Molecular and Cellular Probes 28 (2014) 34–40

Contents lists available at ScienceDirect

Molecular and Cellular Probes

journal homepage: www.elsevier.com/locate/ymcpr

Assessment of marker proteins identified in whole cell extracts for bacterial speciation using liquid chromatography electrospray ionization tandem mass spectrometry

Jennifer Kooken [a,1], Karen Fox [a], Alvin Fox [a], David Wunschel [b,*]

[a] Department of Pathology, Microbiology and Immunology, School of Medicine, University of South Carolina, Columbia, SC 29208, USA
[b] Chemical and Biological Signature Sciences, Pacific Northwest National Laboratory, PO Box 999 MS P7-50, Richland, WA 99354, USA

ARTICLE INFO

Article history:
Received 10 July 2013
Received in revised form
16 August 2013
Accepted 17 August 2013
Available online 29 August 2013

Keywords:
Staphylococcus
tandem mass spectrometry
bacterial identification
protein markers

ABSTRACT

Staphylococcal strains (CoNS) were speciated in this study. Digests of proteins released from whole cells were converted to tryptic peptides for analysis. Liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI MS/MS, Orbitrap) was employed for peptide analysis. Data analysis was performed employing the open-source software X!Tandem which uses sequenced genomes to generate a virtual peptide database for comparison to experimental data. The search database was modified to include the genomes of the 11 *Staphylococcus* species most commonly isolated from man. The number of total peptides matching each protein along with the number of peptides specifically matching to the homologue (or homologues) for strains of the same species were assessed. Any peptides not matching to the species examined were considered conflict peptides. The proteins typically identified with the largest percentage of sequence coverage, number of matched peptides and number of peptides corresponding to only the correct species were elongation factor Tu (EF Tu) and enolase (Enol). Additional proteins with consistently observed peptides as well as peptides matching only homologues from the same species were citrate synthase (CS) and 1-pyrroline-5-carboxylate dehydrogenase (1P5CD).

Protein markers, previously identified from gel slices, (aconitate hydratase and oxoglutarate dehydrogenase) were found to provide low confidence scores when employing whole cell digests. The methodological approach described here provides a simple yet elegant way of identification of staphylococci. However, perhaps more importantly the technology should be applicable universally for identification of any bacterial species.

1. Introduction

In bacteriology, proteomics is primarily utilized in identifying proteins expressed by an organism under specific growth conditions, not for chemotaxonomic characterization. In contrast MALDI TOF MS (matrix assisted time of flight/deionization mass spectrometry) protein profiling [1,4,11,19] has become an established technique for identification of bacteria particularly with relevance to clinical microbiology [22,23]. MALDI TOF MS is used for rapid determination of a mass pattern of proteins for determination of species identity using a database of mass profiles of previously characterized species; these proteins are generally not identified [12]. Alternatively, there have been a handful of reports identifying bacteria with potentially greater confidence utilizing sequence differences among peptides released by tryptic digestion ("tryptic peptides"). Experimental spectra of tryptic peptides (generated using liquid chromatography electrospray ionization tandem mass spectrometry [LC-ESI MS/MS] analysis) are compared to virtual spectra generated from genomic databases. Custom software has been used successfully primarily by one group [8,14,15]. However more recently standard proteomics software has been also used [6].

Microbiological testing in clinical settings is still largely based on biochemical characteristics. Many of these tests are routinely used for accurate identification of many pathogenic or opportunistic species but for less studied species the results are often less than optimal [20]. Numerous variants of the polymerase chain reaction (PCR) and/or restriction enzymes are commonly used in more advanced reference laboratories in species identification. However in developing new genetic markers two conserved

* Corresponding author. Tel.: +1 509 371 6852.
E-mail addresses: DavidWunschel@Frontier.com, David.Wunschel@pnnl.gov (D. Wunschel).
[1] Current address: Diagnos-Techs Clinical & Research Laboratory, Kent, WA 98032, USA.

0890-8508/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.mcp.2013.08.002

genetic regions (surrounding a variable region) are required for primers to provide initiation for amplification and for the variable regions to provide the information for discriminating closely related species; selection of a gene for assessing sequence variation can be somewhat arbitrary. Whole genome sequencing and annotation is still a technically demanding and expensive alternative [13], although this situation is changing rapidly. Accordingly genes that are present universally in bacteria, most commonly16S rDNA, are still widely employed. However, it has become clear over the years that 16S rDNA sequences are too conserved amongst many closely related species and only genus-level identification is provided [24].

Our group has studied members of the genera *Staphylococcus* including coagulase negative staphylococci (CoNS) and *Micrococcus* [9,10,18]. Commercial biochemical tests do not provide adequate identification of CoNS. Variation in sequences of 16S rRNA, but not sodA sequence, correlates with MALDI TOF MS [7] suggesting that genus-level identification is being achieved. *SodA* sequence is the current gold standard for identification of CoNS. Others have also found MALDI TOF MS inadequate for CoNS identification [2]. However, in a companion study to the current work excellent correlation with variation in sodA sequence and peptide marker sequences were found [18]. MALDI TOF MS or LC–MS/MS analysis of gel bands identified aconitate hydratase and oxoglutarate dehydrogenase as the dominant proteins present in these gel bands. LC–MS/MS analysis of tryptic peptides, released from whole cells, also provided speciation, correlating with sodA sequence. However, different marker proteins (including elongation factor Tu) were found in whole cell digests. The current work further explores differences in the utility of the different marker proteins.

2. Materials and methods

2.1. Strains analyzed

Staphylococcus aureus ATCC 12598, *Staphylococcus capitis* ATCC 27841, *Staphylococcus chonii* ATCC 29972, *Staphylococcus chromogenes* ATCC 43764, *Staphylococcus epidermidis* ATCC 12228, *Staphylococcus lugdunensis* ATCC 49576,CNS 1, CNS5, CNS18, CNS20, ASO2 C21, ASO15 C28, ASO15C 40Y, ASO5 C106, cow924RR, cow970RR, MUS5949, MUS5951, 09-304-034, 900-200-150, 990RR, ASO1 C8, ASO2 C53, ASO2 C63, ASO3 C6, ASO3 C59W, ASO15 C64, ASO15 C84, ASUNO2-15 C6Y and ASUNO2-15 C11Y.

2.2. Culture conditions and sample preparation

Bacteria were grown on nutrient agar plates at 37 °C for 24–48 h. The isolates were identified as staphylococci by Gram stain morphology and fermentation characteristics. Colonies were harvested from plates using 0.1 M NaCl, 50 mM Tris–HCl, 0.5 mM phenylmethylsulfonyl fluoride [PMSF] and placed in a FastPrep®-24 (MP Biomedicals, Solon, OH) for 6 ms × 30 s with 5 min on ice between each cycle for a total of six cycles. The samples were then centrifuged at 4 °C for 1 h at 10,000 *g*. The supernatants were removed and placed at −70 °C for two freeze–thaw cycles to eliminate DNA. Each strain was generally analyzed once. Reproducibility was assessed for several species by analyzing multiple strains. For any aberrant results, the analysis was repeated again.

2.3. Protein separation, tryptic digestion and MS analysis

Bacteria were harvested (after growing as confluent lawns) from plates using 0.1 M NaCl, 50 mM Tris–HCl, 0.5 mM phenylmethylsulfonyl fluoride [PMSF] and placed in a FastPrep®-24 bead beater (MP Biomedicals, Solon, OH) for 6 ms × 30 s with 5 min on ice between each cycle for a total of six cycles to release proteins. The samples were then centrifuged at 4 °C for 1 h at 10,000 *g*. The supernatants were removed and placed at −70 °C for two freeze–thaw cycles to eliminate DNA.

2.4. Peptide preparation

Bacteria samples were thawed, vortexed briefly to re-suspend. Fifty μl of supernatant was transferred to labeled low protein binding 1.5 ml microfuge tubes. Fifty μl of freshly made 8 M urea, 1 μl of β-mercapto ethanol, 24 μl of water, and 25 μl of 200 mM Tris–HCl pH 8.0 were added tubes were vortexed briefly and incubated at 60 °C for 1 h in a Thermomixer (ThermoFisher Scientific) shaking at 300 RPM. The tubes were centrifuged briefly after the incubation to collect the evaporate on top of the lids. Eight hundred μl of 50 mM ammonium bicarbonate was added to each tube to reduce the urea concentration to below 1 M. Two μl of trypsin gold at a concentration of 2 μg/μl was added to each tube and briefly vortexed to mix. The samples were then incubated at 37 °C for 15 h in a Thermomixer shaking at 300 RPM.

2.5. Purification of peptides

Solid phase extraction (SPE) was performed with a vacuum manifold using Strata C-18 T solid phase extraction columns (Phenomenex, Torrence, CA) and following the manufacturer's protocol. Briefly, 1 ml of 100% methanol was added to activate the resin, followed by a conditioning step of 1 ml 0.1% TFA water, then addition of the samples. The samples are washed with 5% acetonitrile in 0.1% TFA water, and finally elution of the samples with 80% acetonitrile in 0.1% TFA water into labeled clean low protein binding 1.5 ml microfuge tubes. The desired flow rate for vacuuming steps is 0.5 ml/min with the vacuum seal released after each solution. Samples were dried down to near completeness (5–10 μl remaining) with a Thermo speed vac. 25 μl of 0.5% formic acid was added to each sample, using the pipettor gently wash the sides of the tube to recover as much of the sample as possible. The samples were then transferred to labeled HPLC vials with 200 μl inert glass inserts and capped with screw caps.

2.6. MS–MS analysis of the peptides

Peptides were separated using an Agilent 1200 HPLC with a 40 cm long, 0.15 mm ID fused silica column packed with Jupiter 5 μm C-18 resin. Column was made in house. A 50 min gradient was established by changing the relative concentrations of a two solvent systems where "A" is 5% acetonitrile, 0.1% formic acid in H_2O and "B" is 95% acetonitrile, 0.1% formic acid in H_2O at a flow rate of 2 μl per minute. The separation had a 10 min isocratic step at 5% B and a gradient from 5% to 60% B over 50 minutes. Eluate from the HPLC was directly transferred into an LTQFT Orbitrap system (Thermo electron, Billercia MA). The electrospray conditions used were: 2.5 kV spray voltage, 200 °C and 10 V ion transfer tube voltage. The ion injection time was set for automatic gain control with a maximum injection time of 200 ms for 5 × 10e7 charges in the trap. The MS parent scan was performed using the Orbitrap mass analyzer using a resolution setting of 30,000. Dynamic parent ion selection was performed where the top five most abundant ions were selected for MS–MS in the linear quadrupole ion trap using a 3 *m*/*z* mass window.

2.7. Database searches

Searches employing MS–MS data were performed using the open-source software X!Tandem (www.thegpm.org/tandem) [5].