

Preface

Proteomics of microbial pathogens

Infectious diseases still plague mankind. According to the World Health Report 2004, 19.1% of the deaths estimated in 2002 were caused by infectious diseases. Aids, tuberculosis and malaria each contributed more than 2% to this figure. In June 2006, 387 completely sequenced genomes (<http://www.genomesonline.org/>) have been published in total, 352 of them from bacteria, an important prerequisite for the analysis of the proteomes of these organisms. In total 940 ongoing bacterial genome projects were reported.

The first successful proteome studies revealed vaccine candidates with promising results in animal models. Immunoproteomics resulted in the detection of antigens which may be used for diagnostics and vaccine candidate prediction. So it can be assumed that proteomics will make a marked contribution to the improvement of worldwide health within the next few years.

Here we look at some of the trends in this field. As there are so many microorganisms currently under investigation, it is not possible to present a comprehensive overview of microbial proteomics. Proteomics technology has been automated within recent years: spot picking, digestion, LC-MS/MS and database searches have increased throughput but produced new bottlenecks in quality control and data evaluation. Microorganisms are ideal models for the application of these new technologies. Bacteria with genomes containing 600 to 7000 predicted genes present a medium-sized complexity which can be used to apply proteomic techniques with a good chance of obtaining an overview of a substantial part of the proteome in combination with prefractionation procedures. Standardization is now an important theme in proteomic technology but the multiple properties of organisms and proteins make standardizing sample preparation nearly impossible. Even related bacteria need different procedures for sample preparation, as outlined in this book in the example of *Mycobacterium leprae*. It may be estimated that in one biological situation more than 50% of the predicted proteins may be identified for genomes such as *Mycoplasma pneumoniae* containing less than 1000 genes, 30% for those containing less than 2500 genes and only 10% for those containing more than 4000 genes. Subfractionation contributes to the number of accessible proteins, but in the future throughput has to be increased further to allow the presentation of the proteome in a kind of film with changing environ-

mental conditions. Only then may more complete proteomes become accessible. Bioinformatics accompanies proteomics through all the technological steps, allowing the data obtained to be stored in a database. A microbial proteomics database system was set up at the Max Planck Institute for Infection Biology (<http://www.mpiib-berlin.mpg.de/2D-PAGE/>) and by June 2006 it contains 18 bacterial species and 4889 identified spots. Peptide mass fingerprinting data are stored for *Helicobacter pylori* and isotopic labelling results are represented for *Mycobacterium tuberculosis* LC/MS data. Proteomics of microorganisms allow the scientist to start with a hypothesis-free global approach and focus early on the hypotheses elaborated from this first step. In the first few years we learned that posttranslational modifications play a more important role than expected in bacteria, and the resulting protein species composition may be directly visualized by 2-DE/MS but not by LC/MS which has other advantages such as higher throughput and sensitivity potentials. At the moment, for example, the impact of more than 10 ESAT-6 protein species in *Mycobacterium tuberculosis* remains unclear. Proteome analysis at the protein species level is a task for the future.

We wish to thank the authors for their contributions, the referees for their prompt reviewing of the manuscripts and the publishers for their help in producing this book. We also take this opportunity to thank the "Bundesministerium für Bildung und Forschung" in Germany for financing the project "New methods to access the complete proteome of bacteria" and the European Union for support in developing the European Bacterial Proteome Database within the project "Comparative analysis of proteome modulation in human pathogenic bacteria for the identification of new vaccines, diagnostics and antibacterial drugs" (QLRT-1999-31536). Several articles in this book were supported by these two initiatives.

Peter R. Jungblut
Max Planck Institute for Infection Biology
Core Facility Protein Analysis
Berlin

Michael Hecker
Ernst-Moritz-Arndt-Universität Greifswald
Institut für Mikrobiologie
Greifswald