Top-down proteomics: the next step in clinical microbiology.

In the last decade, the introduction of MALDI-TOF Mass Spectrometry (MS) for rapid microbial identification has revolutionized the field of clinical microbiology. The approach has been widely embraced by hospitals as it is faster, more accurate, and less expensive than conventional phenotypic or genotypic methods. However, it suffers from important limitations. The discriminatory power of the technique is insufficient to differentiate closely related bacteria or sub-species and more importantly resistance and virulence cannot be addressed. There is therefore a crucial need for innovative analytical approaches allowing an efficient and more accurate bacterial identification based on protein analysis.

Top-down proteomics is an emerging technology based on the analysis of intact proteins using very high-resolution mass spectrometry [1]. It provides the highest molecular precision for analyzing primary structures by examining proteins in their intact state, leading to more straightforward and reliable results than the classical bottom-up approach based on protein enzymatic digestion.

Top-down proteomics is particularly suited to the analysis of bacterial proteins, which are of small size (< 30 kDa) and produced in large amount by bacterial pathogens [2,3].

In order to use top-down proteomics for clinical microbiology applications [4], we set up an integrated platform in which all steps have been carefully optimized: bacterial lysis, protein extraction, LC-MS/MS analysis of intact proteins and data processing. For this last point, a new software tool, which branches from an existing one [5], but tailored towards top-down proteomics data, has been developed. This new software, based on machine learning, can rapidly cluster the thousands of MS/MS spectra obtained in top-down LC-MS/MS experiments, compare datasets obtained from various bacterial pathogens and identify discriminative spectra.

Using this integrated top-down platform, we show that it is now possible to differentiate closely-related pathogens that are impossible to distinguish with MALDI-TOF MS, in only a few hours after bacterial culture. We also highlight the great potential of top-down approaches to delineate complete protein sequences (including C-terminal and N-terminal extremities) and detect single nucleotide polymorphisms.

References