Review on Identification of Proteins of Microorganisms by Two-dimensional gel electrophoresis-Mass spectrometry (2-DE-MS)

P. Muralidhar Reddy¹,², Yen-Peng Ho¹* and Vadde Ravinder²*
¹Department of Chemistry, National Dong Hwa University, Hualien -97401, Taiwan
²Department of Chemistry, Kakatiya University, Warangal-506 009, A.P., India
*¹E-mail ypho@mail.ndhu.tw , *²E-mail: ravichemku@rediffmail.com,

Abstract: Mass spectrometry (MS) has greatly enhanced research in the field of microbial proteomics. Microbial characterization, including detection, differentiation, and identification, are the most important tasks of microbial proteomic laboratories worldwide. 2-DE is a well-established technique for separation and quantitation of a large set of proteins and has much to contribute to the experimental analysis of microbial organisms. Proteins separated by 2-DE subsequently rely on MS techniques for identification and further characterization. This review will briefly summarize the basic concepts of 2-DE, and 2-DE/MS technology and its specific methods of detection of proteins of microorganisms, the proteome of cell membrane and intracellular or cytoplasmic proteome turnover in biomarker discovery. Clearly, it is not possible to cover all of the literatures that were published in these years. Therefore we focused on the success stories as well as some of the challenges that are currently being faced.

Keywords: Microorganisms, Mass spectrometry, Protein, Proteomics, Two-dimensional gel electrophoresis

Introduction
The last decade in life sciences was deeply influenced by the development of the “Omics” technologies (genomics, transcriptomics, proteomics, and metabolomics), which aim for a global view on biological systems. The recent evolution towards rapid protein identification has made large-scale proteome analysis possible¹². The invention of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) by O’Farrell³, made an extensive separation of hundreds or even thousands of protein species possible and therefore brought the goal of a preferably complete proteome analysis more rapidly. The comparison of protein spots between gel images includes detection of protein spots, matching spots between images from different gels, and quantitation of spots between matched images. Accuracy in spot detection and matching determines the quality of results obtained. It is impractical to quantify and analyze large numbers of spots manually; therefore, a number of sophisticated automated image analysis software tools have been developed commercially⁴. Furthermore, due to the development of soft ionization methods⁵⁶, mass spectrometry (MS) was entering protein science. Combined with the introduction of database search algorithms⁷⁹, proteins separated on 2D-gels could now be identified in a significantly higher throughput compared with the hitherto performed N-terminal sequencing via Edman degradation¹⁰. The increase in throughput, the partial automation, and the higher reproducibility of 2D-PAGE analysis recently made it a very attractive tool to study cellular functions on a molecular level.

Mass spectrometry has the potential for identification of low intensity spots of silver-stained 2-DE gels¹¹. Protein and peptide separation, mass spectrometry techniques, and the combination of these technologies with powerful bioinformatics tools. Proteomics has been extensively reviewed elsewhere¹¹⁻¹⁹. Proteomics is concerned with the identification of proteins and classification of where, when and how much they are expressed in cells. In addition to all the proteomics types of mass spectrometry analyses, mass spectrometry can be used to probe many functional aspects of individual proteins. In the world of protein mass spectrometry, there is not one, all-purpose workflow. Some identify proteins using peptide mass fingerprinting, while others sequence using tandem mass spectrometry (MS/MS)²⁰⁻²⁵. Furthermore, MS has...
become a powerful tool to identify bacterial biomarkers\textsuperscript{1,2,26-31}. Biomarkers are compounds that potentially can be used for early diagnostic or disease/treatment surveillance purposes. The bottleneck in MS based microorganism (peptide and protein) identification is now at the stage of data analysis and verification of results. The development of two so-called ‘mild ionization’ techniques in biological research are matrix assisted laser desorption ionization (MALDI) and electrospray ionization (ESI), provided practitioners of mass spectrometry with the means to generate gas-phase ions from moderate to high molecular weight thermally labile compounds. Both MALDI and ESI have been used to analyze lipids, DNA, proteins and small molecules of pathogens\textsuperscript{1,28-32}. On the other hand, MALDI/ESI coupled instruments can be operated in MS/MS mode, and therefore, give more structurally related information of microbial molecules, including protein sequences\textsuperscript{1,22,25,33}. This review highlights on the critical overview of the basic concepts of 2-DE, and 2-DE/MS technology and its specific methods of detection of proteins of microorganisms, the proteome of cell membrane and intracellular or cytoplasmic proteome turnover in biomarker discovery. 2-DE in combination with mass spectrometry has the potential to resolve the proteome of an organism down to the level of protein species. Clearly, it is not possible to cover all of the literatures that were published in these years. Therefore we focused on the success stories as well as some of the challenges that are currently being faced.

**Basic concepts of 2D gel electrophoresis**

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) is the most common technique used in proteomics. It has been the technique of choice for analyzing the protein composition of cells, tissues and fluids as well as for studying the changes in global patterns of gene expression. The biggest advantage of this technique is its compatibility with mass spectrometry which can then be used to identify the proteins\textsuperscript{1,11}. The technique provides high-resolution separation of the individual proteins of a proteome\textsuperscript{11}. 2-Dgels can separate hundreds of proteins on a single gel. The process of moving from a cellular homogenate to identification of a proteomic signature; clearly 2D-DE is central to this process this proteomic process. In 2D-DE, proteins are separated by two distinct properties: isoelectric point (pI) and molecular weight (MW). The isoelectric point is the pH at which the protein has no net charge. In the first dimension, which is known as isoelectric focusing (IEF), proteins are applied to a mylar strip containing an immobilized pH gradient (called IPG strips) and the proteins migrate to their isoelectric point when an electric field is applied. As proteins migrate to their respective isoelectric points, they pick up or lose protons. As they continue to migrate, the net charge on the proteins and their mobility decreases and eventually they come to a point where their net charge is zero and they stop moving. This is their isoelectric point. IPG strips are available in different pH ranges (examples: pH 3-10, 4-7, 3-6, 5-8, 7-10) and various lengths (7, 11 and 17 cm). Proteins separated by IEF are further separated orthogonally by SDS-PAGE (Sodium dodecyl sulfate-polyacrylamide gel electrophoresis). The transition from first to second dimension involves 2 steps. The first step involves the equilibration of IPG strips in a buffer containing SDS, a detergent that imparts negative charge to the protein that is proportional to the mass of a protein. In addition, the proteins that have been separated by IEF are also reduced and alkylated using DTT (dithiothreitol) and iodoacetamide (IAA). DTT reduces the disulfide bonds (R-S-S-R) found in proteins. The resulting –SH groups are then alkylated with iodoacetamide, which attaches an alkyl group to the –SH (R-S-CH2-CONH2) to prevent reformation of the disulfide bonds.

The treated IPG strips are then embedded in the top of an SDS-PAGE gel. An electric current is then applied to the system, causing the proteins to migrate from the IPG strip into the SDS-PAGE gel, where they are separated by size. Proteins move through the porous gel with small molecules moving more rapidly than the larger ones. Polyacrylamide is a cross-linked polymer of acrylamide. Acrylamide is the material used for preparing gels for electrophoresis and separates proteins by their molecular size. Ammonium persulfate is the polymerizing agent which when added to acrylamide gel mixed with bisacrylamide, forms a network of cross-linked polymer. The pore size in the gel is inversely related to the amount of acrylamide used. Gels containing a high percentage of acrylamide are used to resolve small proteins, whereas low percentage acrylamide gels are used to resolve large protein. For example the gel containing 7% acrylamide will have larger pores in the gel and will separate larger proteins than a 12% polyacrylamide gel, which would have small pores and thus would separate small size proteins. Typical polyacrylamide concentrations range from 5% to 25%. Gels can also be prepared with a gradient in polyacrylamide concentration (e.g. 8-16%), which broadens the molecular weight range for protein separation on a single gel.

**Identification of Proteins of Microorganisms by 2-DE-MS (Two-dimensional gel electrophoresis-Mass spectrometry)**

2-DE is a well-established technique for separation and quantitation of a large set of proteins and has much to contribute to the experimental analysis of microbial organisms. It allows for the visualization of hundreds of proteins at a time. Proteins separated by 2-DE subsequently rely on MS techniques for identification and further characterization. 2-DE/MS mapping strategy has been applied effectively to investigate the intracellular or cytoplasmic proteome of *S. aureus*\textsuperscript{34-36}. Furthermore, it was also used with tryptic peptide mass mapping via MALDI-TOF-MS to study the regulatory networks in pathogenicity and extracellular protein expression in *S. aureus*\textsuperscript{37-40}. By using mass spectrometry

\textsuperscript{1}Vadde Ravinder et al / Int.J. PharmTech Res. 2009, 1(3) 765
to identify the proteins, one can produce reference map. The establishment of protein reference maps is a crucial starting point for many physiological studies that may follow. The reference map of cytoplasmic proteins from two S. aureus strains COL and 8325 has also been published by Cordwell et al. 35 Nowdays, many bacterial protein maps are available; among those maps several are from pathogenic bacteria, including Chlamydia pneumoniae 41, M. tuberculosis 42, and Helicobacter pylori 43. Recently, the first proteome reference map for B. pseudomallei was constructed by Wongtrakoongate et al. 44 They analyzed and identified a total of 67 intracellular B. pseudomallei protein fractions (seven-unique proteins were identified from 88 spots presented on the gel) at a stationary phase using 2D-GE-MALDI-TOF-MS. Based on this they differentiated the protein markers of closely related species, B. pseudomallei by comparing with B. thailandensis. The 2D-gel proteome map of H. influenzae displays approximately 500 identified proteins, representing about 30% of all predicted open reading frames. Those proteins were mainly identified by peptide mass fingerprinting, using MALDI-TOF-MS, and some additional proteins were identified by amino acid composition analyses 45,46.

Similarly, Len et al. 46 published map of Streptococcus mutans contains 416 identified proteins and could be considered “comprehensive” by current 2-DE standards, the results clearly show that significantly more proteins could be visualized with SYPRO Ruby than could be analyzed by MALDI-TOF-MS. Although 95% of a bacterial genome is expressed as protein products 47, the theoretical resolving power of 2-DE is still estimated to be approximately 75% of the proteome 48. Wilkins et al. 49,50 have used 2-D PAGE-MALDI-TOF-MS to identify Streptococcus oralis and Streptococcus mutans proteins with altered expression during growth at pH 5.2 and 7.0. Proteins which were differentially expressed were excised, digested with trypsin using an in-gel protocol, and analyzed to generate a distinctive peptide mass fingerprint (PMF) which is characterized sufficiently discriminating to allow the unique identification of unknown proteins. In general, if the genome has not been sequenced, the peptide mass fingerprint can be compared with very closely related species from the same genera Peptide mass fingerprints from 2-DE protein spots obtained from Streptococcus oralis and S. mutans have been used to search the annotated genomic database of Streptococcus pneumoniae and Streptococcus pyogenes 49,50. Unfortunately, successful protein identification relies on a high degree of peptide sequence identity between species. Moreover, subcellular proteomics is the proteome analysis of the macromolecular architecture of a cell 51. In terms of the E. coli proteome, subcellular proteomics based on 2-DE-MS/MS can be used to assign various proteins to the cytosol, periplasm, inner membrane, or outer membrane by biochemical fractionation; this method was used to assemble the largest proteome database to date 52. Analysis of 2,160 spots revealed 575 unique ORF entries, including 151 hypothetical ORF entries, 76 proteins of completely unknown functions, and 222 proteins currently not assigned in the SWISS-PROT database. Of the 575 different entries identified, 241 (42%) were found to exist in more than 1 form, at an average of 7.5 forms per entry. These findings indicate that proteomics involving sample fractionation and 2-DE can be a valuable research technique 51.

Bacillus subtilis, the best analyzed representative of the Gram-positive bacteria has been established as a model system for functional genomics of bacteria 53. In B. subtilis, two major classes of signal peptides can be distinguished on the basis of the SPase recognition sequence 55. A first series of predictions concerning the composition of the so-called secretome of B. subtilis, which, by definition, includes both the secreted proteins and the protein secretion machineries. Antelmann et al. 54 accomplished this by using two-dimensional (2D) gel electrophoresis and subsequent MS of extracellular complement of the B. subtilis secretome. Using different growth conditions and a hyper-secreting mutant, ~200 extracellular proteins were visualized, of which 82 were identified. In addition a significant number of membrane proteins, predicted lipoproteins, and even proteins without a signal peptide were also identified in the growth medium.

To obtain a global view of staphylococci proteins in the cell, the highly sensitive 2-DE is a well established technique and has to be complemented with the identification of proteins by tryptic peptide mass fingerprinting (PMF) or even more complicated MS techniques to characterize post-translationally modified peptides or proteins that were poorly amenable to PMF. S. aureus strains COL and 8325 were grown to mid-exponential phase in tryptone soy broth at 37°C. Neutral or weakly acidic proteins (pH 4–7) were separated, as were more alkaline proteins (pH 6–11). Total 377 proteins from cells grown under these standard conditions were analyzed using MALDI-TOF-MS, corresponding to 266 expressed ORFs, or approximately 12% of the S. aureus proteome 35. In recent years, 2-DE reference maps of the total or the extracellular proteins from various S. aureus laboratory-adapted reference strains have been published 35,40,57-59. Several groups investigated diverse subpopulations of proteins of S. aureus, e.g., secreted or cell wall associated proteins 40,60, or they have focused on the comparison of wild type and mutant strains 61. Furthermore, cytoplasmic and membrane proteins of S. aureus were analyzed by gel-free proteomics approaches 58,59,62.

Microbial membrane proteins constitute an important facet in physiological proteomics due to their important functions in signal transduction pathways, ion transport, cell interaction, energy conservation and their role in virulence regarding pathogenic bacteria 63-65. However, proteome analysis of membrane and cell surface proteins is complex due to their intrinsic hydrophobic nature, alkaline pI and the number of
transmembrane spanning regions. These properties lead to difficulties in solubilizing proteins for 2-DE. For bacterial proteins, many are known to be soluble in SDS sample buffers but cannot be solubilized with non-ionic detergents compatible with 2-DE. Recently, however, several new solubilising agents have been developed and have allowed the visualization of significantly more hydrophobic, membrane-associated proteins. Despite these difficulties, reference maps of membrane proteins have been published for *Escherichia coli* 63 *Acinetobacter radioresistens* 66 *Bacillus subtilis* 67 *Caulobacter crescentus* 68, *Pseudomonas aeruginosa* 69, *Helicobacter pylori* 70 *Corynebacterium glutamicum* 71 *Borrelia burgdorferi* 72 and *Staphylococcus aureus* 76. Molloy et al. describe a proteomic approach to study the protein composition of the *E. coli* outer membrane. They introduced a new isolation method of sequential extractions with increasing concentrations of sodium carbonate in analyzing *E. coli* outer membrane proteins using 2-DE-MS. This led to the successful identification of 21 out of 26 of the predicted integral outer membrane proteins 65. Similarly, Lai et al. 67 identified more than 200 *E. coli* membrane proteins by use of the above method, after modifying it to minimize non membrane protein contamination. Using this method (high-resolution 2-DE-MS), they identified 36 protein spots that were enriched in minicell membranes and 15 proteins spots enriched in rod cell membranes. The largest database of *E. coli* membrane proteins constructed and reported by Fountoulakis and Gasser 44, who identified 394 different gene products using a method identical to that described by Molloy et al. 65. Similarly, the membrane proteins of *B. subtilis* are separated by one-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (1D-SDS-PAGE), enzymatically digested and their peptides resolved via reverse phase chromatography followed by MS/MS (MALDI-MS-MS/NanoLC-ESI-MS-MS) analysis. With a result of 268 identified proteins which contain membrane-spanning domains this technique was successfully applied to start the dissection of the *B. subtilis* membrane proteome of exponentially growing cells 73,74.

Approximately one third of the proteins were predicted to be integral membrane and membrane associated, carrying one to three transmembrane domains. Nandakumar et al. 60 profiled cell membrane and cell wall-associated proteins of the *S. aureus* using various cell envelope solubilization strategies and identified numerous transmembrane domain proteins, nine proteins harboring signal sequence motifs and one protein with a cell wall anchor motif. Similarly, Sellman et al. reported the identification of cell wall associated-proteins which were isolated from *S. epidermidis* (0-47) after grown in serum by 2-DE-MS 75. Vyvytyska et al analyzed surface proteins derived from *S. aureus* (*S. aureus* strain 8325–4 and the methicillin-resistant *S. aureus* COL strain) grown in vitro under various conditions. 2-DE immunoblotting using pools of sera from patients suffering from infections or healthy individuals revealed a number of highly reactive protein spots. Twenty-one spots were isolated by preparative 2-DE, and analyzed by MALDI-MS and MS/MS sequencing of tryptic peptides allowed the identification of 15 *S. aureus* proteins including several known surface proteins. Similar approach was made to generate a comprehensive view of exoproteins of community-associated MRSA (CA-MRSA) strains. 250 unique exoproteins were identified by 2-D gel electrophoresis coupled with automated direct infusion-tandem mass spectrometry (ADI-MS/MS) analysis 76. Recently, Planchon et al. 77 reported the first proteomic analysis of cell envelope proteins of *S. xylosus* C2a, a coagulase-negative staphylococci. By coupling conventional 2-DE-MALDI-TOF MS and bioinformatics analysis, the approach developed by them allows fractionating, resolving and analyzing a significant and important set of cell envelope proteins from *S. xylosus* C2a.

Gatlin et al. 62 have successfully isolated cell envelope proteins of *S. aureus* by combining anion chromatography with either iso-electrofocusing (IEF) and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) or LC-MS/MS. They conducted proteomic analysis for two isogenic vancomycin-intermediate *S. aureus* (VISA) strains, (HIP5827 and VP32) and identified 144 proteins. Of these proteins, 48 contained predicted cell wall localization or export signal motifs, including 14 with distinct covalent peptidoglycan-anchor sites, 4 of which are uncharacterized to date. The remaining 96 proteins, devoid of recognizable motifs, were repeatedly profiled in the VISA cell envelope fractions. Bioinformatic motif searches were applied to explore the association of the identified proteins with the bacterial cell envelope. The analysis can provide information about the expression of virulence factors, which might help to evaluate the pathogenicity of clinical isolates of *S. aureus*. Kohler et al. 38 analyzed the cytoplasmic proteome of *S. aureus* COL by using the 2-DE approach combined with MALDI-TOF-MS/MS and a gel-free system using 2D-LC-MS/MS. They identified 1123 cytoplasmic proteins. Of these, 473 proteins were identified by 2D gel setup and the remaining 650 proteins were identified by the gel-free approach. A comprehensive 2-DE reference map was established that provides the basis for analyzing the global regulation of metabolism in growing and non-growing cell of *S. aureus* COL. The correlation of proteomic and transcriptomic analysis of *S. aureus* strain N315 during the post-exponential phase of growth was performed to study a sequenced strain at the system level. Total protein and membrane protein extracts were prepared and analyzed by using 2-DE, SDS-PAGE combined with microcapillary LC-MALDI-MS/MS, and multidimensional liquid chromatography. The use of a co-solvent during the isoelectric focusing 78 allowed the analysis of the most soluble membrane and membrane-associated proteins by 2-DE. Presence of high amount of detergents during SDS-PAGE separation allowed strongly hydrophobic protein fractions enriched after
phase-partitioning to be analyzed. Gene-expression data revealed that 97% of the 2596 ORFs were detected during the postexponential phase. At the protein level, 23% of these ORFs (591 proteins) were identified. Correlation of the two datasets revealed that 42% of the identified proteins (248 proteins) were amongst the top 25% of genes with highest mRNA signal intensities, and 69% of the identified proteins (406 proteins) were amongst the top 50% with the highest mRNA signal intensities. The comparative assessment and integration of both genome/proteome-wide techniques, especially when quantitative, will improve the knowledge about bacterial growth, virulence and relation with its environment.

The use of the combined SDS-PAGE and MALDI-TOF MS analysis provide a rapid tool for the characterization of some important toxins and virulence factors secreted by *S. aureus* (MSSA ATCC 29213 and MRSA ATCC 43300). Secreted proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, enzymatically digested with trypsin and analyzed by MALDI-TOF/MS 37. When grown at 42 °C, alpha- and beta-hemolysins were found to accumulate in *S. aureus* (MSSA ATCC 29213 and MRSA ATCC 43300) supernatants while the concentration of protein A was slightly decreased. The detection of low abundant exoproteins like enterotoxins required more sensitive methods like Western blot analysis. Resch and coworkers 79 reported a new approach using high-resolution 2D-PAGE and mass spectrometry to compare the proteins synthesized by biofilm and planktonic cells of wild-type *S. aureus* (strain 113). The proteins were separated at pH ranges of 4–7 or 6–11. The protein patterns revealed significant differences in 427 protein spots; from these, they identified 258 non-redundant proteins using ESI-MS/MS, expressed at significantly different levels in biofilm and planktonic cells. A database of the sequences of six openly available *S. aureus* strains and three staphylococcal phages 80 was searched for data corresponding to the ESI-MS/MS spectra. The RNA expression profiling largely supports the proteomic data 81. The results were mapped to metabolic pathway maps based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) 78,80. In recent studies Zecconi and co-workers focused on 2-DE reference map of surface proteins isolated following lysostaphin treatment from a *S. aureus* strain 1673 (in isotonic conditions), selected for its virulence characteristics and for promoting bovine mastitis. The most abundant protein components were identified by MALDI-TOF MS 82.

Oxidative stress is a crucial environmental stimulus for many pathogenic bacteria because, during infection, bacteria are frequently exposed to reactive oxygen species produced by the non-specific immune response of the host 34. The high-resolution 2-DE technique combined with MALDI-TOF MS was used to analyze the oxidative stress response in *S. aureus* COL. During these studies, a shift in the isoelectric point (pl) of some proteins triggered by oxidative stress and the nature of the protein modification was analyzed by Hecker and co-workers 83. Brady et al. 84 have identified the antigens present during an osteomyelitis infection by utilizing a rabbit model of tibial osteomyelitis and an in vitro biofilm growth system. By employing 2-DE and immunoblotting with sera from these infected rabbits followed by MALDI-TOF analysis, they were able to identify in vivo-expressed *S. aureus* (strain MRSA-M2) antigens. The up-regulation of these biofilm antigens was also globally confirmed by microarray analyses.

**Acknowledgements**

We thank the National Science Council of the Republic of China for financially supporting part of the work mentioned in this review.

**References**


