Journal of Chromatography A, 1218 (2011) 8777–8790

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Review

Comprehensive chromatographic separations in proteomics

P. Donato[a,b], F. Cacciola[b,c], L. Mondello[b,c], P. Dugo[b,c,*]

[a] Università Campus Bio-Medico, Via Álvaro del Portillo 21, 00128 Roma, Italy
[b] Dipartimento Farmaco-chimico, Università di Messina, viale Annunziata, 98168 Messina, Italy
[c] Chromaleont s.r.l. A spin-off of the University of Messina, Via Industriale, 143, 98123 Messina, Italy

ARTICLE INFO

Article history:
Available online 30 May 2011

Keywords:
Proteomics
Multidimensional liquid chromatography
Mass spectrometry

ABSTRACT

In such a complicated field as proteomic analysis, scientists are more and more challenged in implementing separation systems capable to provide enhanced separation power, as well as sensitivity of detection for adequate identification and, to a lesser extent, quantification of the separated compounds. To address such issues, several combinations of different separation modes have been investigated in comprehensive liquid chromatographic platforms, in which the entire sample eluted from the first dimension is subjected to a secondary chromatographic separation. The different applications exploited for comprehensive LC analysis of intact or digested proteins are the focus of this review, in which advantages and disadvantages of the different columns combinations, interfaces, and operating modes are pointed out. The combination with mass spectrometry as part of the total system is stressed, and illustrated in more detail. Theoretical concerns and practical requirements will be briefly discussed, as well.

Contents

1. Introduction

The last decade has witnessed an exponential increase of the number of multidimensional liquid chromatography (LC) applications, addressed to attain enhanced resolving power for the separation of highly complex samples. The field of proteomics has undoubtedly represented one of the major driving forces towards the implementation of such platforms, which are defined "comprehensive" when the entire sample eluted from the first dimension (D1) is subjected to a secondary chromatographic separation (second dimension, D2).

A number of different LC separation modes has been exploited in D1, attempting to deliver a certain degree of orthogonality to D2 which usually consisted in reversed phase (RP) due to its amenability of direct linkage to mass spectrometer (MS).

From the detection standpoint, the advent of electrospray (ESI) and matrix assisted laser desorption ionization (MALDI) techniques, as well as the high resolution and tandem MS capabilities

* Corresponding author at: Dipartimento Farmaco-chimico, Università di Messina, viale Annunziata, 98168 Messina, Italy. Tel.: +39 090 6766541; fax: +39 090 358220.
E-mail address: pdugo@unime.it (P. Dugo).

0021-9673/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2011.05.070

of the modern mass analyzers, have definitely concurred to make LC–MS and LC–MS/MS emerge to a central role in modern proteomics.

This review will illustrate the different applications of comprehensive LC (2D LC) in the field of proteome analysis, providing a critical discussion of *pros* and *cons* of the different stationary phase combinations, interfaces, and operating modes. Theoretical concerns and practical requirements will be introduced, as well.

2. Tools for proteomic analysis

The term "proteomics" was coined in 1995 by Wilkins and co-workers [1,2] as the study of a proteome expressed by a specific genome, encompassing all proteins expressed in a cell at any given time, including protein isoforms as well as co- and post-translational modified (PTM) forms.

Expression proteomics deals with the characterization i.e., the identification and quantification of proteins in cells, tissues, or biological fluids; that could mean such a complex mixture as up to 20,000 proteins in a single cell population (serum proteome). The wide dynamic range of protein expression within the proteome brings in added difficulty to their analysis; just to make an example, the concentration ratio between albumin and the least abundant species in serum can be as high as 1:10,000 [3]. The enormous complexity, variability, and dynamic range make proteome analysis a more challenging task with respect to that of the genome, which shows little (if any) variations between cells and tissues [4], and whose dynamic range may vary to a lesser extent, i.e., only 5 orders of magnitude for DNA [5].

The word "characterization" in turn encompasses the determination of the function(s) of all expressed proteins (functional proteomics), and the assessment of their cellular localization and post-translational modifications. Since proteins rarely act alone at the biochemical level, functional proteomics also involves the assay of proteins interactions, rather to perform a given cellular task, or as key players in a number of diseases. Furthermore, protein expression will be ultimately affected by environmental, biological, pharmacological, and disease factors, which will determine statistically significant variations in the diversity and extent of their production.

The diversity and extent of proteome complexity has placed great demand for highly efficient analytical platforms, using a combination of protein separation and identification techniques. Fractionation of a proteome sample may be carried out mainly at the protein or peptide levels. While intact proteins are difficult to handle, their digestion with a proteolytic enzyme (usually, trypsin) on the other hand dramatically increases sample complexity. Considering that trypsin yields, as an average, 30 peptides per protein and taking plasma as an example, which contains approximately 30,000 different proteins, the proteolytic digestion will in fact result in as many as 900,000 peptides, not taking into account any processing or modification. Whereas proteins can be very diverse in their chemical nature, peptides show a uniform behaviour ("peptidomics", rather than "proteomics") making the protein analysis more affordable.

Two main fields of separation science which have ever applied to the analysis of proteins and peptides are liquid chromatographic and electrophoretic techniques, both allowing to resolve complex mixtures employing different separation mechanisms according to the chemical and physical properties of the solute [6].

From the detection standpoint, mass spectrometric (MS) and tandem MS based on fast atom bombardment (FAB) ionization have dominated the field of protein and peptide analysis for over a decade, overcoming the limitations on the ionization side that, for more than half a century, have excluded the use of MS to investigate large molecules, including proteins [7]. Later on, the development of electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) has exponentially increased the capability of MS for biomolecule analysis, allowing to generate ions of over a wide range of molecular masses [8,9].

Meanwhile, new analyzers have come to replace the older sector machines, to be used in conjunction with these new, simpler, sensitive, and more versatile ionization methods: ion trapping instruments, quadrupoles, and time of flight (TOF). Later on, a variety of hybrid instruments have become commercially available, providing the capability of high resolution, high sensitivity, and high mass accuracy over a wide dynamic range; among these, ion mobility TOF, quadrupole time of flight (QqTOF), ion trap-TOF (IT-TOF), linear ion trap-Fourier transform ion cyclotron resonance (FT-ICR). These developments have definitely concurred to bring mass spectrometry to a central role in present-day proteome research, overcoming most of the limitations associated with two-dimensional gel electrophoresis (2D-GE) and Edman degradation techniques, which have represented classical approaches to amino acid sequencing.

Since increasingly complex biological molecules are studied by MS and/or tandem MS, the need for more powerful and highly resolving separation methods has grown over the past years. 2D gel electrophoresis (2D GE) has dominated the field of protein separation for a long time, given its excellent resolving power for intact proteins (≈2000). In a classical gel-based experiment, high orthogonality can in fact be achieved for complex mixtures of proteins, which are separated according to their p*I* and molecular weight. Excised proteins are then subjected to enzymatic digestion (usually with trypsin) into peptides, and subsequently analyzed by MS, usually via ESI or MALDI interface [10–15].

More recently, LC has been proposed as an alternative technique in order to overcome most of the issues associated with this technique, consisting in difficulty of automation, low accessibility of membrane-bound proteins, problematic detection of proteins characterized by large molecular weight, high p*I*, strong hydrophobicity, or low abundance. Major drawbacks also consisted in lack of reliable quantitation and high reproducibility [16–20].

As a consequence, much effort has been placed in the development of multidimensional LC (MDLC), especially in the comprehensive mode (LC × LC) [21–46], and several review papers have recently dealt with such a topic [16,47–50,51–55]. MDLC combines two or more forms of LC to increase the peak capacity, and thus the resolving power in order to better fractionate peptides before entering the MS detector. In particular, enhanced separation methods prior to MS fulfils two requirements: it minimizes ion suppression and improve ionization efficiency, as it dramatically simplifies the complexity of peptide ions prior to MS detection. Ion suppression phenomena may in fact arise from highly abundant peptides, which would obscure the detection of other low-abundance peptides present in the mixture. This effect is likely to be encountered in proteomics, due to the wide dynamic range of protein expression within the proteome.

The phenomenon of ion suppression is clearly visible from the LC–MS plot depicted in Fig. 1, where the regions of ion suppression and the peptides causing the ion suppression are marked and numbered.

Gel-free, LC based separation techniques come along with the benefits of higher throughput, relative speed, capability of quantitation, easiness of full automation, and the likelihood of straightforward hyphenation to mass spectrometry and, as a consequence, have gained ever wider acceptance [56].