The handle [http://hdl.handle.net/1887/66032](http://hdl.handle.net/1887/66032) holds various files of this Leiden University dissertation.

**Author:** Fleurbaaij, F.
**Title:** Novel applications of mass spectrometry-based proteomics in clinical microbiology
**Issue Date:** 2018-09-27
CHAPTER 4

Typing *Pseudomonas aeruginosa* isolates with ultrahigh resolution MALDI–FTICR mass spectrometry

Frank Fleurbaaij\(^1\), Margriet E.M. Kraakman\(^1\), Eric C. J. Claas\(^1\), Cornelis W. Knetsch\(^1\), Hans C. van Leeuwen\(^1\), Yuri E.M. van der Burgt\(^2\), Karin Ellen Veldkamp\(^1\), Margreet C. Vos\(^3\), Wil Goessens\(^3\), Bart J. Mertens\(^4\), Ed J. Kuijper\(^1\), Paul J. Hensbergen\(^3\), Simone Nicolardi\(^2\)

*Analytical Chemistry, 2016, 88 (11), pp 5996–6003*

\(^1\)Department of Medical Microbiology, Section Experimental Bacteriology, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands
\(^2\)Center for Proteomics and Metabolomics, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands
\(^3\)Department of Medical Microbiology and Infectious Disease, Erasmus MC, 3015 CN Rotterdam, The Netherlands
\(^4\)Department of Medical Statistics, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands
Abstract

The introduction of standardized matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) platforms in the medical microbiological practice has revolutionized the way microbial species identification is performed on a daily basis. To a large extent this is due to the ease of operation. Acquired spectra are compared to profiles obtained from cultured colonies present in a reference spectra database. It is fast, reliable and costs are low compared to previous diagnostic approaches. However, the low resolution and dynamic range of the MALDI-TOF profiles have shown limited applicability for the discrimination of different bacterial strains, as achieved with typing based on genetic markers. This is pivotal in cases where certain strains are associated with e.g. virulence or antibiotic resistance. Ultrahigh resolution MALDI-FTICR MS allows the measurement of small proteins at isotopic resolution and can be used to analyze complex mixtures with increased dynamic range and higher precision than MALDI-TOF MS, while still generating results in a similar timeframe.

Here we propose to use ultrahigh resolution 15T MALDI-FTICR MS to discriminate clinically relevant bacterial strains. We used a collection of well characterized *Pseudomonas aeruginosa* strains, featuring distinct antibiotic resistance profiles, and isolates obtained during hospital outbreaks. Following cluster analysis based on amplification fragment length polymorphism (AFLP), these strains were grouped into three different clusters. The same clusters were obtained using protein profiles generated by MALDI-FTICR MS. Subsequent intact protein analysis by ESI-CID-FTICR MS was applied to identify protein isoforms that contribute to the separation of the different clusters, illustrating the additional advantage of this analytical platform.
Introduction

The introduction and the development of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) strategies as a profiling tool for the analysis of bacterial peptides and proteins has revolutionized applications in clinical microbiology1,2. With the implementation of MALDI-TOF MS systems for in-vitro diagnostic applications, this technology has been introduced worldwide as a routine identification tool for characterizing bacteria and yeasts in clinical microbiology laboratories. The accurate identification of unknown microorganisms is achieved by comparing MALDI-TOF MS profiles obtained from cultured colonies to a database of reference spectra, without any a priori knowledge3. MALDI-TOF MS-based methods are high-end analytical tools that provide accurate data within a few min measurement time. These advantages, compared to conventional analysis4-6 are the main drivers behind the success of MALDI-TOF MS and the establishment of the technique as part of controlling infectious diseases and outbreaks in hospitals7-10.

One of the strongest points of the approach is the straightforward sample preparation and data acquisition protocol which does not require well-trained and experienced operators as required in other methods for identification of bacteria11. Bacterial colonies, cultured on solid media, are directly deposited onto a MALDI target plate and covered with a MALDI matrix solution (e.g. an α-cyano-4-hydroxycinnamic acid solution). To improve the quality of the spectra, a protein extraction procedure (e.g. using a formic acid solution) can be performed either before spotting or directly on the MALDI target plate. When bacteria are cultured on a support other than solid media, such as blood culture bottles, a more elaborate sample preparation is required to extract the bacterial proteins and generate the MALDI-TOF MS spectra12,13. Currently, MALDI-TOF MS is used to identify a variety of microorganisms from different sources including skin, faeces and body fluids14.

In general, only high-abundant proteins such as ribosomal proteins are detected using direct MALDI-TOF MS profiling of bacterial extracts15. This usually limits the identification of microorganisms to a genus or species level. However, bacterial species can be genetically extremely heterogeneous and the typing of bacterial strains with specific phenotypic characteristics (e.g. associated with pathogenicity, antibiotic resistance and/or virulence) is, in many cases, equally important16. In case of an outbreak where a multidrug-resistant bacterium can quickly spread in the hospital, a prompt identification and localization of the source is required to stop transmission, and efficiently treat the infection. A few examples have shown that MALDI-TOF MS can also be used for typing bacterial strains17-20 but, in general, this application is strongly limited by the low dynamic range and low resolution of MALDI-TOF MS instruments.
Genotyping is nowadays the most common strategy to discriminate bacterial strains. For this purpose different methods are available, each with its pros and cons\(^2\). In our medical center, amplified fragment length polymorphism (AFLP) is routinely applied on bacterial isolates when an outbreak with a specific strain is suspected\(^23\). Moreover, whole-genome sequencing provides complete information on genetic diversity of bacterial strains. This method is expensive however, and although continuous developments in technology keep reducing time of analysis, it is still time-consuming\(^24\).

In 2003, Jones and co-workers pioneered in using MALDI-Fourier transform ion cyclotron resonance (FTICR) MS for analysis of bacterial proteins up to 10 kDa, resulting in higher mass accuracy and resolution but lower dynamic range than MALDI-TOF MS\(^2\). Since then, with further development of FTICR systems and availability of higher magnetic fields, the quality of protein profiles generated using MALDI-FTICR MS has drastically improved\(^2\). Recently we have shown full isotopic resolution of proteins up to about 17 kDa that were measured on a MALDI-FTICR MS system equipped with a 15T magnet\(^2\). Likewise, complex protein mixtures (e.g. human serum proteins) have been analyzed in more detail with a dynamic range, and mass measurement precision and accuracy that are superior to those obtained with MALDI-TOF MS\(^2\)

The aim of this study is to use ultrahigh resolution MALDI-FTICR MS to type bacterial strains by profiling bacterial proteins in more detail than MALDI-TOF MS. To this end, different clinical isolates of *Pseudomonas aeruginosa*, including multidrug-resistant types, associated with hospital outbreaks\(^2\) were analyzed using 15T MALDI-FTICR MS. Cluster analysis of the MALDI-FTICR MS data resulted in a similar classification as observed with AFLP. Intact protein analysis by collision-induced dissociation (CID) was subsequently performed by direct infusion ESI-FTICR MS/MS to identify proteoforms that differ between clusters, signifying the additional power of the analysis platform.

**Materials and Methods**

**Bacterial isolates**

*Pseudomonas aeruginosa* strains were obtained from the Department of Medical Microbiology at the Leiden University Medical Center (LUMC, Leiden, the Netherlands) and the Department of Medical Microbiology & Infectious Diseases at the Erasmus Medical Center (Rotterdam, the Netherlands) (see Table S1 in Supporting Information). Two of these bacterial strains (i.e. KML 01 and KML 04) were previously identified during the outbreak described by Knoester and co-workers\(^2\).
Antimicrobial susceptibility testing was performed using the VITEK²® system (BioMérieux, Marcy-l’Étoile, France). Susceptibility results were interpreted using EUCAST criteria (http://www.eucast.org/clinical_breakpoints/).

Table 1. P. Aeruginosa strains used for this study

<table>
<thead>
<tr>
<th>P. Aeruginosa Strain Nr.</th>
<th>Place of Collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>KML 01*</td>
<td>LUMC</td>
</tr>
<tr>
<td>KML 02</td>
<td>LUMC</td>
</tr>
<tr>
<td>KML 03</td>
<td>LUMC</td>
</tr>
<tr>
<td>KML 04*</td>
<td>LUMC</td>
</tr>
<tr>
<td>KML 05</td>
<td>LUMC</td>
</tr>
<tr>
<td>KML 06</td>
<td>LUMC</td>
</tr>
<tr>
<td>KML 07</td>
<td>LUMC</td>
</tr>
<tr>
<td>KML 08</td>
<td>LUMC</td>
</tr>
<tr>
<td>KML 09**</td>
<td>LUMC</td>
</tr>
<tr>
<td>KML 10</td>
<td>LUMC</td>
</tr>
<tr>
<td>KML 11</td>
<td>Erasmus MC</td>
</tr>
<tr>
<td>KML 12</td>
<td>Erasmus MC</td>
</tr>
<tr>
<td>KML 13</td>
<td>Erasmus MC</td>
</tr>
<tr>
<td>KML 14</td>
<td>Erasmus MC</td>
</tr>
<tr>
<td>KML 15</td>
<td>Erasmus MC</td>
</tr>
<tr>
<td>KML 16</td>
<td>Erasmus MC</td>
</tr>
<tr>
<td>KML 17</td>
<td>Erasmus MC</td>
</tr>
<tr>
<td>KML 18</td>
<td>Erasmus MC</td>
</tr>
</tbody>
</table>

* Bacterial strains reported by Knoester et al. (29)
** No AFLP data available

AFLP analysis

Nucleic acids of *P. aeruginosa* were extracted using the Qiasymphony SP robot with the Total Nucleic Acid kit (Qiagen, Hilden, Germany). For AFLP typing, the DNA was digested using EcoRI and MseI restriction enzymes and subsequently the corresponding adaptors were ligated²³. After amplification using labeled primers, the fluorescent amplified fragments were separated by the genetic analyzer ABI PRISM ® 3500xL (Applied Biosystems, Foster City, U.S.A.). AFLP patterns were analyzed using BioNumerics software, version 7.1 (Applied Maths, St.-Martens-Latem, Belgium). Similarity between normalized AFLP patterns (range 60 to 600 bp) was calculated with the Pearson product-moment correlation coefficient and UPGMA algorithms.
**Bacterial protein extraction and MALDI spotting**

*P. aeruginosa* clinical isolates and a reference *P. aeruginosa* strain (KML 10) were incubated on sheep blood agar (BioMérieux, Marcy-l’Étoile, France) overnight at 37 °C. Cultured colonies were sampled with an inoculation loop, transferred in an Eppendorf tube and suspended in 150 µL deionized water. Then, 450 µL of ethanol was added, vortexed and the suspension was centrifuged at 13,000 x g in a benchtop centrifuge for two min. The supernatant was removed and the pellet reconstituted in 25 µL of 70% formic acid. Subsequently, 25 µL of acetonitrile was added and mixed using a pipette. The suspension was then centrifuged again for 2 min. One µL of the supernatant was spotted and mixed with 1 µL of α-cyano-4-hydroxycinnamic acid (HCCA) MALDI-matrix solution (saturated solution in 50% acetonitrile, 5% trifluoracetic acid) directly onto either a ground-steel MALDI target plate for MALDI-FTICR MS analysis or a polished steel MALDI target plate for MALDI-TOF MS Biotyper analysis.

The bacterial cultures were stored at 4°C and bacterial protein extraction was repeated from new colonies after 24 hours and 48 hours.

For LC-MS experiments, the bacterial proteins from the same *P. aeruginosa* strain were extracted in triplicate (i.e. from three different colonies). The bacterial protein extracts from these triplicates were mixed and concentrated to near dryness using a speed-vacuum system. Then, 25 µL of LC mobile phase A (see below) was added, vortexed and centrifuged for 5 min at 16,220 rpm (Mod. 5427 R; Eppendorf, Nijmegen, the Netherlands). Twenty µL of the supernatant was then transferred into a glass vial and 4 µL were injected into the LC system.

**MALDI–TOF mass spectrometry (Biotyper)**

All isolates were analyzed using a MALDI Biotyper system (BrukerDaltonics, Germany) and the Biotyper software 3.0. This software is used to compare a MALDI Biotyper spectrum from an unknown bacterium with a database of reference spectra. This comparison is based on pattern recognition and leads to a score value with the following interpretation: a score between 0.000 and 1.699 corresponds to a poor (unreliable) identification; a score between 1.700 and 1.999 results in a good genus identification; a score between 2.000 and 2.299 results in a secure genus identification and good species identification; a score between 2.300 and 3.000 corresponds to an excellent species identification.

For cluster analysis, raw spectra were pre-processed using the default method for mean spectrum profile (MSP) creation available in the Biotyper 3.0 software. The reference peak list of each MSP was pairwise matched to each other, the resulting matching scores were used to generate a MSP dendrogram. The distance level of the MSP dendrogram was normalized to a maximum value of 1.000.
MALDI- and ESI–FTICR mass spectrometry

Both MALDI- and ESI-FTICR MS experiments were performed on a Bruker 15T solariX XR FTICR mass spectrometer controlled by ftmsControl software and equipped with a CombiSource and a novel ParaCell (Bruker Daltonics, Bremen, Germany). A Bruker Smartbeam-II Laser System was used for irradiation at a frequency of 500 Hz using the “medium” predefined shot pattern. Ions were generated and measured as previously reported with some modifications. Briefly, each MALDI-FTICR spectrum was generated from 200 laser shots and in the m/z-range from 3497.3 to 30,000 with 512K data points (i.e. transient length of 3.986 s). The ParaCell parameters were as follows: the DC bias RX0, TX180, RX180, and TX0 were 5.13, 5.20, 5.27, 5.20 V respectively; the trapping potentials were set at 4.5 V and 4.4 V; the excitation power and sweep step time 55 % and 13.5 μs. The transfer time of the ICR cell was 2.4 ms and the quadrupole mass filter was set at m/z 3,300.

Both LC-ESI-FTICR MS and direct infusion (DI) experiments were performed in the m/z-range from 306.80 to 3,000 with 1M data points. The ParaCell parameters were as follows: the DC bias RX0, TX180, RX180, and TX0 were 1.430, 1.500, 1.570, 1.500 V, respectively; the trapping potentials were set at 1.5 V and 1.5 V; the excitation power and sweep step time 22 % and 15 μs. The transfer time to the ICR cell was 1.1 ms. LC-ESI-FTICR MS experiments were performed using the quadrupole filter set at m/z 322, an accumulation time in the hexapole of 0.250 s and three averaged scans per retention time point. Direct infusion ESI-FTICR MS experiments were carried out at an infusion rate of 2 μL/min. Precursors ions were isolated through a quadrupole (Q), accumulated in the hexapole collision cell and fragmented by collision-induced dissociation (CID). Accumulation times and collision energies were optimized for each precursor ion. DataAnalysis Software 4.2 (Bruker Daltonics) was used for the visualization and interpretation of both MALDI- and ESI-(CID)-FTICR spectra.

Reversed-Phase Liquid Chromatography

An Ultimate3000 LC system (Thermo Scientific, Breda, The Netherlands) was coupled to the ESI source of the 15T Solarix XR FTICR MS system (Bruker Daltonics, Bremen, Germany). Four μL of the concentrated bacterial protein extract (i.e. from three replicate extractions) were injected onto a Luna RPC18 column (150 × 1 mm, 3 μm, Phenomenex, Utrecht, The Netherlands). Proteins were separated using a linear gradient starting from 0% B to 40% B in 30 min, then 100% B for 4 min and re-equilibration at 0% B for 11 min and a flow rate of 50 μL/min. The mobile phases consisted of 0.5% formic acid in 5% acetonitrile and 94.95% milliQ water for A and 0.5% formic acid in 94.95% acetonitrile and 5% milliQ water for B. Fractions were collected manually every 1 min in an Eppendorf tube. Each fraction was diluted with 50 μL of an aqueous solution of 0.05% formic and 50% acetonitrile solution prior to DI-ESI-FTICR MS experiments.
MALDI–FTICR–MS spectra processing
Each bacterial protein extract from the 18 *P. Aeruginosa* strains was spotted in duplicate and measured by MALDI-FTICR MS. All spectra were visually inspected and for each duplicate the spectrum with the highest signal intensity was selected for further processing. An isotope cluster was identified when at least three isotopic peaks were detected with a S/N higher than 10. For each detected isotope cluster in the spectra the m/z-value of the three most abundant isotopic peaks was included in a so-called reference file. For each isotope peak a certain m/z-window was also defined. Using the in-house developed Xtractor software, all the intensities of the data points within the defined m/z-windows were summed and the resulting peak intensity was assigned to the corresponding m/z-value. In this way, all isotope clusters were accurately quantified even when partial overlaps occurred. The signal intensity of the three most abundant isotopic peaks were then summed and assigned to the most intense isotopic peak of the corresponding peptide or protein. The resulting list of m/z-values and peak intensities was used for further statistical analysis.

Data analysis and statistics
Peak intensities were median-standardized and adjusted for the interquartile range prior to subsequent analysis. Hierarchical linkage clustering was used to generate dendrograms using the Euclidean distance measure and average linkage. Classification between groups was verified using the Random Forest classifier. Classifications were generated using leave-one-out cross-validation and verified using the out-of-bag classifications provided by the classifier. Importance measures generated by the Random Forest were saved for subsequent inspection and identification of important discriminatory proteins.

Results and discussion
Rapid identification and typing of bacteria is pivotal for infection management and treatment control. Our aim was to develop a new MS-based profiling method for typing of bacterial strains using ultrahigh resolution MALDI-FTICR MS which provides superior performance as compared to MALDI-TOF MS. For this purpose, we selected a set of well-defined clinical isolates of *P. aeruginosa*, an opportunistic pathogenic Gram-negative bacterium, commonly causing disease in immunocompromised patients. Moreover, this bacterium is known as a nosocomial pathogen, with hospital equipment acting as a possible reservoir, sometimes leading to an outbreak. The collection of *P. aeruginosa* isolates was initiated after an outbreak which occurred at the intensive care unit of the LUMC. The collection used in this study features two strains from this outbreak (KML 01 and KML 04) as well as additional strains collected at later time points at the LUMC and Erasmus MC. An additional reference *P. aeruginosa* strain was also included.
First of all, the isolates were analyzed using MALDI-TOF MS on the Biotyper platform. For such standard analyses, ions are measured between $m/z$ 2000-20,000 in low resolution spectra in which ions are detected as relatively broad peaks (Figure 1A and C). The number of peaks with a signal-to-noise (S/N) above 3 ranged between 71 (lowest) to 100 (highest) in the best spectrum. Moreover, as confirmed by isotopic resolution measurement by MALDI-FTICR MS, the majority of peaks in the $m/z$-range from 2,000 to 5,000 were doubly charged ions of the major species observed at higher $m/z$-values. Despite these limitations, all different strains were correctly identified as *P. aeruginosa* with high identification scores ranging from 2.231 to 2.477. When different reference strains are present in the database, the MALDI-TOF MS spectrum of an identified bacterium is matched against these reference spectra and an identification score is given. For example, Oumeraci and co-workers mapped 5547 *P. aeruginosa* isolates to five individual Biotyper database entries. In our study, the *P. aeruginosa* strains were matched against the following reference database entries (i.e. different strains): ATCC 27853 THL, DSM 50071 QC DSM, DSM 1117 DSM, DSM 50071T HAM, 19955_1 CHB, LMG 8029 LMG, 8147_2 CHB, A07_08_Pudu FLR and DSM 1128 DSM. For each analyzed *P. aeruginosa* strain an identification score higher than 2.000 was obtained for more than one reference strain. For example, a good match to all nine *P. aeruginosa* entries was obtained for the strain KML 05 with identification scores ranging from 2.043 to 2.352. The ranking of these identifications was not consistent between duplicate measurements demonstrating that an unambiguous identification of the most representative reference strain was not possible using MALDI-TOF MS analysis.

The bacterial protein extracts from the same *P. aeruginosa* strains were subsequently measured by ultrahigh resolution MALDI-FTICR MS. Peptides and proteins were detected in the $m/z$-range from 3,497 to about 15,000. The high resolving power is reflected by the different protein signals being isotopically resolved in MALDI-FTICR MS measurements whereas these are indistinguishable in MALDI-TOF MS analysis (Figure 1B and D). Moreover, the S/N is enhanced in the FTICR MS spectra. When the number of distinguishable features between MALDI-TOF MS and MALDI-FTICR MS analysis of bacterial extracts are compared (Figure 1C and D), an improved data quality is noticed with MALDI-FTICR. In total, 279 unique isotope clusters were observed in the 18 MALDI-FTICR MS spectra obtained from the different *P. aeruginosa* strains. These include doubly charged ions which were also observed, but to a higher extent, in the corresponding MALDI-TOF MS spectra. A database search for bacterial species identification using MALDI-FTICR MS spectra was not possible since these spectra are not compatible with the MALDI Biotyper software. Thus, the spectra were further processed (see Materials and Methods) to perform clustering analysis.
Figure 1: MALDI-TOF and MALDI-FTICR bacterial protein profiles of *P. aeruginosa*.

A) Typical low resolution MALDI-TOF spectrum obtained in the clinical laboratory using a Biotyper system; B) Ultrahigh resolution MALDI-FTICR spectrum acquired using a 15T solariX XR FTICR MS system. Insets: enlargements illustrating the difference in resolving power between the spectra generated by MALDI-TOF and MALDI-FTICR MS. C and D: Enlargement of overlaid spectra of multiple isolates, illustrating the higher dynamic range and resolving power provided by MALDI-FTICR MS.

We tested the possibility of using MALDI-TOF and MALDI-FTICR MS data for differentiation of the different *P. aeruginosa* isolates. To benchmark the genetic diversity, the collection was first characterized using AFLP (Figure 2A). Three different clusters appeared upon using the AFLP method. It has been shown previously that isolates in cluster 1 are resistant to ciprofloxacin, while strains in cluster 2 are susceptible. The third cluster has previously not been identified. The non-outbreak related reference strain (KML 10) was positioned most closely to cluster 1.
Figure 2: Side-by-side comparison of clustering dendrograms based on AFLP, MALDI-TOF Biotyper and MALDI-FTICR MS profiling of P. aeruginosa clinical isolates.

A) Using AFLP genotyping, three distinct clusters were identified. B) Partial and inaccurate clustering was obtained from MALDI-TOF MS data. C) Cluster analysis based on MALDI-FTICR MS analysis resulted in the same three groups identified using AFLP data. * No AFLP data available. KML 10 is a reference P. aeruginosa strain.

Cluster analysis based on the MALDI-TOF MS spectra revealed some clustering of strains belonging to cluster 1 but overall no clear separation as achieved with AFLP was obtained (Figure 2B). Importantly, the improved quality of bacterial protein profiles obtained using ultrahigh resolution MALDI-FTICR MS, such as the higher dynamic range, allowing the use of peak intensity values to classify the different P. aeruginosa strains, revealed a clustering of the P. aeruginosa isolates into the same three groups as obtained using the AFLP data (compare Figure 2A and C). The main advantage of typing bacteria based on MALDI-FTICR MS data is that this analysis can be performed at similar speed as the identification using MALDI-TOF MS. This means that in principle the typing of a novel isolate, or the classification of a collection of clinical isolates in an outbreak scenario, can be performed using MALDI-FTICR MS the same day of the identification of bacteria using MALDI-TOF MS. Successful and reliable same day typing can be an invaluable tool for infection management. For example, in our set discrimination between cluster 1 and 2 is essential because of the difference in susceptibility towards ciprofloxacin.
Within the clusters, some differences between AFLP and MALDI-FTICR MS were observed, which probably derive from two main factors: 1) AFLP and MALDI-FTICR MS provide limited and different information on genetic variation (i.e. AFLP at the DNA level and MALDI-FTICR MS at the protein level); 2) the relative abundance of proteins as observed with MALDI-FTICR MS is affected by growing conditions. In fact, it is well known that colonies from replicate cultures or even from the same culture plate can be very heterogeneous. The effect of this factor can be seen on the similarity between replicate cultures of the strains KML 03 and KML 15. Moreover, changes in similarity were observed in replicate measurements performed on three different days (Figure S1 in Supporting Information); importantly, however, the overall clustering into three groups remained the same. Further studies are needed to evaluate whether and how the changes in protein expression between colonies or growing conditions would affect a clustering analysis based on both MALDI-FTICR MS data and a pattern recognition method (i.e. without use of peak intensities).

As mentioned above, the identification of bacteria based on MALDI-FTICR MS data was not possible because no database of MALDI-FTICR reference spectra is available. The implementation of such a database would probably allow a more accurate and reliable identification of bacterial species. However, considering the cost of such expensive technology, it is more realistic to develop the platform towards the typing of bacteria (e.g. during an outbreak) in parallel to the validated MALDI-TOF based species identification which is already implemented in the clinic.

**Top-down mass spectrometry identification of proteins important for the classification of* P. aeruginosa* strains**

The discriminative power of each protein of the different *P. aeruginosa* strains can be visualized by plotting their relative importance in the classification (Figure 3A). Proteins with a discriminative value higher than 4 were considered as major contributors. Then, a search was performed to trace these discriminating proteins in the corresponding LC-ESI-FTICR MS data. A protein observed at \( m/z \) 7354.0032 in the MALDI-FTICR MS spectra (Figure 3A, blue dot and inset), was found in the LC-ESI-FTICR MS data from strain KML 18. The LC fraction containing this protein was further analyzed by direct infusion ESI-FTICR-MS/MS aiming for top-down identification. In the ESI-FTICR MS spectrum the protein was detected at \( m/z \) 1051.4163 \([\text{M+7H}^+\] and \( m/z \) 920.1132 \([\text{M+8H}^+\] The latter ion was selected for CID fragmentation and the resulting spectra were manually annotated, which resulted in a sequence tag TATTQA(I/L)Q(I/L)(I/L)QN. Following BLAST searches in the UniProt database, the sequence tag TATTQALQLQN fully matched with the C-terminal region of an uncharacterized protein from *P. aeruginosa* (UniProt, entry A0A0M2DCU1). Moreover, the ion at \( m/z \) 433.2774, from where the sequence tag started, corresponds to the \( y_4 \) fragment containing the four C-terminal amino acids (TGKK, Figure 3B). The identified protein contains a putative signal
Typing *Pseudomonas aeruginosa* isolates with ultra-high resolution MALDI-FTICR mass spectrometry

**Figure 3**: Top-down identification of two highly discriminatory proteins between different *P. aeruginosa* isolates.

A) Plot illustrating the contribution of individual proteins on the separation of the *P. aeruginosa* clinical isolates into different clusters and the workflow for subsequent top-down protein identification. The inset shows the signal in MALDI-FTICR MS analysis of the two selected species between m/z 7320-7370. B and C: Top-down ESI-CID-FTICR MS analysis of the two different proteins highlighted in A. The insets show the precursor ion signal in ESI-FTICR MS. Based on the CID spectra a possible sequence tag was generated. BLAST searches in the UniProt database led to the identification of two different isoforms of the same hypothetical protein with theoretical m/z-values of 7353.8493 [M+H]+ and 7323.8388 [M+H]+, respectively, including a disulphide bond in both.
peptide of 19 amino acids and the mature protein therefore starts at Ala-20 (Figure S2A in Supporting Information). Taking into account a disulfide bridge between the two cysteines, the theoretical \( m/z \)-value of this protein species is 7353.8493 [M+H]+, which fits with the MALDI-FTICR and ESI-FTICR MS data. The full protein sequence allowed for the assignment of additional b- and y-ions (Figure S2A and Figure S3 in Supporting Information). The fact that no fragment ions were observed between the two cysteines hints to the presence of a disulfide bridge.

Interestingly, the initial sequence tag resulted in the identification of two isoforms of this protein only differing in position 86 (Figure S2B in Supporting Information). Either a Thr (as in the protein identified above) or an Ala is found in this position. Another discriminating protein (Figure 3A, orange dot and inset) at \( m/z \) 7323.9827 is present in the MALDI-FTICR MS data of different clinical isolates. This is 30.021 Da lower than the protein identified above and corresponds nicely to an Thr-Ala substitution (theoretical change 30.016 Da). This was confirmed with top-down MS/MS analysis of this protein species, selecting the ion at \( m/z \) 916.3668 (M+8H)\(^8\) in the corresponding chromatographic fraction from the extract of another clinical isolate. The fragmentation spectrum contained the same b-ions (e.g. b\(_5\)) but, the y-ions starting from y\(_4\) were shifted with the expected mass (Figure 3C).

Subsequently, we checked all our MALDI-FTICR MS spectra and found that the presence of the two isoforms is mutually exclusive. In fact, the Thr-86 variant, shown in blue in Figure 3, was found exclusively in the protein profiles of strains of cluster 1, whereas the Ala-86 variant was found only in the strains of clusters 2 and 3. This further emphasizes that this protein contributes to the differentiation of the clusters. No clear protein homologues of this protein with an annotated function could be found in other species. Using the I-TASSER server for protein function prediction\(^39\), some indications for a function in metal binding were retrieved. This warrants further investigation however, especially since under this scenario the cysteine residues would not be conserved. Of note, in one of the spectra (KML 08), none of the two isoforms was resolved. This strain was still clearly clustered within cluster 1, demonstrating that other features were also important for the classification.
Conclusions

The standardization of MALDI-TOF MS platforms has paved the way for routine applications of mass spectrometry in clinical microbiology laboratories. Due to intrinsic low resolution and moderate sensitivity, bacterial typing using MALDI-TOF MS is, as also shown in this paper, limited. In this study we used ultrahigh resolution MALDI-FTICR MS to analysis in more detail bacterial proteins from *P. aeruginosa* isolates. This is a novel application of ultrahigh resolution profiling, which maintains all the speed advantages of MALDI-TOF bacterial protein profiling, while being able to reliably discriminate between different strains of *P. aeruginosa*. In fact, using this new method we were able to cluster 18 *P. aeruginosa* clinical isolates in a similar manner as was obtained using AFLP analysis. The additional power of the FTICR platform was further demonstrated by the *de-novo* sequencing of two proteoforms of a hypothetical protein that contributes significantly to the discrimination of the different clusters. However, further investigation is needed to assess whether any of the analyzed bacterial proteins from cluster 1 can be linked to the antibiotic resistance.

In the future when expanding on this proof-of-principle experiment, we believe that this approach can be used for the rapid typing of clinical isolates, especially in the context of hospital outbreaks where fast and appropriate infection control management is essential.

Acknowledgements

This research was financially supported by The Netherlands organisation of scientific research (NWO, ZonMW grant number 50-51700-98-142).
References

Nicolardi, S.; van der Burgt, Y. E. M.; Dragan, I.; Hensbergen, P. J.; Deelder, A. M. J. Proteome Res. 2013, 12, 2260-2268.


Figure S1. Clustering of *P. aeruginosa* clinical isolates following analysis on different days.

Day-to-day repeatability experiment of the clustering analysis based on MALDI-FTICR MS profiles. On three consecutive days extracts were prepared and measured. Clustering in three groups was consistent in these independent experiments.
Figure S2. Signal peptide analysis and isoform comparison of the P. aeruginosa hypothetical protein identified by topdown ESI-CID-FTICR MS.

A: Predicted cleavage site of the signal peptide of the hypothetical protein identified by topdown ESI-CID-FTICR MS (Figure 3). B: Annotation of the signal peptide, the de novo obtained sequence variation at position 86 of the hypothetical protein where either a Thr or Ala is present. Sequences were obtained from the Uniprot database.
Figure S3. Top-down identification of two highly discriminatory proteins between different *P. aeruginosa* isolates.

Enlargement of ESI-CID-FTICR spectra of the discriminant proteins highlighted in blue (A) and in orange (B) in Figure 3A.
Typing Pseudomonas aeruginosa isolates with ultrahigh resolution MALDI-FTICR mass spectrometry