Novel liquid-chromatography columns for proteomics research

Michiel H.M. van de Meent, Gerhardus J. de Jong

The enormous interest in proteomics research in recent years has inspired many developments in peptide chromatography. Different strategies have been developed to cope with the vast complexity of proteomics samples, trying to provide sufficient degree of separation to be able to exploit fully the potential of protein identification by mass spectrometry (MS). As reversed-phase liquid chromatography (RPLC) coupled to MS is still the method of choice for the analysis of protein digests, many efforts focus on the development of high-efficiency RP methods (e.g., monolithic columns and ultra-high-performance LC). This can also increase the speed and the sensitivity of the analysis of protein digests.

As RPLC-MS alone is unlikely to provide sufficient resolution to unravel the composition of highly complex samples comprehensively, multidimensional methods will remain essential in proteome research. In this area, hydrophilic interaction chromatography (HILIC) seems to be a promising alternative to the traditional strong cation-exchange-based methods. Also, HILIC has found application in the analysis of post-translational modifications (e.g., phosphorylation and glycosylation).

This review describes recent developments in LC methods for proteomics research, focusing on advances in column technology and the application of novel column materials. Illustrative examples show the possibilities of the new columns in proteomics research.

Keywords: Column material; Column technology; Hydrophilic interaction chromatography (HILIC); Liquid chromatography (LC); Mass spectrometry (MS); Peptide chromatography; Post-translational modification (PTM); Protein identification; Proteomics research; Reversed-phase liquid chromatography (RPLC)

Michiel H.M. van de Meent, Gerhardus J. de Jong*
Division of Biomedical Analysis, Department of Pharmaceutical Sciences, Utrecht University, P.O. Box 80082, 3508 TB, Utrecht, The Netherlands

*Corresponding author.
E-mail: G.J.deJong@uu.nl

1. Introduction

Proteomics is the area of biomedical research that aims to characterize the full protein complement of a cell, a tissue or an organism at a certain point in time under given conditions. The general analytical strategy in proteomics comprises sample preparation, separation, identification and interpretation. However, a given proteome may comprise hundreds to tens of thousands of proteins, spanning up to ten orders of magnitude in concentration [1]. The heterogeneity between proteins in a biological sample adds extra complexity that needs to be addressed by the technologies used to analyze the proteome.

Separation strategies for proteomics can generally be divided into two approaches:

(1) separation of proteins; and,
(2) separation of peptides after protein digestion.

Due to its high resolving power, two-dimensional gel electrophoresis (2D-GE) is the method of choice for protein separation. However, 2D-GE has a limited dynamic range and shows poor performance for very large, very small and very hydrophobic proteins. Also, automation, reproducibility and quantification remain difficult. This has prompted researchers to explore alternatives {e.g., combining 1D SDS-PAGE protein fractionation with enzymatic digestion and a liquid chromatography (LC)-based peptide separation [2]}. Protein digestion has the advantage of reducing the physicochemical complexity of the sample, as peptides are more homogenous than intact proteins. However, cutting the proteins into pieces obviously increases the numerical complexity of the sample as an average protein may yield around 20 peptides after digestion.

LC-based methods coupled to mass spectrometry (MS) are the most widely used methods for peptide analysis because of their high resolution and ease of automation. The efficiency of the peptide separation is important for MS-based identification, so continual attention is paid to optimization of LC systems.

0165-9936/$ - see front matter © 2011 Elsevier Ltd. All rights reserved. doi:10.1016/j.trac.2011.06.012

Different approaches can be used to improve the sequence coverage of protein mixtures. Moreover, analysis times are often relatively long and new columns can be used to increase the speed. The availability of various separation modes offers a high degree of flexibility, while coupling multiple (orthogonal) LC modes [e.g., strong cation exchange (SCX) and reversed phase (RP)] provides a great increase in separation power [3].

Recently, novel column types [e.g., columns packed with relatively small (<2 μm) particles for ultra-high performance LC (UPLC), monolithic columns and hydrophilic interaction chromatography (HILIC) materials] have been introduced in proteomics. In this review, we focus on the use of these materials for the analysis of protein digests. We discuss the suitability and the potential of these columns in both 1D and 2D methods and illustrate them with relevant applications. We compare the systems on the basis of parameters relevant to the routine analysis of highly complex samples (i.e. efficiency of separation, speed and sensitivity).

2. Reversed-phase materials

2.1. Ultra-high-performance liquid chromatography (UPLC)

Due to the complexity of the samples, digest analysis is typically performed using gradient-LC methods. The performance of these systems is usually expressed as the peak capacity (PC), which is defined as the maximum number of peaks that will fit within the applied elution window with a resolution of 1.0. It is calculated by dividing the length of the gradient by the average peak width. PC is proportional to the square root of plate number N and therefore depends on column length and particle size. PC also increases with longer gradients, but ultimately reaches a maximum when the increase in peak width due to the shallow gradient cancels out the advantage of a longer gradient time. Increase in length and decrease in particle size cause an increase in column backpressure, which limits their application. This limitation can be overcome using LC equipment designed to cope with these increased pressures. Table 1 gives an overview of the use of UPLC systems in proteomics.

The use of capillary RP columns packed with small C18-modified silica particles, operated at very high pressures, was pioneered by the group of Jorgenson in the late 1990s [4]. With a system using 30-μm id columns packed with 1.0 μm non-porous particles, operating at pressures in excess of 2500 bar, analysis of fluorescent-labeled ovalbumin tryptic peptides using laser-induced fluorescence detection was performed in approximately 30 min (Fig. 1) [4]. A peak capacity of about 300 was achieved with peak widths increasing from 3–4 s at the beginning of the chromatogram to about 12 s at the end. This is a consequence of the exponential gradient that was used. This UPLC technique was also combined with MS and was used to

Table 1. Overview of ultra-high performance LC systems for proteomics applications

Sample	Column	Stationary Phase	Analysis time	Remarks	Ref.
One-dimensional systems					
Ovalbumin digest	27 cm × 30 μm ID	1.0 μm nonporous C18-bonded silica	30 min	LIF detection, labeled sample	[4]
Various protein digests	22 cm × 150 μm ID	1.5 μm nonporous C18-bonded silica	30 min		[5]
Human Serum	10 cm × 2.1 mm ID	1.7 μm Acquity BEH C_{18}	5 min		[6]
β-casein digest	5 cm × 2.1 mm ID	1.7 μm Acquity BEH C_{18}	7.5 min		[17]
Protein kinase α digest	15 cm × 75 μm ID	1.7 μm Acquity BEH C_{18}	30 min		[7]
Botulinum neurotoxin digest products	5 cm × 2.1 mm ID	1.7 μm Acquity BEH C_{18}	2 min		[8]
5-protein digest	15–50 cm × 75 μm ID	1.7 and 3.0 μm C18-bonded silica	24-432 min		[9]
Shewanella oneidensis lysate	40 cm × 50 μm ID	1.4 μm porous C18-bonded silica	120–720 min		[10]
S. oneidensis lysate	20 cm × 50 μm	0.8 μm porous C18-bonded silica	8–50 min		[11]
Yeast lysate	87 cm × various ID	3 μm porous C18-bonded silica	180 min		[12]
Yeast lysate	150 cm × 50 μm ID	3 μm porous C18-bonded silica	200 min	Two parallel columns	[14]
Yeast lysate	80 cm × 150 μm ID	1.5, 2, 3 μm porous C18-bonded silica	180 min	Multiple columns in parallel	[15]
Yeast lysate	100 cm × 75 μm ID	3 μm porous C18-bonded silica	120 min	Two parallel columns, SPE preconcentration	[18]
Multidimensional systems					
Yeast lysate	5 cm × 100 μm ID 50 cm × 50 μm ID	5 μm Partisphere SCX 3 μm Aqua C18	7–27 h	Ultra-high-pressure MudPIT	[16]
Human plasma lysate	80 cm × 320 μm ID 85 cm × 30 μm ID	3 μm polysulfoethyl aspartamide 3 μm porous C18-bonded silica	About 80 h	Off-line system; 200 min SCX gradient, 15 fractions, 300 min RP gradient	[19]