# Application of high-field asymmetric waveform ion mobility separation to LESA mass spectrometry of bacteria

Kocurek, Klaudia; May, Robin; Cooper, Helen

*DOI:*
10.1021/acs.analchem.9b00307



*Document Version*
Peer reviewed version

*Citation for published version (Harvard):*
Kocurek, K, May, R & Cooper, H 2019, 'Application of high-field asymmetric waveform ion mobility separation to LESA mass spectrometry of bacteria', *Analytical Chemistry*, vol. 91, no. 7, pp. 4755-4761. https://doi.org/10.1021/acs.analchem.9b00307

# Application of high-field asymmetric waveform ion mobility separation to LESA mass spectrometry of bacteria

Klaudia I. Kocurek,[1] Robin C. May[1,2] and Helen J. Cooper[1]*

[1]School of Biosciences, and [2]Institute of Microbiology and Infection, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

## Abstract

We have previously demonstrated the analysis of intact proteins directly from bacterial colonies (including Gram-negative and Gram-positive clinical isolates) grown on agar media by liquid extraction surface analysis mass spectrometry (LESA MS). Several challenges were identified in that work including (1) interference of background signal derived from the nutrient media (*Escherichia coli*), (2) a high density of protein peaks leading to the isolation of multiple protein precursor ions in a single window and consequent acquisition of composite tandem mass spectra (*Pseudomonas aeruginosa*), and (3) the overabundance of secreted peptides suppressing peaks corresponding to proteins (*Staphylococcus aureus*). Here we present the coupling of high-field asymmetric waveform ion mobility spectrometry (FAIMS) separation into the LESA MS protocol, with the aim of resolving the aforementioned challenges and thus improving the capabilities of LESA MS for bacterial characterisation. The results show that inclusion of FAIMS expands the range of detected proteins through separation of background peaks from protein signal, as well as through resolution of overlapping protein peaks which could not previously be isolated by LESA MS alone.

## Introduction

Liquid extraction surface analysis mass spectrometry (LESA MS)[1] has been extensively applied to the analysis of intact proteins from biological substrates, including dried blood spots,[2-4] tissue sections[5-9] and bacterial colonies.[7,10,11] LESA is a liquid junction-based technique. A droplet of solvent is deposited on the sample surface by a robotic pipette system and a liquid junction is maintained between the surface and the pipette tip. Analytes diffuse into the droplet, following which it is re-aspirated into the pipette tip and introduced into the mass spectrometer via chip-based nanoelectrospray ionisation.

Intact protein mass spectra generated by LESA from biological substrates are complex, containing peaks corresponding to multiple proteins and proteoforms in a wide range of charge states, as well as other classes of analytes, such as lipids or small metabolites, which may suppress less abundant protein peaks. We previously demonstrated the analysis of intact proteins directly from Gram-negative and Gram-positive bacterial colonies grown on agar media.[7,10,11] Several challenges pertaining to the sampling and analysis of bacteria were identified in our prior work.[11] Firstly, particularly pronounced in *E. coli*, multiple highly abundant peaks corresponding to components of the solvent system as well as the agar media were seen in the low *m/z* range, obstructing signal corresponding to putative small proteins. Secondly, in *P. aeruginosa*, multiple chimeric tandem mass spectra were acquired from precursor ions detected in the higher *m/z* range where the isotopic distributions of protein precursor ions would overlap and prevent clean isolation. Lastly, in *S. aureus*, the mass spectra were dominated by small secreted toxins which may have suppressed peaks corresponding to larger proteins.

High-field asymmetric waveform ion mobility spectrometry (FAIMS) separation has previously been applied in conjunction with LESA mass spectrometry for the analysis of proteins in tissue sections and dried blood spots.[5,7,9] Preliminary work suggested that FAIMS may also benefit LESA MS of bacteria.[7] FAIMS separates ions based on differences in their mobility in high and low electric

fields.[12] Ions traverse the device between two parallel electrodes aided by a carrier gas. An asymmetric waveform is applied to one of the electrodes, giving rise to an oscillating electric field which consists of a high field of short duration (known as the dispersion field, DF) and a low field of long duration; the net electric field in one cycle is zero. The differential mobility of ions in the high and low electric fields results in deflection from their trajectories through the device and leads ultimately to their neutralisation at one of the electrode surfaces. A DC compensation voltage, giving rise to the compensation field (CF), is superimposed on the asymmetric waveform to selectively correct the trajectories of a subset of ions of particular differential mobility and transmit them through the device. Here, we describe the coupling and optimisation of FAIMS to LESA MS to address the specific challenges described above in three bacterial species (*E. coli*, *P. aeruginosa* and *S. aureus*).

## Experimental section

### Materials

Acetonitrile and HPLC-grade water were purchased from J.T.Baker, Deventer, Netherlands. Formic acid and pre-mixed, powdered lysogeny broth (LB) (Miller composition) were purchased from Sigma Aldrich, Gillingham, UK. Bacteriological agar was purchased from VWR International, Leuven, Belgium.

*Escherichia coli* K-12 were obtained from the collection of the Institute of Microbial Infection (IMI), University of Birmingham, UK. *Pseudomonas aeruginosa* PS1054 and *Staphylococcus aureus* MSSA476 were provided by Mark A. Webber from the library of clinical isolates of the Queen Elizabeth Hospital Birmingham. 5 ml lysogeny broth (LB) cultures of *Escherichia coli* K-12, *Pseudomonas aeruginosa* PS1054 and *Staphylococcus aureus* MSSA476 were incubated overnight with shaking (200 rpm). 1 µl aliquots of the liquid culture were deposited onto LB agar plates and incubated for 24 hours to yield sample colonies amenable to LESA (approx. 3-5 mm diameter).

## LESA

LESA sampling was performed by use of a TriVersa NanoMate (Advion, Ithaca, NY, USA) as described previously.[11] The solvent system was composed of acetonitrile, HPLC-grade water and formic acid; the proportions used were 40:60:1 for *E. coli* and *P. aeruginosa* and 50:45:5 for *S. aureus*. Petri dishes containing colonies of interest were placed directly onto the sampling tray of the NanoMate alongside half a 96-well plate holding the solvent system. Briefly, 3 µl of sampling solvent was withdrawn from a well and 2 µl was deposited onto the colony of interest; the pipette tip was put in contact with the colony. Contact was maintained for 3 (*E. coli*) or 15 seconds (*P. aeruginosa* and *S. aureus*), following which the solvent was re-aspirated.

## FAIMS

FAIMS was performed by use of a chip-based ultraFAIMS[13,14] device (Owlstone, Cambridge, UK) in one-dimensional (1D) mode and in static mode. A dispersion field (DF) of 270 Td was used for all experiments. For 1D analyses, compensation field (CF) values were ramped from 0.0 to 4.0 Td over a period of 240 seconds. Data for static field analyses were collected for a minimum of one minute at each setting. The static CF values were determined following 1D analysis and are stated in the text.

## Mass spectrometry

Mass spectrometry was performed by use of an Orbitrap Elite (Thermo Scientific, Bremen, Germany) at a resolution of 120,000 at *m/z* 400. Samples were introduced into the mass spectrometer via nanoelectrospray ionisation at 1.75 kV and 0.3 psi. The inlet capillary temperature was increased to 300 °C. Full scan mass spectra were acquired in the *m/z* 600-2000 range.

Precursors were selected for fragmentation in isolation windows of 5 *m/z*. MS/MS experiments were performed by CID in the ion trap with use of helium gas at a normalised collision energy of 35%. Tandem mass spectra were acquired in the orbitrap for a minimum of five minutes. Automatic gain control targets were $1 \times 10^6$ charges for full-scan mass spectra and $5 \times 10^5$ charges for MS/MS spectra.

## Protein identification

Protein identification was carried out as described previously[11] using ProSightPC 4.1 Alpha software (Thermo Scientific, Bremen, Germany). Briefly, MS/MS spectra were loaded into ProSightPC in profile mode and deconvoluted by the THRASH algorithm at default settings. For each bacterial species, a custom intact protein database was created using UniProt proteome files as the basis (Uniprot ID UP000000625, 4306 protein entries for *E. coli*, UP000008816, 2889 protein entries for *S. aureus* and UP000002438, 5563 protein entries for *P. aeruginosa*). An absolute mass search was subsequently created for each species. Putative hits were optimised in ProSightPC's Sequence Gazer. Fragment assignment was generated automatically in ProSightPC from data deconvoluted by the THRASH algorithm.

# Results and discussion

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

In the absence of FAIMS, the low *m/z* region (600-800) of LESA mass spectra of *E. coli* is characterised by singly charged peaks corresponding to extraction solvent, metabolites, putative lipids and components of the underlying agar media. Peaks corresponding to protein ions in this *m/z* region, while occasionally discernible, are largely obscured by the peaks corresponding to singly charged species, preventing the isolation of protein peaks and their identification. It was proposed that these small molecules could be filtered out by FAIMS, leaving behind an undisturbed distribution of protein peaks in the same *m/z* range. Thus, smaller proteins and peptides, as well as higher charge states of larger species, should be detectable.

A one-dimensional FAIMS experiment was performed at DF 270 Td to optimise compensation field values for the transmission of proteins of interest in the low *m/z* range. DF 270 Td was selected to maximise the separation capability of the device.[7,15] Representative mass spectra acquired during the one-dimensional analysis are shown in **Supporting Figure S-1**. At CF = 0.0 to 1.2 Td, transmission of singly charged species was observed. Transmission of protein ions was observed from CF = 1.2 Td

onwards with the emergence of peaks corresponding to an unidentified 18.1 kDa protein (*m/z* 865.40 21+, 908.77 20+, 956.50 19+, 1009.69 18+, 1068.90 17+, 1135.97 16+, 1211.30 15+, 1297.67 14+ and 1397.57 13+; $MW_{obs}$ = 18155.29 Da). A distribution of peaks of low abundance corresponding to the histone-like protein H-NS (*m/z* 856.67 18+, 906.94 17+, 963.56 16+, 1027.67 15+, 1100.93 14+, 1185.53 13+, 1284.58 12+ and 1401.36 11+; $MW_{obs}$ = 15399.89 Da) was detected between CF = 2.2 Td and CF = 3.0 Td. Peaks corresponding to YahO (*m/z* 963.99 8+, 1101.42 7+ and 1284.82 6+; $MW_{obs}$ = 7702.87 Da), one of the proteins consistently detected in every LESA mass spectrum acquired from *E. coli* K-12,[11] emerged at approximately CF = 2.2 Td and persisted until CF = 3.8 Td; HU-α (*m/z* 734.10 13+, 795.12 12+, 867.30 11+, 953.92 10+, 1060.03 9+ and 1192.53 8+; $MW_{obs}$ = 9529.17 Da) and HU-β (*m/z* 710.39 13+, 769.42 12+, 839.18 11+, 923.00 10+, 1025.45 9+ and 1153.63 8+; $MW_{obs}$ = 9219.96 Da) exhibited a similar pattern of transmission. A correlation between the strength of the compensation field and the mass-to-charge ratio of transmitted ions was observed, exemplified by the distributions of HU-α and HU-β charge states. Charge states 8+ and 9+ of both proteins emerged at CF = 2.2 Td, followed by 10+ at CF = 2.4 Td, 11+ at CF = 2.6 Td, 12+ at CF = 2.9 Td and finally 13+ at CF = 3.0 Td. In the CF range 3.0–3.5 Td, singly charged peaks in the low *m/z* region were almost entirely removed and protein peaks were observed without interference of lower molecular weight analytes. Between CF = 3.5 Td and CF = 4.0 Td, the protein signal rapidly decreased in intensity.

Mass spectra in static FAIMS field mode were subsequently acquired in the range of CF = 3.0—3.5 Td, where the greatest reduction in the abundance of singly charged background peaks was observed. **Figure 1** shows a comparison of a LESA-FAIMS mass spectrum of *E. coli* K-12 acquired with no field applied versus two LESA-FAIMS mass spectra acquired in static field mode at DF = 270 Td-CF = 3.2 Td and DF = 270 Td-CF = 3.5 Td, respectively. In the mass spectrum acquired without FAIMS, as seen in **Figure 1a**, only three protein peaks were observed in the *m/z* 600-800 region; the total number of protein peaks observed in the full displayed mass range (*m/z* 600-1600) was 23.

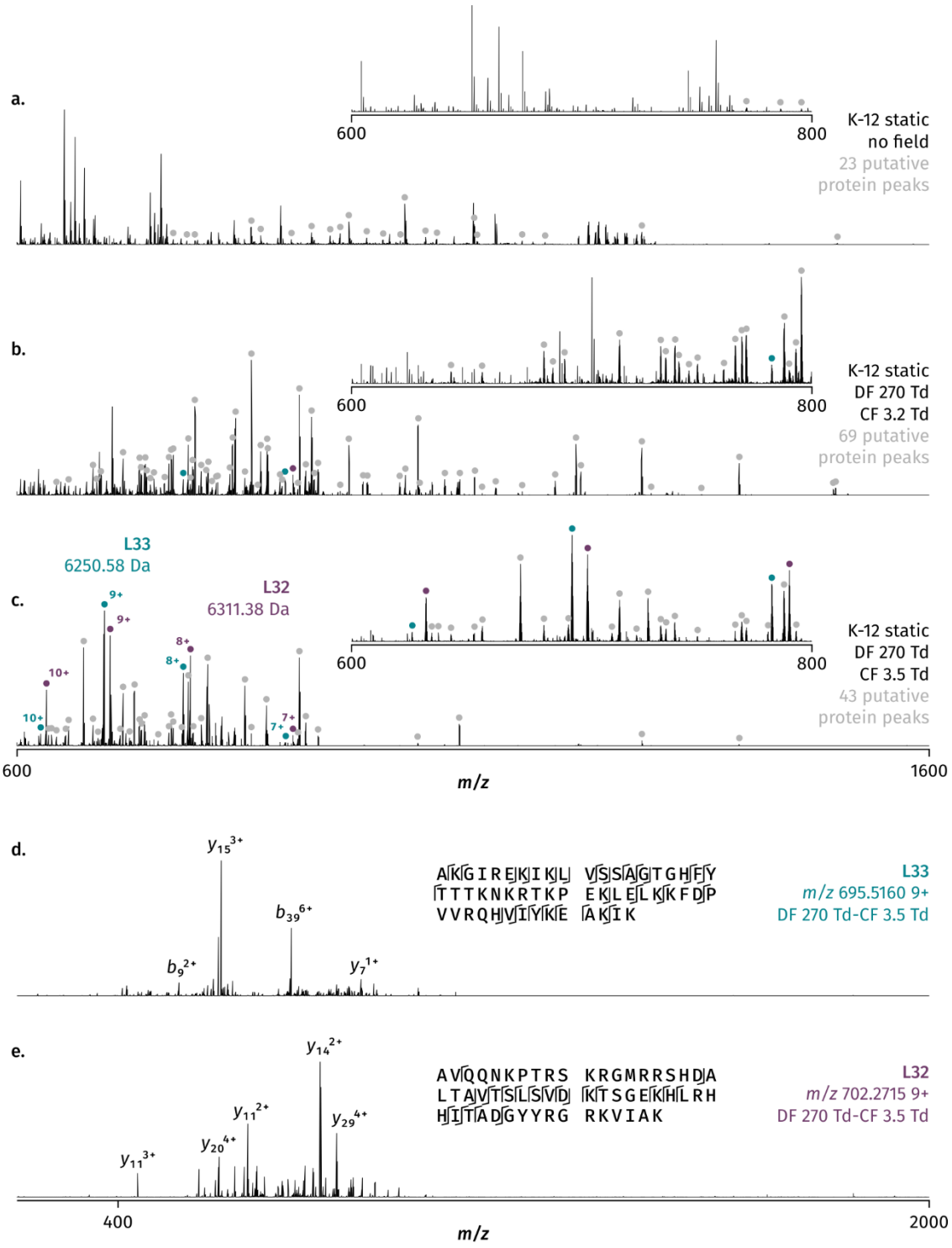


**Figure 1**. LESA-FAIMS mass spectra of *E. coli* K-12. **a**. No FAIMS field applied. **b**. DF 270 Td-CF 3.2 Td. **c**. DF 270 Td-CF 3.5 Td. Each mass spectrum was summed across 1 minute of data acquisition. The *m/z* 600-800 range of each mass spectrum is shown as an inset. MS/MS spectra of two ribosomal proteins isolated by use of FAIMS at DF 270 Td-CF 3.5 Td are shown below. **d**. L33 ($MW_{obs}$ = 6250.58 Da); **e**. L32 ($MW_{obs}$ = 6311.38 Da).

**Figure 1b** shows a mass spectrum acquired in static field mode at DF = 270 Td-CF = 3.2 Td. A total of 69 protein peaks were observed across the full mass range shown; 21 protein peaks were detected in the *m/z* 600-800 range. Singly charged background peaks were eliminated in the *m/z* 700-800 range but still constituted the majority of peaks in the *m/z* 600-700 range.

At CF = 3.5 Td, 28 protein peaks were detected in the *m/z* 600-800 range (**Figure 1c**). Singly charged peaks were absent. A total of 43 protein peaks were observed across the full mass range shown. Four peaks (*m/z* 695.92 9+, 702.27 9+, 728.54 7+ and 916.37 7+) were selected for fragmentation and identification at these settings. **Table 1** lists all the proteins isolated and identified by use of FAIMS for *E. coli* and *P. aeruginosa* (see following section); the number of peaks identified was considered a representative sample and should not be viewed as indication of technical limitations. The peak at *m/z* 695.92 (9+; $MW_{obs}$ = 6250.58 Da) was identified as the ribosomal protein L33 (**Figure 1d**). This protein was not detected previously in *E. coli* by LESA MS. The intact mass and fragmentation pattern were consistent with the cleavage of initiator methionine and the presence of methylalanine at the N-terminus, documented by prior literature using classical peptide sequencing methods.[16] Fragment assignments for all proteins identified here are supplied in **Supporting Tables S-1—13**. The peak at *m/z* 702.27 (9+; $MW_{obs}$ = 6311.38 Da) was identified as L32 (**Figure 1e**), another structural constituent of the ribosome which has not previously been identified by LESA MS. Again, the initiator methionine was not present. The peak at *m/z* 728.54 (7+; $MW_{obs}$ = 5092.75 Da) was confirmed as SRA, a sigma-associated factor previously identified by LESA MS in the absence of FAIMS.[11] Similarly, the peak at *m/z* 916.37 (7+; $MW_{obs}$ = 6406.63 Da) was identified as ribosomal protein L30.

| Protein | MW (obs.) | ProSight E-score | CF (Td) | Notes |
|---|---|---|---|---|
| *Escherichia coli* K-12 | | | | |
| SRA | 5092.75 | 8.2E-30 | 3.5 | Ribosome-associated |
| L33 | 6250.58 | 2.2E-40 | 3.5 | Ribosomal |
| L32 | 6311.38 | 1.7E-31 | 3.5 | Ribosomal |
| L30 | 6406.63 | 1.2E-06 | 3.5 | Ribosomal |
| RaiA | 12647.50 | 7.8E-12 | 2.2 | Ribosome-associated inhibitor |
| YgiW | 11968.91 | 1.8E-49 | 3.0 | Biofilm formation on agar |
| *Pseudomonas aeruginosa* PS1054 | | | | |
| L33 | 6041.32 | 3.3E-14 | 3.5 | Ribosomal |
| L32 | 6670.37 | 1.3E-10 | 2.5 | Ribosomal |
| CsrA | 6904.68 | 5.3E-20 | 3.2 | Carbon storage regulator |
| PA4739 | 8556.48 | 6.6E-14 | 3.2 | Stress response; missing N-terminal sequence AQPISLA- |
| HU-β | 9080.99 | 3.1E-14 | 3.2 | DNA-binding |
| L11 | 14898.00 | 1.7E-08 | 1.0 | Large ribosomal |
| Pilin | 14983.82 | 3.9E-12 | 2.5 | Pilus formation |

**Table 1**. A list of proteins identified on the basis of tandem mass spectra acquired by the LESA-FAIMS approach. Proteins previously identified by LESA MS in the absence of FAIMS are shown in grey.

Two further peaks were isolated and identified in static field mode at other FAIMS settings. The peak at *m/z* 1497.25 (8+; $MW_{obs}$ = 11968.91 Da), isolated at DF = 270 Td-CF = 3.0 Td, was identified as YgiW, a periplasmic protein involved in biofilm formation on inanimate substrates (i.e. agar media); the intact mass confirmed the loss of a 20 amino acid signal peptide previously documented in the literature in a large-scale study using gel electrophoresis followed by Edman sequencing.[17] The peak at *m/z* 1054.88 (12+; $MW_{obs}$ = 12647.50 Da), isolated at DF 270 Td—CF 2.2 Td, was identified as RaiA, a well-characterised ribosome-associated inhibitor which increases the accuracy of translation under stress conditions, such as low temperature and high growth density in stationary phase.[18,19] Its detection here is consistent with the high density of colonies used for LESA sampling.

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

In the case of *P. aeruginosa*, the coupling of FAIMS to LESA MS was spurred by the generation of multiple chimeric tandem mass spectra in earlier work.[11] It was hypothesised that FAIMS could

separate the overlapping precursor peaks and yield clean tandem mass spectra for subsequent protein identification.

As for *E. coli K-12*, a one-dimensional FAIMS (DF 270 Td, CF 0.0—4.0 Td) experiment was performed to elucidate the pattern of transmission of proteins of interest. Representative mass spectra from the analysis are displayed in **Supporting Figure S-2**. No protein peaks were observed between CF = 0.0 and 1.0 Td. At a CF of 1.0 Td and above, a clear envelope of five charge states (13+ to 17+; *m/z* 1146.93, 1065.01, 994.07, 932.00 and 877.36, respectively) corresponding to L11, a large ribosomal protein ($MW_{obs}$ = 14898.00 Da; $MW_{calc}$ = 14897.81 Da), was observed (see below for details on identification).

Increasing the compensation field led to the gradual disappearance of the L11 charge state distribution and the subsequent emergence of a distribution of peaks (*m/z* 856.86 10+, 951.95 9+, 1070.82 8+, 1223.51 7+ and 1427.43 6+) corresponding to PA4739 ($MW_{obs}$ = 8557.50 Da, $MW_{calc}$ = 8557.51 Da), a protein previously characterised by direct LESA mass spectrometry,[11] at CF values between 1.2 and 3.2 Td. At the highest compensation fields, a large population of smaller proteins was observed, among which PA2146 (*m/z* 606.08 9+, 681.71 8+, 778.96 7+, 908.62 6+, 1090.1367 5+; $MW_{obs}$ = 5445.65 Da) was the most abundant. The total transmission of proteins was highest at approximately CF = 3.0-3.2 Td, offering good coverage of medium to small proteins, including putative ribosomal proteins.

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

Tandem mass spectra of PA4739 acquired during the previous study[11] (in the absence of FAIMS) were found to contain fragments larger than the precursor ion in addition to the fragments corresponding to PA4739, indicating the presence of an additional precursor in the isolation window. The precursor of PA4739 isolated without FAIMS at *m/z* 1070.82 (8+) is shown in **Figure 2a**. At DF =

270 Td-CF = 2.5 Td, the co-isolating 14+ precursor ion was revealed at *m/z* 1071.35, $MW_{obs}$ = 14984.81 Da (**Figure 2b**). Its fragmentation and subsequent analysis by ProSightPC identified it as fimbrial protein (pilin), a structural protein involved in the formation of the *Pseudomonas* pilus and thus contributing to pathogenicity (**Figure 2c**).[20] Fragments of masses consistent with the cleavage of an N-terminal propeptide (six amino acids) and methylation of the N-terminal phenylalanine were detected and assigned by ProSightPC; the experimentally derived intact mass was, however, significantly different to the theoretical value ($MW_{obs}$ = 14983.79 Da, $MW_{calc}$ = 14870.71 Da; Δ111.07 Da). *b* and *y* ions of expected masses covering the entire sequence were identified by supplementary manual verification, suggesting that the mass difference may have been due to a labile post-translational modification which was lost upon fragmentation. No reports of such putative modifications were found in the literature. Whilst some of the high molecular weight fragments observed in the original tandem mass spectrum (without FAIMS) of PA4739 (*m/z* 1127.42 12+, 1163.69 11+) could subsequently be assigned to pilin ($y_{131}^{12+}$ and $y_{123}^{11+}$, respectively), the fragment detected at *m/z* 1172.50, 11+ could not be assigned to pilin and remained unidentified. This observation suggests that a third, unidentified protein may have been present in this *m/z* region.

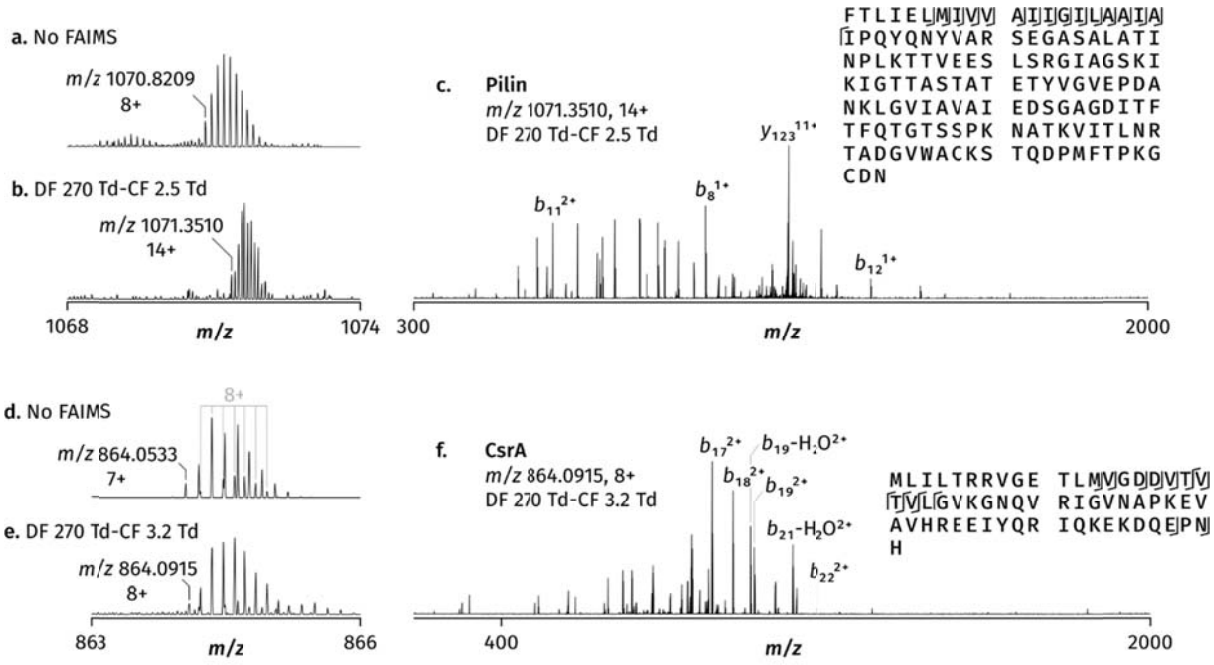


**Figure 2**. Isolation and identification of pilin and CsrA in *P. aeruginosa*. **a**. *m/z* 1071, no FAIMS applied; only the precursor of PA4739 (8+) is detected. **b**. DF 270 Td-CF 2.5 Td. The 14+ precursor (pilin) is isolated. The mass spectrum acquired without FAIMS is outlined in grey for comparison. **c**. Tandem mass spectrum of pilin (14+) isolated by application of FAIMS. **d**. *m/z* 864, no FAIMS applied; both precursors are detected simultaneously. **e**. DF 270 Td-CF 3.2 Td. The 8+ precursor (CsrA) is isolated. **f**. Tandem mass spectrum of CsrA (8+) isolated by application of FAIMS.

The second instance involved PA4739 in the 11+ charge state and PA2146 detected in the 7+ charge state at *m/z* 778 (**Supporting Figure S-3**). Without FAIMS, both charge states were detected simultaneously; the isotopic distribution of the 11+ charge state was entirely encompassed by the isotopic distribution of the 7+ charge state. The 11+ peak (*m/z* 778.9604; $MW_{obs}$ = 8557.48 Da) was isolated at DF = 270 Td-CF = 3.2 Td and confirmed by intact mass and fragmentation pattern as PA4739. The 7+ peak (*m/z* 779.9576; $MW_{obs}$ = 5445.65 Da), corresponding to PA2146, was isolated at DF = 270 Td-CF = 3.5 Td. No fragments originating from the PA2146 precursor were detected in the tandem mass spectrum of PA4739 acquired from the 11+ charge state using FAIMS, indicating complete separation of the two precursors. Whilst neither protein was identified here for the first time, isolation of these two specific charge states was not previously possible.

The third instance involved two proteins detected at *m/z* 864. Without FAIMS, a clear peak corresponding to the 7+ charge state of the 50S ribosomal protein L33 (*m/z* 864.06, $MW_{obs}$ = 6041.35

Da was observed (**Figure 2d**). (This protein was also detected and identified without FAIMS following fragmentation of the 9+ charge state at *m/z* 672.27[11]) A low abundance precursor of charge state 8+ (*m/z* 864.09) could, however, be discerned underneath the 7+ distribution of L33; this peak was favoured at CF = 3.2 Td and could thus be isolated (**Figure 2e**) and identified as CsrA ($MW_{obs}$ = 6904.69 Da), a translational regulator involved in the regulation of carbon storage, amongst numerous other functions (**Figure 2f**). CsrA was detected in multiple chimeric tandem mass spectra acquired previously without FAIMS, as identified previously by ProSightPC 3.0 at a low confidence level. Here, it was detected for the first time with sufficient clarity to be unambiguously identified (ProSightPC expectation score of $5.3 \times 10^{-20}$, indicating high confidence; 22% sequence coverage).

In addition to proteins identified in chimeric mass spectra, further proteins previously not identified by LESA alone were identified here in static field mode. L11, the ribosomal protein mentioned above, was isolated at DF = 270 Td-CF = 1.0 Td (*m/z* 994.07, 15+; $MW_{obs}$ = 14898.00 Da). This protein was detected at a very low abundance without FAIMS and was not previously identified by LESA as its abundance was too low to permit isolation. The application of FAIMS yielded a clear envelope of five charge states (13+ to 17+), of which 15+ was the most abundant. Fragmentation required the use of HCD (NCE 30%) rather than CID due to the extremely poor cleavage efficiency obtained by CID. In addition to pilin, isolated at DF = 270 Td-CF = 2.5 Td, another peak (*m/z* 953.92 7+; $MW_{obs}$ = 6670.37 Da) was isolated at the same static field settings and identified as the ribosomal constituent L32. The observed intact mass was 14.01 Da higher than the theoretical mass and within 1.2 ppm of a theoretical mass consistent with the presence of methylation; fragment assignment confirmed the presence of monomethylation at the N-terminus of the protein, not present in the database.

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

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

When FAIMS was not applied, δ-haemolysin was consistently observed in two charge states, 4+ (*m/z* 759.67) and 3+ (*m/z* 1012.56); in several mass spectra, the 2+ charge state at *m/z* 1518.33 was also observed at low abundance. The 4+ charge state at *m/z* 759.67 was the most abundant and usually constituted the base peak of the mass spectrum. At DF 270 Td and CF ranging from 0.0 to 4.0 Td, all three charge states could be observed. Representative mass spectra from the one-dimensional FAIMS experiment are shown in **Figure 3**. A single mass spectrum acquired in the absence of FAIMS is shown in **Figure 3a**. The first peak corresponding to δ-haemolysin (*m/z* 1012.56 3+; $MW_{obs}$ = 3034.64 Da) was detected at CF = 0.5 Td. The peak remained detectable until CF = 4.0 Td and constituted the base peak of the mass spectrum from CF = 2.0 Td to 2.7 Td (**Figure 3b**); it was subsequently overtaken in abundance by the 4+ charge state (*m/z* 759.67; **Figure 3d**). The only exception occurred in one scan at CF = 2.7 Td where the related peptide CAG44330 exceeded the abundance of both charge states of δ-haemolysin (*m/z* 969.03 4+ and 1291.71 3+, $MW_{obs}$ = 3872.10 Da; **Figure 3c**). Peptides previously unseen by LESA alone (*m/z* 1142.65 4+, 1192.18 4+; $MW_{obs}$ = 4566.58 and 4764.71 Da, respectively), most likely belonging to the phenol soluble modulin β class of secreted toxins as inferred from their intact masses, were also observed in one mass spectrum, as were nine peaks corresponding to six proteins of higher molecular weight (*m/z* 1070.95 6+, $MW_{obs}$ = 6014.63 Da; *m/z* 1091.24 10+, 1212.16 9+ and 1363.93 8+, $MW_{obs}$ = 10899.38 Da; *m/z* 1114.06 10+, $MW_{obs}$ = 11130.56 Da; *m/z* 1164.47 7+, $MW_{obs}$ = 8145.26 Da; *m/z* 1198.59 11+ and 1318.75 10+, $MW_{obs}$ = 13172.43 Da; *m/z* 1233.74 7+, $MW_{obs}$ = 8624.09 Da); none, however, could be isolated and fragmented before their signal disappeared and the detection of the peaks in subsequently acquired mass spectra could not be reproduced. Thus, whilst the strong signal of δ-haemolysin could, to some extent, be reduced under the right FAIMS conditions, the original goal of its complete elimination was not achieved and the identification of previously undetected proteins was not accomplished. These results suggest that the efficiency of cell lysis, affected by the structure of the cell wall, may still be the primary factor limiting the number of proteins observed in this case despite prior

optimisation of sampling parameters for Gram-positive species.[11] Future work will address these challenges to improve analyte extraction from a wider range of clinically relevant, robust species.

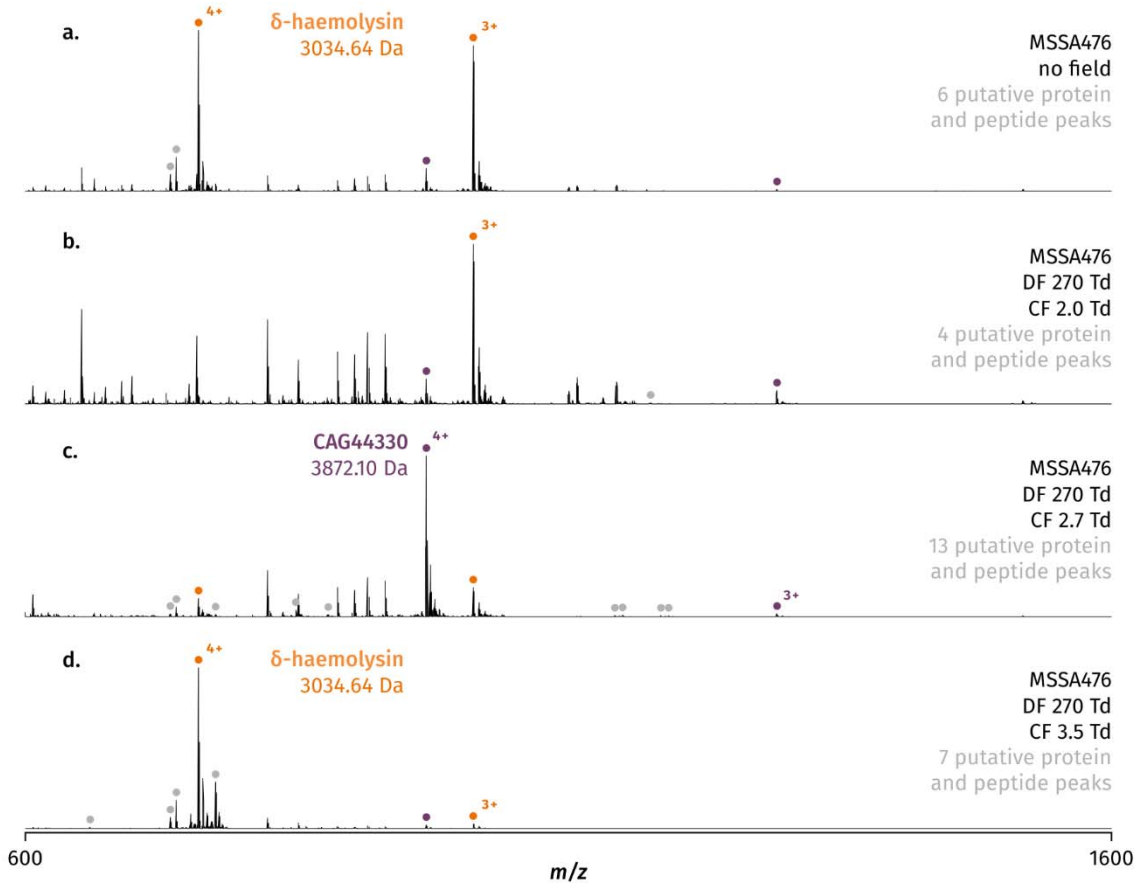


**Figure 3**. Representative LESA-FAIMS-MS single scan spectra of *S. aureus* MSSA476 acquired during a one-dimensional FAIMS experiment at DF 270 Td. The compensation field (CF) ramp ranged from CF 0.0 Td to CF 4.0 Td over 240 seconds. Putative protein and peptide peaks (charge states 3+ and above)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

## Conclusions

The coupling of FAIMS to the LESA MS workflow was shown here to expand the range of proteins identifiable in the bacterial species under investigation. Highly abundant background peaks in the *m/z* 600-800 region of *E. coli* mass spectra were successfully removed by use of FAIMS, revealing additional protein peaks. Three examples of chimeric mass spectra were resolved in *P. aeruginosa*, leading to the identification of proteins which could not otherwise be identified using LESA alone. Through isolation of individual protein peaks in the low *m/z* range without background interference,

FAIMS could potentially be used to target previously inaccessible, high charge states of proteins of interest, thus yielding improved sequence coverage.

The removal of δ-haemolysin from mass spectra of *S. aureus* was not successful. The wide band of transmission for δ-haemolysin (CF 0.5 Td to CF 4.0 Td), even at a high dispersion field (270 Td), precluded its complete elimination from mass spectra of this species, although a reduction of its abundance was achieved under certain FAIMS conditions. Whilst no new proteins were identified by application of FAIMS to *S. aureus* in this study, a number of previously undetected protein and peptide peaks were nevertheless observed.

Potential utility of the workflow in a clinical setting relies on the increased number and signal-to-noise ratio of putatively diagnostic protein peaks; in particular, the large selection of small (up to 20 kDa) proteins transmitted at DF = 270 Td-CF = ~3.2 Td (for *E. coli* and *P. aeruginosa* shown here) is analogous to the range of proteins detected by MALDI-TOF-MS, albeit with greater MS/MS capabilities suited for the in-depth analysis of important biomarkers (such as those expressed by resistant sub-strains). A FAIMS device operated in static mode at these or equivalent settings (depending on the type and manufacturer of the device) could thus constitute an excellent ion filter, focusing the proteomic characterisation of clinical isolates on the most informative subset of analytes without extending the analysis time. It would thus be beneficial to investigate the performance of LESA-FAIMS-MS in such context in the future, also taking into account novel developments in FAIMS technology.

## Acknowledgements

H.J.C. is an EPSRC Established Career Fellow (EP/L023490/1). K.I.K. is in receipt of an EPSRC studentship in collaboration with the National Physical Laboratory. RCM was supported by project MitoFun, funded by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement No. 614562 and by a Wolfson Research Merit Award from the Royal Society. The Advion TriVersa NanoMate and Thermo Fisher Orbitrap Elite

mass spectrometer used in this research were funded through Birmingham Science City Translational Medicine, Experimental Medicine Network of Excellence Project with support from Advantage West Midlands. Supplementary data supporting this research are openly available from the University of Birmingham data archive at DOI: 10.25500/edata.bham.00000311.

## Supporting Information.

Additional figures depicting the results of one-dimensional FAIMS experiments on *E. coli* (**Figure S-1**) and *P. aeruginosa* (**Figure S-2**), as well as the separation of PA4739 and PA2146 in *P. aeruginosa* at *m/z* 778 by use of FAIMS (**Figure S-3**). Fragment assignments for proteins listed in **Table 1** supplied in **Tables S-1** to **S-13**.

## References


(1) Kertesz, V.; Van Berkel, G. J. *Journal of Mass Spectrometry* 2010, *45*, 252-260.
(2) Griffiths, R. L.; Dexter, A.; Creese, A. J.; Cooper, H. J. *Analyst* 2015, *14*, 6879-6885.
(3) Edwards, R. L.; Creese, A. J.; Baumert, M.; Griffiths, P.; Bunch, J.; Cooper, H. J. *Analytical Chemistry* 2011, *83*, 2265-2270.
(4) Edwards, R. L.; Griffiths, P.; Bunch, J.; Cooper, H. J. *Journal of the American Society for Mass Spectrometry* 2012, *23*, 1921-1930.
(5) Griffiths, R. L.; Creese, A. J.; Race, A. M.; Bunch, J.; Cooper, H. J. *Anal. Chem.* 2016, *88*, 6758-6766.
(6) Griffiths, R. L.; Sisley, E. K.; Lopez-Clavijo, A. F.; Simmonds, A. L.; Styles, I. B.; Cooper, H. J. *International Journal of Mass Spectrometry* 2017.
(7) Sarsby, J.; Griffiths, R. L.; Race, A. M.; Bunch, J.; Randall, E. C.; Creese, A. J.; Cooper, H. J. *Analytical Chemistry* 2015, *87*, 6794-6800.
(8) Quanico, J.; Franck, J.; Dauly, C.; Strupat, K.; Dupuy, J.; Day, R.; Salzet, M.; Fournier, I.; Wisztorski, M. *Journal of Proteomics* 2013, *79*, 200-218.
(9) Griffiths, R. L.; Simmonds, A. L.; Swales, J. G.; Goodwin, R. J. A.; Cooper, H. J. *Anal. Chem.* 2018.
(10) Randall, E. C.; Bunch, J.; Cooper, H. J. *Analytical Chemistry* 2014, *86*, 10504-10510.
(11) Kocurek, K. I.; Stones, L.; Bunch, J.; May, R. C.; Cooper, H. J. *Journal of the American Society for Mass Spectrometry* 2017, *28*, 2066-2077.
(12) Buryakov, I. A.; Krylov, E. V.; Nazarov, E. G.; Rasulev, U. K. *Int. J. Mass Spectrom. Ion Process.* 1993, *128*, 143-148.
(13) Shvartsburg, A. A.; Smith, R. D.; Wilks, A.; Koehl, A.; Ruiz-Alonso, D.; Boyle, B. *Analytical Chemistry* 2009, *81*, 6489-6495.
(14) Shvartsburg, A. A.; Tang, K. Q.; Smith, R. D.; Holden, M.; Rush, M.; Thompson, A.; Toutoungi, D. *Analytical Chemistry* 2009, *81*, 8048-8053.
(15) Shvartsburg, A. A.; Tang, K. Q.; Smith, R. D. *Journal of the American Society for Mass Spectrometry* 2004, *15*, 1487-1498.
(16) Wittmannliebold, B.; Pannenbecker, R. *FEBS Lett.* 1976, *68*, 115-118.
(17) Link, A. J.; Robison, K.; Church, G. M. *Electrophoresis* 1997, *18*, 1259-1313.
(18) Agafonov, D. E.; Spirin, A. S. *Biochem. Biophys. Res. Commun.* 2004, *320*, 354-358.
(19) Vila-Sanjurjo, A.; Schuwirth, B. S.; Hau, C. W.; Cate, J. H. D. *Nat. Struct. Mol. Biol.* 2004, *11*, 1054-1059.
(20) Pasloske, B. L.; Finlay, B. B.; Paranchych, W. *FEBS Lett.* 1985, *183*, 413-416.

## For TOC only