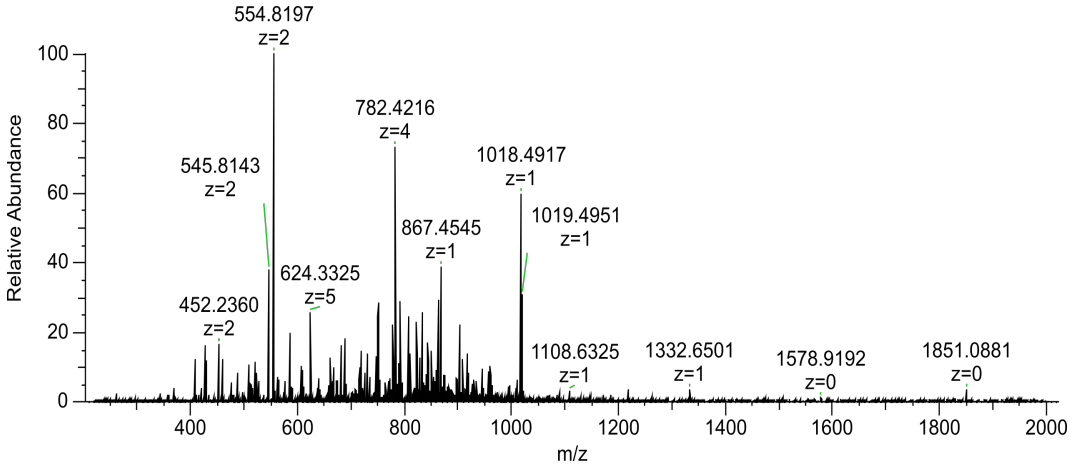


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
598.3359	598.3381	1	b5-H2O	-0.0022	-3.6769
655.7141	655.7166	3	y17-H2O	-0.0075	-3.8126
661.3897	661.3921	3	y17	-0.0072	-3.6287
694.9029	694.9059	4	y24-H2O	-0.0120	-4.3171
699.4061	699.4085	4	y24	-0.0096	-3.4315
765.4185	765.4214	2	b13	-0.0058	-3.7888
830.9385	830.9417	2	b14	-0.0064	-3.8511
855.4834	855.4865	2	y15-NH3	-0.0062	-3.6237
863.9964	863.9998	2	y15	-0.0068	-3.9352
881.8206	881.8239	3	b22	-0.0099	-3.7422
885.9013	885.9049	5	y38	-0.0180	-4.0636
887.4806	887.4837	2	b15	-0.0062	-3.4930
904.3410	904.3434	6	y46-H2O	-0.0144	-2.6539
907.1779	907.1812	6	y46	-0.0198	-3.6376
934.1179	934.1166	5	y40-NH3	0.0065	1.3917
944.9937	944.9972	2	b16	-0.0070	-3.7037
959.5366	959.5388	5	y41-H2O	-0.0110	-2.2928
963.1376	963.1409	5	y41	-0.0165	-3.4263
967.0366	967.0398	6	y49	-0.0192	-3.3091
982.3481	982.3524	5	y42-NH3	-0.0215	-4.3772
985.7552	985.7577	5	y42	-0.0125	-2.5361
1007.9634	1007.9641	5	y43-NH3	-0.0035	-0.6945
1011.3688	1011.3694	5	y43	-0.0030	-0.5933
1028.7150	1028.7191	7	y61	-0.0287	-3.9855
1033.5798	1033.5831	5	y44-NH3	-0.0165	-3.1928
1059.3920	1059.3916	5	y45-NH3	0.0020	0.3776
1062.7936	1062.7970	5	y45	-0.0170	-3.1991
1064.5569	1064.5607	2	b18-H2O	-0.0076	-3.5695
1073.5620	1073.5660	2	b18	-0.0080	-3.7259
1085.0062	1085.0106	5	y46-NH3	-0.0220	-4.0553
1108.0146	1108.0160	5	y47-NH3	-0.0070	-1.2635

1111.4173	1111.4213	5	y47	-0.0200	-3.5990
1116.4467	1116.4498	6	y57-NH3	-0.0186	-2.7767
1137.6093	1137.6134	2	b19	-0.0082	-3.6040
1162.1416	1162.1438	6	y59	-0.0132	-1.8931
1167.3923	1167.3940	4	y40-NH3	-0.0068	-1.4562
1171.6751	1171.6793	1	y10	-0.0042	-3.5846
1175.6552	1175.6535	6	y60-H2O	0.0102	1.4460
1178.6494	1178.6552	6	y60	-0.0348	-4.9209
1186.4511	1186.4499	5	b50	0.0060	1.0114
1192.6332	1192.6374	2	b20-H2O	-0.0084	-3.5216
1197.1638	1197.1666	6	y61-NH3	-0.0168	-2.3389
1199.4134	1199.4177	4	y41-NH3	-0.0172	-3.5851
1227.6865	1227.6887	4	y42-NH3	-0.0088	-1.7920
1258.1807	1258.1848	2	b21	-0.0082	-3.2587
1300.2067	1300.2112	4	b44-H2O	-0.0180	-3.4610
1304.7080	1304.7138	4	B44	-0.0232	-4.4454
1322.2279	1322.2322	2	b22	-0.0086	-3.2521
1332.7188	1332.7178	4	b45-NH3	0.0040	0.7503
1336.9692	1336.9744	4	b45	-0.0208	-3.8894
1386.7483	1386.7535	2	b23	-0.0104	-3.7498
1400.7882	1400.7930	1	b12	-0.0048	-3.4266
1451.2686	1451.2748	2	b24	-0.0124	-4.2721
1511.4811	1511.4800	3	b38	0.0033	0.7278
1529.8292	1529.8356	1	b13	-0.0064	-4.1835
1660.8712	1660.8761	1	b14	-0.0049	-2.9503
1773.9542	1773.9601	1	b15	-0.0059	-3.3259
1888.9792	1888.9871	1	b16	-0.0079	-4.1821

UPF0337 protein (gene EF_1180)

Uniprot accession number: Q836D5
Charge state: 11+
Observed monoisotopic mass: 8873.5100 Da
Sequence:
ADLKGRFDDAKDKVEGTAKEAQGKVTDDKGKELEGKAQSTFADVKDKARDAG
DDLKEGAEKLTDKVKEGFEDLKDKFSKDK

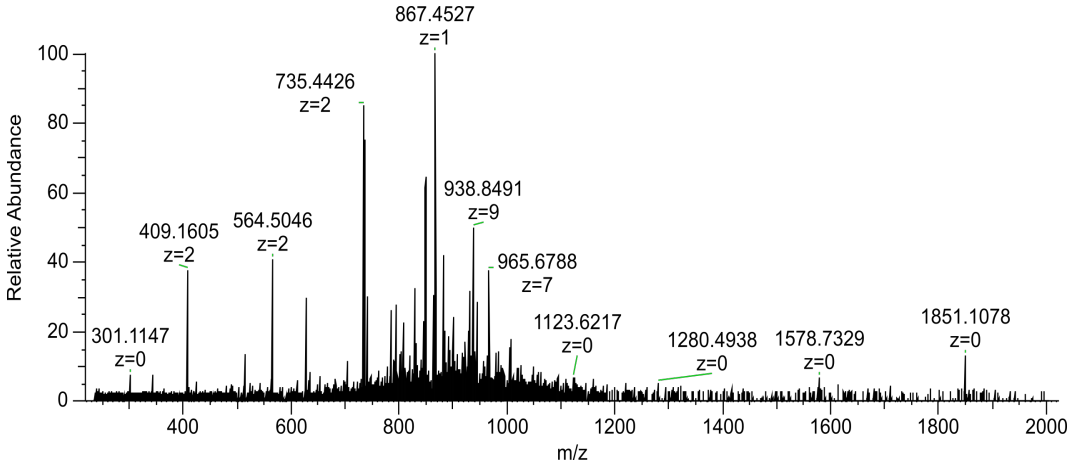


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
428.2122	428.2504	1	b4	-0.0382	-5.8377
452.2360	452.2378	2	b8	-0.0036	-2.7640
487.5859	487.5879	3	b13	-0.0060	-1.7091
498.2776	498.2796	2	y8	-0.0040	-2.5086
509.7492	509.7513	2	b9	-0.0042	-2.4522
520.6087	520.6107	3	b14	-0.0060	-1.6007
554.8197	554.8217	2	y9	-0.0040	-2.2530
563.6228	563.6249	3	b15	-0.0063	-1.4785
624.3325	624.3352	1	y5	-0.0027	-4.0043
640.5819	640.5845	4	b24	-0.0104	-0.9757
661.3426	661.3453	5	b31	-0.0135	-0.7560
665.3484	665.3516	4	b25	-0.0128	-0.9394
666.8283	666.8308	2	b12	-0.0050	-1.8745
716.8758	716.8787	6	y38	-0.0174	-0.5812
719.3672	719.3703	4	b27	-0.0124	-0.8688
721.8750	721.8789	4	y25	-0.0156	-0.8658
730.8752	730.8782	2	b13	-0.0060	-1.7103
748.1232	748.1270	4	b28	-0.0152	-0.8354
782.1707	782.1737	4	y27	-0.0120	-0.7991
790.7713	790.7753	11	y79	-0.0440	-0.2874
791.6122	791.6149	5	y35	-0.0135	-0.6316
822.2677	822.2711	7	b54	-0.0238	-0.4343
841.7242	841.7263	7	y53	-0.0147	-0.4243
843.4195	843.4227	2	y14	-0.0064	-1.4821
855.1819	855.1852	8	b64	-0.0264	-0.3654
858.1564	858.1588	7	y54	-0.0168	-0.4162
863.8798	863.8832	9	b72	-0.0306	-0.3215
867.4545	867.4571	1	y7	-0.0026	-2.8820
873.7424	873.7480	10	b80	-0.0560	-0.2861
903.4649	903.4683	1	b8	-0.0034	-2.7671
907.4660	907.4702	2	y15	-0.0084	-1.3775

916.6606	916.6640	5	b43	-0.0170	-0.5455
919.1513	919.1553	3	y24	-0.0120	-0.9066
929.6402	929.6443	6	y50	-0.0246	-0.4482
959.1551	959.1484	6	b54	0.0402	-0.4344
962.1661	962.1695	3	y25	-0.0102	-0.8661
964.2927	964.2954	5	y43	-0.0135	-0.5185
997.1636	997.1669	3	b28	-0.0099	-0.8357
1018.4917	1018.4952	1	b9	-0.0035	-2.4546
1108.6325	1108.6361	1	y9	-0.0036	-2.2550
1217.6236	1217.6273	1	b11	-0.0037	-2.0532
1332.6501	1332.6543	1	b12	-0.0042	-1.8760

DNA-binding protein HU

Uniprot accession number: Q834T3
Charge state: 11+
Observed monoisotopic mass: 9519.1340 Da
Sequence:
ANKAELIENVASSTGLTKKDATAAVDAVFSTIQETLAKGEKVQLIGFGNFEVRERA
ARKGRNPQTGQEIQIAASKVPAFKPGKALKDAVK



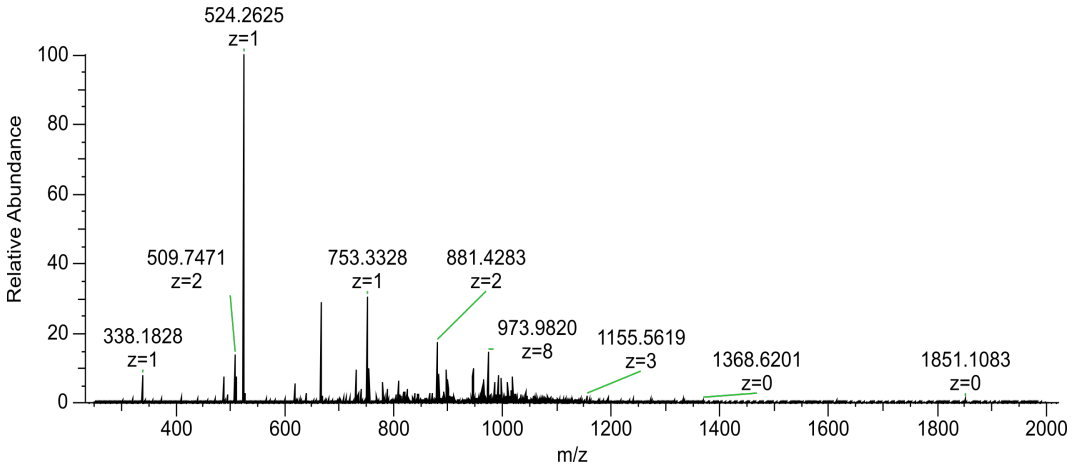
Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
627.3436	627.3461	1	b6	-0.0025	-3.9850
735.4427	735.4456	2	y14	-0.0058	-1.6996
740.4271	740.4301	1	b7	-0.0030	-3.3764
784.9765	784.9798	2	y15	-0.0066	-1.5924
844.8431	844.8473	8	y62	-0.0336	-0.3699
930.6203	930.6253	9	y79	-0.0450	-0.2985
938.5102	938.5183	9	y80	-0.0729	-0.2960

944.3824	944.3861	7	y61	-0.0259	-0.3782
965.3935	965.3959	7	y62	-0.0168	-0.3699
979.5453	979.5485	7	y63	-0.0224	-0.3646
989.6929	989.6967	7	y64	-0.0266	-0.3609
1006.1261	1006.1291	7	y65	-0.0210	-0.3550
1020.2806	1020.2817	7	y66	-0.0077	-0.3500
1082.5814	1082.5840	1	b10	-0.0026	-2.3093

Klebsiella pneumoniae KP257

CscD domain-containing protein

Uniprot accession number: A6TGV4
Charge state: 9+
Observed monoisotopic mass: 8303.0402 Da
Sequence:
MNKDEIGGNWKQFKGKAKEQWGKLTDDDMTVI
EGKRDQLVGKIQERYGYEKDQAEKEVSDWEHKNDYRW
Signal peptide 1-17 cleaved

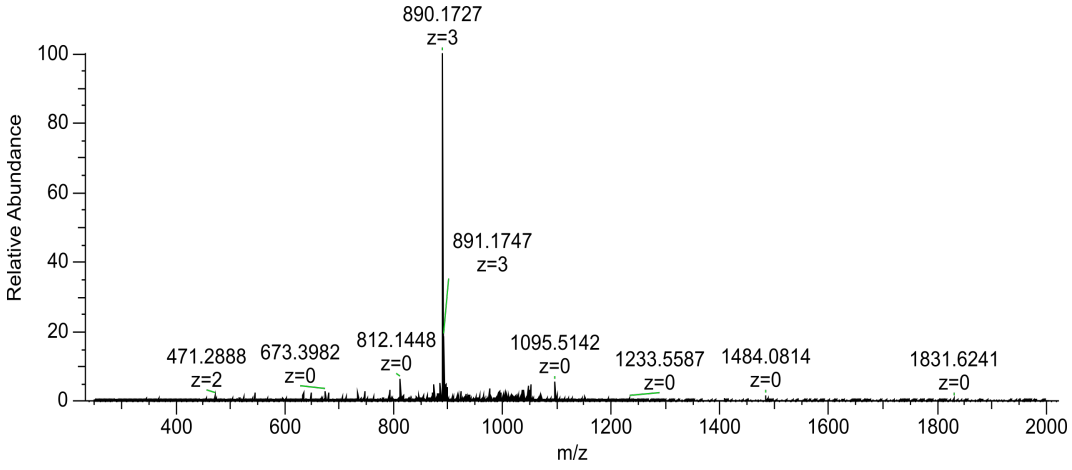


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
509.7471	509.7463	2	y7	0.0016	1.5694
524.2625	524.2616	1	y3	0.0009	1.7167
618.2560	618.2552	1	b5	0.0008	1.2940
667.3085	667.3073	2	y9	0.0024	1.7983
731.3404	731.3393	1	b6	0.0011	1.5041
740.6763	740.6750	3	y17	0.0039	1.7552
753.3328	753.3315	1	y5	0.0013	1.7257
768.3382	768.3367	2	y11	0.0030	1.9523

788.3922	788.3907	6	y38	0.0090	1.9026
804.9036	804.9021	6	y39	0.0090	1.8636
812.9020	812.9003	4	b28	0.0068	2.0913
817.8726	817.8710	2	y12	0.0032	1.9563
821.7437	821.7434	6	y40	0.0018	0.3651
835.4242	835.4216	3	b21	0.0078	3.1122
838.9157	838.9143	2	b14	0.0028	1.6688
843.5844	843.5835	6	y41	0.0054	1.0669
845.6619	845.6604	4	b29	0.0060	1.7738
881.4283	881.4264	1	y6	0.0019	2.1556
897.4438	897.4421	5	y36	0.0085	1.8943
945.8691	945.8674	5	y38	0.0085	1.7973
965.6823	965.6811	5	y39	0.0060	1.2426
973.4801	973.4792	8	b66	0.0072	0.9245
985.8920	985.8907	5	y40	0.0065	1.3186
992.2394	992.2387	8	y66	0.0056	0.7055
996.7831	996.7829	7	b60	0.0014	0.2006
1008.2471	1008.2505	8	y67	-0.0272	-3.3722
1015.0155	1015.0116	6	b52	0.0234	3.8423
1018.4875	1018.4853	1	y7	0.0022	2.1601

30S ribosomal protein S16

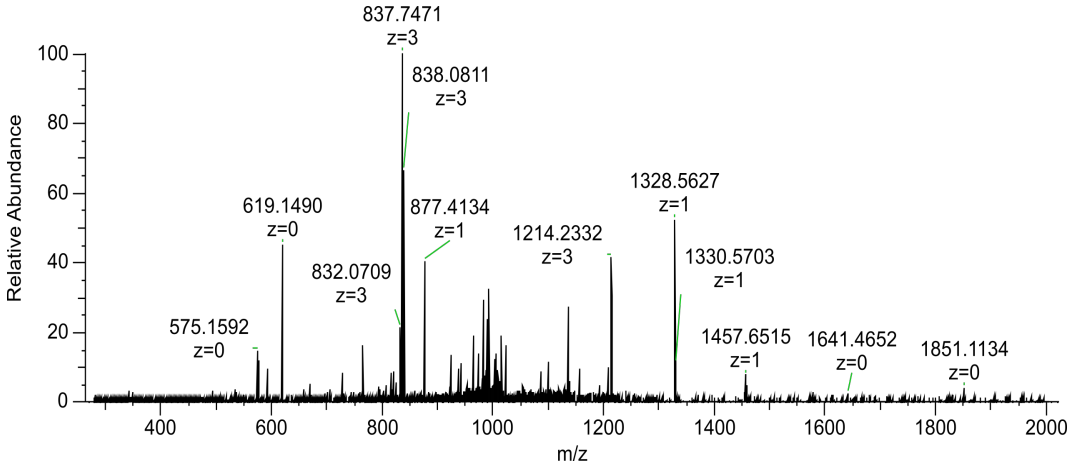
Uniprot accession number: A6TCL7
Charge state: 10+
Observed monoisotopic mass: 9084.8622 Da
Sequence:
MVTIRLARHGAKKRPFYQVVVTDSRNARNGRFIERVGFFNPIANGAEEETR
LDLDRIAHWVGQGATVSDRVAALIKAANKAA



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
540.3307	540.3315	3	b14	-0.0024	-1.4806
545.3036	545.3042	1	y6	-0.0006	-1.1003
648.9084	648.9092	2	y13	-0.0016	-1.2328
673.3986	673.3991	1	y7	-0.0005	-0.7425
751.4208	751.4216	4	y29	-0.0032	-1.0646
782.2627	782.2641	6	b40	-0.0084	-1.7897
798.4390	798.4396	6	b41	-0.0036	-0.7515
817.2847	817.2869	6	b42	-0.0132	-2.6918
829.1270	829.1265	6	b43	0.0030	0.6030
872.7359	872.7366	4	y33	-0.0028	-0.8021
889.8383	889.8391	3	b23	-0.0024	-0.8990
893.4819	893.4849	10	b80	-0.0300	-3.3576
902.6218	902.6237	7	b55	-0.0133	-2.1050
917.9172	917.9167	7	y59	0.0035	0.5447
933.9986	933.9980	6	b49	0.0036	0.6424
937.2569	937.2579	4	y35	-0.0040	-1.0669
1014.8732	1014.8746	6	b53	-0.0084	-1.3795
1019.0418	1019.0445	8	y74	-0.0216	-2.6495
1024.2937	1024.2950	8	b73	-0.0104	-1.2692
1030.0464	1030.0439	4	y39	0.0100	2.4271
1038.4304	1038.4305	8	b74	-0.0008	-0.0963
1047.8026	1047.8032	4	y40	-0.0024	-0.5726
1052.5620	1052.5660	8	b75	-0.0320	-3.8002
1100.3381	1100.3374	4	y42	0.0028	0.6362
1334.2555	1334.2550	2	b23	0.0010	0.3747

Uncharacterised protein (gene yciG)

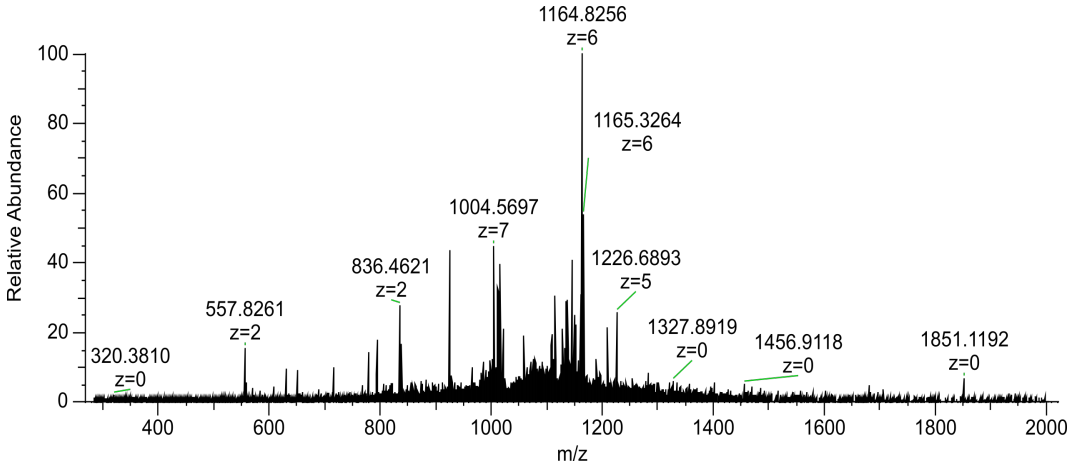
Uniprot accession number: A6T7L1
Charge state: 6+
Observed monoisotopic mass: 6147.8875 Da
Sequence:
AEHRGGSGNFAEDREKASEAGRKGGQHSGGNFKNDPQRASEAGKKGGQNS
HGGGRKSDNS



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
837.4128	837.4116	3	y25	0.0036	1.4330
877.4134	877.4122	1	y9	0.0012	1.3677
965.0732	965.0732	5	y47	0.0000	0.0000
989.1400	989.1403	6	b58	-0.0018	-0.3033
992.3099	992.3080	6	y58	0.0114	1.9147
1213.5364	1213.5345	1	b12	0.0019	1.5657
1328.5627	1328.5614	1	b13	0.0013	0.9785
1457.6515	1457.6476	1	y15	0.0039	2.6755

50S ribosomal protein L29

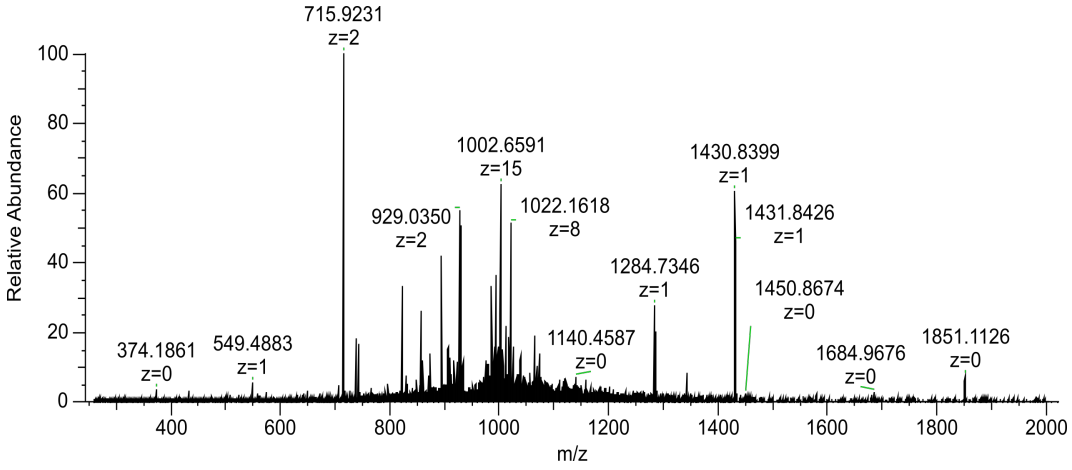
Uniprot accession number: A6TEW4
Charge state: 7+
Observed monoisotopic mass: 7239.0382 Da
Sequence:
MKAKELREKSVEELNAELLNLLREQFNLRMQAASGQLQQTHLLKQVRRDVA
RVKTLLTQKAGA



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
557.8260	557.8237	2	b9	0.0046	4.1232
650.8765	650.8739	2	b11	0.0052	3.9946
715.3985	715.3952	2	b12	0.0066	4.6128
728.4562	728.4539	2	y14	0.0046	3.1574
779.9196	779.9165	2	b13	0.0062	3.9748
836.4620	836.4585	2	b14	0.0070	4.1843
882.8727	882.8735	5	b38	-0.0040	-0.9061
893.4833	893.4800	2	b15	0.0066	3.6934
993.5233	993.5199	2	b17	0.0068	3.4222
1004.1397	1004.1362	7	b60	0.0245	3.4856
1014.2888	1014.2844	7	b61	0.0308	4.3380
1077.6154	1077.6125	5	y47	0.0145	2.6911
1109.6266	1109.6237	6	y58	0.0174	2.6135
1152.4829	1152.4800	6	y60	0.0174	2.5163
1163.0584	1163.0539	5	y51	0.0225	3.8691
1164.3238	1164.3195	6	y61	0.0258	3.6931
1208.6802	1208.6761	5	y53	0.0205	3.3921
1226.0877	1226.0825	5	y54	0.0260	4.2412
1455.9068	1455.9006	1	y14	0.0062	4.2585

DNA-binding protein HU-α

Uniprot accession number: A6TGQ7
Charge state: 10+
Observed monoisotopic mass: 9471.1692 Da
Sequence:
MNKTQLIDVIADKADLSKAQAKAALESTLAAITESLKEGDAVQLVGFGTFKVNH
RAERTGRNPQTGKEIKIAAANVPAFVSGKALKDAVK

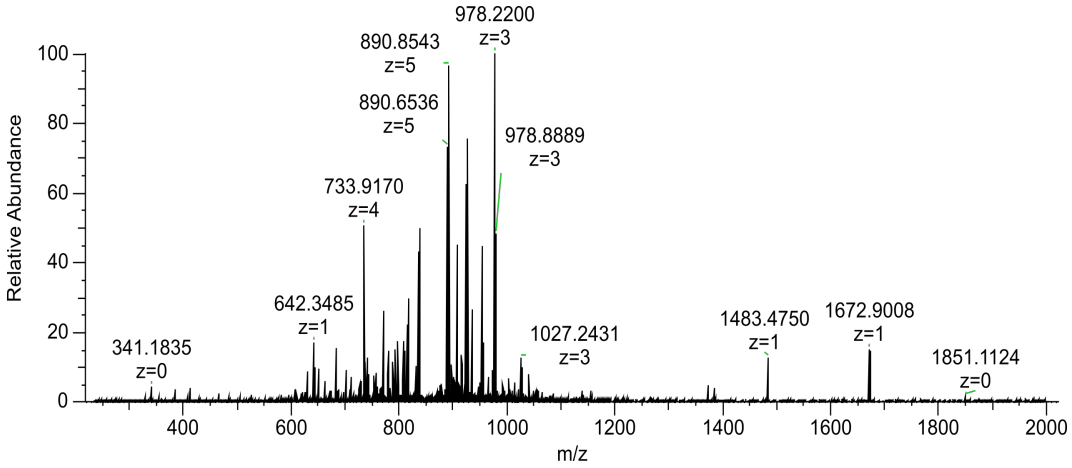


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
475.2339	475.2333	1	b4	0.0006	1.2625
715.9232	715.9219	2	y14	0.0026	1.8158
744.4628	744.4614	1	y7	0.0014	1.8806
765.4573	765.4561	2	y15	0.0024	1.5677
795.1065	795.1053	3	b22	0.0036	1.5092
822.4793	822.4776	2	y16	0.0034	2.0669
829.4620	829.4600	1	b7	0.0020	2.4112
857.9978	857.9962	2	y17	0.0032	1.8648
893.5162	893.5147	2	y18	0.0030	1.6788
904.2813	904.2796	9	y78	0.0153	1.8799
929.0348	929.0333	2	y19	0.0030	1.6146
976.5423	976.5451	5	y46	-0.0140	-2.8673
979.5359	979.5330	8	b74	0.0232	2.9606
985.5773	985.5753	2	y20	0.0040	2.0293
1006.1749	1006.1720	8	b76	0.0232	2.8822
1016.6116	1016.6099	1	y10	0.0017	1.6722
1038.2718	1038.2690	7	b68	0.0196	2.6968
1043.5575	1043.5554	1	b9	0.0021	2.0124
1056.4163	1056.4158	6	y60	0.0030	0.4733
1070.9196	1070.9183	3	b31	0.0039	1.2139
1072.7257	1072.7231	7	b70	0.0182	2.4237
1262.7485	1262.7467	1	y12	0.0018	1.4255
1342.7072	1342.7035	1	b12	0.0037	2.7556
1430.8393	1430.8366	1	y14	0.0027	1.8870

Uncharacterised protein (gene KPN_00497)

Uniprot accession number: A6T5S6
Charge state: 9+

Observed monoisotopic mass: 7698.9845 Da
Sequence:
AQLITKEEVKHFKLTKVGPISVGPSGGEFSSPSDLHDQLSKLADEKGGKYY
VITAAREHGPNFEATAEVYK
Signal peptide 1-19 cleaved
R→K mutation at position 49

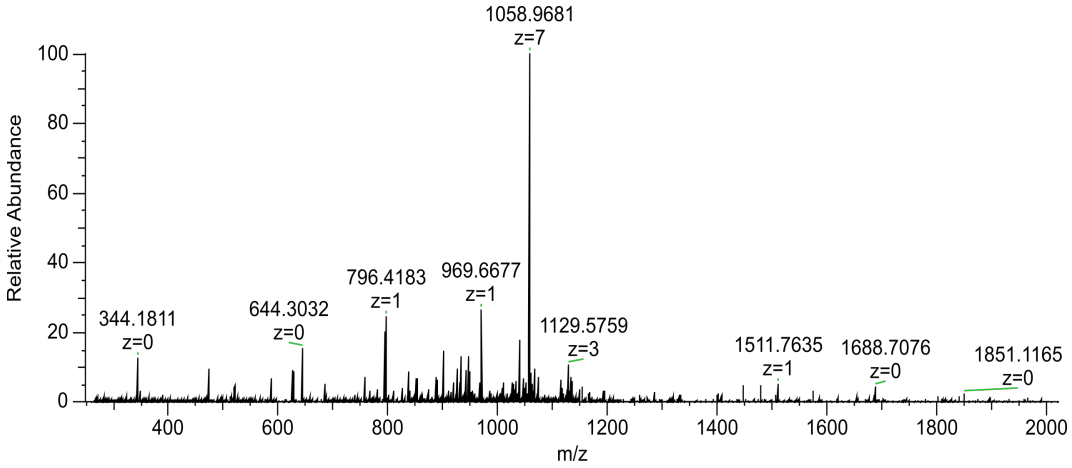


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
663.3250	663.3222	2	y12	0.0056	4.2212
672.8948	672.8916	4	b25	0.0128	4.7556
687.1501	687.1470	4	b26	0.0124	4.5114
701.4053	701.4024	4	b27	0.0116	4.1346
733.6660	733.6630	4	b28	0.0120	4.0891
742.2117	742.2087	6	y40	0.0180	4.0420
757.8874	757.8845	4	y27	0.0116	3.8264
770.4334	770.4301	4	b29	0.0132	4.2833
781.4126	781.4090	1	y7	0.0036	4.6071
791.1918	791.1881	4	b30	0.0148	4.6765
795.7338	795.7308	6	y43	0.0180	3.7701
816.4881	816.4843	3	b22	0.0114	4.6541
830.6303	830.6266	5	y37	0.0185	4.4545
835.4953	835.4915	3	b23	0.0114	4.5482
890.4531	890.4490	5	y40	0.0205	4.6044
907.8595	907.8554	5	y41	0.0205	4.5161
925.2659	925.2618	5	y42	0.0205	4.4312
934.8716	934.8674	3	b27	0.0126	4.4926
954.6797	954.6755	5	y43	0.0210	4.3994
977.8861	977.8816	3	b28	0.0135	4.6018
991.8923	991.8883	5	y45	0.0200	4.0327
1026.9091	1026.9044	3	b29	0.0141	4.5769
1040.1145	1040.1095	5	y48	0.0250	4.8072
1140.6677	1140.6623	1	b10	0.0054	4.7341

Acinetobacter baumannii AYE

Uncharacterised protein (gene ABAYE1298)

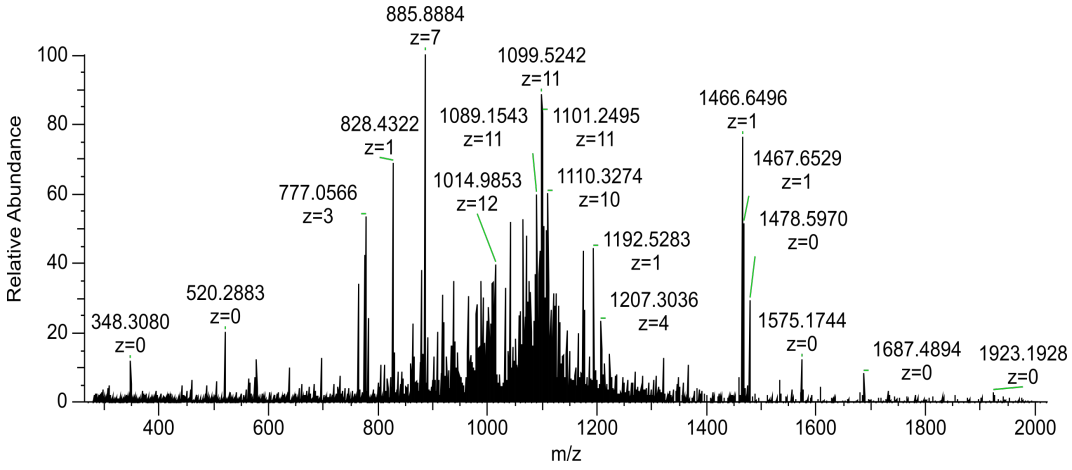
Uniprot accession number: B0V801
Charge state: 9+
Observed monoisotopic mass: 8713.3637 Da
Sequence:
MNTLNINDIKKHADVIASCGTKVGTVDHLEGENQLKLTKDENDQHHLIPTSWIG
EVKEDQVILNKNSEEVKENWQAI



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
888.4559	888.4574	1	y7	-0.0015	-1.6883
889.4547	889.4535	6	b48	0.0072	1.3491
941.4726	941.4739	5	b43	-0.0065	-1.3808
1032.7765	1032.7781	8	y73	-0.0128	-1.5492
1073.7964	1073.8000	8	y76	-0.0288	-3.3526
1046.9146	1046.9137	8	y74	0.0072	0.8597
1125.9684	1125.9677	5	y48	0.0035	0.6217
1128.5739	1128.5754	3	y29	-0.0045	-1.3291
1134.2387	1134.2402	6	b61	-0.0090	-1.3225
1153.0863	1153.0875	6	b62	-0.0072	-1.0407
1166.2678	1166.2701	3	y30	-0.0069	-1.9721
1285.6933	1285.6933	1	b11	0.0000	0.0000
1332.6435	1332.6430	1	y11	0.0005	0.3752
1446.6844	1446.6859	1	y12	-0.0015	-1.0369

Uncharacterised protein (gene ABAYE2274)

Uniprot accession number: B0VBH8
Charge state: 12+
Observed monoisotopic mass: 12309.7539 Da
Sequence:
TDQNTDHAQLVAGDHNYSDSRWRESYLTRPYYQEAQLTTPDLDYDRDFSAAY
ELGHRARSESKEGTQFKDMEGSLQQKWEELKAESRLKWEHAKQAIKDAWDDM

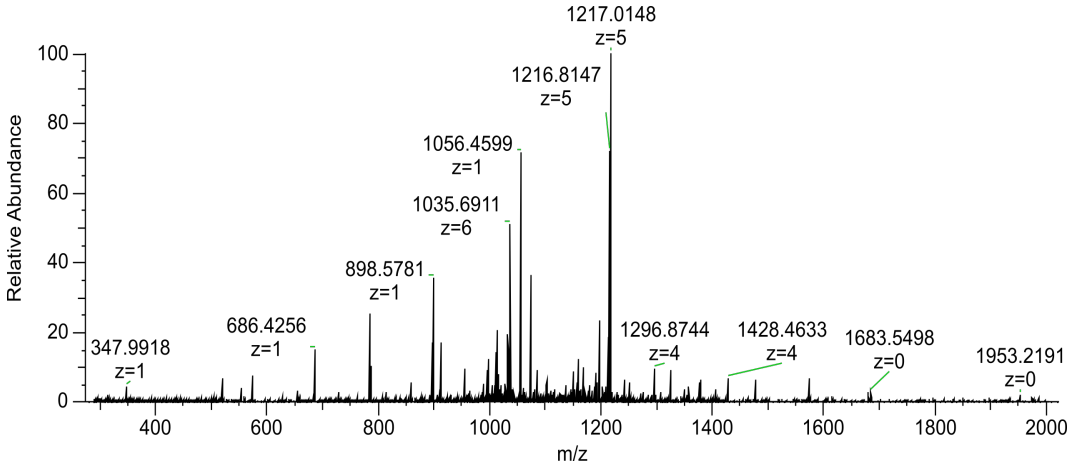


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
811.3106	811.3097	B7	0.0009	1.1660
1123.4891	1123.4894	B10	-0.0003	-0.2563
1465.6493	1465.6433	B14	0.0059	4.0515
2081.8510	2081.8675	B19	-0.0165	-7.9140
4295.9052	4295.9281	B36	-0.0229	-5.3381
4823.1853	4823.1872	B41	-0.0019	-0.3948
8455.7853	8455.8462	B72	-0.0609	-7.1995
9097.1577	9097.1959	B78	-0.0382	-4.2010
879.3412	879.3433	Y7	-0.0021	-2.3483
992.4244	992.4273	Y8	-0.0029	-2.9554
1063.4650	1063.4644	Y9	0.0006	0.5473
1191.5166	1191.5230	Y10	-0.0064	-5.3654
1319.6129	1319.6180	Y11	-0.0051	-3.8579
1527.7159	1527.7140	Y13	0.0019	1.2745
2327.1421	2327.1480	Y19	-0.0060	-2.5740
2768.3935	2768.4068	Y23	-0.0133	-4.8002
4229.0299	4229.0309	Y35	-0.0010	-0.2469
6040.9574	6040.9700	Y51	-0.0126	-2.0914
7486.5120	7486.5799	Y63	-0.0679	-9.0677
7486.5346	7486.5799	Y63	-0.0453	-6.0523
7698.6510	7698.6596	Y65	-0.0086	-1.1193
11016.1012	11016.1722	Y92	-0.0710	-6.4468

11087.2128	11087.2093	Y93	0.0035	0.3134
11186.1847	11186.2777	Y94	-0.0930	-8.3164
12093.6962	12093.6925	Y102	0.0037	0.3088

Uncharacterised protein (gene ABAYE1876)

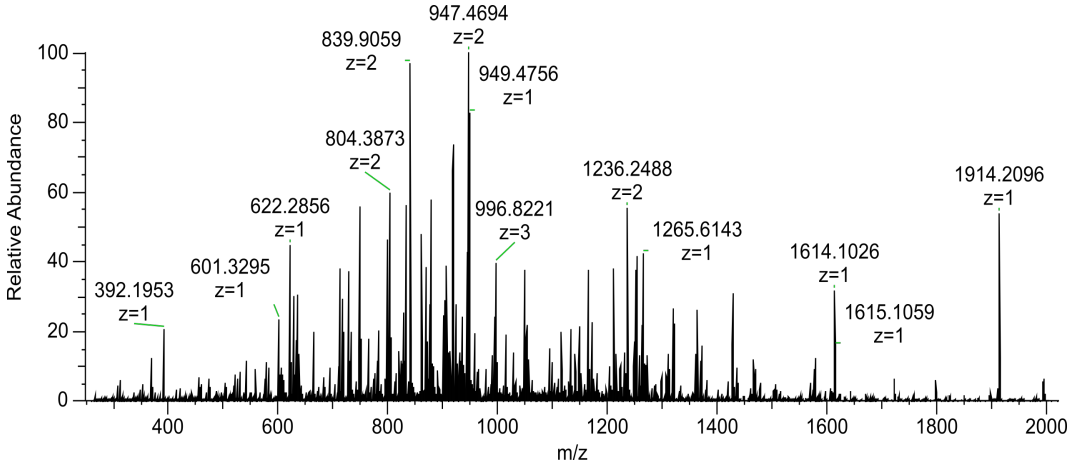
Uniprot accession number: B0V4B5
Charge state: 6+
Observed monoisotopic mass: 6323.1751 Da
Sequence:
MLKSILVLGTGIAIGMCMYKKKQKNSKSFSTDSDTDSLISKDNNKSKDDNS
LDNAVQV



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
685.4206	685.4197	B6	0.0009	1.3656
784.4873	784.4881	B7	-0.0007	-0.9216
897.5724	897.5721	B8	0.0003	0.3454
954.5918	954.5936	B9	-0.0017	-1.8291
4120.0951	4120.1080	B38	-0.0129	-3.1310
5978.9244	5978.9737	B55	-0.0493	-8.2533
5978.9260	5978.9737	B55	-0.0477	-7.9800
6078.0274	6078.0421	B56	-0.0147	-2.4166
6078.0407	6078.0421	B56	-0.0014	-0.2325

Entericidin B

Uniprot accession number: B0V4V9
Charge state: 4+
Observed monoisotopic mass: 3903.1785 Da
Sequence:
MMKKVLVASVMVAFVLTGCNTFKGFGQDVSKAGDAVTNTAQKTENKM
Signal peptide 1-18 cleaved
Lipid attached to the N-terminal cysteine: lipid-CNTFKGFGQDVSKAGDAVTNTAQ
KTENKM

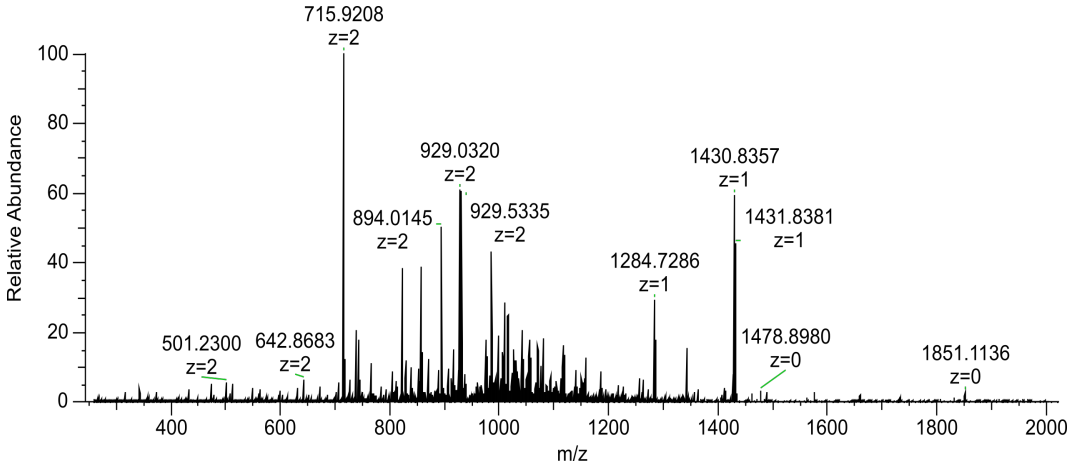


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
392.1953	392.1962	1	y3	-0.0009	-2.2948
525.7653	525.7660	2	y9	-0.0014	-0.8559
582.7869	582.7875	2	y10	-0.0012	-0.7722
602.9712	602.9720	3	y17	-0.0024	-0.4975
622.2856	622.2865	1	y5	-0.0009	-1.4463
633.3107	633.3114	2	y11	-0.0014	-0.7106
665.0046	665.0055	3	y19	-0.0027	-0.4511
682.8443	682.8456	2	y12	-0.0026	-0.6590
703.3470	703.3478	3	y20	-0.0024	-0.4265
718.3632	718.3641	2	y13	-0.0018	-0.6264
747.6178	747.6186	4	y28	-0.0032	-0.3010
750.3803	750.3815	1	y6	-0.0012	-1.1994
765.0400	765.0411	3	y22	-0.0033	-0.3921
775.8759	775.8776	2	y14	-0.0034	-0.5800
804.3873	804.3883	2	y15	-0.0020	-0.5594
839.9059	839.9069	2	y16	-0.0020	-0.5358
875.7684	875.7694	3	y25	-0.0030	-0.3426
878.4387	878.4400	1	y7	-0.0013	-1.0245
903.9538	903.9544	2	y17	-0.0012	-0.4978

Enterobacter cloacae S11

DNA-binding protein

Uniprot accession number: A0A094ZZI7
Charge state: 10+
Observed monoisotopic mass: 9499.1592 Da
Sequence:
MNKTQLIDVIADKADLSKVQAKAALESTLAAITESLKEGDAVQLVGFGTFKVNHRA
ERTGRNPQTGKEIKIAAANVPAFVSGKALKDAVK

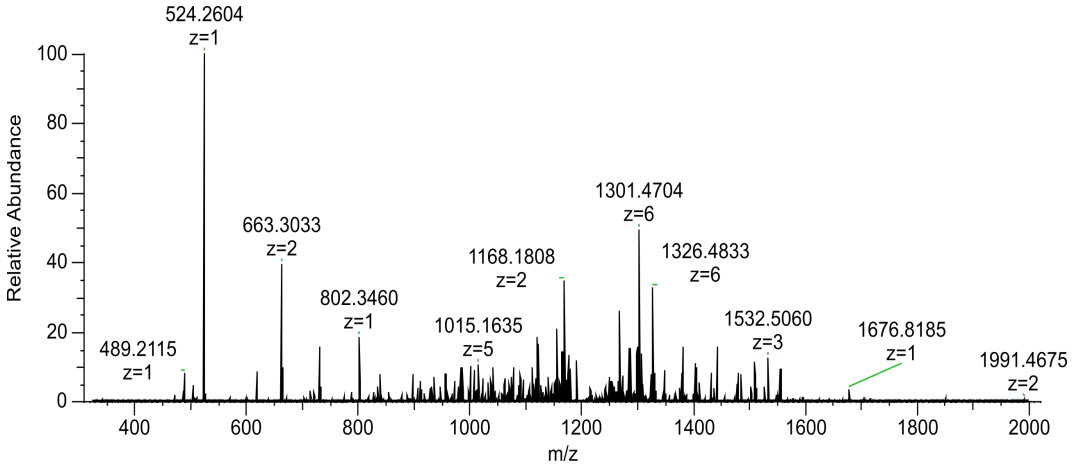


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
432.2445	432.2453	4	y4	-0.0032	-1.8508
475.2327	475.2333	1	b4	-0.0006	-1.2625
477.6166	477.6170	3	y14	-0.0012	-0.8375
508.8079	508.8086	2	y10	-0.0014	-1.3758
560.3397	560.3402	1	y5	-0.0005	-0.8923
603.2910	603.2919	1	b5	-0.0009	-1.4918
631.8757	631.8770	2	y12	-0.0026	-2.0574
673.4230	673.4243	1	y6	-0.0013	-1.9304
715.9208	715.9219	2	y14	-0.0022	-1.5365
737.7781	737.7790	3	y22	-0.0027	-1.2199
744.4601	744.4614	1	y7	-0.0013	-1.7462
765.4546	765.4561	2	y15	-0.0030	-1.9596
771.4194	771.4214	2	b14	-0.0040	-2.5926
792.6863	792.6891	4	b30	-0.0112	-3.5323
804.4473	804.4491	3	b22	-0.0054	-2.2376
810.4465	810.4483	4	b31	-0.0072	-2.2210
822.4765	822.4776	2	y16	-0.0022	-1.3374
851.8063	851.8071	3	b24	-0.0024	-0.9392

854.1449	854.1459	6	y48	-0.0060	-1.1708
857.9950	857.9962	2	y17	-0.0024	-1.3986
889.5009	889.5018	3	b25	-0.0027	-1.0118
893.5133	893.5147	2	y18	-0.0028	-1.5668
907.3909	907.3942	9	y78	-0.0297	-3.6368
911.1706	911.1716	6	y52	-0.0060	-1.0975
915.9125	915.9135	5	y43	-0.0050	-1.0918
929.0320	929.0333	2	y19	-0.0026	-1.3993
932.6777	932.6787	6	y53	-0.0060	-1.0722
956.7308	956.7315	5	y45	-0.0035	-0.7317
976.5450	976.5451	5	y46	-0.0005	-0.1024
983.0321	983.0370	8	b74	-0.0392	-4.9846
985.5736	985.5753	2	y20	-0.0034	-1.7249
999.1596	999.1619	5	y47	-0.0115	-2.3019
1009.6728	1009.6759	8	b76	-0.0248	-3.0703
1015.0033	1015.0072	9	y87	-0.0351	-3.8423
1024.7717	1024.7737	5	y48	-0.0100	-1.9517
1025.7272	1025.7289	6	y58	-0.0102	-1.6574
1029.2375	1029.2400	9	y88	-0.0225	-2.4290
1032.9018	1032.9040	3	b29	-0.0066	-2.1299
1042.1463	1042.1480	7	y70	-0.0119	-1.6312
1053.5652	1053.5674	4	b40	-0.0088	-2.0881
1056.4165	1056.4158	6	y60	0.0042	0.6626
1068.2537	1068.2553	6	y61	-0.0096	-1.4978
1071.3253	1071.3267	4	b41	-0.0056	-1.3068
1076.7270	1076.7276	7	b70	-0.0042	-0.5572
1080.2600	1080.2620	3	b31	-0.0060	-1.8514
1118.4460	1118.4493	6	y64	-0.0198	-2.9505
1223.6620	1223.6642	3	b35	-0.0066	-1.7979
1227.6728	1227.6766	1	b11	-0.0038	-3.0953
1261.3575	1261.3588	3	b36	-0.0039	-1.0306
1342.7035	1342.7035	1	b12	0.0000	0.0000
1428.0990	1428.0999	3	b41	-0.0027	-0.6302
1430.8357	1430.8366	1	y14	-0.0009	-0.6290
1461.1238	1461.1227	3	b42	0.0033	0.7528

CsbD family protein

Uniprot accession number: A0A0M2G4R3
Charge state: 7+
Observed monoisotopic mass: 8322.0209 Da
Sequence:
MNKDEIGGNWKQFKGKAKEQWGKLTDDDMTVIEGKRDQLVGKIQERYGYEKDQ
AENEVKDWETRNDYRW

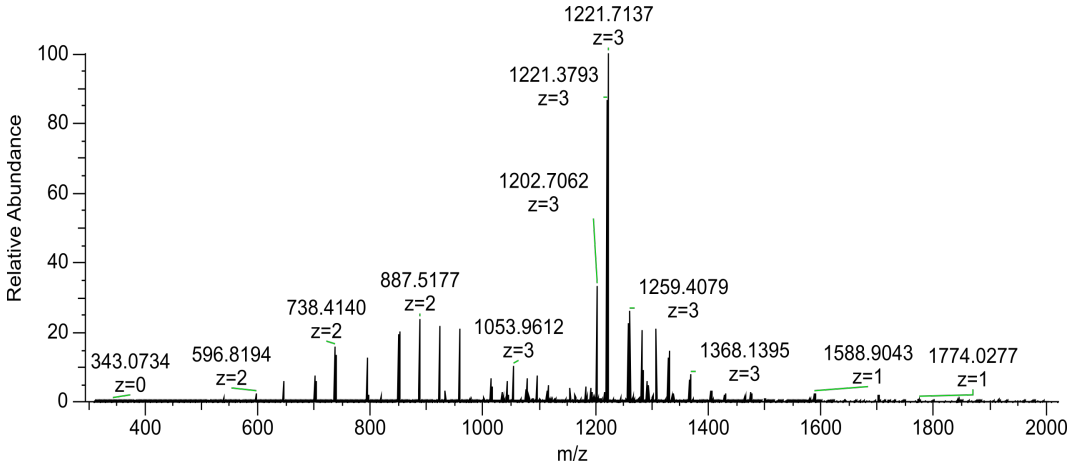


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
489.2115	489.2126	1	b4	-0.0011	-2.2485
505.7428	505.7438	2	y6	-0.0020	-1.9773
524.2604	524.2616	1	y3	-0.0012	-2.2889
618.2536	618.2552	1	b5	-0.0016	-2.5879
663.3033	663.3047	2	y9	-0.0028	-2.1106
720.8162	720.8182	2	y10	-0.0040	-2.7746
731.3375	731.3393	1	b6	-0.0018	-2.4612
753.3292	753.3315	1	b5	-0.0023	-3.0531
774.8642	774.8668	2	b13	-0.0052	-3.3554
788.3587	788.3607	1	b7	-0.0020	-2.5369
834.3975	834.3999	2	y12	-0.0048	-2.8763
897.1245	897.1271	3	b23	-0.0078	-2.8981
931.4707	931.4725	2	b16	-0.0036	-1.9324
934.8195	934.8218	3	b24	-0.0069	-2.4604
955.9406	955.9426	2	y14	-0.0040	-2.0922
1006.8435	1006.8467	3	b26	-0.0096	-3.1782
1020.4618	1020.4639	2	y15	-0.0042	-2.0579
1031.0361	1031.0386	2	b18	-0.0050	-2.4247
1050.8115	1050.8149	3	y24	-0.0102	-3.2356
1055.9795	1055.9825	2	y16	-0.0060	-2.8410
1083.5291	1083.5313	3	b28	-0.0066	-2.0304
1095.5593	1095.5599	2	b19	-0.0012	-0.5477
1120.0118	1120.0098	2	y17	0.0040	1.7857
1126.3001	1126.3023	4	y36	-0.0088	-1.9533
1154.8540	1154.8575	7	y67	-0.0245	-3.0307
1217.8101	1217.8125	5	b52	-0.0120	-1.9707
1236.8598	1236.8629	4	y40	-0.0124	-2.5063
1241.5690	1241.5727	2	y19	-0.0074	-2.9801
1252.6267	1252.6288	2	b21	-0.0042	-1.6765
1256.7851	1256.7876	6	y62	-0.0150	-1.9892
1266.2865	1266.2912	6	y63	-0.0282	-3.7116

1269.6217	1269.6231	4	y41	-0.0056	-1.1027
1273.5964	1273.5994	1	b11	-0.0030	-2.3555
1285.1352	1285.1385	6	y64	-0.0198	-2.5678
1300.8019	1300.8041	6	b66	-0.0132	-1.6913
1306.6447	1306.6456	6	y65	-0.0054	-0.6888
1325.8160	1325.8168	6	y66	-0.0048	-0.6034
1330.2500	1330.2515	5	y55	-0.0075	-1.1276
1347.1623	1347.1659	6	y67	-0.0216	-2.6723
1349.3193	1349.3191	3	y32	0.0006	0.1482
1409.4257	1409.4262	4	y46	-0.0020	-0.3548
1676.8185	1676.8213	1	b14	-0.0028	-1.6698

UPF0391 membrane protein SAMEA2054040_04753

Uniprot accession number: A0A094XPL9
Charge state: 5+
Observed monoisotopic mass: 5687.2876 Da
Sequence:
fMFRWGIIFLVIALIAAALGFGGLAGTAAWAAKIVFVVGIILFLVSLFTGRRRP



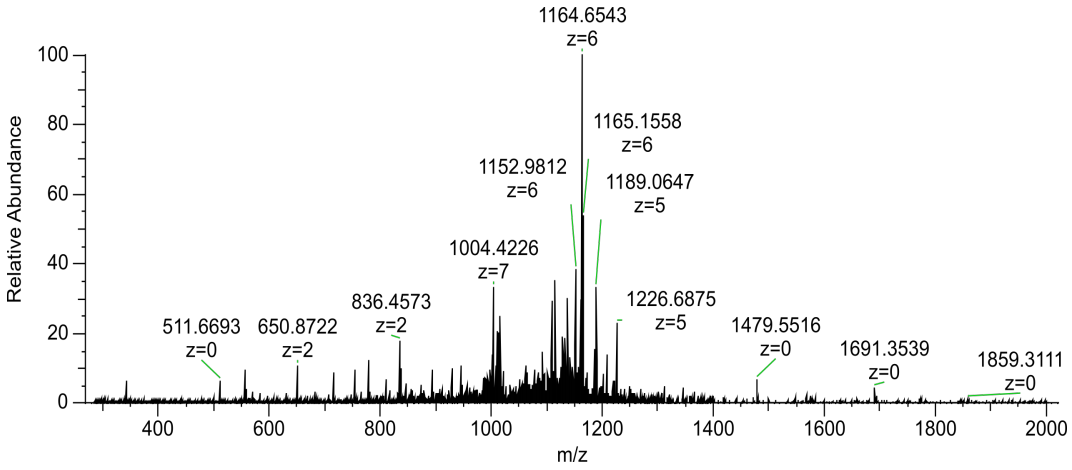
Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
818.3909	818.3898	B6	0.0011	1.3771
931.4730	931.4738	B7	-0.0008	-0.8578
1078.5418	1078.5422	B8	-0.0004	-0.4015
1191.6255	1191.6263	B9	-0.0008	-0.6571
1290.6937	1290.6947	B10	-0.0010	-0.7663
1290.6972	1290.6947	B10	0.0025	1.9571
1403.7746	1403.7788	B11	-0.0042	-2.9948
1403.7814	1403.7788	B11	0.0026	1.8628

1474.8163	1474.8159	B12	0.0004	0.2990
1474.8174	1474.8159	B12	0.0016	1.0659
1587.8932	1587.8999	B13	-0.0067	-4.2263
1587.9029	1587.8999	B13	0.0030	1.8628
1700.9814	1700.9840	B14	-0.0026	-1.5285
1772.0213	1772.0211	B15	0.0002	0.1185
1772.0239	1772.0211	B15	0.0028	1.6044
1843.0459	1843.0582	B16	-0.0124	-6.7024
1914.0915	1914.0953	B17	-0.0038	-1.9910
2027.1712	2027.1794	B18	-0.0082	-4.0366
2084.2011	2084.2008	B19	0.0003	0.1420
2231.2551	2231.2692	B20	-0.0142	-6.3439
2345.3062	2345.3122	B22	-0.0060	-2.5485
2458.3709	2458.3962	B23	-0.0253	-10.3092
2529.4142	2529.4333	B24	-0.0191	-7.5539
3645.0714	3645.0461	B35	0.0254	6.9631
1787.0897	1787.0934	Y15	-0.0037	-2.0693
1943.1847	1943.1832	Y17	0.0014	0.7395
2042.2440	2042.2516	Y18	-0.0076	-3.7189
2529.5684	2529.5675	Y22	0.0009	0.3744
2600.5821	2600.6046	Y23	-0.0225	-8.6345
2671.6311	2671.6417	Y24	-0.0106	-3.9710
2857.7243	2857.7210	Y25	0.0033	1.1457
2928.7468	2928.7581	Y26	-0.0113	-3.8692
2928.7570	2928.7581	Y26	-0.0011	-0.3869
2999.7880	2999.7952	Y27	-0.0073	-2.4232
2999.7917	2999.7952	Y27	-0.0035	-1.1697
3100.8153	3100.8429	Y28	-0.0276	-8.9031
3157.8527	3157.8644	Y29	-0.0117	-3.7031
3157.8599	3157.8644	Y29	-0.0044	-1.4032
3228.8829	3228.9015	Y30	-0.0186	-5.7589
3603.0661	3603.0969	Y34	-0.0308	-8.5352
3660.1084	3660.1183	Y35	-0.0099	-2.7125
3844.2189	3844.2395	Y37	-0.0206	-5.3701
3915.2697	3915.2766	Y38	-0.0069	-1.7562
3986.2815	3986.3137	Y39	-0.0323	-8.0932
4099.3685	4099.3978	Y40	-0.0293	-7.1430
4212.4563	4212.4818	Y41	-0.0256	-6.0720

50S ribosomal protein L29

Uniprot accession number: A0A2T4XV45
Charge state: 7+
Observed monoisotopic mass: 7238.0086 Da
Sequence:

MKAKELREKSVEELNAELLNLLREQFNLRMQAASGQLQQTHLL
KQVRRNVARVKTLLTQKAGA

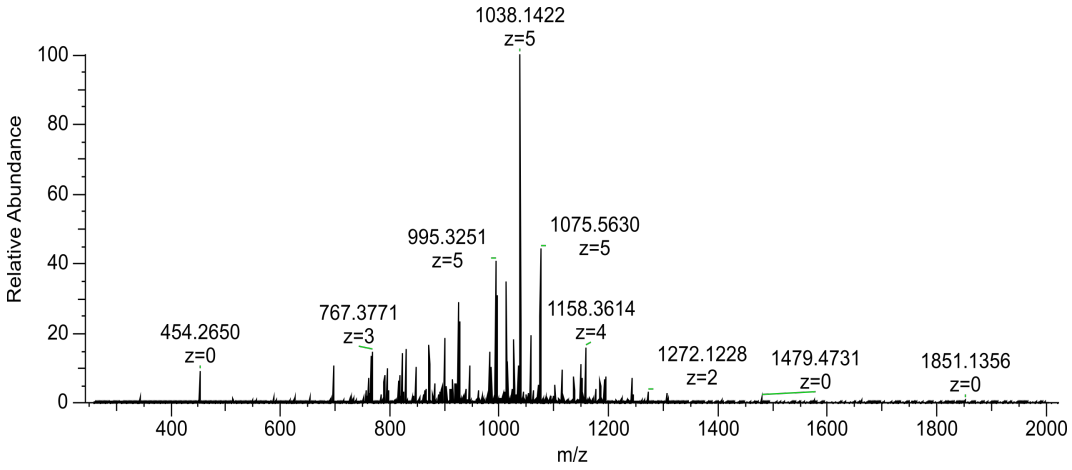


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
1113.6275	1113.6328	B9	-0.0054	-4.8113
1113.6299	1113.6328	B9	-0.0029	-2.5987
1428.7783	1428.7759	B12	0.0025	1.7162
1557.8142	1557.8185	B13	-0.0043	-2.7404
1670.9064	1670.9025	B14	0.0039	2.3209
1855.9661	1855.9826	B16	-0.0164	-8.8379
1985.0148	1985.0251	B17	-0.0104	-5.2211
7148.9156	7148.9771	B62	-0.0615	-8.6037
5382.0189	5382.0422	Y47	-0.0233	-4.3318
5453.0588	5453.0793	Y48	-0.0205	-3.7623
5567.1257	5567.1222	Y49	0.0035	0.6291
5809.2177	5809.2489	Y51	-0.0312	-5.3701
5938.2590	5938.2915	Y52	-0.0324	-5.4640
6037.3280	6037.3599	Y53	-0.0319	-5.2882
6124.3528	6124.3919	Y54	-0.0391	-6.3908
6650.6697	6650.7146	Y58	-0.0450	-6.7633
6779.7214	6779.7572	Y59	-0.0359	-5.2905
6907.8102	6907.8522	Y60	-0.0420	-6.0844
6978.8647	6978.8893	Y61	-0.0246	-3.5304

DUF1471 domain-containing protein

Uniprot accession number: A0A176X7A3
Charge state: 8+
Observed monoisotopic mass: 7667.8618 Da
Sequence:

ADMISKDEAHHFKLEYLGNVSVGASGGQISSPSDLHNKLSKLADEKGGKYY
VIIAAREHGPNFQAVAEVFK



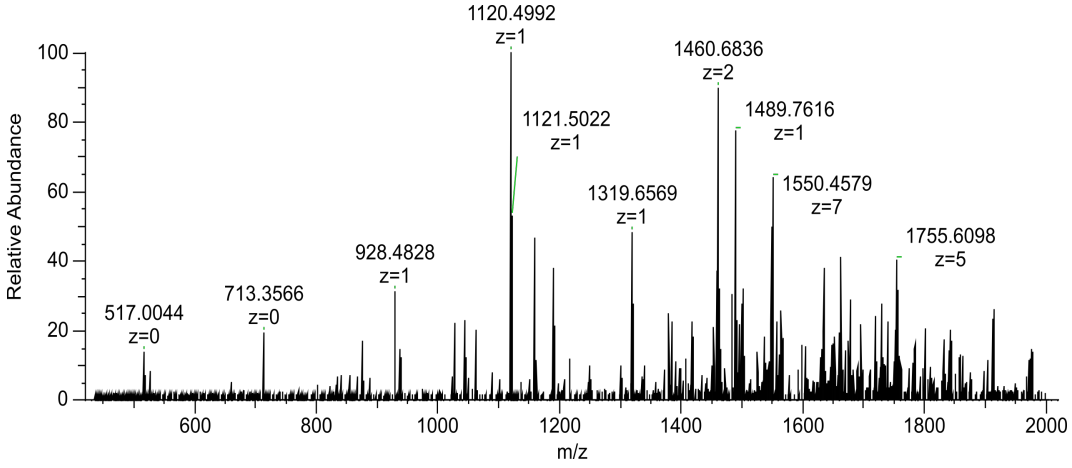
Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
1234.5386	1234.5401	B11	-0.0014	-1.1672
2027.9663	2027.9775	B17	-0.0112	-5.5104
2084.9866	2084.9989	B18	-0.0123	-5.8911
2199.0238	2199.0419	B19	-0.0181	-8.2259
2199.0402	2199.0419	B19	-0.0016	-0.7399
2298.1068	2298.1103	B20	-0.0035	-1.5139
2298.1121	2298.1103	B20	0.0019	0.8189
2385.1302	2385.1423	B21	-0.0121	-5.0852
2385.1381	2385.1423	B21	-0.0042	-1.7513
2484.1862	2484.2107	B22	-0.0245	-9.8442
2484.1983	2484.2107	B22	-0.0124	-4.9879
2541.2108	2541.2322	B23	-0.0214	-8.4069
2612.2625	2612.2693	B24	-0.0067	-2.5748
2612.2630	2612.2693	B24	-0.0063	-2.4128
2699.2873	2699.3013	B25	-0.0140	-5.1999
2941.3874	2941.4028	B28	-0.0154	-5.2220
762.4261	762.4276	Y7	-0.0014	-1.8953
1022.5212	1037.5546	Y9	-0.0057	-5.4499
1305.6717	1305.6717	Y12	0.0000	-0.0260
1798.9056	1798.9114	Y16	-0.0058	-3.2181
1869.9359	1869.9485	Y17	-0.0126	-6.7611
1983.0188	1983.0326	Y18	-0.0138	-6.9641
2096.1054	2096.1167	Y19	-0.0112	-5.3623
4726.4629	4741.4867	Y43	0.0039	0.8323
4726.4916	4741.4867	Y43	0.0326	6.8743
4911.5468	4926.5667	Y45	0.0077	1.5728
4968.5748	4983.5882	Y46	0.0143	2.8793
5055.5979	5070.6202	Y47	0.0054	1.0682

5055.5988	5070.6202	Y47	0.0063	1.2356
5126.6100	5141.6573	Y48	-0.0196	-3.8111
5126.6446	5141.6573	Y48	0.0150	2.9168
5183.6706	5198.6788	Y49	0.0195	3.7516
5183.6744	5198.6788	Y49	0.0233	4.4831
5282.7108	5297.7472	Y50	-0.0087	-1.6434
5282.7279	5297.7472	Y50	0.0084	1.5851
5369.7320	5384.7792	Y51	-0.0195	-3.6218
5369.7570	5384.7792	Y51	0.0055	1.0205
5468.8225	5483.8476	Y52	0.0026	0.4742

Biofilm model – *Candida glabrata*

Uncharacterised protein (gene HSP 12)

Uniprot accession number: Q6FPF6
Charge state, fragmentation method: 7+, CID
Observed monoisotopic mass: 11084.2050 Da
Sequence:
aSDAGRKNFSDKLNEGLTPDSQKSTWDKGKEFVTDETDKL
AGKFQGEENKGVAQGMHDSAQKGADEANAESYADTAREYMDAAKSKLN
DAVEYVSKSVHGGEK



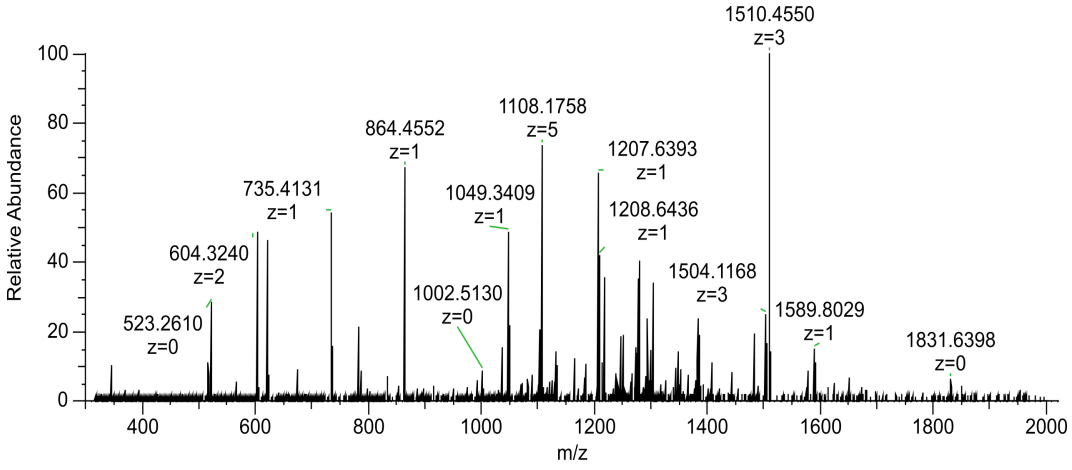
Observed mass (monoisotopic) *m/z*	Theoretical mass (monoisotopic) *m/z*	Fragment	Mass error [Da]	Mass error [ppm]
1119.4901	1119.4945	B10	-0.0044	-3.8937
1603.7575	1603.7590	B14	-0.0015	-0.9596
1874.9124	1874.9122	B17	0.0002	0.1115
2086.9821	2086.9919	B19	-0.0098	-4.7111
2430.1688	2430.1775	B22	-0.0087	-3.5742

3047.4488	3047.4584	B27	-0.0096	-3.1607
3823.8333	3823.8289	B34	0.0044	1.1562
4168.9372	4168.9461	B37	-0.0089	-2.1396
4481.1740	4481.1622	B40	0.0118	2.6243
6951.2740	6951.2917	B64	-0.0177	-2.5398
10156.7078	10156.7426	B93	-0.0348	-3.4274
712.3476	712.3504	Y7	-0.0028	-3.9756
927.4793	927.4774	Y9	0.0019	2.0809
1189.6101	1189.6091	Y11	0.0010	0.8288
1318.6487	1318.6517	Y12	-0.0030	-2.2569
1417.7152	1417.7201	Y13	-0.0049	-3.4689
1488.7599	1488.7572	Y14	0.0027	1.8163
1717.8236	1717.8271	Y16	-0.0035	-2.0444
1830.8975	1830.9112	Y17	-0.0137	-7.4597
2316.1875	2316.2073	Y22	-0.0198	-8.5463
3182.5743	3182.5666	Y29	0.0077	2.4248
4132.9096	4132.9283	Y38	-0.0187	-4.5319
4790.2167	4790.2365	Y45	-0.0198	-4.1340
6142.7984	6142.8144	Y58	-0.0160	-2.6024
6915.2591	6915.2739	Y65	-0.0149	-2.1513

Biofilm model – *Pseudomonas aeruginosa* PS1054

Uncharacterised protein (gene PA0039)

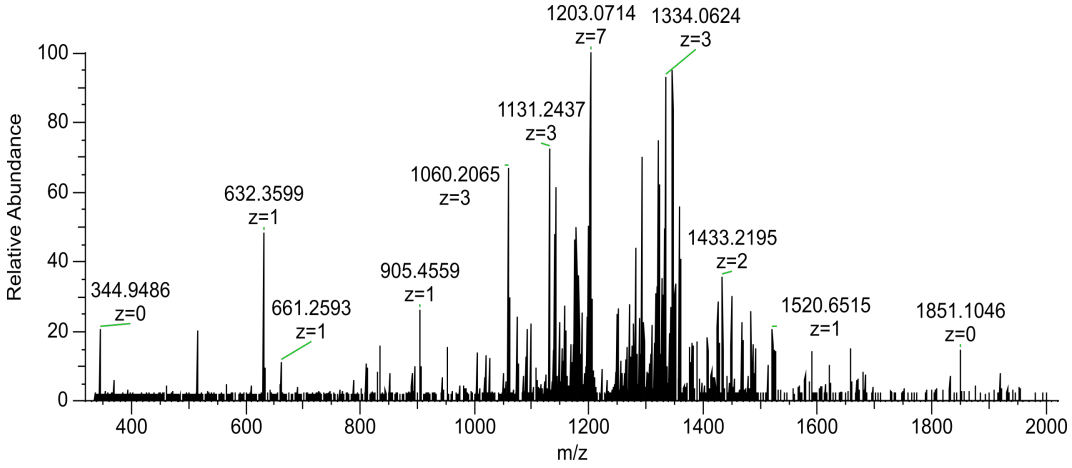
Uniprot accession number: Q9I793
Charge state, fragmentation method: 5+, CID
Observed monoisotopic mass: 5731.9696 Da
Sequence:
AKPCEELKAEIDAKIKANGVPAYTLEIVDKGSVTDKKVVGTCD
GGTKEIVYQRG



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
4525.3157	4525.3447	B43	-0.0290	-6.4079
4868.5087	4868.5303	B47	-0.0216	-4.4352
4997.5815	4997.5729	B48	0.0086	1.7258
5110.6215	5110.6569	B49	-0.0355	-6.9375
5209.6998	5209.7253	B50	-0.0255	-4.9022
621.3219	621.3235	Y5	-0.0016	-2.5108
734.4070	734.4075	Y6	-0.0005	-0.7258
863.4516	863.4501	Y7	0.0015	1.7187
1206.6340	1206.6357	Y11	-0.0016	-1.3650
1206.6343	1206.6357	Y11	-0.0014	-1.1511
5532.8103	5532.8483	Y52	-0.0380	-6.8666
5532.8259	5532.8483	Y52	-0.0224	-4.0534

BON domain-containing protein

Uniprot accession number: Q9HV60
Charge state, fragmentation method: 7+, CID
Observed monoisotopic mass: 8557.5023 Da
Sequence:
ANDTMQKTEEAVSDTWITSKVKSSLIANKNVSGVDIKVETNKGVVSLSG
NVKSDAERDLAIETAKGIKGVKAVSADGLKSVE

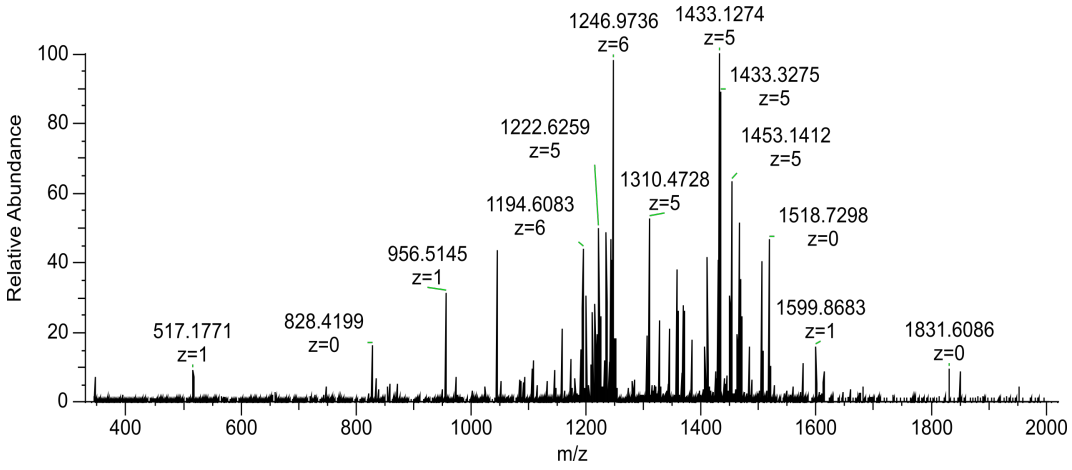


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
1317.5825	1317.5871	B12	-0.0046	-3.4669
1519.6385	1519.6460	B14	-0.0075	-4.9604
2335.0930	2335.1002	B21	-0.0072	-3.0714
2863.4207	2863.4273	B26	-0.0066	-2.3189

2934.4589	2934.4644	B27	-0.0056	-1.8917
3048.5050	3048.5074	B28	-0.0024	-0.7810
3176.5844	3176.6023	B29	-0.0179	-5.6362
3176.5919	3176.6023	B29	-0.0104	-3.2695
3290.6216	3290.6452	B30	-0.0236	-7.1764
3389.7058	3389.7137	B31	-0.0078	-2.3031
3476.7301	3476.7457	B32	-0.0156	-4.4916
3747.8554	3747.8625	B35	-0.0071	-1.8920
3860.9389	3860.9466	B36	-0.0077	-1.9894
4560.3050	4560.3381	B42	-0.0331	-7.2609
4716.4001	4716.4280	B44	-0.0278	-5.8994
4815.4870	4815.4964	B45	-0.0093	-1.9356
7653.0334	7653.0574	B73	-0.0240	-3.1368
7926.1627	7926.1534	B76	0.0093	1.1730
8096.2477	8096.2590	B78	-0.0113	-1.3950
8224.3531	8224.3539	B79	-0.0009	-0.1053
8410.4495	8410.4544	B81	-0.0049	-0.5825
631.3547	631.3541	Y6	0.0006	1.0279
904.4477	904.4502	Y9	-0.0025	-2.7221
1074.5570	1074.5557	Y11	0.0013	1.1921
1358.7421	1358.7405	Y14	0.0016	1.1761
2854.5621	2854.5763	Y28	-0.0142	-4.9857
3454.8514	3454.8630	Y34	-0.0116	-3.3640
3741.9857	3742.0112	Y37	-0.0254	-6.7985
3841.0842	3841.0796	Y38	0.0046	1.2030
3997.1751	3997.1694	Y40	0.0056	1.4075
4568.4704	4568.4660	Y45	0.0044	0.9631
4809.6029	4809.6450	Y47	-0.0421	-8.7541
5167.7861	5167.7939	Y51	-0.0077	-1.4948
5266.8371	5266.8623	Y52	-0.0251	-4.7729
5380.8771	5380.9052	Y53	-0.0281	-5.2257
6750.6809	6750.7345	Y66	-0.0536	-7.9449
6936.7998	6936.8138	Y67	-0.0140	-2.0112

Probable cold-shock protein

Uniprot accession number: Q9I4H8
Charge state, fragmentation method: 6+, CID
Observed monoisotopic mass: 7588.8487 Da
Sequence:
ADREVGTVKWFNDAKGYGFIQRDSGPDVFVHYRAIRGEGHRSLVEGQKVEFSVIQ
GQKGLQAEDVSKV



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
955.5095	955.5087	B9	0.0009	0.9168
1517.7194	1517.7263	B13	-0.0068	-4.4988
2653.2892	2653.3037	B23	-0.0144	-5.4442
3009.4468	3009.4369	B27	0.0100	3.3126
5990.0055	5990.0152	B53	-0.0097	-1.6147
6103.0821	6103.0993	B54	-0.0171	-2.8053
6231.1311	6231.1578	B55	-0.0267	-4.2851
6544.3049	6544.3328	B58	-0.0280	-4.2712
6913.4949	6913.5341	B62	-0.0391	-5.6585
7157.5943	7157.6036	B64	-0.0092	-1.2904
7256.6645	7256.6720	B65	-0.0075	-1.0363
7343.6761	7343.7040	B66	-0.0279	-3.7967
1044.5476	1044.5451	Y10	0.0024	2.3359
1357.7190	1357.7201	Y13	-0.0012	-0.8470
6071.1231	6071.1517	Y55	-0.0286	-4.7052
6071.1304	6071.1517	Y55	-0.0213	-3.5043

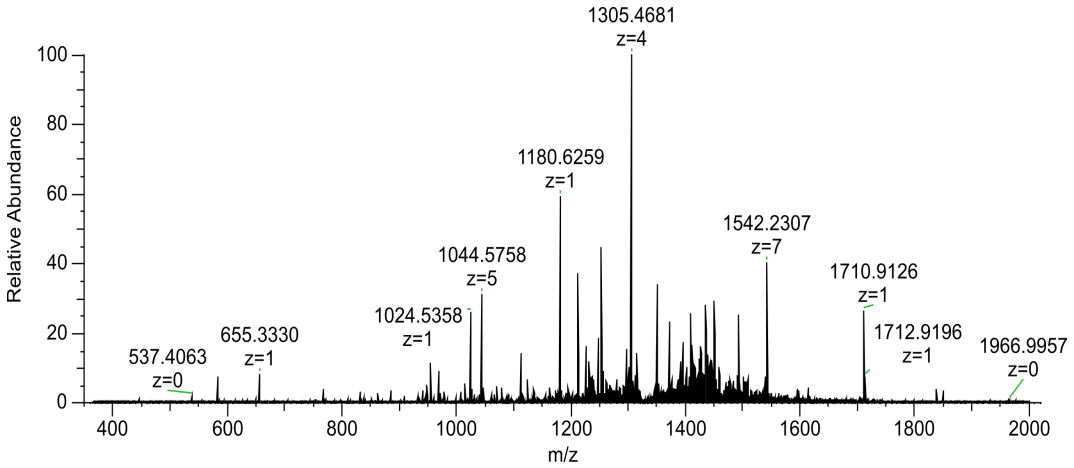
B.2 Chapter 4

Blood agar substrate

Horse haemoglobin subunit β

Uniprot accession number: P02062
Charge state, fragmentation method: 12+, CID
Observed monoisotopic mass: 15998.3319 Da
Sequence:
VQLSGEEKAAVLALWDKVNEEEVGGEALGRLLVVYPWTQRFFDSFGDLSNPGAV

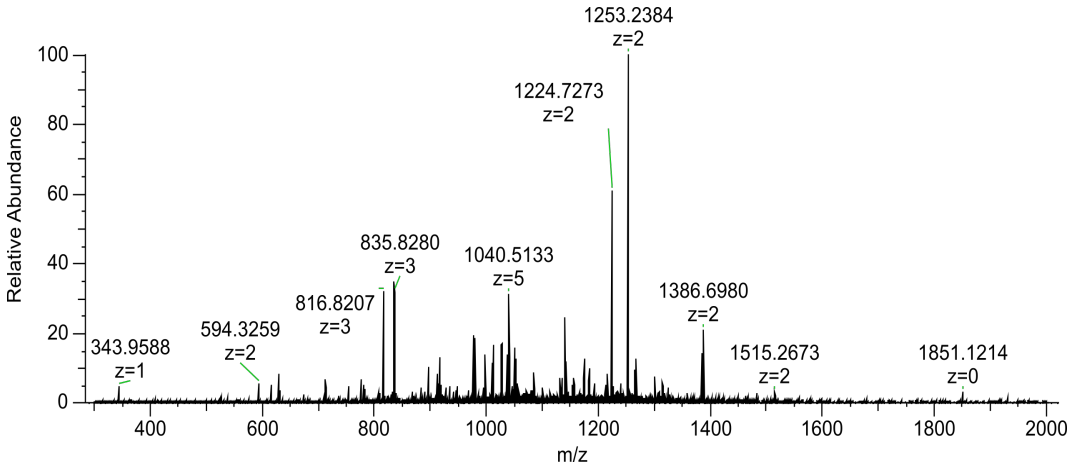
MGNPKVKAHGKKVLHSFGEGVHHLDNLKGTFAALSELHCDKLHVDPENFRLLG
NVLVVVLARHFGKDFTPELQASYQKVVAGVANALAHKYH



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
655.3330	655.3311	1	y5	0.0019	2.8993
768.4177	768.4151	1	y6	0.0026	3.3836
832.1138	832.1116	3	y23	0.0066	2.6439
942.4924	942.4891	1	b9	0.0033	3.5014
953.4958	953.4953	1	y8	0.0005	0.5244
969.5414	969.5384	2	b18	0.0060	3.0943
978.5288	978.5261	2	y18	0.0054	2.7593
1024.5358	1024.5323	1	y9	0.0035	3.4162
1044.1747	1044.1713	5	y47	0.0170	3.2562
1112.5986	1112.5946	1	b11	0.0040	3.5952
1123.6047	1123.6007	1	y10	0.0040	3.5600
1134.6204	1134.6160	2	y21	0.0088	3.8780
1162.6408	1162.6379	5	y52	0.0145	2.4943
1173.6158	1173.6200	6	y64-NH3	-0.0252	-3.5787
1180.6259	1180.6222	1	y11	0.0037	3.1339
1211.6381	1211.6328	4	b43	0.0212	4.3743
1251.6637	1251.6593	1	y12	0.0044	3.5153
1304.9664	1304.9623	4	y47	0.0164	3.1419
1313.1747	1313.1700	4	b47	0.0188	3.5791
1350.7325	1350.7277	1	y13	0.0048	3.5536
1371.7268	1371.7217	2	y25	0.0102	3.7180
1409.8052	1409.7999	1	b14	0.0053	3.7594
1449.8012	1449.7961	1	y14	0.0051	3.5177
1541.3730	1541.3662	7	b99	0.0476	4.4117
1710.9127	1710.9061	1	b16	0.0066	3.8576
1615.1798	1615.1746	3	b43	0.0156	3.2195

***Klebsiella pneumoniae* KP257**
DUF1471 domain-containing protein (gene KPN_00497)

Uniprot accession number: A6T5S6
Charge state, fragmentation method: 7+, CID
Observed monoisotopic mass: 7698.9838 Da
Sequence:
AQLITKEEVKHFKLTKVGPISVGPSGGEFSSPSDLHDQLSKLADEKGGKYYVITA
AREHGPNFEATAEVYK
R→K substitution at position 49

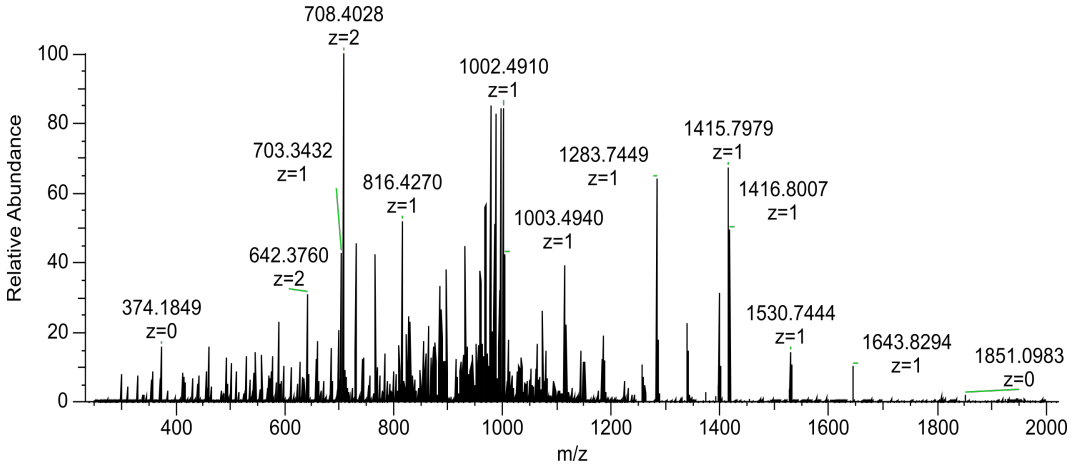


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
1011.5631	1011.5600	B9	0.0031	3.0448
1139.6528	1139.6550	B10	-0.0022	-1.9339
1423.7861	1423.7823	B12	0.0038	2.6345
1551.8872	1551.8773	B13	0.0099	6.3974
1664.9675	1664.9613	B14	0.0062	3.6944
1766.0142	1766.0090	B15	0.0051	2.9077
1894.1101	1894.1040	B16	0.0061	3.2348
1993.1723	1993.1724	B17	-0.0001	-0.0256
2050.1864	2050.1938	B18	-0.0074	-3.6216
2260.3277	2260.3307	B20	-0.0029	-1.3038
2347.3587	2347.3627	B21	-0.0040	-1.7151
2418.4404	2446.4311	B22	-0.0047	-1.9040
2446.4323	2446.4311	B22	0.0012	0.5024
2446.4408	2446.4311	B22	0.0097	3.9756
2503.4464	2503.4526	B23	-0.0062	-2.4814
2503.4606	2503.4526	B23	0.0080	3.2092
2687.5513	2687.5374	B25	0.0139	5.1731
2744.5501	2744.5588	B26	-0.0087	-3.1641
2801.5889	2801.5803	B27	0.0086	3.0797
2930.6325	2930.6229	B28	0.0096	3.2911

3077.6836	3077.6913	B29	-0.0076	-2.4840
3164.7350	3164.7233	B30	0.0116	3.6809
3550.8691	3550.8671	B34	0.0020	0.5725
1989.9525	1989.9544	Y18	-0.0019	-0.9392
2365.1660	2365.1702	Y21	-0.0042	-1.7627
2528.2334	2528.2335	Y22	-0.0002	-0.0767
2500.2801	2528.2335	Y22	0.0326	12.8991
2770.3718	2798.3776	Y25	-0.0197	-7.0252
2898.4592	2926.4725	Y26	-0.0272	-9.3058
4148.1174	4176.1028	Y37	0.0007	0.1777
4621.2976	4649.2786	Y42	0.0052	1.1085
4954.4190	4982.4110	Y45	-0.0060	-1.2004
5195.4868	5223.5173	Y48	-0.0444	-8.4926
5195.5277	5223.5173	Y48	-0.0035	-0.6664
5252.5165	5280.5387	Y49	-0.0361	-6.8431
5351.6098	5379.6071	Y50	-0.0113	-2.0992
7273.6794	7301.7060	Y67	-0.0406	-5.5604

***Pseudomonas aeruginosa* PS1054**
DNA-binding protein HU-α

Uniprot accession number: P05384
Charge state, fragmentation method: 10+, CID
Observed monoisotopic mass: 9081.0762 Da
Sequence:
MNKSELIDAIAASADIPKAVAGRALDAVIESVTGALKAGDSVVLVGFGTFAVKERAARTG
RNPQTGKPIKIAAAKIPGFKAGKALKDAVN



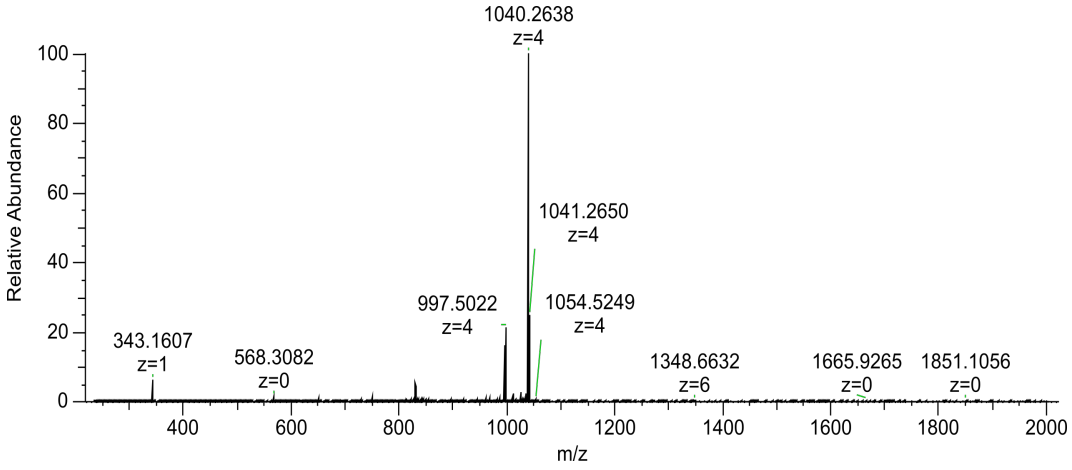
Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
460.2097	460.2104	B4	-0.0007	-1.5167
589.2516	589.2530	B5	-0.0014	-2.3810
702.3381	702.3371	B6	0.0010	1.4309
815.4222	815.4211	B7	0.0010	1.2779
930.4486	930.4481	B8	0.0005	0.5771
1001.4867	1001.4852	B9	0.0016	1.5737
1001.4884	1001.4852	B9	0.0032	3.1933
1114.5652	1114.5692	B10	-0.0040	-3.5834
1185.6052	1185.6063	B11	-0.0012	-1.0037
1256.6416	1256.6434	B12	-0.0019	-1.4913
1529.7410	1529.7395	B15	0.0015	0.9564
1529.7418	1529.7395	B15	0.0022	1.4617
1642.8179	1642.8236	B16	-0.0057	-3.4441
3120.6426	3120.6488	B31	-0.0063	-2.0076
3320.7409	3320.7649	B33	-0.0240	-7.2357
3448.8178	3448.8235	B35	-0.0056	-1.6348
3561.9060	3561.9075	B36	-0.0016	-0.4458
3933.0674	3933.0880	B40	-0.0206	-5.2419
4218.2572	4218.2569	B43	0.0003	0.0747
7553.1535	7553.1670	B75	-0.0135	-1.7936
7666.2348	7666.2511	B76	-0.0163	-2.1275
7763.3288	7763.3038	B77	0.0250	3.2164
8095.4440	8095.4887	B80	-0.0447	-5.5180
8351.5996	8351.6422	B83	-0.0426	-5.1041
8663.8412	8663.8583	B86	-0.0171	-1.9739
8948.9317	8948.9908	B89	-0.0591	-6.6036
658.3632	658.3650	Y6	-0.0018	-2.7553
729.4038	729.4021	Y7	0.0017	2.3362
985.5567	985.5556	Y10	0.0010	1.0583
1414.7879	1414.7932	Y14	-0.0053	-3.7702
1414.7888	1414.7932	Y14	-0.0044	-3.1135
1527.8758	1527.8773	Y15	-0.0015	-0.9660
1655.9727	1655.9722	Y16	0.0005	0.2736
1727.0124	1727.0093	Y17	0.0031	1.7695
1869.0743	1869.0836	Y19	-0.0093	-4.9687
1869.0847	1869.0836	Y19	0.0011	0.6030
1982.1569	1982.1676	Y20	-0.0107	-5.3941
1982.1652	1982.1676	Y20	-0.0024	-1.1962
3714.1455	3714.1645	Y36	-0.0189	-5.0978
4446.5340	4446.5451	Y43	-0.0111	-2.4943
4749.6828	4749.7034	Y46	-0.0205	-4.3177
4749.6963	4749.7034	Y46	-0.0070	-1.4757
4862.7753	4862.7874	Y47	-0.0122	-2.5010
4862.7833	4862.7874	Y47	-0.0041	-0.8507
5147.9251	5147.9563	Y50	-0.0312	-6.0591

5319.9825	5320.0047	Y52	-0.0222	-4.1726
5319.9863	5320.0047	Y52	-0.0184	-3.4613
5391.0097	5391.0418	Y53	-0.0321	-5.9469
5519.1244	5519.1367	Y54	-0.0124	-2.2407
5632.1859	5632.2208	Y55	-0.0349	-6.1979
5861.3204	5861.3270	Y58	-0.0067	-1.1380
5861.3246	5861.3270	Y58	-0.0025	-0.4231
6459.6200	6459.6597	Y64	-0.0397	-6.1409
7666.2348	7666.3317	Y76	-0.0969	-12.6429
7824.3704	7824.4008	Y78	-0.0304	-3.8853
7895.4040	7895.4379	Y79	-0.0340	-4.3008
7966.4481	7966.4750	Y80	-0.0269	-3.3818
7966.4544	7966.4750	Y80	-0.0207	-2.5947
8150.5490	8150.5962	Y82	-0.0472	-5.7932
8265.6180	8265.6232	Y83	-0.0051	-0.6221
8491.7304	8491.7913	Y85	-0.0609	-7.1722
8707.8376	8707.8659	Y87	-0.0283	-3.2491
8835.9442	8835.9609	Y88	-0.0166	-1.8821

Labskin
Human skin proteins

β-defensin 4A

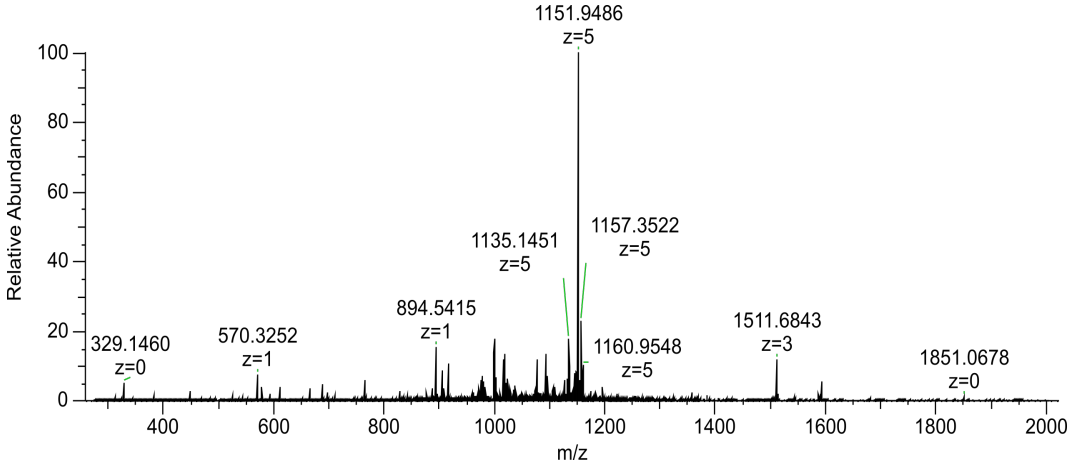
Uniprot accession number: O15263
Charge state, fragmentation method: 5+, CID
Observed monoisotopic mass: 4325.1296 Da
Sequence:
GIGDPVTCLKSGAICHPVFCPRRYKQIGTCGLPGTKCCKKP
This protein contains 3 disulfide bonds.



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
343.1607	343.1612	1	b4	-0.0005	-1.4570
996.7506	998.264	4	y37	-0.0068	-1.7030
1025.5071	1027.0207	4	y38	-0.0076	-1.8500
1039.7628	1041.2761	4	y39	-0.0064	-1.5366

Elafin

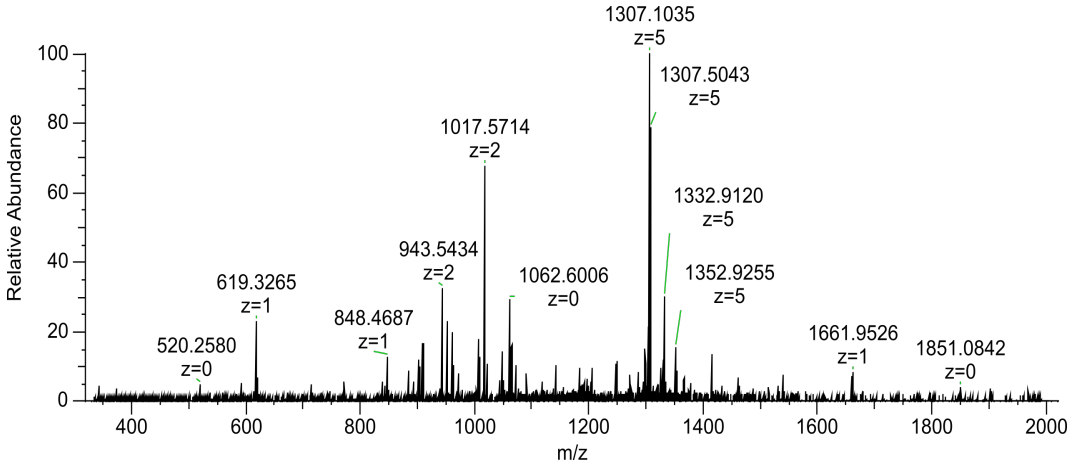
Uniprot accession number: P19957
Charge state, fragmentation method: 6+, CID
Observed monoisotopic mass: 5994.8239 Da
Sequence:
AQEPVKGPVSTKPGSCPIILIRCAMLNPPNRCLKDTDCPGIKKCCEGSCGMACFV
PQ
This protein contains 4 disulfide bonds.



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
329.1459	329.1456	1	b3	0.0003	3.0382
611.8436	611.8431	2	b12	0.0010	0.8172
906.5047	906.5043	1	b9	0.0004	1.1031
1151.3477	1151.3470	5	b55	0.0037	0.1737
1160.1514	1160.1522	5	y55	-0.0038	0.1724

Ubiquitin

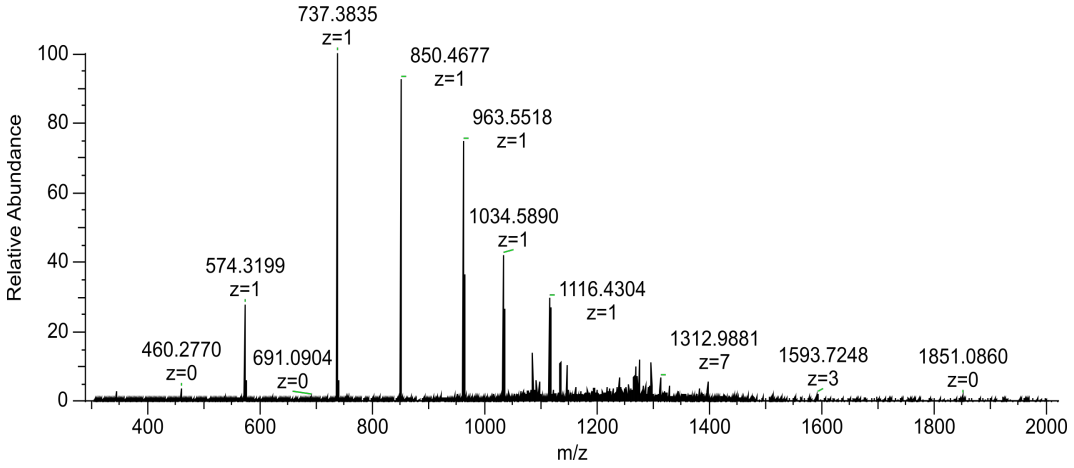
Uniprot accession number: P62987
Charge state, fragmentation method: 7+, CID
Observed monoisotopic mass: 8559.6029 Da
Sequence:
MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTL
SDYNIQKESTLHLVLRLRGG



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
619.3265	619.3272	1	b5	-0.0007	-1.1303
848.4687	848.4699	1	b7	-0.0012	-1.4143
894.0087	894.0104	2	b16-H2O	-0.0034	-1.9015
903.0151	903.0157	2	b16	-0.0012	-0.6644
909.5058	909.5069	3	y24	-0.0033	-1.2094
943.5434	943.5446	2	b17-H2O	-0.0024	-1.2718
952.5491	952.5499	2	b17	-0.0016	-0.8399
961.553	961.5539	1	b8	-0.0009	-0.9360
1008.0652	1008.0659	2	b18-H2O	-0.0014	-0.6944
1017.0699	1017.0712	2	b18	-0.0026	-1.2782
1049.0997	1049.1	2	y18	-0.0006	-0.2860
1141.3695	1141.3702	4	y40	-0.0028	-0.6133
1270.6899	1270.6881	5	b57	0.009	1.4166
1306.5037	1306.505	5	y58	-0.0065	-0.9950
1332.3128	1332.3136	5	y59	-0.004	-0.6005
1352.1224	1352.1273	5	y60	-0.0245	-3.6239
1436.7702	1436.7731	3	b39	-0.0087	-2.0184
1461.8475	1461.8498	1	b13	-0.0023	-1.5733

S100-A6

Uniprot accession number: P06703
Charge state, fragmentation method: 9+, CID
Observed monoisotopic mass: 10027.3088 Da
Sequence:
aACPLDQAIGLLVAIFHKYSGREGDKHTLSKKELKELIQKELTIGSKLQDAEIARLM
EDLDRNKDQEVNFQEYVTFLGALALIYNEALK

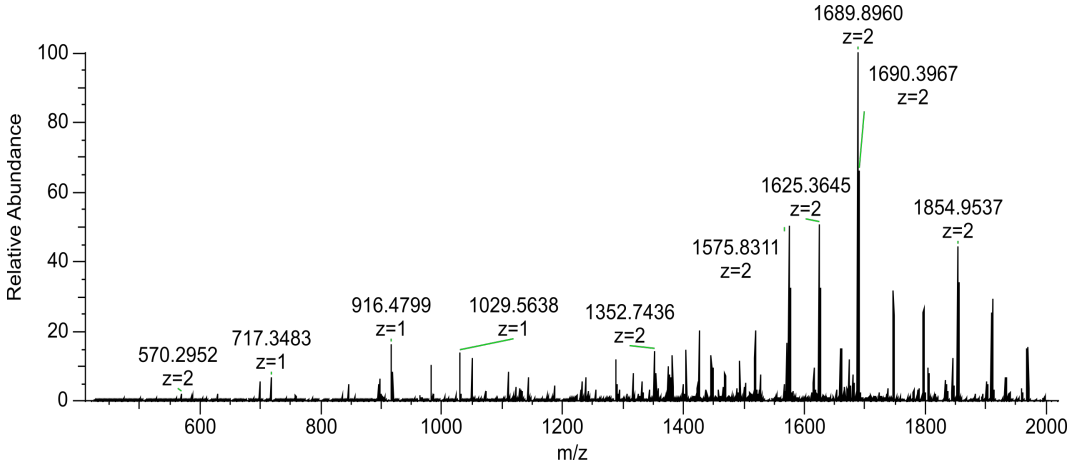


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
460.2768	460.2766	1	y4	0.0002	0.4345
574.3199	574.3195	1	y5	0.0004	0.6965
720.3569	720.3563	1	y6-NH3	0.0006	0.8329
737.3834	737.3828	1	y6	0.0006	0.8137
850.4676	850.4669	1	y7	0.0007	0.8231
963.5517	963.551	1	y8	0.0007	0.7265
1034.5889	1034.5881	1	y9	0.0008	0.7733
1147.6736	1147.6721	1	y10	0.0015	1.3070
1235.2237	1235.2192	7	b75	0.0315	3.6431
1275.7313	1275.7307	1	y12	0.0006	0.4703
1312.129	1312.1261	7	b81	0.0203	2.2102
1328.2859	1328.2809	7	b82	0.035	3.7643

S100-A8

Uniprot accession number: P05109
Charge state, fragmentation method: 7+, CID
Observed monoisotopic mass: 10827.6373 Da
Sequence:

MLTELEKALNSIIDVYHKYSLIKGNFHAVYRDDLKKLLETECPQYIRKKGADVWF
KELDINTDGAVNFQEFLILVIKMGVAAHKKSHEESHKE

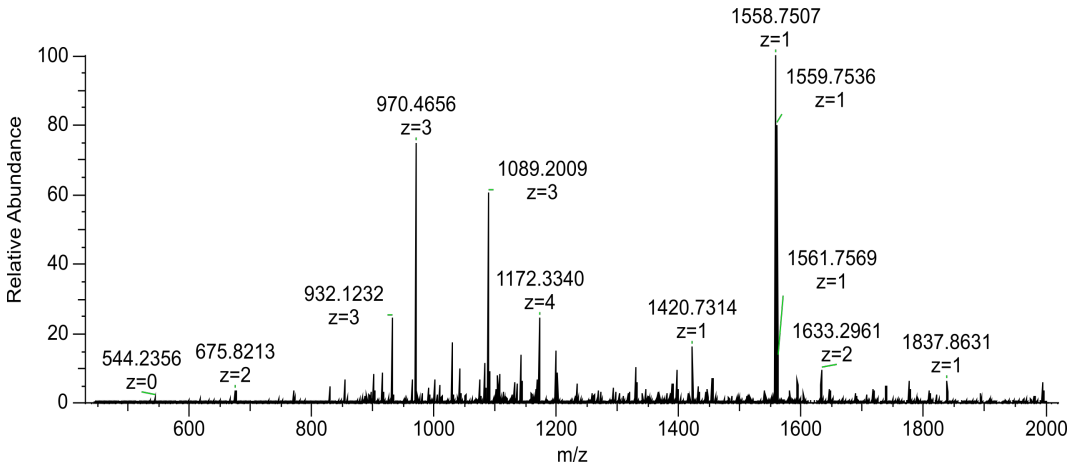


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
588.3055	588.3062	1	b5	-0.0007	-1.1899
629.2883	629.2889	1	y5	-0.0006	-0.9535
699.3375	699.3382	1	b6-H2O	-0.0007	-1.0009
717.3483	717.3488	1	b6	-0.0005	-0.6970
845.4428	845.4437	1	b7	-0.0009	-1.0645
895.3898	895.3904	1	y7	-0.0006	-0.6701
898.4693	898.4703	1	b8-H2O	-0.0010	-1.1130
902.9408	902.9416	2	y16	-0.0016	-0.8860
916.4799	916.4808	1	b8	-0.0009	-0.9820
966.9883	966.9891	2	y17	-0.0016	-0.8273
982.4214	982.4225	1	y8	-0.0011	-1.1197
986.0151	986.0164	2	b17	-0.0026	-1.3184
1023.5305	1023.5311	2	y18	-0.0012	-0.5862
1029.5638	1029.5649	1	b9	-0.0011	-1.0684
1050.0630	1050.0639	2	b18	-0.0018	-0.8571
1073.0645	1073.0653	2	y19	-0.0016	-0.7455
1110.5165	1110.5174	1	y9	-0.0009	-0.8104
1122.5902	1122.5903	2	b19-H2O	-0.0002	-0.0891
1126.5829	1126.5813	1	b10-NH3	0.0016	1.4202
1129.6064	1129.6073	2	y20	-0.0018	-0.7967
1143.6069	1143.6078	1	b10	-0.0009	-0.7870
1175.1095	1175.1116	2	b20	-0.0042	-1.7871
1186.1483	1186.1494	2	y21	-0.0022	-0.9274
1231.6506	1231.6536	2	b21	-0.0060	-2.4357
1238.6110	1238.6124	1	y10	-0.0014	-1.1303
1242.6909	1242.6914	2	y22	-0.0010	-0.4024
1279.6824	1279.6824	2	b22-NH3	0.0000	0.0000

1288.1940	1288.1956	2	b22	-0.0032	-1.2420
1293.3352	1293.3364	3	b33	-0.0036	-0.9278
1316.2244	1316.2256	2	y23	-0.0024	-0.9117
1331.3104	1331.3108	5	b56	-0.0020	-0.3005
1343.7224	1343.7239	1	b12	-0.0015	-1.1163
1352.2413	1352.2431	2	b23	-0.0036	-1.3311
1357.1204	1357.1194	5	b57	0.0050	0.7369
1375.6702	1375.6713	1	y11	-0.0011	-0.7996
1379.7358	1379.7362	5	b58	-0.0020	-0.2899
1393.7414	1393.7373	5	a59-NH3	0.0205	2.9417
1402.7396	1402.7416	5	b59	-0.0100	-1.4258
1425.3564	1425.3584	5	b60	-0.0100	-1.4032
1444.7749	1444.7762	2	y25	-0.0026	-0.8998
1448.1655	1448.1670	5	b61	-0.0075	-1.0358
1456.8064	1456.8080	1	b13	-0.0016	-1.0983
1464.9732	1464.9712	5	b62-NH3	0.0100	1.3652
1487.7855	1487.7798	5	b63-H2O	0.0285	3.8312
1491.3821	1491.3819	5	b63	0.0010	0.1341
1502.3055	1502.3043	2	b26-H2O	0.0024	0.7988
1518.3092	1518.3104	2	y26	-0.0024	-0.7904
1517.1941	1517.1949	5	y65-H2O	-0.0040	-0.5273
1566.3248	1566.3266	2	y27-H2O	-0.0036	-1.1492
1571.8330	1571.8349	1	b14	-0.0019	-1.2088
1575.3296	1575.3319	2	y27	-0.0046	-1.4600
1611.5238	1611.5280	3	b41	-0.0126	-2.6062
1624.8634	1624.8661	2	y28	-0.0054	-1.6617
1651.3737	1651.3793	2	y29-H2O	-0.0112	-3.3911
1660.3842	1660.3846	2	y29	-0.0008	-0.2409
1673.8339	1673.8354	1	y15	-0.0015	-0.8961
1679.8877	1679.8901	2	y30-H2O	-0.0048	-1.4287
1688.8931	1688.8954	2	y30	-0.0046	-1.3618
1737.4078	1737.4035	2	y31-H2O	0.0086	2.4750
1746.4062	1746.4088	2	y31	-0.0052	-1.4888
1787.9233	1787.9274	2	y32-H2O	-0.0082	-2.2932
1796.9300	1796.9327	2	y32	-0.0054	-1.5026
1804.8740	1804.8759	1	y16	-0.0019	-1.0527
1833.9653	1833.9667	1	b16	-0.0014	-0.7634
1844.9484	1844.9489	2	y33-H2O	-0.0010	-0.2710
1853.9505	1853.9541	2	y33	-0.0072	-1.9418
1881.9840	1881.9874	2	b32	-0.0068	-1.8066
1901.4879	1901.4909	2	y34-H2O	-0.0060	-1.5777
1910.4927	1910.4962	2	y34	-0.0070	-1.8320
1932.9689	1932.9709	1	y17	-0.0020	-1.0347
1939.4993	1939.5009	2	b33	-0.0032	-0.8250
1959.0029	1959.0044	2	y35-H2O	-0.0030	-0.7657
1968.0081	1968.0096	2	y35	-0.0030	-0.7622

S100-A7

Uniprot accession number: P31151
Charge state, fragmentation method: 7+, CID
Observed monoisotopic mass: 11360.4787 Da
Sequence:
aSNTQAERSIIGMIDMFHKYTRRDDKIDKPSLLTMMKENFPNFLSACDKKGTNYL
ADVFEKKDKNEDKKIDFSEFLSLLGDIATDYHKQSHGAAPCSGGSQ

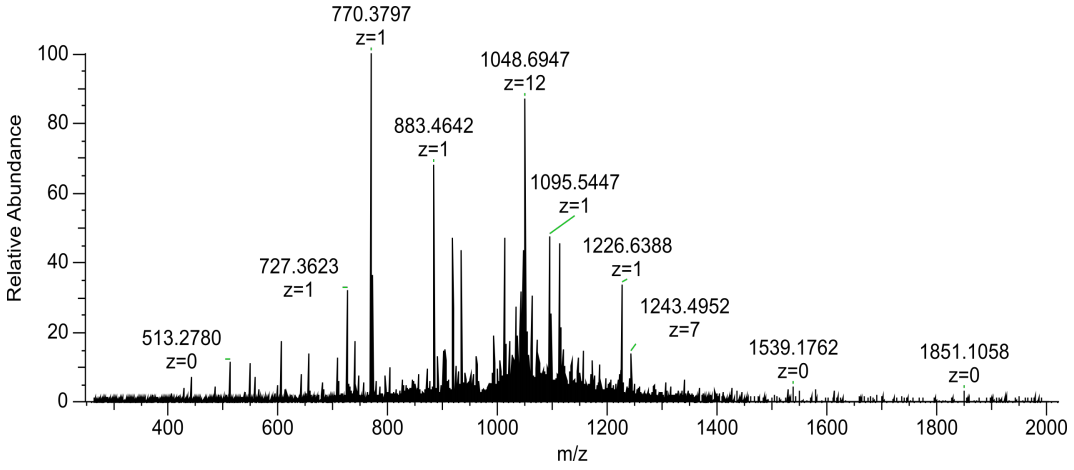


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
600.3040	600.3044	2	b11	-0.0008	-0.6663
829.3789	829.3799	1	b7	-0.0010	-1.2057
893.4472	893.4478	3	b22	-0.0018	-0.6716
898.4001	898.4013	1	b8-H2O	-0.0012	-1.3357
916.4107	916.4119	1	b8	-0.0012	-1.3095
925.7859	925.7866	3	b23-H2O	-0.0021	-0.7561
931.7890	931.7902	3	b23	-0.0036	-1.2878
951.4882	951.4896	4	b32	-0.0056	-1.4714
964.1285	964.1290	3	b24-H2O	-0.0015	-0.5186
970.1314	970.1325	3	b24	-0.0033	-1.1339
976.7503	976.7516	4	b33	-0.0052	-1.3309
1005.0085	1005.0090	4	b34-H2O	-0.0020	-0.4975
1009.5107	1009.5117	4	b34	-0.0040	-0.9906
1012.8297	1012.8308	3	b25	-0.0033	-1.0861
1029.4950	1029.4960	1	b9	-0.0010	-0.9713
1037.7661	1037.7692	4	b35-H2O	-0.0124	-2.9872
1042.2708	1042.2718	4	b35	-0.0040	-0.9594
1044.8549	1044.8500	3	b26-NH3	0.0147	4.6897
1051.5101	1051.5115	2	b18	-0.0028	-1.3314

1074.2942	1074.2955	4	b36	-0.0052	-1.2101
1082.8630	1082.8643	3	b27-H2O	-0.0039	-1.2005
1088.8666	1088.8678	3	b27	-0.0036	-1.1021
1102.0526	1102.0535	4	b37-H2O	-0.0036	-0.8167
1106.5549	1106.5562	4	b37	-0.0052	-1.1748
1114.5845	1114.5851	1	a10	-0.0006	-0.5383
1125.8945	1125.8906	3	b28-NH3	0.0117	3.4639
1131.5640	1131.5661	3	b28	-0.0063	-1.8558
1135.0643	1135.0669	4	b38	-0.0104	-2.2906
1142.5790	1142.5800	1	b10	-0.0010	-0.8752
1167.3289	1167.3314	4	b39-H2O	-0.0100	-2.1416
1171.8326	1171.8340	4	b39	-0.0056	-1.1947
1199.6003	1199.6015	1	b11	-0.0012	-1.0003
1268.3156	1268.3171	3	b32	-0.0045	-1.1827
1330.6420	1330.6403	1	b12	0.0017	1.2776
1345.6789	1345.6798	3	b34	-0.0027	-0.6688
1383.3561	1383.3565	3	b35-H2O	-0.0012	-0.2892
1389.3584	1389.3600	3	b35	-0.0048	-1.1516
1397.1814	1397.1816	2	b23	-0.0004	-0.1431
1426.3889	1426.3828	3	b36-NH3	0.0183	4.2766
1445.6883	1445.6898	2	b24-H2O	-0.0030	-1.0376
1454.6930	1454.6951	2	b24	-0.0042	-1.4436
1558.7507	1558.7530	1	b14	-0.0023	-1.4755
1632.7957	1632.7981	2	b27	-0.0048	-1.4699
1689.7908	1689.7935	1	b15	-0.0027	-1.5978

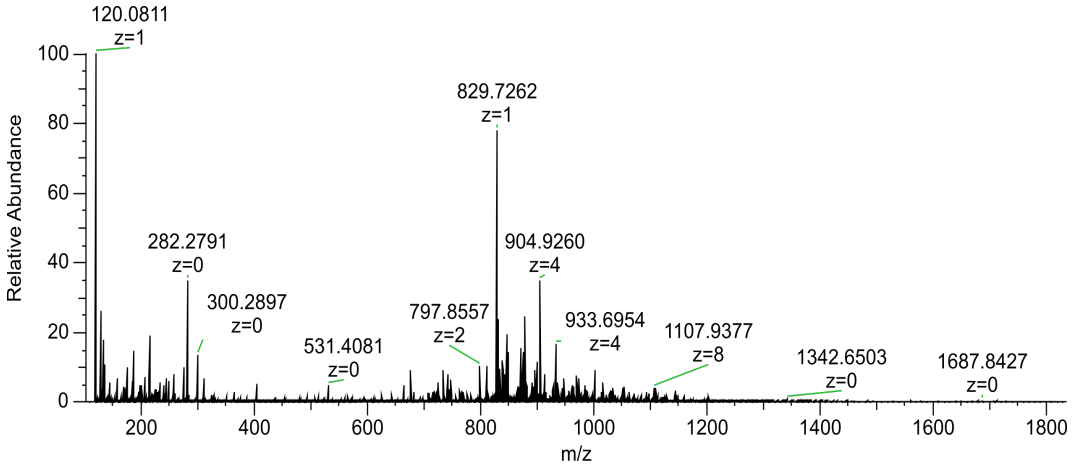
S100-A9

Uniprot accession number: P06702
Charge state, fragmentation method: 13+, CID
Observed monoisotopic mass: 12682.2879 Da
Sequence:
aSQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIE
HIMEDLDTNADKQLSFEEFIMLMARLTWASHEKMHEGDEGPGHHHKPGLGEGTP



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
442.2354	442.2352	2	b7	0.0004	0.4522
557.2805	557.2804	2	b9	0.0002	0.1794
656.3365	656.3362	1	b5	0.0003	0.4571
739.3689	739.3684	2	y15	0.0010	0.6763
770.3797	770.3791	1	b6	0.0006	0.7788
872.4136	872.4099	4	y32	0.0148	4.2411
883.4642	883.4632	1	b7	0.0010	1.1319
900.6791	900.6809	4	y33	-0.0072	-1.9985
917.8127	917.8119	3	b23	0.0024	0.8716
925.0856	925.0818	5	b39	0.0190	4.1077
933.4422	933.4411	4	y34	0.0044	1.1784
1012.5065	1012.5058	1	b8	0.0007	0.6914
1032.1521	1032.1502	11	y98	0.0209	1.8408
1047.0259	1047.0276	12	y108	-0.0204	-1.6236
1051.2571	1051.2544	11	b98	0.0297	2.5684
1113.5549	1113.5535	1	b9	0.0014	1.2572
1172.3187	1172.3234	7	b69	-0.0329	-4.0091
1226.6388	1226.6375	1	b10	0.0013	1.0598
1242.6362	1242.6356	7	b73	0.0042	0.4828
1306.006	1306.0031	6	b66	0.0174	2.2205

Charge state, fragmentation method: 16+, HCD
Sequence:
MTC(S-nitroso)KMSQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNF
LKKENKNEKVIEHIMEDLDTNADKQLSFEEFIMLMARLTWASHEKMHEGDEGPG
HHHKPGLGEGTP

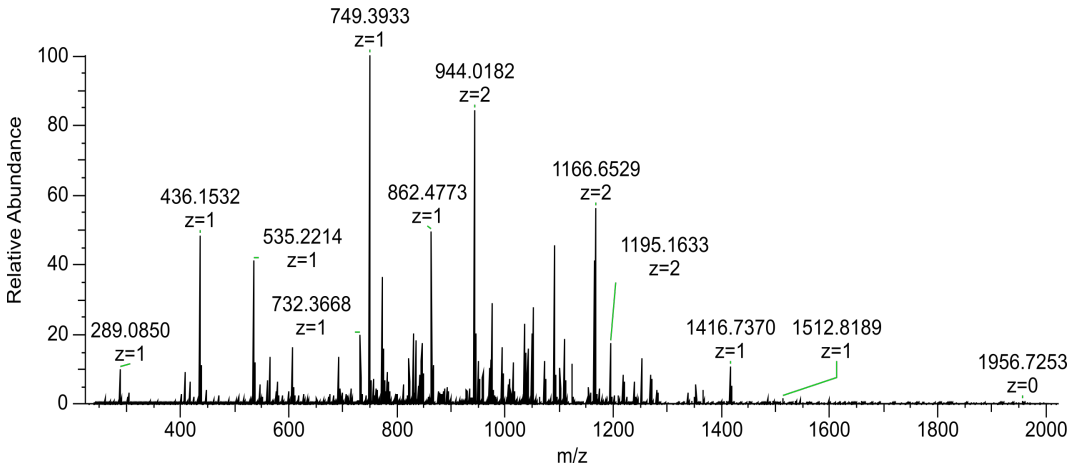


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
811.4214	811.4215	9	b61	-0.0009	-0.1232
833.9141	833.9139	4	b28	0.0008	0.2398
869.6997	869.7013	15	y112	-0.0240	-1.8397
882.2125	882.2116	13	y99	0.0117	1.0202
891.0102	891.0109	9	b67	-0.0063	-0.7856
898.6829	898.6832	13	y101	-0.0039	-0.3338
912.7241	912.7233	8	b61	0.0064	0.8765
936.7151	936.7180	12	y97	-0.0348	-3.0959
946.2240	946.2216	12	y98	0.0288	2.5364
955.6444	955.6453	12	y99	-0.0108	-0.9418
961.7126	961.7121	4	y35	0.0020	0.5199
973.7455	973.7475	8	b65	-0.0160	-2.0539
986.2777	986.2803	9	b75	-0.0234	-2.6362
1002.2629	1002.2614	8	b67	0.0120	1.4966
1014.8962	1014.8923	8	b68	0.0312	3.8428
1029.1497	1029.1477	8	b69	0.0160	1.9434
1032.1533	1032.1502	11	y98	0.0341	3.0034
1038.0276	1038.0274	8	b70	0.0016	0.1927
1042.9705	1042.9684	7	b61	0.0147	2.0135
1052.4039	1052.4057	8	b71	-0.0144	-1.7104
1059.1252	1059.1233	7	b62	0.0133	1.7939
1084.4221	1084.4249	8	b73	-0.0224	-2.5820
1098.5570	1098.5604	8	b74	-0.0272	-3.0950
1109.4412	1109.4394	8	b75	0.0144	1.6224
1121.5314	1121.5256	4	y40	0.0232	5.1715
1128.8688	1128.8653	7	b66	0.0245	3.1005
1145.2948	1145.2977	7	b67	-0.0203	-2.5321
1159.7321	1159.7331	7	b68	-0.0070	-0.8623
1202.6068	1202.6055	7	b71	0.0091	1.0810

S. aureus NCTC13435

Phenol-soluble modulin α3 peptide

Uniprot accession number: P0C805
Charge state, fragmentation method: 3+, CID
Observed monoisotopic mass: 2633.4082 Da
Sequence:
fMEFVAKLFKFFKDLLGKFLGNN



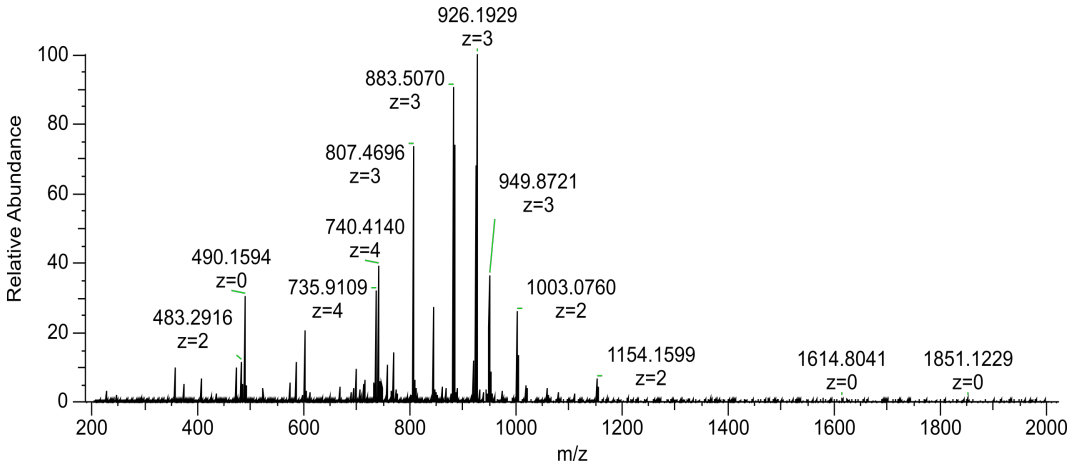
Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
289.0850	289.0853	1	b2	-0.0003	-1.0378
408.1582	408.1588	1	a3	-0.0006	-1.4700
417.2087	417.2092	1	y4	-0.0005	-1.1984
535.2214	535.2221	1	b4	-0.0007	-1.3079
547.2503	547.2511	1	y5-NH3	-0.0008	-1.4619
564.2770	564.2776	1	y5	-0.0006	-1.0633
606.2585	606.2592	1	b5	-0.0007	-1.1546
609.8450	609.8457	2	y11	-0.0014	-1.1478
626.3327	626.3334	2	b10-H2O	-0.0014	-1.1176
635.3378	635.3387	2	b10	-0.0018	-1.4166
683.3788	683.3799	2	y12	-0.0022	-1.6096
692.3720	692.3726	1	y6	-0.0006	-0.8666
708.8722	708.8729	2	b11	-0.0014	-0.9875
716.3439	716.3436	1	b6-H2O	0.0003	0.4188
732.3668	732.3675	1	y7-NH3	-0.0007	-0.9558
740.0755	740.076	3	b18	-0.0015	-0.6756
748.3992	748.4008	2	y13-NH3	-0.0032	-2.1379
749.3933	749.3941	1	y7	-0.0008	-1.0675
756.9132	756.9141	2	y13	-0.0018	-1.1890
772.9193	772.9203	2	b12	-0.0020	-1.2938
777.7695	777.7706	3	b19	-0.0033	-1.4143

782.7834	782.7839	3	y20	-0.0015	-0.6387
796.7766	796.7778	3	b20	-0.0036	-1.5061
821.4275	821.4285	2	b13-H2O	-0.0020	-1.2174
830.4324	830.4338	2	b13	-0.0028	-1.6859
834.7906	834.7921	3	b21	-0.0045	-1.7969
840.7944	840.7956	3	b21-H2O	-0.0036	-1.4272
847.4372	847.4382	1	b7	-0.0010	-1.1800
862.4770	862.4781	1	y8	-0.0011	-1.2754
886.9731	886.9758	2	b14	-0.0054	-3.0441
894.4955	894.4958	2	y15	-0.0006	-0.3354
934.5107	934.5126	2	b15-H2O	-0.0038	-2.0331
943.5165	943.5179	2	b15	-0.0028	-1.4838
951.0367	951.0378	2	y16	-0.0022	-1.1566
958.5339	958.5356	1	y9-NH3	-0.0017	-1.7735
972.0273	972.0286	2	b16	-0.0026	-1.3374
975.5609	975.5622	1	y9	-0.0013	-1.3326
994.5053	994.5067	1	b8	-0.0014	-1.4077
1006.0787	1006.08	2	y17-H2O	-0.0026	-1.2921
1015.0839	1015.0853	2	y17	-0.0028	-1.3792
1036.0747	1036.0761	2	b17	-0.0028	-1.3513
1041.6006	1041.5986	2	y18-H2O	0.0040	1.9201
1050.6027	1050.6039	2	y18	-0.0024	-1.1422
1073.5610	1073.5626	1	y10-NH3	-0.0016	-1.4904
1090.5876	1090.5891	1	y10	-0.0015	-1.3754
1100.1361	1100.1381	2	y19	-0.0040	-1.8180
1109.6088	1109.6103	2	b18	-0.0030	-1.3518
1122.6003	1122.6016	1	b9	-0.0013	-1.1580
1157.1456	1157.147	2	b19-H2O	-0.0028	-1.2099
1166.1508	1166.1523	2	b19	-0.0030	-1.2863
1173.6710	1173.6723	2	y20	-0.0026	-1.1076
1194.6613	1194.6631	2	b20	-0.0036	-1.5067
1201.6556	1201.6575	1	y11-NH3	-0.0019	-1.5811
1218.6828	1218.6841	1	y11	-0.0013	-1.0667
1251.6852	1251.6845	2	b21	0.0014	0.5592
1269.6680	1269.67	1	b10	-0.0020	-1.5752
1365.7502	1365.7525	1	y12	-0.0023	-1.6841
1416.7360	1416.7384	1	b11	-0.0024	-1.6940
1659.8578	1659.8603	1	b13	-0.0025	-1.5062

δ-hemolysin

Uniprot accession number: Q2FWM8
Charge state, fragmentation method: 4+, CID
Observed monoisotopic mass: 3004.6109 Da
Sequence:

fMAQDIISTIGDLVKWIIDTVNKFTKK

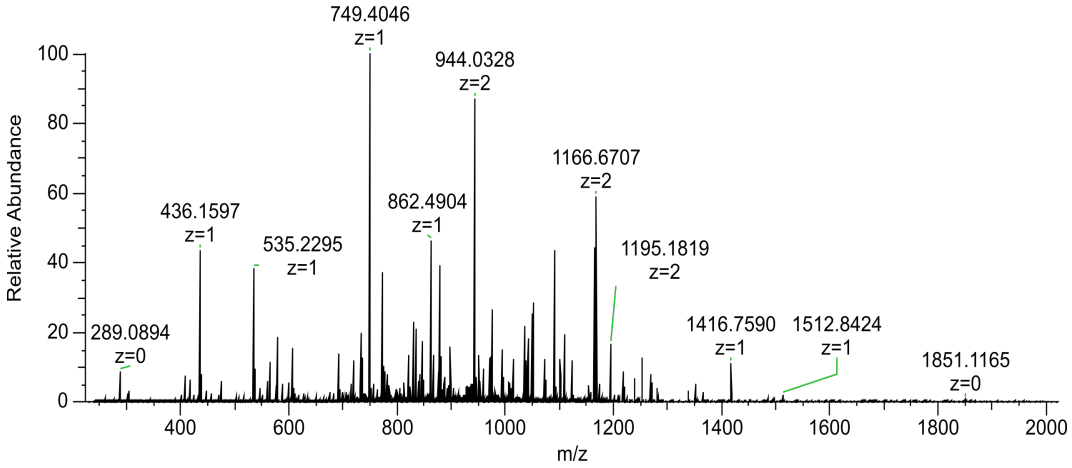


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
483.2916	483.2926	2	y8	-0.0020	-2.0691
694.6458	694.6468	4	y24	-0.0040	-1.4396
706.7484	706.7489	3	y18	-0.0015	-0.7075
712.4053	712.4061	4	y25	-0.0032	-1.1230
763.438	763.4386	3	y20-H2O	-0.0018	-0.7859
769.4409	769.4421	3	y20	-0.0036	-1.5596
807.1355	807.1368	3	y21	-0.0039	-1.6106
844.8302	844.8315	3	y22	-0.0039	-1.5388
860.5101	860.5114	2	y14	-0.0026	-1.5107
883.1725	883.1738	3	y23	-0.0039	-1.4720
919.8554	919.8565	3	y24-H2O	-0.0033	-1.1958
925.8585	925.86	3	y24	-0.0045	-1.6201
943.5343	943.5355	3	y25-H2O	-0.0036	-1.2718
949.5374	949.5391	3	y25	-0.0051	-1.7903
953.8447	953.849	3	b25	-0.0129	-4.5081
974.5653	974.5669	2	y16	-0.0032	-1.6418
1003.076	1003.0777	2	y17	-0.0034	-1.6948
1059.6185	1059.6197	2	y18	-0.0024	-1.1325
1110.142	1110.1435	2	y19	-0.0030	-1.3512
1153.6583	1153.6596	2	y20	-0.0026	-1.1268
1210.1981	1210.2016	2	y21	-0.0070	-2.8921
1366.2148	1366.2224	2	b24	-0.0152	-5.5628

S. aureus MSSA476

Phenol-soluble modulin α3 peptide

Uniprot accession number: P0C805
Charge state, fragmentation method: 3+, CID
Observed monoisotopic mass: 2633.4094 Da
Sequence:
fMEFVAKLFKFFKDLLGKFLGNN

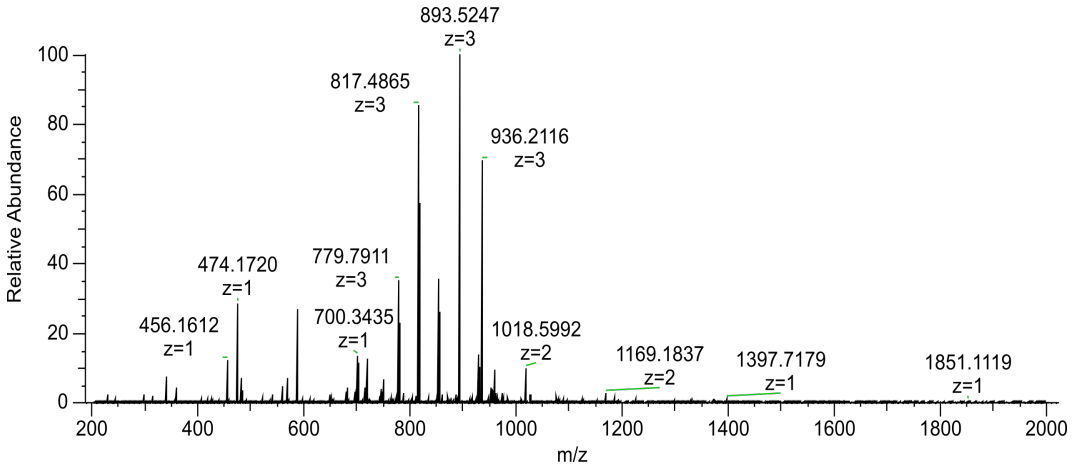


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
289.0854	289.0853	1	b2	0.0001	0.3459
304.1255	304.1252	1	y3	0.0003	0.3288
406.1537	406.1537	1	b3	0.0000	0.2462
535.2221	535.2221	1	b4	0.0000	0.1868
564.2777	564.2776	1	y5	0.0001	0.1772
606.2594	606.2592	1	b5	0.0002	0.1649
692.3727	692.3726	1	y6	0.0001	0.1444
749.3943	749.3941	1	y7	0.0002	0.1334
756.9142	756.9141	2	y13	0.0002	0.0661
772.9206	772.9203	2	b12	0.0006	0.0647
777.7716	777.7706	3	b19	0.0030	0.0429
782.7842	782.7839	3	y20	0.0009	0.0426
834.7925	834.7921	3	b21	0.0012	0.0399
847.4384	847.4382	1	b7	0.0002	0.1180
862.4785	862.4781	1	y8	0.0004	0.1159
943.5181	943.5179	2	b15	0.0004	0.0530
951.0382	951.0378	2	y16	0.0008	0.0526
972.0294	972.0286	2	b16	0.0016	0.0514
975.5626	975.5622	1	y9	0.0004	0.1025
994.5068	994.5067	1	b8	0.0001	0.1006
1015.0861	1015.0853	2	y17	0.0016	0.0493

1036.0768	1036.0761	2	b17	0.0014	0.0483
1050.6043	1050.6039	2	y18	0.0008	0.0476
1090.5897	1090.5891	1	y10	0.0006	0.0917
1100.1385	1100.1381	2	y19	0.0008	0.0454
1109.6108	1109.6103	2	b18	0.0010	0.0451
1122.6024	1122.6016	1	b9	0.0008	0.0891
1166.1528	1166.1523	2	b19	0.0010	0.0429
1194.664	1194.6631	2	b20	0.0018	0.0419
1218.6842	1218.6841	1	y11	0.0001	0.0821
1251.6855	1251.6845	2	b21	0.0020	0.0399
1269.6707	1269.6700	1	b10	0.0007	0.0788
1365.7532	1365.7525	1	y12	0.0007	0.0732
1416.7394	1416.7384	1	b11	0.0010	0.0706
1512.8216	1512.8209	1	y13	0.0007	0.0661

δ-hemolysin

Uniprot accession number: Q6G7S2
Charge state, fragmentation method: 4+, CID
Observed monoisotopic mass: 3034.6413 Da
Sequence:
fMAQDIISTISDLVKWIIDTVNKFTKK



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
231.0800	231.0798	1	b2	0.0002	0.8655
359.1358	359.1384	1	b3	-0.0026	0.5569
383.2347	383.2345	2	y6	0.0004	0.2609
432.7686	432.7687	2	y7	-0.0002	0.2311

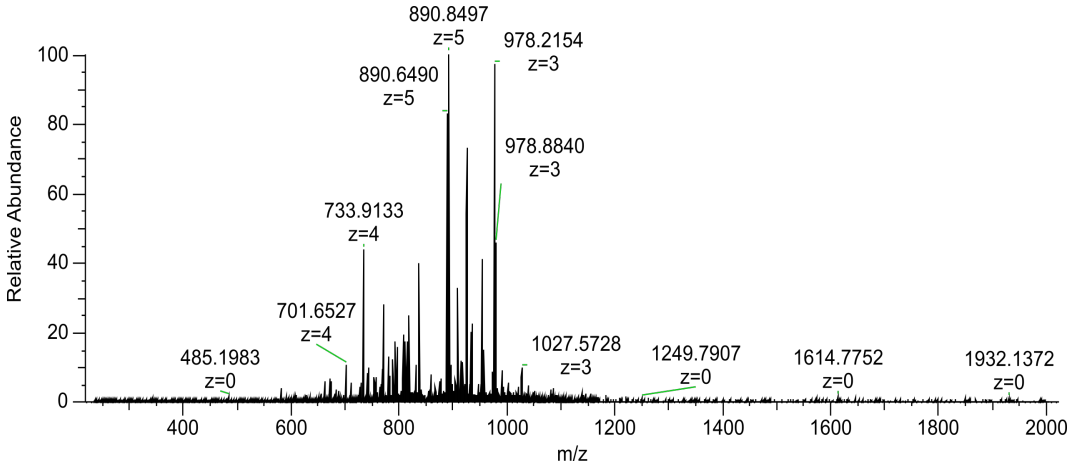
474.1654	474.1653	1	b4	0.0001	0.4218
483.2927	483.2926	2	y8	0.0002	0.2069
523.3239	523.3239	1	y4	0.0000	0.3822
574.0101	574.0101	3	y14	0.0000	0.1161
587.2496	587.2494	1	b5	0.0002	0.3406
597.3482	597.3481	2	y10	0.0002	0.1674
611.7049	611.7047	3	y15	0.0006	0.1090
650.0472	650.0471	3	y16	0.0003	0.1026
653.8905	653.8901	2	y11	0.0008	0.1529
679.0580	679.0577	3	y17	0.0009	0.0982
700.3339	700.3334	1	b6	0.0005	0.2856
702.1498	702.1495	4	y24	0.0012	0.0712
719.9091	719.9088	4	y25	0.0012	0.0695
723.1416	723.1412	4	b25	0.0016	0.0691
746.9301	746.9298	2	y12	0.0006	0.1339
779.4462	779.4457	3	y20	0.0015	0.0855
787.3658	787.3655	1	b7	0.0003	0.2540
817.1409	817.1403	3	y21	0.0018	0.0816
854.8354	854.8350	3	y22	0.0012	0.0780
860.5119	860.5114	2	y14	0.0010	0.1162
893.1780	893.1773	3	y23	0.0021	0.0746
935.8642	935.8635	3	y24	0.0021	0.0712
959.5436	959.5426	3	y25	0.0030	0.0695
974.5679	974.5669	2	y16	0.0020	0.1026
1018.0839	1018.0830	2	y17	0.0018	0.0982
1074.6256	1074.6250	2	y18	0.0012	0.0931
1080.6055	1080.6048	1	y9	0.0007	0.1851
1088.5320	1088.5292	1	b10	0.0028	0.1837
1125.1496	1125.1488	2	y19	0.0016	0.0889
1168.6655	1168.6648	2	y20	0.0014	0.0856
1193.6899	1193.6888	1	y10	0.0011	0.1675
1225.2078	1225.2069	2	y21	0.0018	0.0816
1281.7509	1281.7489	2	y22	0.0040	0.0780
1415.7102	1415.7087	1	b13	0.0015	0.1413

K. pneumoniae KP257

Uncharacterised protein (gene KPN_00497)

Uniprot accession number: A6T5S6
Charge state, fragmentation method: 9+, CID
Observed monoisotopic mass: 7698.9926 Da
Sequence:
AQLITKEEVKHFKLTKVGPISVGPSGGEFSSPSDLHDQLSKLADEKGGKYYV
ITAAREHGPNFEATAEVYK

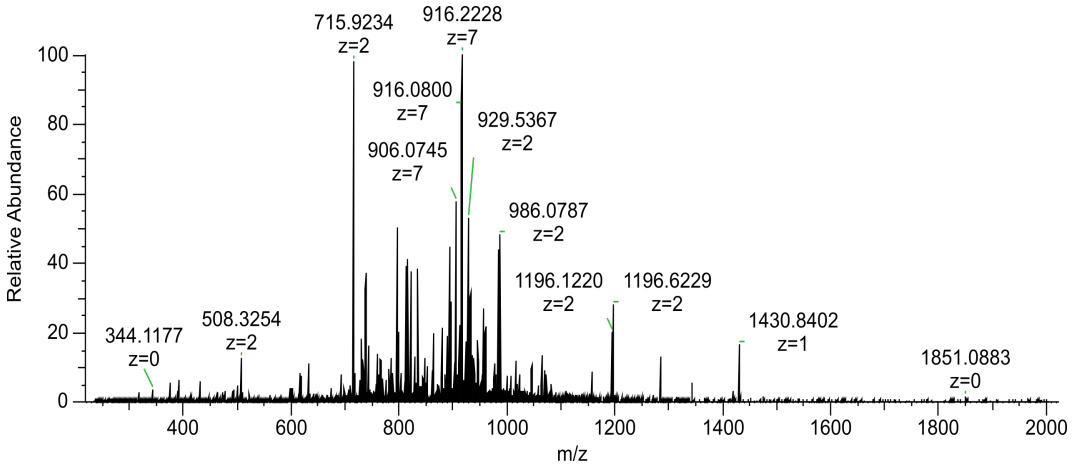
R→K mutation at position 49



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
663.3220	663.3222	2	y12	-0.0004	-0.3015
672.8912	672.8916	4	b25	-0.0016	-0.5944
687.1468	687.1470	4	b26	-0.0008	-0.2911
701.4017	701.4024	4	b27	-0.0028	-0.9980
733.6627	733.6630	4	b28	-0.0012	-0.4089
742.2078	742.2087	6	y40	-0.0054	-1.2126
757.8837	757.8845	4	y27	-0.0032	-1.0556
770.4297	770.4301	4	b29	-0.0016	-0.5192
783.4615	783.4615	3	b21	0.0000	0.0000
792.1875	792.1881	4	b30	-0.0024	-0.7574
795.7296	795.7308	6	y43	-0.0072	-1.5080
816.4836	816.4843	3	b22	-0.0021	-0.8573
830.6256	830.6266	5	y37	-0.0050	-1.2039
835.4908	835.4915	3	b23	-0.0021	-0.8378
876.4287	876.4294	6	y49	-0.0042	-0.7987
890.4485	890.4490	5	y40	-0.0025	-0.5615
896.8527	896.8531	3	b25	-0.0012	-0.4460
907.8550	907.8554	5	y41	-0.0020	-0.4406
915.8595	915.8602	3	b26	-0.0021	-0.7643
925.2609	925.2618	5	y42	-0.0045	-0.9727
934.8668	934.8674	3	b27	-0.0018	-0.6418
954.6753	954.6755	5	y43	-0.0010	-0.2095
977.8811	977.8816	3	b28	-0.0015	-0.5113
1003.2921	1003.2926	5	y46	-0.0025	-0.4984
1026.9036	1026.9044	3	b29	-0.0024	-0.7790
1040.1075	1040.1095	5	y48	-0.0100	-1.9229

DNA-binding protein HU-α

Uniprot accession number: A6TGQ7
Charge state, fragmentation method: 11+, CID
Observed monoisotopic mass: 9471.1664 Da
Sequence:
MNKTQLIDVIADKADLSKAQAKAALESTLAAITESLKEGDAVQLVGFGTFKVNHR
AERTGRNPQTGKEIKIAAANVPAFVSGKALKDAVK



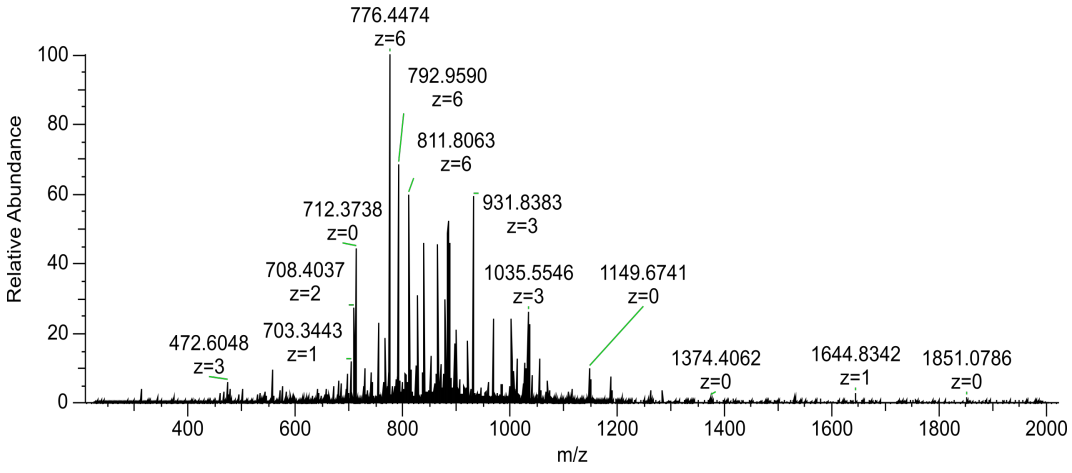
Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
317.2190	317.2183	1	y3	0.0007	2.2067
432.2461	432.2453	1	y4	0.0008	1.8508
673.4255	673.4243	1	y6	0.0012	1.7819
692.6325	692.6310	4	b26	0.0060	2.1657
715.9234	715.9219	2	y14	0.0030	2.0952
737.7804	737.7790	3	y22	0.0042	1.8976
744.4628	744.4614	1	y7	0.0014	1.8806
765.4579	765.4561	2	y15	0.0036	2.3515
785.6842	785.6812	4	b30	0.0120	3.8183
787.9400	787.9405	6	y44	-0.0030	-0.6346
797.4453	797.4441	6	y45	0.0072	1.5048
813.9572	813.9555	6	y46	0.0102	2.0886
822.4791	822.4776	2	y16	0.0030	1.8238
828.9367	828.9349	2	b15	0.0036	2.1715
832.8059	832.8028	6	y47	0.0186	3.7224
846.4700	846.4701	7	y56	-0.0007	-0.1181
879.3417	879.3401	7	y58	0.0112	1.8195
885.4785	885.4769	2	b16	0.0032	1.8069
895.4965	895.4950	7	y59	0.0105	1.6751
905.6448	905.6431	7	y60	0.0119	1.8771
915.7934	915.7913	7	y61	0.0147	2.2931
929.0348	929.0333	2	y19	0.0030	1.6146

931.9471	931.9462	7	y62	0.0063	0.9657
956.7343	956.7315	5	y45	0.0140	2.9266
958.8165	958.8147	7	y64	0.0126	1.8773
961.3092	961.3066	9	y83	0.0234	2.7047
976.5473	976.5451	5	y46	0.0110	2.2528
985.5775	985.5753	2	y20	0.0044	2.2322
1016.6117	1016.6099	1	y10	0.0018	1.7706
1023.5612	1023.5602	3	b29	0.0030	0.9770
1043.5589	1043.5554	1	b9	0.0035	3.3539
1046.5634	1046.5596	4	b40	0.0152	3.6309
1064.3213	1064.3189	4	b41	0.0096	2.2550
1070.9203	1070.9183	3	b31	0.0060	1.8676
1342.7067	1342.7035	1	b12	0.0032	2.3833
1430.8402	1430.8366	1	y14	0.0036	2.5160

P. aeruginosa PS1054

DNA-binding protein HU-β

Uniprot accession number: P05384
Charge state, fragmentation method: 11+, CID
Observed monoisotopic mass: 9081.0514 Da
Sequence:
MNKSELIDAIAASADIPKAVAGRALDAVIESVTGALKAGDSVVLVGFGTFAVKER
AARTGRNPQTGKPIKIAAAKIPGFKAGKALKDAVN

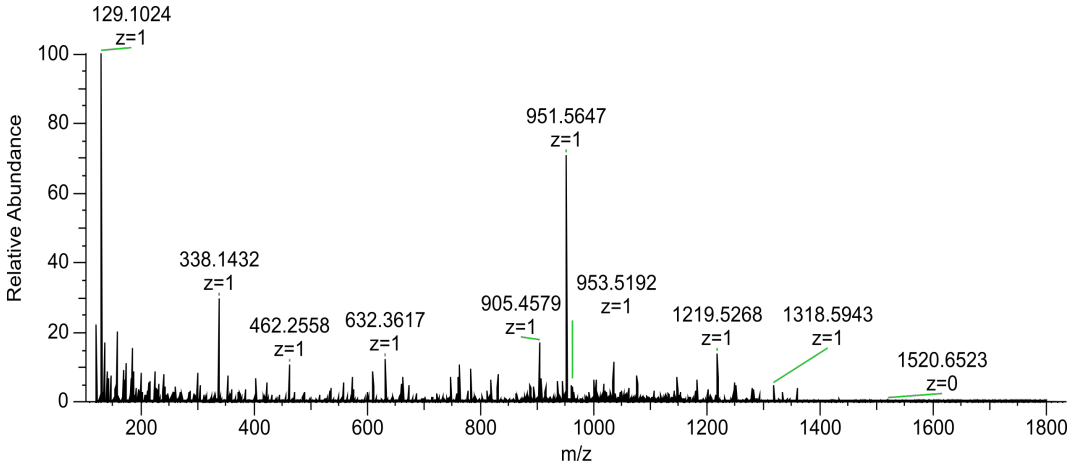


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
472.6048	472.6050	3	y14	-0.0006	-0.4232
679.5373	679.5363	7	y46	0.0070	1.4716

703.3443	703.3443	1	b6	0.0000	0.0000
709.8438	709.8438	7	y48	0.0000	0.0000
730.4093	730.4094	1	y7	-0.0001	-0.1369
733.6706	733.6732	8	y58	-0.0208	-3.5438
742.0976	742.0981	6	y43	-0.0030	-0.6738
764.9459	764.9459	2	y15	0.0000	0.0000
776.1131	776.1131	6	y45	0.0000	0.0000
811.4720	811.4719	6	y47	0.0006	0.1232
838.3399	838.3397	7	y58	0.0014	0.2386
859.4926	859.4924	7	y59	0.0014	0.2327
859.0006	859.0000	6	y50	0.0036	0.6985
864.9256	864.9255	7	y60	0.0007	0.1156
870.3824	870.3852	9	y78	-0.0252	-3.2170
878.2783	878.2782	9	b87	0.0009	0.1139
883.3595	883.3602	7	y61	-0.0049	-0.7924
886.1714	886.1712	9	y80	0.0018	0.2257
896.0079	896.0077	10	y89	0.0020	0.2232
899.5161	899.5150	7	y62	0.0077	1.2229
919.4035	919.4099	9	y83	-0.0576	-6.9610
931.5038	931.5040	3	b28	-0.0006	-0.2147
959.2887	959.2867	8	b76	0.0160	2.0849
969.1986	969.1987	3	b29	-0.0003	-0.1032
1006.0375	1006.0373	4	b41	0.0008	0.1988
1012.2126	1012.2129	3	b30	-0.0009	-0.2964
1030.8038	1030.8044	4	b42	-0.0024	-0.5821
1041.2237	1041.2236	3	b31	0.0003	0.0960
1055.5713	1055.5715	4	b43	-0.0008	-0.1895
1107.9314	1107.9289	3	b33	0.0075	2.2565
1115.5769	1115.5765	1	b10	0.0004	0.3586
1150.6162	1150.6151	3	b35	0.0033	0.9560
1188.3099	1188.3098	3	b36	0.0003	0.0842

Uncharacterised protein (gene PA4739)

Uniprot accession number: Q9HV60
Charge state, fragmentation method: 9+, HCD
Observed monoisotopic mass: 8557.5098 Da
Sequence:
ANDTMQKTEEAVSDTWITSKVKSSLIANKNVSGVDIKVETNKGVV
SLSGNVKSDAERDLAIETAKGIKGVKAVSADGLKSVE

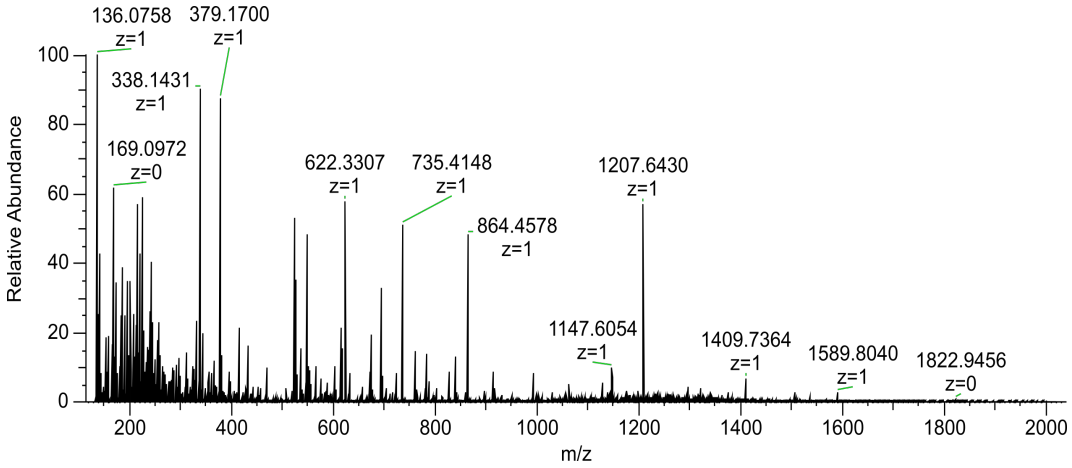


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
402.1617	402.1619	1	b4	-0.0002	-0.4973
462.2558	462.2558	1	y4	0.0000	0.0000
574.7481	574.7481	2	b10	0.0000	0.0000
610.2668	610.2666	2	b11	0.0004	0.3277
632.3614	632.3617	1	y6	-0.0003	-0.4744
661.2613	661.2610	1	b6	0.0003	0.4537
680.3778	680.3775	2	y14	0.0006	0.4409
714.6516	714.6514	4	y28	0.0008	0.2799
747.3890	747.3883	1	y7	0.0007	0.9366
789.3562	789.3560	2	b7	0.0004	0.2534
797.1895	797.1899	4	y31	-0.0016	-0.5018
802.6152	802.6148	6	y47	0.0024	0.4984
811.3545	811.3541	2	b15	0.0008	0.4930
818.4260	818.4254	1	y8	0.0006	0.7331
864.7235	864.7231	4	y34	0.0016	0.4626
886.4811	886.4811	4	y35	0.0000	0.0000
890.4035	890.4036	1	b8	-0.0001	-0.1123
914.7004	914.7005	5	y45	-0.0005	-0.1093
936.5104	936.5101	4	y37	0.0012	0.3203
961.2763	961.2772	4	y38	-0.0036	-0.9363
962.9354	962.9363	5	y47	-0.0045	-0.9346
965.3976	965.3979	7	y66	-0.0021	-0.3108
991.9791	991.9807	7	y67	-0.0112	-1.6129
1000.2982	1000.2997	4	y40	-0.0060	-1.4996
1004.5258	1004.5259	1	y10	-0.0001	-0.0995
1006.4153	1006.4161	7	y68	-0.0056	-0.7949
1019.4463	1019.4462	1	b9	0.0001	0.0981
1019.8858	1019.8857	3	y30	0.0003	0.0981
1022.8474	1022.8485	7	y69	-0.0077	-1.0754
1034.5665	1034.5661	5	y51	0.0020	0.3866
1035.2812	1035.2816	7	y70	-0.0028	-0.3864

1054.3810	1054.3798	5	y52	0.0060	1.1381
1062.5839	1062.5840	3	y31	-0.0003	-0.0941
1075.5632	1075.5630	1	y11	0.0002	0.1859
1077.1882	1077.1883	5	y53	-0.0005	-0.0928
1107.2813	1107.2824	6	y65	-0.0066	-0.9934
1118.3568	1118.3567	4	y44	0.0004	0.0894
1133.6237	1133.6212	3	y33	0.0075	2.2053
1143.1242	1143.1238	4	y45	0.0016	0.3499
1148.4891	1148.4888	1	b10	0.0003	0.2612
1152.6307	1152.6283	3	y34	0.0072	2.0822
1152.6307	1152.6283	3	y34	0.0072	2.0822
1175.1481	1175.1475	4	y46	0.0024	0.5106
1181.6393	1181.6390	3	y35	0.0009	0.2539
1203.6595	1203.6579	1	y12	0.0016	1.3293
1219.5268	1219.5259	1	b11	0.0009	0.7380
1248.3438	1248.3444	3	y37	-0.0018	-0.4806
1250.2941	1250.2948	3	b35	-0.0021	-0.5599
1281.3661	1281.3672	3	y38	-0.0033	-0.8585
1292.9544	1292.9558	4	y51	-0.0056	-1.0828
1318.5943	1318.5944	1	b12	-0.0001	-0.0758
1333.3969	1333.3971	3	y40	-0.0006	-0.1500
1359.7469	1359.7478	1	y14	-0.0009	-0.6619
1520.6523	1520.6533	1	b14	-0.0010	-0.6576
1621.6995	1621.7010	1	b15	-0.0015	-0.9250

Uncharacterised protein (gene PA0039)

Uniprot accession number: Q9I793
Charge state, fragmentation method: 5+, HCD
Observed monoisotopic mass: 5731.9826 Da
Sequence:
AKPCEELKAEIDAKIKANGVPAYTLEIVDKGSVTDKKVVGTCDGGTKEIVYQRG

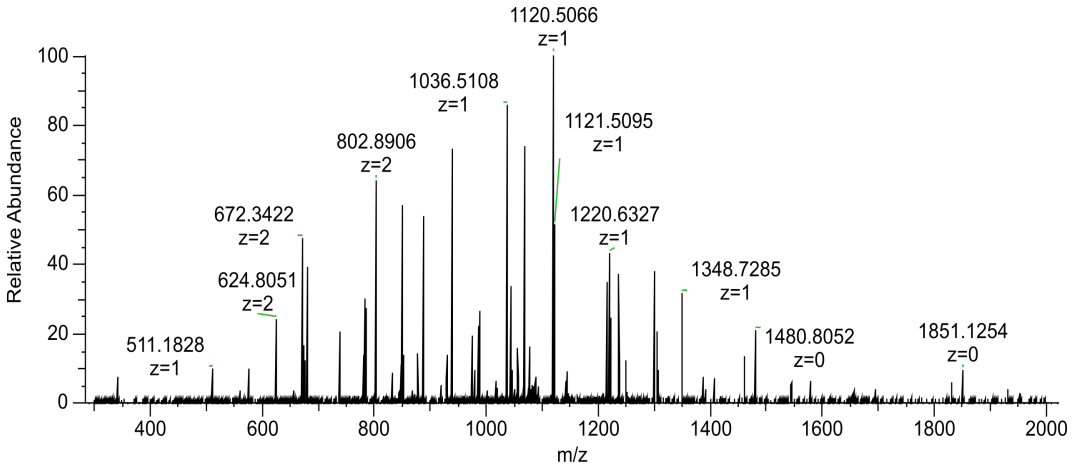


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
523.2625	523.2623	1	y4	0.0002	0.3822
604.3256	604.3251	2	y11	0.0010	0.1655
622.3312	622.3307	1	y5	0.0005	0.3214
735.4154	735.4148	1	y6	0.0006	0.2720
864.4583	864.4574	1	y7	0.0009	0.2314
992.5531	992.5524	1	y8	0.0007	0.2015
1207.6430	1207.6430	1	y11	0.0000	0.1656
1218.1350	1218.1402	4	b47	-0.0206	-4.2360
360.1985	360.1990	1	y3	-0.0005	-1.3881

Candida glabrata

Uncharacterised protein (gene HSP12)

Uniprot accession number: Q6FPF6
Charge state, fragmentation method: 5+, CID
Observed monoisotopic mass: 5516.6756 Da
Sequence:
aSDAGRKNFSDKLNEGLTPDSQKSTWDKGKEFVTDETDKLAGKFQGEENK
Truncated protein



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Charge	Fragment	Mass error [Da]	Mass error [ppm]
576.2653	576.2624	1	y5	0.0029	5.0324
624.8051	624.8020	2	b11	0.0062	2.3207
674.8678	674.8646	2	y12	0.0064	2.1486
681.3471	681.3440	2	b12	0.0062	2.1281
738.3689	738.3655	2	b13	0.0068	1.9638

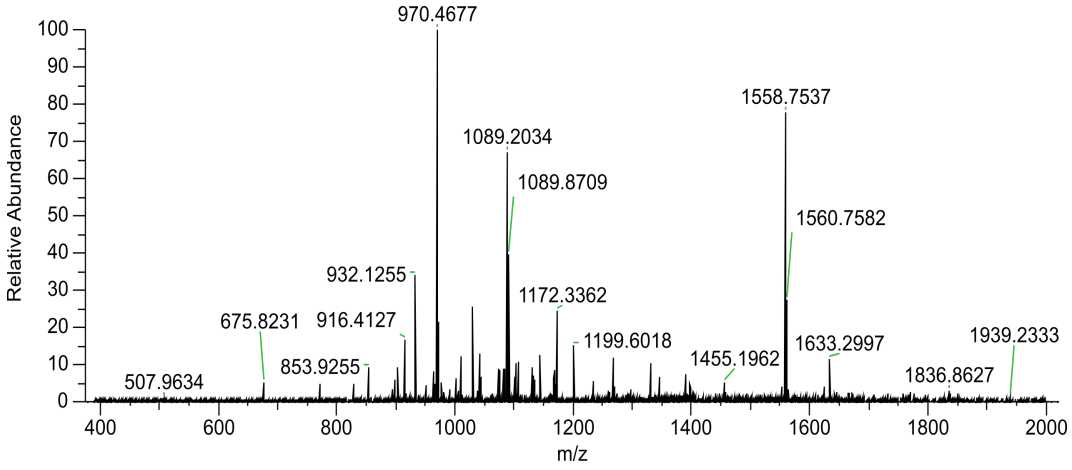
802.8906	802.8868	2	b14	0.0076	1.8060
831.4015	831.3975	2	b15	0.0080	1.7441
847.4272	847.4232	2	y15	0.0080	1.7111
851.3932	851.3894	1	y7	0.0038	3.4062
887.9437	887.9396	2	b16	0.0082	1.6330
938.4681	938.4634	2	b17	0.0094	1.5451
979.4883	979.4843	1	y8	0.0040	2.9607
1036.5108	1036.5058	1	y9	0.0050	2.7979
1044.5081	1044.5033	2	y19	0.0096	1.3882
1055.5297	1055.5255	5	y47	0.0210	0.5495
1120.5066	1120.5018	1	b10	0.0048	2.5881
1143.0551	1143.0502	2	y20	0.0098	1.2685
1214.9281	1214.9235	3	y32	0.0138	0.7957
1220.6327	1220.6270	1	y11	0.0057	2.3758
1235.6143	1235.6085	2	y22	0.0116	1.1735
1248.6061	1248.6060	3	y33	0.0003	0.7742
1299.6618	1299.6559	2	y23	0.0118	1.1157
1305.3144	1305.3079	3	y35	0.0195	0.7406
1348.7285	1348.7219	1	y12	0.0066	2.1502
1460.6959	1460.6890	2	b26	0.0138	0.9927
1544.2547	1544.2489	2	y27	0.0116	0.9390

B.3 Chapter 5

Human skin proteins

Psoriasin (S100-A7)

Uniprot accession number: P31151
Charge state, fragmentation method: 8+, CID
Observed monoisotopic mass: 11358.5074 Da
Sequence:
aSNTQAERSIIGMIDMFHKYTRRDDKIDKPSLLTMMKENFPNFLSACDKKGTNY
LADVFEKKDKNEDKKIDFSEFLSLLGDIATDYHKQSHGAAPCSGGSQ

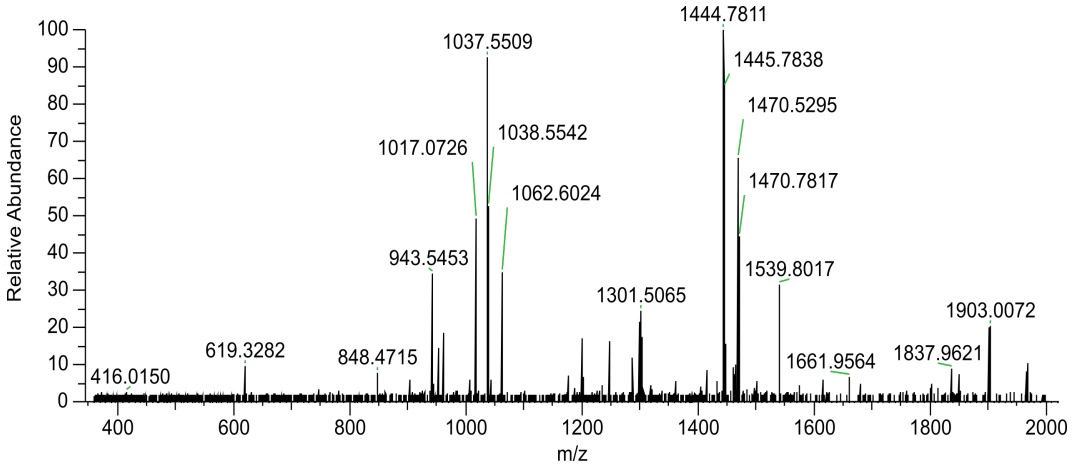


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
915.4059	915.4046	B8	0.0013	1.4474
1028.4947	1028.4887	B9	0.0061	5.8844
1141.5791	1141.5727	B10	0.0021	1.8198
1141.5791	1141.5727	B10	0.0064	5.5853
1198.5921	1198.5942	B11	-0.0064	-5.3217
1198.5921	1198.5942	B11	-0.0021	-1.7354
1329.6308	1329.6347	B12	-0.0082	-6.1525
1329.6308	1329.6347	B12	-0.0039	-2.9196
1557.7537	1557.7457	B14	0.0037	2.3479
1557.7537	1557.7457	B14	0.0080	5.1074
2792.3466	2792.3486	B23	-0.0063	-2.2610
2792.3466	2792.3486	B23	-0.0020	-0.7216
2907.3675	2907.3756	B24	-0.0124	-4.2532
2907.3675	2907.3756	B24	-0.0081	-2.7747
2907.3730	2907.3756	B24	-0.0068	-2.3477
2907.3730	2907.3756	B24	-0.0025	-0.8692
3263.5722	3263.5815	B27	-0.0136	-4.1686
3263.5743	3263.5815	B27	-0.0116	-3.5475
3263.5722	3263.5815	B27	-0.0093	-2.8515
3263.5743	3263.5815	B27	-0.0073	-2.2304
3391.6740	3391.6765	B28	-0.0068	-1.9909
3391.6740	3391.6765	B28	-0.0025	-0.7235
3801.9041	3801.9294	B32	-0.0297	-7.7988
3801.9058	3801.9294	B32	-0.0279	-7.3343
3801.9041	3801.9294	B32	-0.0254	-6.6682
3801.9058	3801.9294	B32	-0.0236	-6.2037
3902.9695	3902.9771	B33	-0.0119	-3.0522
3902.9695	3902.9771	B33	-0.0076	-1.9508
4033.9909	4034.0176	B34	-0.0310	-7.6840
4033.9915	4034.0176	B34	-0.0304	-7.5366
4033.9909	4034.0176	B34	-0.0267	-6.6185

4033.9915	4034.0176	B34	-0.0261	-6.4710
4165.0420	4165.0581	B35	-0.0203	-4.8839
4165.0420	4165.0581	B35	-0.0160	-3.8518
4165.0491	4165.0581	B35	-0.0132	-3.1749
4165.0491	4165.0581	B35	-0.0089	-2.1428
4293.1538	4293.1530	B36	-0.0035	-0.8126
4293.1538	4293.1530	B36	0.0008	0.1887
4422.2058	4422.1956	B37	0.0058	1.3207
4422.2058	4422.1956	B37	0.0101	2.2928
4536.2147	4536.2385	B38	-0.0282	-6.2062
4536.2147	4536.2385	B38	-0.0239	-5.2585
4683.2950	4683.3070	B39	-0.0163	-3.4740
4683.2950	4683.3070	B39	-0.0120	-2.5561

Ubiqutin

Uniprot accession number: P62979
Charge state, fragmentation method: 6+, CID
Observed monoisotopic mass: 7907.2100 Da
Sequence:
MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTL
SDYNIQKESTLHLV



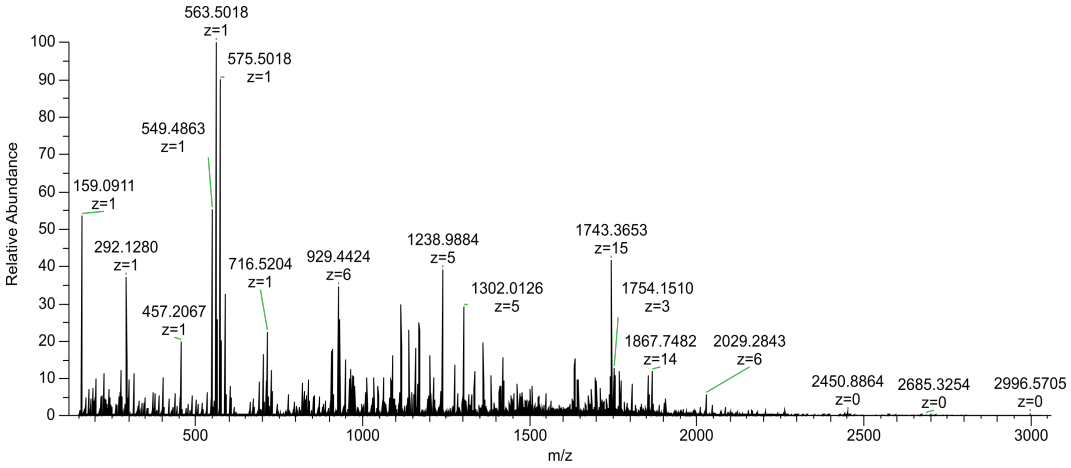
Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
1061.5976	1061.5943	B9	0.0033	3.1180
2032.1215	2032.1278	B18	-0.0064	-3.1297
1443.7718	2096.1854	Y18	-0.0082	-3.8973
2073.0933	2725.4987	Y24	0.0000	-0.0024
5875.0667	6527.4887	Y58	-0.0166	-2.5487

B.4 Chapter 6

Higher molecular weight proteins

Outer membrane porin C (*E. coli* K12)

Uniprot accession number: P06996
Charge state, fragmentation method: 27+, HCD
Observed monoisotopic mass: 38284.0439 Da
Sequence:
AEVYNKDGNKLDLYGKVDGLHYFSDNKDVDGDQTYMRLGFKGETQVTDQLTG
YGQWEYQIQGNSAENENNSWTRVAFAGLKFQDVGSFDYGRNYGVVYDVTSW
TDVLPEFGGDTYGSDNFMQQRGNGFATYRNTDFFGLVDGLNFAVQYQGKNGN
PSGEGFTSGVTNNGRDALRQNGDGVGGSITYDYEGFGIGGAISSSKRTDAQNT
AAYIGNGDRAETYTGGLKYDANNIYLAAQYTQTYNATRVGSLGWANKAQNFEA
VAQYQFDFGLRPSLAYLQSKGKNLGRGYDDEDILKYVDVGATYYFNKNMSTYV
DYKINLLDDNQFTRDAGINTDNIVALGLVYQF



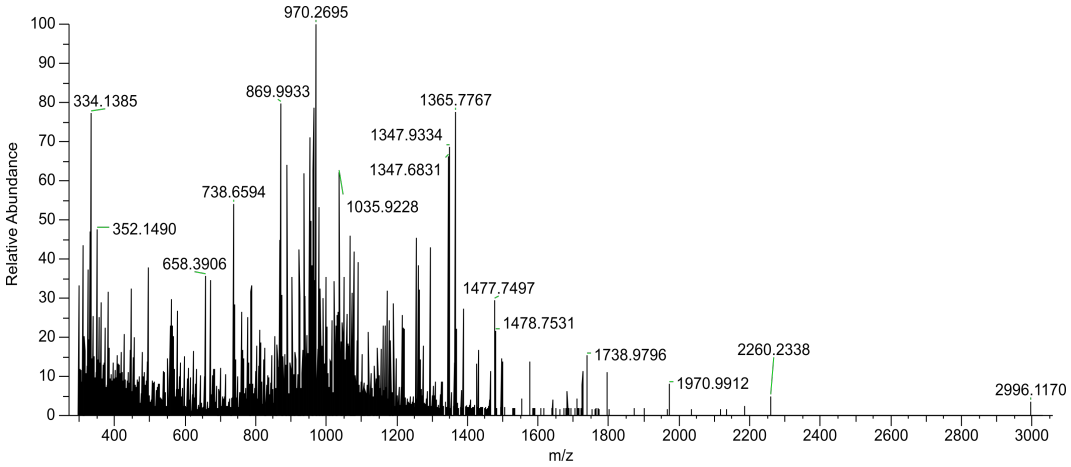
Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
200.0790	200.0797	B2	-0.0007	-3.2587
819.3752	819.3763	B7	-0.0010	-1.2656
1231.6178	1231.6197	B11	-0.0018	-1.4867
1346.6451	1346.6466	B12	-0.0015	-1.1117
1346.6457	1346.6466	B12	-0.0009	-0.6498
1459.7287	1459.7307	B13	-0.0020	-1.3461
1622.7939	1622.7940	B14	-0.0001	-0.0493
1807.9024	1807.9104	B16	-0.0080	-4.4482
1906.9699	1906.9788	B17	-0.0090	-4.7111
2022.0006	2022.0058	B18	-0.0051	-2.5381

2192.0992	2192.1113	B20	-0.0121	-5.5070
2329.1568	2329.1702	B21	-0.0135	-5.7776
2329.1598	2329.1702	B21	-0.0104	-4.4810
2492.2152	2492.2335	B22	-0.0183	-7.3629
2639.2908	2639.3019	B23	-0.0111	-4.2246
2841.3459	2841.3609	B25	-0.0150	-5.2897
3083.4707	3083.4988	B27	-0.0282	-9.1315
3198.4972	3198.5257	B28	-0.0285	-8.9116
3198.5105	3198.5257	B28	-0.0153	-4.7753
3297.5718	3297.5942	B29	-0.0224	-6.7910
3297.5787	3297.5942	B29	-0.0154	-4.6758
3412.5986	3412.6211	B30	-0.0225	-6.6011
3412.6074	3412.6211	B30	-0.0137	-4.0010
3584.6404	3584.6695	B32	-0.0291	-8.1279
3813.7496	3813.7758	B34	-0.0262	-6.8591
4263.9634	4263.9807	B37	-0.0173	-4.0570
4377.0530	4377.0647	B38	-0.0117	-2.6751
4581.1380	4581.1546	B40	-0.0167	-3.6351
4581.1423	4581.1546	B40	-0.0124	-2.6963
5124.3858	5124.4199	B45	-0.0341	-6.6519
5124.3873	5124.4199	B45	-0.0326	-6.3621
5223.4416	5223.4883	B46	-0.0467	-8.9490
5223.4460	5223.4883	B46	-0.0423	-8.0908
5223.4526	5223.4883	B46	-0.0357	-6.8391
5324.5039	5324.5360	B47	-0.0320	-6.0184
5324.5132	5324.5360	B47	-0.0227	-4.2693
5439.5275	5439.5629	B48	-0.0354	-6.5101
5439.5386	5439.5629	B48	-0.0243	-4.4680
5439.5476	5439.5629	B48	-0.0153	-2.8122
5567.5776	5567.6215	B49	-0.0439	-7.8788
5567.5890	5567.6215	B49	-0.0325	-5.8362
5680.6514	5680.7056	B50	-0.0541	-9.5258
5680.6625	5680.7056	B50	-0.0430	-7.5769
5680.6664	5680.7056	B50	-0.0392	-6.8951
5781.7062	5781.7532	B51	-0.0470	-8.1288
5781.7259	5781.7532	B51	-0.0274	-4.7326
5781.7284	5781.7532	B51	-0.0249	-4.2985
5838.7277	5838.7747	B52	-0.0470	-8.0527
5838.7356	5838.7747	B52	-0.0391	-6.7016
6001.8039	6001.8380	B53	-0.0341	-5.6834
6001.8076	6001.8380	B53	-0.0305	-5.0776
6058.8366	6058.8595	B54	-0.0229	-3.7811
6186.8715	6186.9181	B55	-0.0466	-7.5286
6186.8846	6186.9181	B55	-0.0334	-5.4056
6186.8975	6186.9181	B55	-0.0205	-3.3196
6372.9464	6372.9974	B56	-0.0509	-7.9931
6372.9488	6372.9974	B56	-0.0486	-7.6186

6372.9750	6372.9974	B56	-0.0224	-3.5177
6501.9794	6502.0400	B57	-0.0606	-9.3160
6501.9891	6502.0400	B57	-0.0509	-7.8263
6501.9900	6502.0400	B57	-0.0500	-7.6871
6665.0539	6665.1033	B58	-0.0494	-7.4153
6665.0565	6665.1033	B58	-0.0468	-7.0182
6793.1093	6793.1619	B59	-0.0525	-7.7353
6793.1233	6793.1619	B59	-0.0386	-5.6806
6793.1270	6793.1619	B59	-0.0349	-5.1341
6906.2121	6906.2459	B60	-0.0338	-4.8995
6906.2213	6906.2459	B60	-0.0246	-3.5626
7034.2817	7034.3045	B61	-0.0229	-3.2494
12050.5312	12050.6060	B106	-0.0748	-6.2085
12163.5989	12163.6901	B107	-0.0912	-7.5002
293.1363	293.1376	Y2	-0.0013	-4.3563
456.1997	456.2009	Y3	-0.0012	-2.6238
725.3744	725.3748	Y6	-0.0004	-0.5487
838.4562	838.4589	Y7	-0.0027	-3.2107
909.4927	909.4960	Y8	-0.0033	-3.5910
1008.5615	1008.5644	Y9	-0.0029	-2.9150
3542.7862	3542.8045	Y31	-0.0183	-5.1521
12701.1489	12701.2359	Y112	-0.0870	-6.8511
17149.3088	17149.3991	Y155	-0.0903	-5.2628
25894.0942	25894.3304	Y237	-0.2362	-9.1201

50S ribosomal protein S16 (*E. coli* K12)

Uniprot accession number: P0ADY7
Charge state, fragmentation method: 16+, HCD
Observed monoisotopic mass: 15302.4269 Da
Sequence:
mMLQPKRTKFRKMHKGRNRGLAQGTDVSFGSFGLKAVGRGRLTARQIEAA
RRAMTRAVKRQGKIWIRVFPDKPITEKPLAVRMGKGKGNVEYWVALIQPG
KVLYEMDGVPEELAREAFKLAAAKLPIKTTFVTKTVM
Hydroxy-arginine (R81)

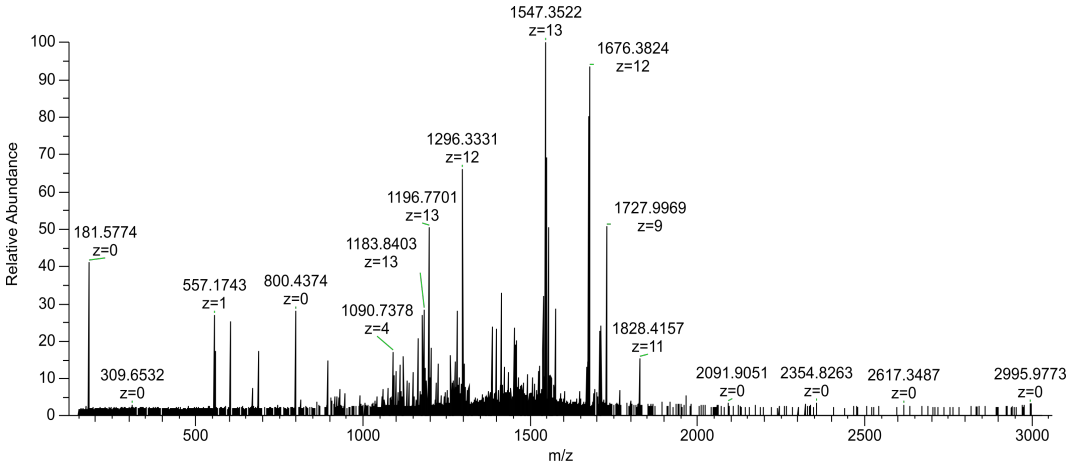


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
2949.5949	2949.6130	B25	-0.0181	-6.1347
2949.6033	2949.6130	B25	-0.0097	-3.2896
3048.6664	3048.6814	B26	-0.0150	-4.9336
3135.6890	3135.7134	B27	-0.0244	-7.7877
3282.7620	3282.7818	B28	-0.0198	-6.0443
3339.7872	3339.8033	B29	-0.0161	-4.8299
5279.8753	5279.9101	B47	-0.0348	-6.5939
7987.3954	7987.4545	B70	-0.0591	-7.4034
7987.4098	7987.4545	B70	-0.0448	-5.6079
8115.4852	8115.5495	B71	-0.0643	-7.9248
8325.6346	8325.6863	B73	-0.0518	-6.2186
8683.8270	8683.8716	B76	-0.0445	-5.1294
10007.5173	10007.6161	B89	-0.0987	-9.8672
10136.5660	10136.6586	B90	-0.0927	-9.1405
10299.6558	10299.7220	B91	-0.0662	-6.4293
10299.6620	10299.7220	B91	-0.0600	-5.8219
10485.7001	10485.8013	B92	-0.1012	-9.6493
10485.7270	10485.8013	B92	-0.0743	-7.0870
10485.7493	10485.8013	B92	-0.0520	-4.9578
10584.7936	10584.8697	B93	-0.0761	-7.1866
10584.8027	10584.8697	B93	-0.0670	-6.3280
10655.8233	10655.9068	B94	-0.0835	-7.8340
10655.8262	10655.9068	B94	-0.0806	-7.5677
10768.9320	10768.9909	B95	-0.0589	-5.4684
10768.9388	10768.9909	B95	-0.0521	-4.8358
10882.0033	10882.0749	B96	-0.0716	-6.5805
10882.0272	10882.0749	B96	-0.0477	-4.3859
11010.0406	11010.1335	B97	-0.0929	-8.4340
578.3080	578.3098	Y5	-0.0018	-3.1056
1026.5405	1026.5420	Y9	-0.0015	-1.4632
1364.7681	1364.7737	Y12	-0.0057	-4.1523

1364.7714	1364.7737	Y12	-0.0023	-1.7204
3102.7315	3102.7514	Y28	-0.0199	-6.4272
3102.7389	3102.7514	Y28	-0.0125	-4.0306
3258.8258	3258.8413	Y30	-0.0155	-4.7489
4291.3135	4291.3363	Y39	-0.0228	-5.3165

Outer membrane porin A (*K. pneumoniae* KP257)

Uniprot accession number: A6T751
Charge state, fragmentation method: 27+, CID
Observed monoisotopic mass: 35990.5745 Da
Sequence:
APKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGG
YQVNPYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDL
DIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIA
TRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPA
PEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVL
GYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGN
TCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQPAA
Signal peptide cleavage (1–21)
Disulfide bond (321–333)
Amidation of C-terminus

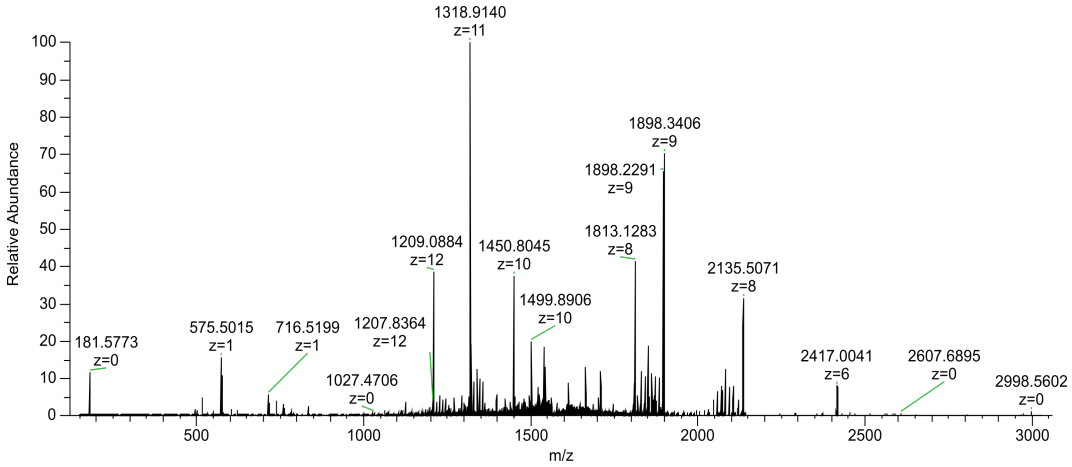


Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
4356.9095	4356.9386	B39	-0.0291	-6.6838
4469.9996	4470.0227	B40	-0.0230	-5.1532
4726.1012	4726.1398	B44	-0.0386	-8.1637
20456.6229	20456.7927	B190	-0.1698	-8.2998

15365.7601	15365.8869	Y143	-0.1268	-8.2523
15365.7620	15365.8869	Y143	-0.1249	-8.1270
15365.7625	15365.8869	Y143	-0.1243	-8.0923
15533.8397	15533.9767	Y145	-0.1371	-8.8242
15533.8478	15533.9767	Y145	-0.1290	-8.3024
15533.8822	15533.9767	Y145	-0.0945	-6.0852
30613.0763	30613.3279	Y285	-0.2516	-8.2197

Probable binding protein component of ABC transporter (*P. aeruginosa* PS1054)

Uniprot accession number: Q9I402
Charge state, fragmentation method: 20+, CID
Observed monoisotopic mass: 30785.6344 Da
Sequence:
DELTGTLKKIKETGTITLGHRDASIPFSYLGTEPGKPIGYSHDLQLKVVEAVK
KELNLPELKVRYNLVTSQTRIPLVQNGTVDIECGSTTNNEERQKQVDFSVGI
FEVGTRLLSKKTANIKDFDDLKGKNVVTTAGTTSERLLKAMNADKKMGMNII
SAKDHGESFMMLESGRAVAFMMDDALLYGEMAKAKKPDDWVVGGTPQSF
EIYGCMVRKGDAAFKKVVDKAITDTYASGEVNKIYDKWFTQPIPPKGLNLNF
PMSEELKKLIASPTDKAAEQM
Signal peptide cleavage (1–23)
Disulfide bond (86–211)



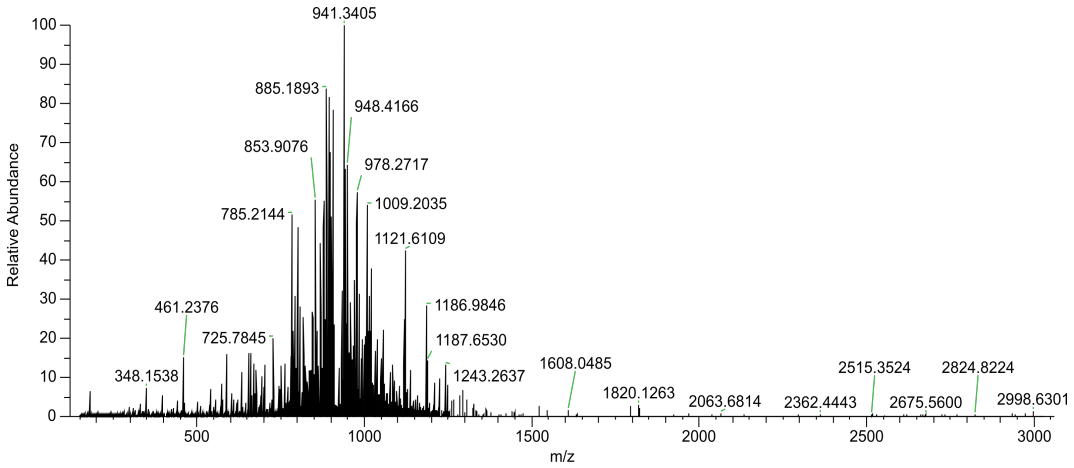
Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
985.5427	985.5444	B9	-0.0017	-1.6772
1098.6229	1098.6285	B10	-0.0056	-5.0891
1226.7180	1226.7234	B11	-0.0054	-4.3808
1614.8822	1614.8828	B15	-0.0006	-0.3709

1727.9610	1727.9669	B16	-0.0059	-3.3890
1829.0045	1829.0146	B17	-0.0100	-5.4920
1942.0877	1942.0986	B18	-0.0109	-5.5986
1999.1150	1999.1201	B19	-0.0051	-2.5551
2478.3325	2478.3442	B23	-0.0117	-4.7189
2565.3592	2565.3762	B24	-0.0170	-6.6380
2565.3638	2565.3762	B24	-0.0124	-4.8344
2678.4508	2678.4602	B25	-0.0094	-3.5102
2678.4534	2678.4602	B25	-0.0068	-2.5533
2775.4929	2775.5130	B26	-0.0201	-7.2437
2775.4996	2775.5130	B26	-0.0134	-4.8420
2922.5636	2922.5814	B27	-0.0178	-6.1066
2922.5702	2922.5814	B27	-0.0112	-3.8257
3009.5954	3009.6134	B28	-0.0180	-5.9852
3009.5965	3009.6134	B28	-0.0169	-5.6280
3172.6530	3172.6768	B29	-0.0237	-7.4785
3172.6539	3172.6768	B29	-0.0228	-7.1983
3285.7483	3285.7608	B30	-0.0126	-3.8204
3285.7522	3285.7608	B30	-0.0086	-2.6256
3342.7659	3342.7823	B31	-0.0164	-4.8962
3342.7671	3342.7823	B31	-0.0151	-4.5304
3572.8521	3572.8726	B33	-0.0204	-5.7186
3572.8581	3572.8726	B33	-0.0145	-4.0612
4624.3512	4624.3812	B43	-0.0301	-6.5025
6329.3991	6329.4314	B58	-0.0323	-5.0992
6555.4830	6555.5267	B60	-0.0437	-6.6646
676.3195	676.3214	Y6	-0.0019	-2.8241
989.4427	989.4488	Y9	-0.0061	-6.1792
1076.4782	1076.4808	Y10	-0.0026	-2.4385
1147.5164	1147.5179	Y11	-0.0015	-1.3368
1260.5956	1260.6020	Y12	-0.0064	-5.0611
1373.6841	1373.6860	Y13	-0.0019	-1.4130
1742.9502	1742.9600	Y16	-0.0099	-5.6616
2316.1610	2316.1705	Y21	-0.0095	-4.0951
2577.2759	2577.2818	Y23	-0.0060	-2.3110
3296.6890	3296.7148	Y30	-0.0258	-7.8305
3296.6973	3296.7148	Y30	-0.0176	-5.3238
3506.8266	3506.8516	Y32	-0.0250	-7.1298
3735.9356	3735.9579	Y34	-0.0223	-5.9714
3882.9990	3883.0263	Y35	-0.0273	-7.0357
4069.0950	4069.1056	Y36	-0.0106	-2.6131
4197.1792	4197.2006	Y37	-0.0213	-5.0824
4588.3436	4588.3749	Y40	-0.0313	-6.8314
5115.5934	5115.6453	Y45	-0.0518	-10.1276
5202.6515	5202.6773	Y46	-0.0258	-4.9584
5273.6689	5273.7144	Y47	-0.0455	-8.6239
5436.7492	5436.7777	Y48	-0.0285	-5.2454

5537.7927	5537.8254	Y49	-0.0327	-5.9128
5652.8171	5652.8523	Y50	-0.0352	-6.2349
5753.8715	5753.9000	Y51	-0.0285	-4.9570
5866.9504	5866.9841	Y52	-0.0337	-5.7462
5937.9777	5938.0212	Y53	-0.0435	-7.3188
6066.0619	6066.1162	Y54	-0.0542	-8.9423
27212.8135	27212.9402	Y246	-0.1267	-4.6559
27341.7572	27341.9828	Y247	-0.2256	-8.2512
27442.9438	27443.0305	Y248	-0.0867	-3.1588
27499.8781	27500.0519	Y249	-0.1739	-6.3230
27612.9677	27613.1360	Y250	-0.1683	-6.0935
27775.9103	27776.1993	Y251	-0.2890	-10.4055
27776.0149	27776.1993	Y251	-0.1845	-6.6414
27862.9161	27863.2314	Y252	-0.3153	-11.3160
28010.0236	28010.2998	Y253	-0.2761	-9.8581
28107.0997	28107.3525	Y254	-0.2528	-8.9936
28107.1792	28107.3525	Y254	-0.1733	-6.1653

Peptidoglycan associated lipoprotein (*P. aeruginosa* PS1054)

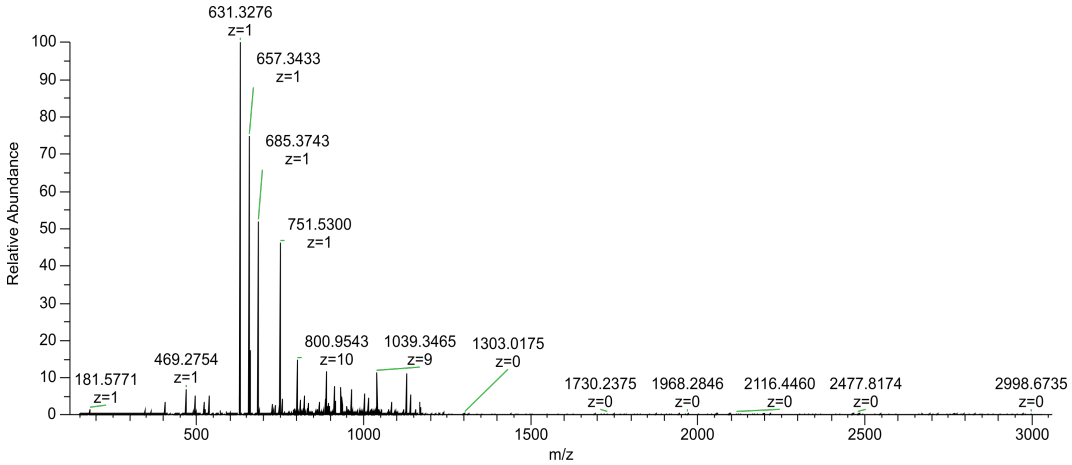
Uniprot accession number: Q9I4Z4
Charge state, fragmentation method: 19+, CID
Observed monoisotopic mass: 17524.1769 Da
Sequence:
CSSKGGDASGEGANGGVDPNAGYGANSGAVDGSLSDEAALRAITTFYFE
YDSSDLKPEAMRALDVHAKDLKGSGQRVVLEGHTDERGTREYNMALGER
RAKAVQRYLVLQGVSPAQLELVSYGKERPVATGHDEQSWAQNRRVELKK
Palmitoylation of N-terminal cysteine (C)
An unknown lipid group attached to the N-terminal cysteine (C)



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
1550.8134	1550.8353	B14	-0.0219	-14.0943
7987.9611	7987.9123	B75	0.0489	6.1172
13711.8547	13711.9199	B126	-0.0652	-4.7570
13711.8800	13711.9199	B126	-0.0399	-2.9110
2918.5131	2918.5223	Y25	-0.0092	-3.1465
3081.5793	3081.5856	Y26	-0.0063	-2.0454
3168.6007	3168.6176	Y27	-0.0170	-5.3610
3168.6056	3168.6176	Y27	-0.0121	-3.8146
3267.6684	3267.6861	Y28	-0.0176	-5.3971
3267.6714	3267.6861	Y28	-0.0146	-4.4756
3380.7431	3380.7701	Y29	-0.0270	-8.0002
3380.7471	3380.7701	Y29	-0.0230	-6.7952
3509.7925	3509.8127	Y30	-0.0202	-5.7644
3750.9242	3750.9553	Y32	-0.0312	-8.3136
3821.9704	3821.9925	Y33	-0.0221	-5.7834
3919.0145	3919.0452	Y34	-0.0307	-7.8279
3919.0242	3919.0452	Y34	-0.0210	-5.3648
4006.0502	4006.0772	Y35	-0.0271	-6.7557
4006.0529	4006.0772	Y35	-0.0244	-6.0798
4162.1477	4162.1671	Y37	-0.0194	-4.6569
5744.0783	5744.1251	Y50	-0.0467	-8.1386
9377.8557	9377.9320	Y83	-0.0763	-8.1340
9377.8717	9377.9320	Y83	-0.0603	-6.4305
11069.6473	11069.7145	Y98	-0.0672	-6.0666
11508.7957	11508.8888	Y101	-0.0931	-8.0866
11508.7996	11508.8888	Y101	-0.0892	-7.7484
11857.9849	11858.0526	Y104	-0.0677	-5.7066
12582.3952	12582.4757	Y111	-0.0805	-6.3982
12784.4516	12784.5347	Y113	-0.0831	-6.5010
12784.4588	12784.5347	Y113	-0.0759	-5.9372
12897.4975	12897.6188	Y114	-0.1213	-9.4035
12984.5254	12984.6508	Y115	-0.1254	-9.6556
13156.5997	13156.6992	Y117	-0.0995	-7.5638
13156.6094	13156.6992	Y117	-0.0898	-6.8280
13255.6575	13255.7676	Y118	-0.1101	-8.3067
13255.6748	13255.7676	Y118	-0.0928	-7.0010
13383.7039	13383.8262	Y120	-0.1223	-9.1355
13383.7196	13383.8262	Y120	-0.1065	-7.9595
13584.7811	13584.9011	Y122	-0.1200	-8.8334
13875.9201	13876.0230	Y125	-0.1030	-7.4196
13932.9500	13933.0445	Y126	-0.0945	-6.7816
14429.1534	14429.2726	Y131	-0.1192	-8.2603

30S ribosomal protein S5 (*E. faecium* E745)

Uniprot accession number: Q3XYX2
Charge state, fragmentation method: 19+, CID
Observed monoisotopic mass: 17335.5197 Da
Sequence:
VYIDPKHLELEDRVVAINRVTKVVKGGRRLRFAALVVVGDKNGHVGFGTG
KAQEVPEAIRKAIEDAKKNLVEVPMVGSTIPHEVIGAFSGGRILMKPAVEGS
GVAAGGPVRAVLELAGVADITSKSLGSNTPINVVRATVEGLKQLKRAEEVA
ELRGKSVEELIG



Observed mass (monoisotopic) m/z	Theoretical mass (monoisotopic) m/z	Fragment	Mass error [Da]	Mass error [ppm]
4111.3731	4111.4010	B36	-0.0279	-6.7865
4481.5602	4481.5862	B40	-0.0260	-5.7950
4917.7644	4917.8045	B44	-0.0400	-8.1400
7113.9232	7113.9583	B65	-0.0350	-4.9259
7995.4159	7995.4917	B73	-0.0758	-9.4808
7995.4390	7995.4917	B73	-0.0527	-6.5966
7995.4465	7995.4917	B73	-0.0452	-5.6571
8092.4758	8092.5445	B74	-0.0687	-8.4887
8092.4874	8092.5445	B74	-0.0571	-7.0540
8223.5149	8223.5850	B75	-0.0700	-8.5150
8223.5317	8223.5850	B75	-0.0532	-6.4722
8322.6106	8322.6534	B76	-0.0428	-5.1379
658.3526	658.3537	Y6	-0.0011	-1.6936
4089.2343	4089.2545	Y38	-0.0202	-4.9410
4289.3450	4289.3706	Y40	-0.0256	-5.9797
6123.3790	6123.4018	Y59	-0.0228	-3.7247
8654.6604	8654.7336	Y85	-0.0732	-8.4581
9012.8522	9012.9188	Y89	-0.0666	-7.3946
9111.9240	9111.9872	Y90	-0.0632	-6.9346

9111.9420	9111.9872	Y90	-0.0452	-4.9639
9339.9993	9340.0805	Y92	-0.0811	-8.6873
16845.1607	16845.3295	Y161	-0.1687	-10.0158