JOURNAL OF CHROMATOGRAPHY LIBRARY - volume 68

emerging technologies in protein and genomic material analysis

edited by
G. Marko-Varga and P. Oroszlan

ELSEVIER

JOURNAL OF CHROMATOGRAPHY LIBRARY – volume 68

emerging technologies in protein and genomic material analysis

edited by

György A. Marko-Varga

Department of Analytical Chemistry, Lund University
Lund, Sweden

and

Peter L. Oroszlan

Zeptosens AG
Witterswil, Switzerland

2003

ELSEVIER

Amsterdam – Boston – Heidelberg – London – New York – Oxford – Paris
San Diego – San Francisco – Singapore – Sydney – Tokyo

ELSEVIER SCIENCE B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

First edition 2003

Library of Congress Cataloging in Publication Data
A catalog record from the Library of Congress has been applied for.

British Library Cataloguing in Publication Data
A catalogue record from the British Library has been applied for.

ISBN: 0-444-50964-X
ISBN: 0301-4770 (Series)

♾ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).
Printed in The Netherlands.

JOURNAL OF CHROMATOGRAPHY LIBRARY – volume 68

emerging technologies in protein and genomic material analysis

This book is dedicated to our dearest;
Carina, Krisztina Alice, Christofer, Julia, Marton, Sebastian, Tillie, Zsofia

Preface

The protein- and gene-expression research areas are developing very rapidly and it is fair to say that analytical chemistry plays a very important role in these developments. New challenges lie ahead, with new discoveries and tools opening up novel dimensions in life-science applications. One could argue that the developments in analytical technology are among the major drivers in the biological sciences of today. Examples include the ground-breaking work on DNA-sequencing using capillary electrophoresis, and mass spectrometers being standard instrumentation in research laboratories throughout the world for protein/DNA sequencing and structure work. The pioneering contribution of mass spectrometry to progress in biological sciences was also highlighted by the December 2002 Nobel prizes that were awarded for the invention of electrospray and laser desorption MS.

In parallel, new discoveries in biology and medicine also set new targets for the scientific community regarding performance of new analytical technologies, systems and methods. Our recognition of the pressing need for advanced technologies to fulfill the requirements of the demanding tasks of functional genomic initiatives is the underlying motivation for all of us working on the cutting edge of bioanalytical technologies. With this book, our intention is to

- compile state-of-the-art knowledge of the field and describe emerging technologies
- provide examples by relevant applications and other case studies
- provide insight into the new frontiers of the field that embrace relevant biological systems
- stimulate young scientists and newcomers to be a part of the ongoing biological revolution, where bioanalytical developments are a cornerstone of these successes

The book is intended to serve as a general reference for researchers and scientists within the bioanalytical field as well as for postgraduate students. Each chapter includes references to the corresponding literature to serve as valuable entry points to anyone wanting to move forward in this field, either as a practitioner or for acquiring state-of-the-art knowledge.

Finally, we would like to express our gratitude to the contributing authors for their time and effort in preparing the chapters. Without the engagement and lively interactions we have had over the last half year to get this project finalized, this book would not have been possible.

GYÖRGY A. MARKO-VARGA AND PETER L. OROSZLAN

List of Contributors

J. Abian
Structural and Biological Mass Spectrometry Unit, Department of Medical Bioanalysis, Instituto de Investigaciones Biomédicas de Barcelona-Consejo Superior de Investigaciones Científicas (IIBB-CSIC), Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Rosselló 161, 7ª Planta, 08036 Barcelona, Spain

P.E. Andren
Biological Mass Spectrometery Center and Department of Pharmaceutical Biosciences, Uppsala University, SE-75124 Uppsala, Sweden and Amersham Biosciences, Björkgatan 30, SE-75184 Uppsala, Sweden

Y. Arntz
Institute of Physics, NCCR Nanoscale Science, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland

J. Bergquist
Department of Analytical Chemistry, Institute of Chemistry, Uppsala University, P.O. Box 531, SE-751 21 Uppsala, Sweden

P. Bertoncini
Institute of Physics, NCCR Nanoscale Science, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland

M. Carrascal
Structural and Biological Mass Spectrometry Unit, Department of Medical Bioanalysis, Instituto de Investigaciones Biomédicas de Barcelona-Consejo Superior de Investigaciones Científicas (IIBB-CSIC), Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Rosselló 161, 7ª Planta, 08036 Barcelona, Spain

R.J.E. Derks
Department of Analytical Chemistry and Applied Spectroscopy, Division of Chemistry, Faculty of Sciences, vrije Universiteit amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands

Ch. Gerber
Institute of Physics, NCCR Nanoscale Science, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland and IBM Zurich Research Laboratory, Säumerstrasse 4, CH-8803 Rüschlikon, Switzerland

W. Grange
Institute of Physics, NCCR Nanoscale Science, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland

M. Grønborg
Protein Research Group, Department of Biochemistry and Molecular Biology, Odense University/University of Southern Denmark, Odense, Denmark

A. Guttman
Torrey Mesa Research Institute, San Diego, CA 92121, USA

M. Hegner
Institute of Physics, NCCR Nanoscale Science, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland

A.C. Hogenboom
Department of Analytical Chemistry and Applied Spectroscopy, Division of Chemistry, Faculty of Sciences, vrije Universiteit amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands

S. Husale
Institute of Physics, NCCR Nanoscale Science, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland

H. Irth
Department of Analytical Chemistry and Applied Spectroscopy, Division of Chemistry, Faculty of Sciences, vrije Universiteit amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands

O. Nørregaard Jensen
Protein Research Group, Department of Biochemistry and Molecular Biology, Odense University/University of Southern Denmark, Odense, Denmark

H.P. Lang
Institute of Physics, NCCR Nanoscale Science, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland and IBM Zurich Research Laboratory, Säumerstrasse 4, CH-8803 Rüschlikon, Switzerland

T. Laurell
Department of Electrical Measurements, Lund Institute of Technology, Lund University, P.O. Box 118, S-221 00 Lund, Sweden

G.A. Marko-Varga
Department of Analytical Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden

J. Nilsson
Department of Electrical Measurements, Lund Institute of Technology, Lund University, P.O. Box 118, S-221 00 Lund, Sweden

P.L. Oroszlan
Zeptosens AG, Benkenstrasse 254, CH-4108 Witterswil, Switzerland

A. Palm
Department of Chemistry, Indiana University, Bloomington, IN 47405, USA

M. Palmblad
Division of Ion Physics, Uppsala University, Box 534, SE-751 21 Uppsala, Sweden

K. Sköld
Biological Mass Spectrometery Center and Department of Pharmaceutical Biosciences, Uppsala University, SE-75124 Uppsala, Sweden

P. Svenningsson
Department of Physiology and Pharmacology, Karolinska Institute, SE-17177 Stockholm, Sweden

M. Svensson
Biological Mass Spectrometery Center and Department of Pharmaceutical Biosciences, Uppsala University, SE-75124 Uppsala, Sweden

J. Zhang
Institute of Physics, NCCR Nanoscale Science, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland

Contents

G.A. Marko-Varga and P.L. Oroszlan (Editors)
Emerging Technologies in Protein and Genomic Material Analysis
Journal of Chromatography Library, Vol. 68

CHAPTER 1

Enabling Bio-analytical Technologies for Protein and Genomic Material Analysis and their Impact on Biology

GYÖRGY A. MARKO-VARGA[a] and PETER L. OROSZLAN[b]

[a]*Department of Analytical Chemistry, Lund University, Lund, Sweden*

[b]*Zeptosens AG, Witterswil, Switzerland*

1.1. BACKGROUND

Although the human genome was made public by the end of June 2000, it was not the complete human map; there were still parts that were not readily available. Currently, there is a plan that the National Human Genome Research Institute will publish the finished draft of the human-genome sequence in a scientific journal by 2003. During the period between the first launch of the human-genome sequences (June 2000) and today, the focus has turned towards the post-genomic area. High capacity and high performance technologies that are able to identify the expression of proteins are currently a research area attracting enormous attention. This is due to the fact that the DNA coding and 'the holy grail' of human life is no further than the nearest computer with an Internet connection. Almost the entire genome can be searched. These gene sequences hold only a limited percentage of the entire proteome that is present in cells. The reason is that, upon transcription and translation, proteins in many cases undergo a post-translational modifying step.

Proteomics has become the new 'hot' research where the protein expression area has received enormous attention scientifically as well as from investors in the stock market [1]. Several hundred million dollars have already been invested by the Pharma and Biotech companies and the business growth until 2005 is expected to be $5 billion [2].

Academia, the pharmaceutical industry and the biotech sector have moved their efforts and positions towards trying to complete the human proteome and what it holds. It seems that this task is not only a major effort to fulfil, but also difficult to define. What exactly the proteome as such holds in terms of gene products is one aspect of addressing the entire map of the human proteome. The other aspect is that the post-translational modifications that occur within the cell reflect a fundamental and very important process whereby additionally a large number of proteins is present. The post-translationally modified proteins may occur as many variants all sharing a larger degree of protein sequences, e.g. enzymes often occur as

References pp. 10

one protein sequence holding an inactive form in a state of rest, which after modification becomes activated. These activation mechanisms include phosphorylations, carbohydrate modifications or simply cleaving a protein sequence from the pre-protein.

Many academic scientists argue that efforts to map the entire proteome are a matter of organisation. This illustrates the need for a single body to coordinate research efforts and advance human knowledge as a whole rather than according to the corporate agenda. Unlike the genome project, where DNA sequencing technology propelled the project towards a single easily understood goal, the proteomics field is driven by an array of emerging technologies having multiple goals. Scientists would like to identify ultimately all human proteins, determine the shapes they assume inside cells and understand their function by observing how they change in different circumstances. The initiative of starting an inventory of protein content in cells was suggested in the 1980s [3].

The Human Proteome Organization (HUPO) is currently trying to provide some help coordinating large proteome investigations. There is also agreement within HUPO that the focus of its work will be on mapping proteins within certain biological materials generated from human clinical studies, with cancer research of primary interest [4].

The 'Human Proteome Index' (HPI) [5] is another initiative from the founders of SwissProt. It is a human protein database to which all human protein sequences can be submitted. Because national and international funding into proteomics research is steadily increasing, sequence information should be stored in a central database accessible to everyone. The number of inactive protein sequences forms one part of the required information. This detailed biological information addresses the actual cell compartment from where the specific protein in question originates and/or where it was found to be present. It is of mandatory importance. An identity by itself, that a protein is expressed, or found to be up- or down-regulated in a certain disease or disease in-vivo or in-vitro model, comprises only the first step in the proteomics process. Unless the expression can be linked to its biological function, resulting limitations in the verification of the biological role will remain.

Sequencing the genome is just the start in a long march towards an understanding of how humans grow and develop and also to the link with disease and disease evolvement. Genes are powerful, but ultimately we need to understand the proteins that carry out bodily functions.

In this context, the evolving 'new science' and its various disciplines are expected to have a large impact also on the life science industry.

1.2. INDUSTRIAL IMPACT

State-of-the-art drug discovery and development—the foremost important consumer of new analytical technologies—is driven largely by chemistry and high-throughput screening. Questions about the biological effect of a candidate drug (i.e. toxicity, efficacy) are addressed only at a rather late phase of the development process. In the contemporary pharmaceutical business, much effort is directed to reducing drug development times significantly in order to expedite marketing incentives. A rational time-saving approach could contribute significantly to the short-term (i.e. fast product launch) as well as