Utility of MALDI-TOF MS as a new tool for \textit{Streptococcus pneumoniae} serotyping

Maria Ercibengoa$^{1,2,3*}$, Marta Alonso$^{1,2}$, Diego Vicente$^{1,2,3}$, Maria Morales$^{2,4}$, Ernesto Garcia$^{2,4}$, Jose Maria Marimon$^{1,2}$

1 Hospital Universitario Donostia–Instituto de Investigación Sanitaria Biodonostia, San Sebastián, Spain, 2 CIBER de Enfermedades Respiratorias–CIBERES, Madrid, Spain, 3 Preventive Medicine and Health Public Department, University of Basque Country UPV/EHU, San Sebastián, Spain, 4 Departamento de Biotecnología Microbiana y de Plantas, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Madrid, Spain

* Current address: Department of Microbiology, Donostia University Hospital–Biodonostia Institute, Paseo Dr. Beguiristain s/n. Donostia–San Sebastian, Spain.

maria.ercibengoaaran@osakidetza.eus

Abstract

Nowadays, more than 95 different \textit{Streptococcus pneumoniae} serotypes are known, being less than one third responsible for the majority of severe pneumococcal infections. After the introduction of conjugate vaccines, a change in the epidemiology of the serotypes causing invasive pneumococcal disease has been observed making the surveillance of circulating serotypes especially relevant. Some recent studies have used matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) technology to identify the most frequent pneumococcal serotypes that cause invasive disease. The objectives of this study were to evaluate the efficacy of previously described discriminatory peaks determined by MALDI-TOF MS for the identification of serotypes 6B, 19F, 19A and 35B using reference and clinical isolates and to try to identify other discriminatory peaks for serotypes 11A, 19F and 19A using transformed pneumococcal strains. Most of the proposed peaks defined in the literature for the identification of serotypes 6B, 19F, 19A, 35B were not found in the spectra of the 10 reference isolates nor in those of the 60 clinical isolates tested corresponding to these four serotypes. The analysis and comparison of the mass spectra of genetically modified pneumococci (transformed strains) did not allow the establishment of new discriminatory peaks for serotypes 11A, 19F, and 19A. MALDI-TOF MS in the usual range of 2,000 to 20,000 $m/z$ did not prove to be a valid technique for direct \textit{S. pneumoniae} serotyping.

Introduction

\textit{Streptococcus pneumoniae} serotyping is not currently of the utmost importance for routine clinical practice unlike in the pre-antibiotic era, when it was critical for the patient’s treatment with monospecific antisera [1]. However, serotyping is of great significance from an epidemiological and preventive perspective because it makes possible to define the distribution of serotypes causing invasive pneumococcal disease (IPD) and contributes to updating the composition of pneumococcal vaccines that include the most frequent \textit{S. pneumoniae} serotypes causing IPD.
The Quellung reaction has been the “gold standard” technique for *S. pneumoniae* serotyping since 1902 [2] in spite of many inconveniences such as being tedious and expensive, including a measure of subjectivity in the interpretation of results and not allowing the analysis of batches of samples [3–4]. Despite these drawbacks, the Quellung reaction continues to be the most frequently used technique for pneumococcal serotyping [5], although other antibody-based alternative methods, such as coagglutination, latex agglutination, ELISA, dot-blot and array techniques have been implemented [6–9]. Additional practices such as counter-immunoelectrophoresis, flow cytometry and immunoblot are cumbersome due to the requirement of special equipment and expertise [10–12].

Since the first works in 2003 on PCR for detecting *S. pneumoniae* serotypes on the basis of the nucleotide sequences of capsular genes [13–14], multiplex PCR has become one of the most used typing techniques, allowing the detection of different numbers of serotypes in the same PCR reaction [15]. However, the similitude of the capsular genes among several serotypes makes impossible to differentiate them in a unique reaction so that other alternatives such as DNA sequencing or verifying the serotypes with the Quellung reaction are necessary to differentiate between them; unfortunately, these approaches demand more work and costs [16–17].

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) has been recently introduced in clinical microbiology laboratories and is considered an excellent method for the identification of most bacterial species. MALDI-TOF MS has also proved to be a reliable tool for other applications, such as antibacterial and antimycotic resistance testing, in addition to some typing applications [18–21]. However, typing applications are extremely new and their use is currently restricted to specialized laboratories [22]. Among these new typing applications, some authors have used MALDI-TOF MS to identify pneumococcal serotypes [23–24].

The main aim of this work was to assess if the discriminatory peaks previously described in the literature for serotypes 6B, 19A, 19F, and 35B [23–24] were present in the spectra of pneumococcal clinical isolates of the same serotypes, regardless of the origin of the sample or the genotype. Another aim was to analyze if the mass spectra obtained by MALDI-TOF MS showed specific discriminatory peaks using three different transformed pneumococci of serotypes 11A, 19F, 19A, and if those peaks were also present in the spectra of clinical isolates of the same serotypes.

Material and methods

Clinical isolates

All *S. pneumoniae* isolates included in this study were obtained from the frozen (–80°C) strain collection available at the Microbiology Department of Donostia University Hospital (Donostia-San Sebastián, Spain). Prior to freezing, all isolates were identified using the optochin sensitivity and bile solubility tests. Pneumococci were serotyped using the Quellung reaction with sera provided by the Statens Serum Institute (Copenhagen, Denmark); serotyping by Quellung was performed in some isolates after determining the capsular type by an in-house designed multiplex PCR [25]. Multi-locus sequence typing (MLST) was performed as previously described [26].

Isolates included in the analysis of reported serotype-defining peaks. To evaluate the specificity of the previously described discriminatory peaks [23–24] for serotypes 6B and 19A, 19F, 35B, the mass spectra of 60 clinical isolates previously serotyped 6B (n = 10), 19A (n = 28), 19F (n = 19), and 35B (n = 3), were studied. Seventeen of these isolates had the

Funding: This work was funded in part by a grant of the Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES), CB06/06/00566 and CB06/06/0003. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

PLOS ONE | https://doi.org/10.1371/journal.pone.0212022 February 12, 2019 2 / 9
same sequence type (ST) to those used for defining discriminatory peaks in other studies (Table 1).

Besides, 5 reference strains of serotype 6B (ATCC 700670, ATCC 700675 ATCC BAA-342, ATCC BAA-658, ATCC 700903), 3 of serotype 19A (ATCC 700678, ATCC 700904, ATCC 700674), and one each of serotypes 19F (ATCC 700905) and of serotype 35B (Statens Serum Institute serotype 35B) were also included [S1 Table].

**S. pneumoniae transformants used in defining serotype-specific peaks.** Three encapsulated and one unencapsulated isogenic *S. pneumoniae* transformants [27–28] were tested. The strains employed in this study were M11 (unencapsulated derivative of the common laboratory strain R6) and P181, P191 and P242 constructed by transformation of strain M11 with DNA of strains 5086 (serotype 19A), 1064 (serotype 19F) and 2963/13(serotype 11A) respectively [27–28]. Transformants selection was made by enriching encapsulated isolates by successive transfer to C+A medium (CpH8 with 0.08% of bovine seroalbumin) supplemented with 0.5 μl/ml anti-R serum that specifically agglutinates non-encapsulated pneumococci) [28]. Serotyping was performed by Quellung reaction. The above mentioned isolates of serotypes 11A 19F and 19A clinical were also studied to define serotype-specific peaks.

**Table 1. Serotypes and sequence types of the isolates used in this study.**

<table>
<thead>
<tr>
<th>Serotype</th>
<th>Sequence type</th>
<th>Number of strains used in the study</th>
</tr>
</thead>
<tbody>
<tr>
<td>6B</td>
<td>90</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>386</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>1624</td>
<td>1</td>
</tr>
<tr>
<td>11A</td>
<td>62</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>3687</td>
<td>1</td>
</tr>
<tr>
<td>19F</td>
<td>63</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>81</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>88</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>157</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>179</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>271</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>549</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1167</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2948</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>6161</td>
<td>1</td>
</tr>
<tr>
<td>19A</td>
<td>193</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>199</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>202</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>276</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>320</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>994</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1201</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2930</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3259</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>4768</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>198</td>
<td>2</td>
</tr>
<tr>
<td>35B</td>
<td>558</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>4679</td>
<td>1</td>
</tr>
</tbody>
</table>

1 Sequence type of the isolates also used in the study of Nakano et al [23]

https://doi.org/10.1371/journal.pone.0212022.t001
MALDI-TOF MS studies

Frozen pneumococcal isolates were thawed and cultured onto Trypticase-soy agar with 5% sheep blood plates and incubated at 37°C for 20–24 h in a 5% CO₂ atmosphere. Bacterial extracts for MALDI-TOF MS identification (Bruker, Daltonics, Germany) were obtained by an ethanol–formic acid extraction method to generate high-quality spectra; a 10 μl loopful of bacterial growth was suspended in a mixture of 300 μl of distilled water and 900 μl of ethanol at 96%. The bacterial suspension was centrifuged at 13,000 r.p.m. for 2 min and the supernatant was discarded. The pellet was resuspended in to 50 μl formic acid–water at 70% and left to stand 15 min; subsequently 50 μl of acetonitrile was added. After centrifugation at 13,000 r.p.m for 2 min, 1 μl of the supernatant that was the product of analysis was transferred to the MALDI target plate and overlaid with 1 μl of matrix solution.

Mass spectra acquisition was performed using a Microflex LT mass spectrometer (Bruker Daltonics, Germany) and the Flex Control software (version 3.4) with default parameter settings. Evaluation of the mass spectra was carried out using the Flex Analysis software and MALDI Biotyper 3.1 software and library (version 3.1.66, Bruker Daltonics Germany, with 7,421 entries, last updated March 2018).

Each isolate was analyzed in triplicate and the resulting average measure of each peak (m/z) was exported to a spreadsheet where a Pearson matrix correlation index was calculated using GraphPad Instat version 3.05 software (GraphPad Software Inc., La Jolla, CA, USA). Cut off intensity was established at 1,000 ua, and a peak tolerance of ±2 Da was admitted as used in other studies [30]. Mass spectra generated by MALDI-TOF were exported in the BTMSP format and visually analyzed using Excel software. Subsequently, the previously discriminatory peaks published by Nakano et al and Pinto et al were evaluated using Microsoft Excel resources”. The same procedure was repeated to try to define discriminatory peaks using the pneumococcal transformants of serotypes 11A, 19F and 19A and its parent unencapsulated isogenic M11 strain. Mass spectrum generated from each isolate was exported and visually analyzed with Microsoft Excel software and the concordance between the peaks of transformants and its parent isogenic unencapsulated strain was evaluated.

Results

In our study we only included clinical isolates of serotypes 11A, 19F and 19A, because they were the isolates corresponding to the serotype of the available transformed strains, and of serotype 6B that was included in the analysis of Pinto et al. [24] and serotype 35B that was also included in the study of Nakano et al. [23] together with serotypes 6B, 19F and 19A.

Almost none of the previously described peaks at 6641, 6920, 8424 m/z identified in > 90% serotype 6B isolates [24] were found among our clinical or reference strains. Only two serotype 6B reference isolates (ATCC BAA658, ATCC 700903) showed the peak at 6270 m/z described in 50–69% of 6B isolates. Also, most of the discriminatory peaks described by Pinto et al. [24] could not be located among clinical isolates. In addition, two serotype 6B ST90 strains showed a peak at 6270 m/z in their mass spectra, whereas the absence of this peak was highlighted as a characteristic of serotype 6B isolates [24].

Similarly, almost none of the peaks used in the initial and best models generated by Clin-ProTools of Nakano et al. [23] for serotypes 6B, 19A, 19F and 35B were found in the reference strains of the corresponding serotypes studied. Only the peak at 6620 m/z was found in the 19F reference strain. When analyzing the 60 clinical isolates serotyped with the Quellung reaction as serotypes 6B, 19F, 19A or 35B, only two serotype 6B, two serotype 19A, one serotype 19F and none of the serotype 35B isolates had any of the discriminatory peaks defined for serotyping [23]. In fact, these discriminatory peaks were not only found in the serotypes from
which they had been described for but also in the spectra of some isolates of other serotypes. Notably, none of the biomarker peaks used by Pinto et al. [24] for serotype 6B corresponded to the peaks described in the models defined by Nakano et al. [23] S1 Table.

To attempt finding new discriminatory peaks not previously described in the literature for the identification of serotypes 11A, 19F, 19A, a comparison between the spectra of the unencapsulated M11 strain and of its 3 transformants P181, P191, P242 as well as 40 clinical isolates of the same serotypes 11A (n = 5), 19F (n = 19), 19A (n = 26) was performed. No distinctive peaks could be observed in the spectra of any of the three transformants that had spectral Pearson correlation index of 0.999 between them and with the spectra of the M11 parent strain. Also, no distinctive peaks could be deduced from the spectra of the transformants and the clinical isolates of the same serotype.

Discussion

This study was focused on analyzing the feasibility of the direct serotyping of *S. pneumoniae* using MALDI-TOF MS because of the typing possibilities and the ease and current availability of the MALDI-TOF technology in many clinical microbiology laboratories. The analysis of fingerprinting could have signified a reliable methodology among species within the *Streptococcus mitis* group (SMG) species, which would make routine real-time typing possible. But, despite these advances, an improvement in the database as well as standardization in the protocols employed is needed for the correct characterization of all the species of SMG [31].

The identification of microorganisms by MS is based on the analysis of their protein content, which might create a problem of reproducibility from the theoretical point of view because bacterial protein synthesis can vary depending on the media and culture conditions [32–33]. These variables must remain constant otherwise the mass spectra could be altered [34–35]. In addition, other factors, such as the matrix chosen, the instrument’s calibration, the methodology mass range, the measurements conditions, and even the operator’s skills, can modify the quality of the peaks obtained in the experiment; hence, the importance of technique standardization to improve the reproducibility and comparison of mass spectra [36–37]. All these variables could have influenced the attainment of the expected results because, as shown in this work, clinical isolates and reference strains of serotypes 6B, 19A, 19F and 35B from our region showed peaks different to those previously defined for serotype identification [23–24]. Besides, peaks defined for serotype 6B were also different in both reports. In the work of Pinto et al. [24] a validation of the biomarker algorithm with isolates from other regions of the world was proposed but, as observed in our work, the peaks defined did not match with those obtained in the spectra of isolates from our region. However, it does not seem that the absence of discriminatory peaks for serotypes 6B, 19A, 19F and 35B could only be the consequence of a different regional clonal composition, since only two of the serotype-specific discriminatory peaks could be found in two of the 10 reference strains tested.

Previous works on pneumococcal serotyping using MALDI-TOF [23–24] inferred serotypes on the basis of the mass-spectral protein profiles as compared to different peak algorithms in a similar way that serotypes could be inferred form the MLST analysis [38]. Depending on the genetic variability of clinical isolates of a certain serotype in a geographical region, these algorithms will have a better or worse performance (sensitivity and specificity). Of note, MALDI-TOF MS has been also used for the typing of other microorganisms such as *Staphylococcus aureus* or *Enterococcus faecium* without finding specific peaks for a consistent and reliable identification of clonal complexes or sequence types [39].

To date, the molecular identification of pneumococcal serotypes has been done by analyzing the genes encoding the capsular polysaccharide. With the notable exceptions of serotypes 3 and
37, which are synthesized by the synthase pathway [40–41], the polysaccharide capsule of all pneumococcal serotypes is synthetized by what is known as the Wzx/Wzy-dependent pathway [42] and the sequences of the wzy gene encoding the polysaccharide polymerase is used for the identification of most serotypes [43]. Hypothetically, differences in the capsular polymerases or synthases could be detected by the MALDI-TOF MS technology, but as demonstrated in this work using transformed strains, specific peaks could not be established for these enzymes. One possible explanation is that the molecular mass of Wzy-type polymerases are over 50 kDa and out of the protein detection range of 2,000 to 20,000 m/z established by default in the Bruker MALDI-TOF-MS for identifying bacterial species. Perhaps, the use of other technologies such as MALDI-TOF/TOF, capable of detecting proteins of higher molecular mass, could aid in differentiating these enzymes [44]. Another approach could be the use of other mass spectra technology such as electrospray ionization ion trap MS, capable of differentiating carbohydrates, which might allow an accurate discrimination between pneumococcal polysaccharide capsules (serotypes).

This study has some limitations, especially the low number of clinical isolates of each serotype studied that were collected in a small region what could have biased the peaks obtained as a consequence of a low diversity. However, it does not seem that the absence of discriminatory peaks is a direct consequence of the low number of clinical isolates tested because the spectra of the 10 reference strains tested also showed a significant lack of discriminatory peaks.

In this work, no differences in the spectra of parent and transformed S. pneumoniae strains of three serotypes were found, so that no specific peaks could be assigned to a given serotype. Also, no specific peak could be determined by analyzing clinical S. pneumoniae isolates of different serotypes. MALDI-TOF MS methodology, in its current form, was not useful for identify Streptococcus pneumoniae serotypes currently, because of its low matching with serological and biological molecular techniques.

Supporting information
S1 Table. (TIF)

Author Contributions
Conceptualization: Marta Alonso.
Data curation: María Ercibengoa.
Investigation: María Ercibengoa, María Morales, Jose María Marimón.
Resources: Ernesto Garcia.
Supervision: Jose María Marimón.
Validation: Diego Vicente, María Morales, Jose María Marimón.
Visualization: Marta Alonso, Diego Vicente, María Morales, Ernesto Garcia.
Writing – original draft: María Ercibengoa.
Writing – review & editing: Jose María Marimón.

References


