LC–tandem MS in proteome characterization

Sarah R. Hart, Simon J. Gaskell

The coupling of high-performance liquid chromatography (HPLC) to mass spectrometry (MS) has been routinely used for a number of years for the analysis of a wide variety of different compounds. In typical proteomic analyses, where enzymatic digestion is used to generate proteolytic peptides, the limited amount of sample restricts the utility of conventional HPLC methods for MS detection. As reduced column diameters and nanolitre per minute flow rates have become increasingly standard, the application of HPLC to the analysis of low-volume, low-abundance samples has now become readily achievable. A number of novel chromatographic methods have increased the utility of such approaches for proteome-wide analyses. However, there remain in proteomic analyses some important challenges, which are being addressed by state-of-the-art methodologies. This article reviews a number of pertinent considerations and technological advances in proteomic analyses using HPLC–tandem MS (MS^2).

Keywords: High-performance liquid chromatography; HPLC; MS^2; Peptide; Proteome; Proteomics; Quantification; Tandem MS

Sarah R. Hart,
Simon J. Gaskell*
Michael Barber Centre for Mass Spectrometry, University of Manchester, Manchester, UK

*Corresponding author.
Tel.: +44 161 306 4474;
Fax: +44 161 306 4421;
E-mail: Simon.Gaskell@manchester.ac.uk.

1. Introduction

1.1. Proteomics

The term proteome was initially coined in 1994 at the first Sienna 2D electrophoresis meeting, and is habitually defined as being the protein complement to the genome [1,2]. Traditionally, mass spectrometry (MS)-based proteomics relied upon the generation of relatively low complexity samples, where proteins were initially separated by 1D or 2D-electrophoresis prior to the identification of peptides using either matrix-assisted laser desorption ionization time of flight MS (MALDI-TOF-MS) by peptide mass fingerprinting [3–5], or using nanoflow electrospray ionization MS (nESI-MS), typically with tandem MS (MS^2) to identify peptides using sequence-specific mass tags [6,7] via automated data-dependent acquisition [8]. The field of proteomics has been revolutionized within the last few years. What was once achievable by only a minority of laboratories running state-of-the-art MS instrumentation, namely confident identification of large numbers of proteins from complex mixtures, is now achievable by a much larger number of proteomics laboratories running instrumentation with relatively simple end-user interfaces [2]. By and large, this revolution has been achieved as a result of the routine application of HPLC separations of complex biological samples to MS analyses [9]. This review does not seek to be a comprehensive account of current HPLC–MS^2-based methods, but uses examples from the work of this laboratory to illustrate a number of the challenges that are currently faced and the advances that are being made to overcome some of these challenges.

In typical proteomic analyses, two or more parallel cell states are compared (e.g., cultured cells prior to and following stimulation with a chemoeffector, or over a time-course of stimulation), either using visualization of electrophoretically separated proteins or via the use of stable isotope-labelling methods [10–13]. Such methods are relatively widespread, and, although novel technical innovations have been made recently, all rely on the same basic principles and will therefore not be discussed here.

There are a variety of reasons why HPLC–MS^2-based identifications have become the methods of choice, not least the ease of automating analyses following the introduction of data-dependent analysis methods and the ability of HPLC-based separations to separate and to analyse highly complex biological samples. The introduction of nanobore-HPLC (nLC) columns and instrumentation capable of maintaining the stability of sub-microlitre flow rates was highly critical to the widespread adoption of LC–MS^2-based methods in proteomics, so generalized workflows

0165-9936/$ - see front matter doi:10.1016/j.trac.2005.04.009

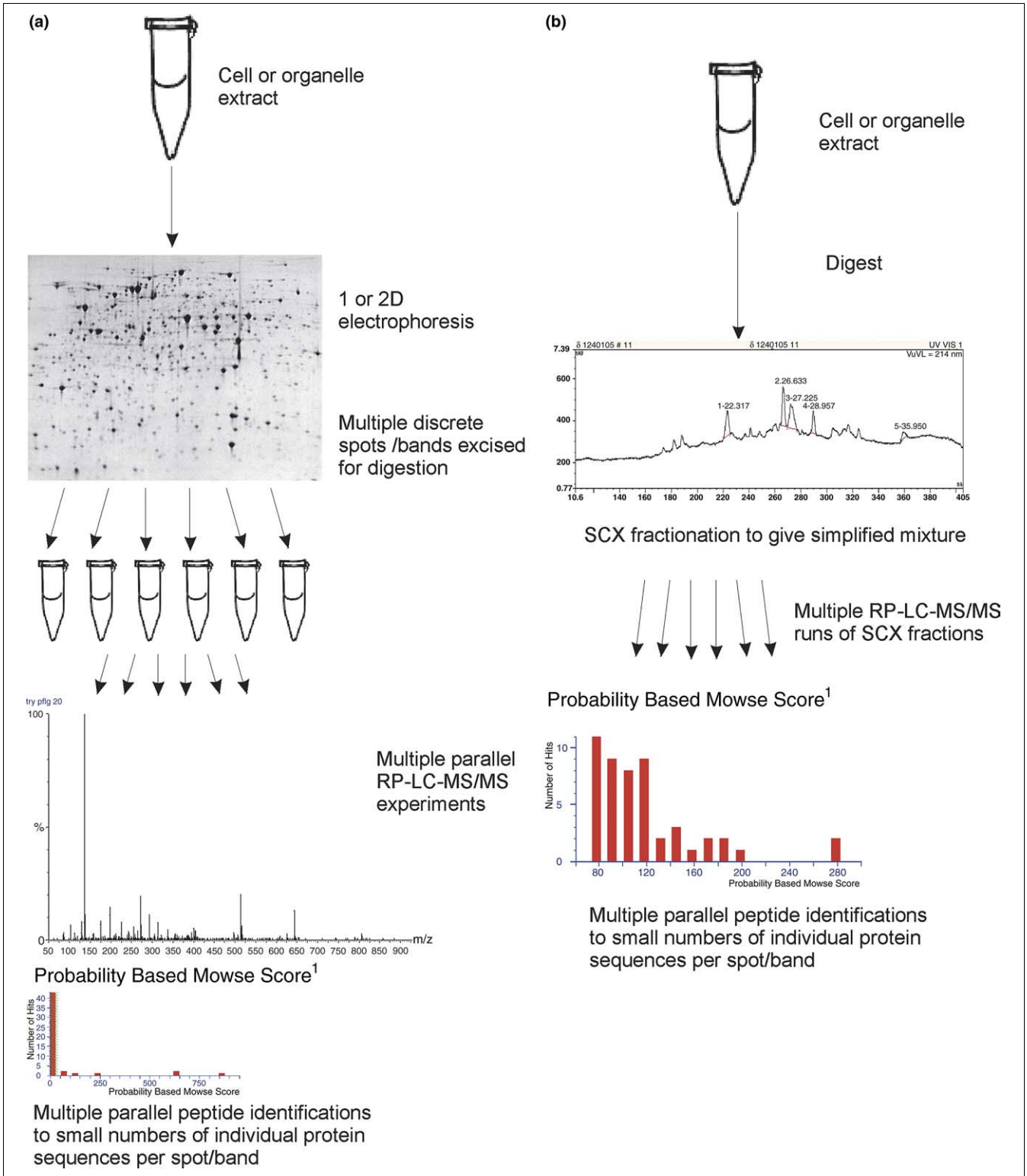

Figure 1. Schematic of general approaches for HPLC–MS^2-based proteomics: (a) gel-based method allows direct correlation of protein expression level with identification between parallel samples; and, (b) the multi-dimensional chromatography approach allows extremes of protein molecular weight and isoelectric point (which are poorly represented using gel-based methods) to be observed. RP-HPLC–MS^2, Reverse-phase high-performance liquid chromatography–tandem MS; SCX, Strong cation-exchange chromatography. [1]Ions score is $-10*\text{Log}(P)$, where P is the probability that the observed match is a random event. Individual ions scores >43 indicate identity or extensive homology ($p < 0.005$). Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.