

Preface

Mass spectrometry proteomics is a relatively young field, but it is rapidly gaining importance in both large and small laboratories around the world, and a steadily increasing number of scientists and students have come into more or less intimate contact with proteomics, through the literature, through collaborations, or through their own practical work. The field relies heavily on bioinformatics to process and analyze the acquired data, and our aim has therefore been to write a book that can serve as a solid introduction to the crossroads where proteomics and bioinformatics meet.

The intention has been to give an introduction to students and beginners in both 'proteoinformatics', experimental proteomics, and researchers who are in more peripheral contact with these fields. Thus, the book includes the basic information needed for bioinformaticians who want to develop algorithms and programs for proteomics. For such work, the informaticians must have some knowledge of molecular biology. They should also understand the function of the instruments, whether it is for writing programs that control the instruments, or for further processing of the data produced. The beginner in practical proteomics, on the other hand, is often confused by the technical and biological complexity of the field, including the theoretical basis underlying proteomics. We hope that this book can contribute to a shortening of the period of confusion. However, we will underline that this book is no replacement for contact with experienced people within bioinformatics or experimental proteomics, and we hope that the book can give the beginner a better basis of knowledge, and thereby improve the reciprocal understanding between bioinformaticians and the experimentalist.

The book concentrates on the identification and characterization problem of proteomics, though quantification and sample comparison are also briefly described. The first chapter provides a fairly detailed introduction to the peptides and proteins that are studied in mass spectrometry proteomics, and their relevant properties. It also gives an overview of the two main approaches used: peptide mass fingerprinting (PMF) and tandem MS (MS/MS). The next four chapters explain techniques and methods used in both of the approaches, including informatics solutions of selected problems. Chapters 6 and 7 describe PMF, with special emphasis on the informatics challenges. Chapters 8 and 9 describe the instruments and principles for tandem MS, and the next four chapters concentrate on algorithms and programs for protein identification and characterization by tandem MS. The last five chapters deal with quantification and some more specific topics.

The application of bioinformatics in mass spectrometry proteomics is a young field, and little theoretical foundations have been developed. It is not the goal of this book to provide such a theoretical foundation, but rather to present the main problems, and extract and systematize some common principles used in solving these problems. The presentation is illustrated by figures and numerous examples.

We have tried to restrict the description of a subject to one section or chapter, but for some of the subjects it is necessary to treat them several times in different contexts. In these cases we have included the necessary cross-references.

We have also included many references and websites, although we are aware that websites can change quickly.

Terms

Many terms used in molecular biology and proteomics do not have unique or commonly accepted definitions. We have mainly tried to follow the IUPAC (International Union of Pure and Applied Chemistry) Compendium of Analytical Nomenclature. IUPAC has an ongoing revision of the terms used in mass spectrometry, and a first draft exists.

IUPAC	http://www.iupac.org/publications/analytical_compendium/
Ongoing revision	http://www.msterms.com/
First draft	http://www.sgms.ch/links/IUPAC_MS_Terms_Draft.pdf

However, for some cases where other terms seem to be more established, we have used those.

Also, new terms are introduced in some of the literature quoted. We have relied partly on these, but in some cases have resorted to defining our own terms. When applicable, we also mention synonymous terms that may be encountered.