Separation methodology to improve proteome coverage depth

Christof Lenz & Henning Urlaub

Published online: 16 May 2014.

Submit your article to this journal ↗

Article views: 862

View related articles ↗

View Crossmark data ↗

Citing articles: 6 View citing articles ↗

EXPERT REVIEWS

Separation methodology to improve proteome coverage depth

Christof Lenz* and Henning Urlaub

Bioanalytical Mass Spectrometry Group, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany
and
Bioanalytics, Institute of Clinical Chemistry, University Medical Center Göttingen, Robert-Koch-Strasse 40, 37075 Göttingen, Germany
*Author for correspondence:
Tel.: +49 055 1201 1060/1478
christof.lenz@mpibpc.mpg.de

So-called 'in-depth proteomics' and its applied separation methodology to improve the proteome coverage depth has become an important issue in mass spectrometric-based proteomics and system-wide cell biology studies. Employing a bottom-up approach and a variety of separation techniques, it allows for identification of proteins with low copy numbers and enables researchers to correlate the number of expressed genes in a cell with the proteome. Here we describe recent advances in this field with emphasis on peptide and protein separation technologies. The discussion is focused both on single injection analyses employing long reversed phase liquid chromatography separations of peptides ('single shot proteomics') and on the combination of orthogonal protein and peptide separation methods to achieve maximum protein coverage. Owing to these improvements, in-depth proteomics has now fully entered the field and is being implemented in an increasing number of laboratories.

KEYWORDS: data dependent acquisition • in-depth proteomics • orthogonal and multidimensional protein and peptide separation • single shot proteomics

In recent years, the main foci in mass spectrometry (MS)-based proteomics have been the issues of protein quantification and the analysis of posttranslational modifications. While these topics still receive a lot of attention, another aspect has more recently entered the field of MS-based proteomics, namely the challenge of identifying as many proteins as possible from a sample – often termed 'in-depth proteomics' – especially in cell biology studies.

In-depth proteome analysis with sufficient numbers of identified proteins promises to allow for correlating the number of gene-coding sequences in a genome with the results of transcriptome analyses, with the aim of understanding variations in the genomes of cell types and the patterns of gene expression and translation, a goal often referred to as 'proteogenomics' [1]. The latter is exemplified by the chromosome-centric human proteome project [2], which aims to identify at least one protein encoded by each of the approximately 20,200 gene-coding sequences present in the Uniprot/Swissprot database.

In-depth proteome analysis presents a daunting challenge for current analytical approaches and technology. Proteome analysis is most often carried out using a 'bottom-up' approach, where proteins are represented by peptide proxies generated by proteolytic digestion. Following the mass spectrometric identification, characterization and quantitation of the peptides, the identity, modification state and quantity of the underlying proteins are then assembled from the experimental data through bioinformatics means. This approach multiplies the number and diversity of analytes that need to be handled by mass spectrometric analysis. More importantly though, protein abundances, for example, in human cells span an enormous range, from 1 to 10^7 copies per cell – a difference of seven orders of magnitude, while the corresponding range of transcribed genes runs only from 1 to 10^4 [3]. In human body fluids, this concentration range can be even wider, for example, in human plasma where it is considered to span 12 orders of magnitude [4,5]. In contrast to transcriptome/RNA-Seq analyses of genes, proteins as the actual gene products cannot be amplified in order to detect them, so the concentration range of proteins in a sample translates more or less directly into the linear dynamic range required from any analytical set-up used. Consequently, MS-based proteomics in body fluids

usually involves additional prefractionation or enrichment steps [6] and is focused on targeted mass spectrometric detection and quantitation of proteins rather than global analysis [7].

In spite of these challenges, significant progress has been made, so that nowadays more than 10,000 proteins can be identified from cultures of mammalian cell lines [8–10]. Above all, this progress is attributed to the technical improvements realized in the latest generation of mass spectrometers, with gains in sensitivity, speed and resolution that allow scientists to deal with highly complex sample mixtures. State-of-the-art electrospray ionization mass spectrometers are capable of sequencing speeds of up to 20 peptide precursors per second in data-dependent acquisition (DDA) mode. As not all sequencing events are successfully converted into a unique peptide sequence, this translates to the identification of at most 8–9 peptide sequences per second [11]. Instrumental improvements in ion sampling and transmission, scanning speed, resolution, mass accuracy and signal processing/digitization allow for impressively complex peptide mixtures to be eluted into the mass spectrometer for in-depth analysis [12–14]. In parallel with improvements in MS hardware, improvements in separation workflows for proteins and peptides prior to MS analysis have also contributed significantly. Even the fastest mass spectrometers are still not able to cope with the demands put up by the complexity of whole proteome samples [15]. Consequently, two separation approaches have been continuously investigated and optimized: On the one hand, a number of groups have optimized the nanoflow reversed-phase liquid chromatography (RP-LC) gradient separation of tryptic peptides that is typically hyphenated to the mass spectrometer. On the other hand, additional dimensions of separation both on the peptide and on the protein level are increasingly being utilized to maximize the depth of analysis.

The performance of gradient separations of peptides is best measured by the achieved peak capacity, that is, the number of peaks that can be separated across the gradient time. For nanoflow RP-LC-based separation of peptides, mainly two factors have been utilized to drive increases in peak capacity: the column length and the size of stationary phase particles [16]. Whereas conservative set-ups were using relatively short (15–120 min) gradients on 50–75 µm inner diameter, 10–15 cm length columns with 3 µm particles for routine operations, many groups are now utilizing longer gradients of up to 10 h on longer columns (25–100 cm) packed with smaller (sub-2 µm) particle sizes to achieve peak capacities in excess of 450 [17–19]. To counter the corresponding increase of backpressure that comes with both increased column length and decreased particle size, ultrahigh pressure liquid chromatography systems capable of up to 1000 bar operating pressure have become the industry standard. In addition, column ovens are used to heat columns and solvent lines, which, besides reducing the back-pressure, also improves the recovery especially of hydrophobic peptides [20].

Another way to increase the performance of RP peptide separations hyphenated to MS includes further miniaturization, for example, by using columns with an inner diameter of 25 µm or smaller. The corresponding decrease in flow rates to the region of 10 nl/min is of great benefit for the nanospray process at the mass spectrometer interface: it enhances the ionization efficiency of peptides and reduces suppression effects between competing analytes, thereby increasing the number of detectable and fragmentable peptides in the MS [21–23]. The technical challenges involved in producing and handling gradient separations at these flow rates have prohibited this approach from becoming routine though.

Owing to these improvements in RP gradient separations, one may ask whether a rather simple workflow for in-depth proteome analysis – comprising hydrolysis of the sample with endoprotease, separation of the peptides by long capillary columns and identification of peptides and thus the proteins by a current ultra-fast and sensitive MS instrument – might already be sufficient to cover the proteome of a cell. Indeed, several of the leading groups in the field of MS-based proteomics have pursued this approach to establish 'single-shot proteomics', that is, an in-depth characterization of the proteome using just a single-injection LC/MS/MS analysis.

However, the current limit for this type of analysis is the identification of approximately 5000 proteins from yeast, and up to 8000 proteins from human cell cultures [8,24,25]. While in studies with yeast this number is sufficient to cover approximately 90% of the expressed proteome [11,26], in human currently only up to 50–60% of the expressed proteome are covered, corresponding mainly to the upper half of the expressed proteins' concentration range.

An interesting perspective on the capabilities and limitations of 'single-shot proteomics' using long reversed separations is given in a review by Zubarev [27], who correlates the sample amount, the column length and the gradient time necessary to reach an analyses depth to cover the percentage of the low-abundance proteins in a proteome. He points out that for reaching a depth of >9000 proteins in a mammalian cell line, at least 1 mg of sample (corresponding to 5 million cells) is required, and also considers the working-time component of in-depth analyses.

From this perspective, 'single-shot proteomics' using current state-of-the-art technology – while presenting a major technological advance – is unlikely to achieve full coverage of the expressed cellular proteome in mammalian systems, even when only protein identification is required. There are two principal reasons for this: First, the dynamic range and sensitivity of the MS and the sample amount required to identify the low-abundance proteins in the range of 10 copies per cell preclude the exclusive use of one-dimensional peptide separation before MS analysis (note that only a maximum of a few µg can be loaded on capillary columns). Second, the proteome consists not only of plain proteins, that is, the gene-coded protein sequences, but also, for example, of proteins that are heavily modified by known and unknown (post)translational events, protein isoforms and naturally truncated proteins. All these further complicate the identification especially of low-abundance

proteins, as the available capacity of the LC–MS system for fragmenting peptide precursors becomes even more heavily utilized. In addition, there is a strong stochastic element in discovery-driven MS analyses as low-abundance peptides are not reliably selected for fragmentation in DDA workflows, hampering the reliable detection of low-abundance proteins [15]. This limitation becomes even more pronounced when the systematic analysis of posttranslational modifications (PTMs) is required [28]. Consequently, additional separation workflows are absolutely indispensable, in order to reduce further the sample complexity of a peptide mixture before it is subjected to the nanoflow RP-LC system coupled to the MS.

Two principal workflows, representing different separation 'philosophies', are frequently applied: (i) addition of further dimensions of separation on the intact protein level, for example, according to apparent protein molecular weight by SDS-PAGE, to isoelectric point by electrophoretic separation (OFFGEL electrophoresis, which is a free-flow isoelectric focusing technique that separates and recovers proteins directly in solution for further analysis [29,30]), or to hydrodynamic radius by size-exclusion chromatography [10] before hydrolysis with endopeptidases; (ii) hydrolysis of the protein sample in solution and subsequent application of additional separations on the peptide level, usually various chromatographic workflows to prefractionate the resulting highly complex peptide mixtures before these are submitted to the nanoflow RP-LC front end of the mass spectrometer.

For protein separation, the approach most frequently taken is still SDS-PAGE. Entire lanes covering a broad range of apparent molecular weight are cut into slices, proteins of the various slices are digested in the gel and extracted peptides from each individual slice are subjected to LC–MS. An advantage of this method is that separation by SDS-PAGE is still the best method in terms of separation range and resolution for entire proteins. In addition, SDS-PAGE is compatible with a very wide range of protein solubilization and sample handling requirements [31,32]. Consequently, when peptides derived from various molecular weight regions are then separated with a capillary column packed with RP material of smaller particle size and operated by ultrahigh pressure liquid chromatography pumps [33], the analysis depth of a sample increases dramatically, and up to 8000 proteins from less than 100 μg of sample material from a human cancer cell line can be identified on a routine basis [34]. Moreover, depending on the scientific task, only proteins from defined molecular weight regions can be investigated and results easily correlated with, for example, western blot analysis. However, although proteins can be concentrated by several precipitation approaches (e.g., acetone, TCA), SDS-PAGE has only a limited loading capacity, and when this is exceeded the separation capability of SDS-PAGE decreases, with high-abundance proteins being smeared out over a large range of apparent molecular weight, hampering the detection of low-abundance proteins in the various molecular weight regions of the PAGE. Therefore, when larger sample amounts are available, further pre-fractionation of proteins by

size-exclusion chromatography or OFFGEL separation, followed by SDS-PAGE or another chromatographic separation step, is often the better option [10]. The alternative 'philosophy' relies solely on the multidimensional chromatographic separation of peptides generated by endopeptidase digestion of proteins in solution. Ideally, the chromatographic dimensions should be chosen in a way that they are highly orthogonal, that is, separate peptides by largely independent physicochemical principles; they can either be coupled on-line (LC × LC) or off-line following fractionation in the first dimension [35]. The first highly successful example of a 2D, online peptide separation was the so-called MudPIT approach, in which peptide mixtures are separated first by strong cation-exchange (SCX) chromatography and then by nanoflow RP chromatography. In the originally described approach, both chromatographic stages were combined in a single capillary emitter directly connected to the mass spectrometer [36]. More robust, and more widespread in use, is the off-line separation of peptide mixtures by SCX chromatography followed by the application of single fractions to the nanoflow RP-LC–MS system. Moreover, anion-exchange chromatography has recently been successfully used as the first dimension of peptide separation [37]. Note that most ion-exchange chromatography protocols operate with salt gradients, which at increasing salt concentrations interfere with the subsequent nanoflow RP-LC–MS analysis, despite the fact that RP capillaries are usually equipped with pre-columns for pre-concentration and desalting of the peptide mixtures. A notable exception is the approach published by Deeb *et al.* [38], who used pH steps at constant salt strength rather than increasing salt concentrations to fractionate by strong anion exchange.

To avoid the 'salt problem' associated with ion-exchange chromatography, many groups have implemented the concept of 'orthogonal' RP separations [39]. Here, peptide mixtures are separated by RP chromatography at neutral or basic pH in the first dimension, then fractionated and the fractions analyzed by nanoflow RP-LC–MS under conventional conditions, that is, at pH 2–3, in the second dimension. The two chromatographic dimensions in this set-up show only limited orthogonality; however, this partially compensated for by the fact that RP separations provide significantly higher chromatographic resolution than ion-exchange separations. In addition, sample transfer from the first to the second dimension is achieved simply by removing excess organic solvent used for the first dimension gradient elution, and the first dimension can be easily up-scaled to analytical or even semi-preparative column dimensions [40].

A completely different, but also very powerful approach is the use of peptide isoelectric focusing (pIEF) in the first dimension either on gel strips or with the above-mentioned OFFGEL system [41,42]. Instead of proteins separated by their isoelectric point on gel strips carrying ampholytes, peptides are separated into defined pH regions according to their pI [43]. When using gel strips with ampholytes, peptides can be eluted from the gel strips, as in the in-gel digestion of proteins, and further analyzed by nanoflow RP-LC–MS. OFFGEL allows for recovering peptides from solution. However, when only pIEF

is used in the first dimension, the separation capacity is comparable to that of the above-mentioned LC workflows. Very recently, two interesting reports described the expansion of the pIEF approach. In a proteogenomic approach, Branca *et al.* [44] used high-resolution pIEF in a very narrow pH range (3.7–4.9) combined with the calculated pI values of all tryptic peptides predicted from genomic sequences that are translated into six open-reading frames. Using this approach, they identified numerous peptides not present in the Peptide Atlas [45] and several novel protein-coding gene loci. Atanassov and Urlaub [46] combined pIEF with in-gel digestion of proteins and showed that the detection of low-abundance proteins is significantly enhanced when peptides derived from individual gel slices after in-gel digestion are further separated by pIEF, extracted from pIEF slices again and then finally subjected to nanoflow LC–MS.

Finally, in addition to the described improvements in MS equipment as well as in the various separation approaches described above, another issue should be taken into account: sample preparation. Lysis of cells and solubilization are critical issues that are often underestimated when comprehensive in-depth MS-based analyses are designed. Of the reagents used so far, SDS has proven to have by far the best solubilization properties. The fact that SDS acts so powerfully in solubilization of protein material is also a factor in favor of the use of SDS-PAGE as a first step in protein separation for in-depth proteome analysis (see above). However, SDS as a non-volatile surfactant is not compatible with MS *per se*, and also not with the above-mentioned LC techniques for peptide separation, so it must be removed from the protein/peptide mixture. This problem was solved by the introduction of FASP (filter-aided sample preparation) where samples (cells, tissue, etc.) are solubilized and denatured in the presence of high concentrations of SDS, which is then exchanged by filtration through urea in order to hydrolyze the samples under these conditions [47]. Urea can be removed by repeated filtration, and peptides are recovered from the filter. This preparation method has proved its feasibility in various in-depth MS-based proteome analyses when combined with the above-mentioned separation techniques for complex mixtures, such as SCX chromatography, strong anion-exchange chromatography, pIEF or the use of long, heated capillary nanoflow RP-LC columns coupled to the mass spectrometer. It is especially useful when dealing with minute sample amounts, for example from formalin-fixed, paraffin-embedded clinical tissue specimens [48,49].

Expert commentary

In-depth proteomics has entered the field of MS-based proteomics in cell biology and is being implemented in an increasing number of laboratories. Impressive improvements in MS instruments, combined with LC systems and capillary columns packed with material of relatively small particle size that allows much sharper peptide separation, yield spectacular numbers of fragmented peptides and thus of proteins identified. Straightforward sample-preparation techniques, along with the use of novel fractionation approaches at the protein and peptide levels, further enhance protein identification. A variety of two-dimensional separation strategies are available today to tackle samples of high complexity and high dynamic range. The choice of strategy depends on the known properties of the analytical system in hand, that is, the expected complexity and dynamic range, and also on very practical questions – especially of sample transfer between separation steps and into the mass spectrometer.

However, the use of additional dimensions of separation greatly increases the instrument time spent on in-depth proteome analyses, requires higher amounts of starting material and – on account of the necessary fractionation – reduces sample-to-sample reproducibility. Consequently many laboratories have developed the highly desirable 'single-shot' proteome analysis, where the one-dimensional separation of peptide mixtures on nanoflow RP-LC linked to the mass spectrometer produces impressive results. While it does not yet allow a comprehensive identification of expressed proteins including low-abundance ones, it allows the analysis of significant numbers of biological and technical replicates with increased reproducibility. As its dynamic range is limited compared with that of methods employing multi-dimensional separation, the single-shot strategy currently seems best suited for the analysis of samples obtained *in vitro* (e.g., cultured cell lysates) rather than those obtained *in vivo* – especially body fluids, which exhibit a high dynamic range of protein expression.

Five-year view

Coming improvements in MS instrumentation working in DDA mode, combined with increasingly powerful separation approaches, will certainly allow us to dig even more deeply into the proteomes of cells or other biological systems. While the identification of, for example, up to 90% of the proteins expressed in yeast is an impressive achievement in itself, it comes at the cost of low sequence coverage and insufficient characterization of, for example, PTMs for most low-abundance proteins. Consequently, the goal of 'in-depth proteomics' is still far from being reached, and even further improvements will still leave a lot of room for focused approaches, for example, for PTM profiling.

A change in paradigm is available through the so-called data-independent acquisition (DIA) strategies. Instead of stochastically selecting peptides for fragmentation analysis by appropriate settings of the mass spectrometer's acquisition software, DIA strategies aim at recording MS/MS libraries of all peptides present in a complex proteome sample [50,51]. These libraries can be re-interrogated after analysis, enabling researchers to perform iterative analysis on the existing data and extracting qualitative and quantitative information independent of the acquisition parameters. While DIA strategies have not yet reached the same depth of analysis as the highly refined DDA approaches, they promise the steepest slope of improvement for in-depth proteomics analysis in the next 5 years.

Financial & competing interests disclosure

Key issues

- 'In-depth proteomics', that is, the attempt to identify all proteins expressed and hence present in a biological system.
- Significant analytical challenges from protein complexity and especially dynamic range, since no protein-amplification technology is available.
- Chromatographic resolution and acquisition speed of the mass spectrometer are maximized to compensate for sample complexity.
- 'Single-shot proteomics' to minimize the impact of additional dimensions of protein/peptide separation on instrument time and reproducibility.
- 90% proteome coverage in yeast and 50–60% in human cell culture are made possible by 'single-shot' analyses.
- Multi-dimensional protein and peptide separations still add an additional level of analysis depth.
- Re-annotations of genomes are performed on the basis of 'in-depth proteomics' studies.

References

Papers of special note have been highlighted as:
• of interest
•• of considerable interest

1. Ansong C, Purvine SO, Adkins JN, et al. Proteogenomics: needs and roles to be filled by proteomics in genome annotation. Brief Funct Genomic Proteomic 2008;7:50-62

2. Paik YK, Jeong SK, Omenn GS, et al. The chromosome-centric human proteome project for cataloging proteins encoded in the genome. Nat Biotechnol 2012;30(3): 221-3

3. Schwanhäusser B, Busse D, Li N, et al. Global quantification of mammalian gene expression control. Nature 2011;473(7347): 337-42

4. Anderson NL. Counting the proteins in plasma. Clin Chem 2010;56(11):1775-6

5. Hortin GL, Sviridov D. The dynamic range problem in the analysis of the plasma proteome. J Proteomics 2010;73(3):629-36

6. Whiteaker JR, Zhang H, Eng JK, et al. Head-to-head comparison of serum fractionation techniques. J Proteome Res 2007;6:828-36

7. Anderson NL, Anderson NG, Pearson TW, et al. A Human proteome detection and quantitation project. Mol Cell Proteomics 2009;8(5):883-6

8. Mann M, Kulak NA, Nagaraj N, Cox J. The coming of age of complete, accurate, and ubiquitous proteomes. Mol Cell 2013; 49(4):583-90

•• **A concise overview of the aims, technical means, state-of-the-art and implications of in-depth proteome analysis.**

9. Beck M, Schmidt A, Malmstroem J, et al. The quantitative proteome of a human cell line. Mol Syst Biol 2011;7:549

10. Nagaraj N, Wisniewski JR, Geiger T, et al. Deep proteome and transcriptome mapping of a human cancer cell line. Mol Syst Biol 2011;7:548

11. Hebert AS, Richards AL, Bailey DJ, et al. The one hour yeast proteome. Mol Cell Proteomics 2013;13(1):339-47

12. Kelstrup CD, Young C, Lavallee R, et al. Optimized fast and sensitive acquisition methods for shotgun proteomics on a quadrupole orbitrap mass spectrometer. J Proteome Res 2012;11:3487-97

13. Andrews GL, Simons BL, Young JB, et al. Performance characteristics of a new hybrid quadrupole time-of-flight tandem mass spectrometer (TripleTOF 5600). Anal Chem 2011;83(13):5442-6

14. Senko MW, Remes PM, Canterbury JD, et al. Novel parallelized quadrupole/linear ion trap/orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates. Anal Chem 2013; 85(24):11710-14

15. Michalski A, Cox J, Mann M. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. J Proteome Res 2011;10: 1785-93

• **An analysis of the challenges of in-depth proteome analysis specifically for high resolution mass spectrometry.**

16. Sandra K, Moshir M, D'hondt F, et al. Highly efficient peptide separations in proteomics. Part 1. Unidimensional high performance liquid chromatography.

J Chromatogr B Analyt Technol Biomed Life Sci 2008;866:48-63

17. Liu H, Finch JW, Lavallee MJ, et al. Effects of column length, particle size, gradient length and flow rate on peak capacity of nano-scale liquid chromatography for peptide separations. J Chromatogr A 2007;1147:30-6

18. Köcher T, Pichler P, Swart R, Mechtler K. Analysis of protein mixtures from whole-cell extracts by single-run nanoLC-MS/MS using ultralong gradients. Nat Protoc 2012; 7(5):882-90

19. Hsieh EJ, Bereman MS, Durand S, et al. Effects of column and gradient lengths on peak capacity and peptide identification in nanoflow LC-MS/MS of complex proteomic samples. J Am Soc Mass Spectrom 2013;24: 148-53

20. Blackler AR, Speers AE, Wu CC. Chromatographic benefits of elevated temperature for the proteomic analysis of membrane proteins. Proteomics 2008;8(19): 3956-64

21. Schmidt A, Karas M, Dülcks T. Effect of different solution flow rates on analyte ion signals in Nano-ESI MS, or: when does ESI turn into Nano-ESI? J Am Soc Mass Spectrom 2003;14:492-500

22. Zhou F, Lu Y, Ficarro SB, et al. Nanoflow low pressure high peak capacity single dimension lc-ms/ms platform for high-throughput, in-depth analysis of mammalian proteomes. Anal Chem 2012;84:5133-9

23. Zhou F, Lu Y, Ficarro SB, et al. Genome-scale proteome quantification by DEEP SEQ mass spectrometry. Nat Commun 2013;4:2171

24. Thakur S, Geiger T, Chatterjee B, et al. Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. Mol Cell Proteomics 2011; 10(8):M110.003699

25. Pirmoradian M, Budamginta H, Chingin K, et al. Rapid and deep human proteome analysis by single-dimension shotgun proteomics. Mol Cell Proteomics 2013; 12(11):3330-8

26. Nagaraj N, Kulak NA, Cox J, et al. System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top orbitrap. Mol Cell Proteomics 2012;11(3):M111.013722

27. Zubarev RA. The challenge of the proteome dynamic range and its implications for in-depth proteomics. Proteomics 2013; 13(5):723-6

•• **An insightful viewpoint on the challenge of dynamic range for in-depth proteome analysis.**

28. Olsen JV, Mann M. Status of large-scale analysis of post-translational modifications by mass spectrometry. Mol Cell Proteomics 2013;12(12):3444-52

• **The authors describe the state-of-the-art for the global analysis of posttranslational protein modifications, for example, phosphorylation.**

29. Ros A, Faupek M, Mees H, et al. Protein purification by Off-Gel electrophoresis. Proteomics 2002;2:151-6

30. Michel PE, Reymond F, Arnaud IL, et al. Protein fractionation in a multicompartment device using Off-Gel isoelectric focusing. Electrophoresis 2003;24:3-11

31. Wilm M, Shevchenko A, Houthaeve T. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 1996;379(6564): 466-9

32. Lundby A, Olsen JV. GeLCMS for in-depth protein characterization and advanced analysis of proteomes. Methods Mol Biol 2011;753:143-55

33. Piersma SR, Warmoes MO, de Wit M, et al. Whole gel processing procedure for GeLC-MS/MS based proteomics. Proteome Sci 2013;11:17-24

34. Wu Z, Gholami AM, Küster B. Systematic identification of the HSP90 regulated proteome. Mol Cell Proteomics 2012;11(6): M111.016675

35. Sandra K, Moshir M, D'hondt F, et al. Highly efficient peptide separations in proteomics. Part 2. Bi- and multidimensional liquid-based separation techniques. J Chromatogr B Analyt Technol Biomed Life Sci 2009;877:1019-39

36. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 2001;19(3): 242-7

37. Ritorto MS, Cook K, Tyagi K, et al. Hydrophilic strong anion exchange (hSAX) chromatography for highly orthogonal peptide separation of complex proteomes. J Proteome Res 2013;12(6):2449-57

38. Deeb SJ, D'Souza RCJ, Cox J, et al. Super-SILAC allows classification of diffuse large b-cell lymphoma subtypes by their expression profiles. Mol Cell Proteomics 2012;11:77-89

39. Toll H, Oberacher H, Swart R, Huber J. Separation, detection, and identification of peptides by ion-pair reversed-phase high-performance liquid chromatography-electrospray ionization mass spectrometry at high and low pH. J Chromatogr A 2005; 1079(1-2):274-86

40. Dwivedi RC, Spicer V, Harder M, et al. Practical implementation of 2D HPLC scheme with accurate peptide retention prediction in both dimensions for high-throughput bottom-up proteomics. Anal Chem 2008;80(18):7036-42

41. Hörth P, Miller CA, Preckel T, Wenz C. Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis. Mol Cell Proteomics 2006; 5(10):1968-74

42. Hubner NC, Ren S, Mann M. Peptide separation with immobilized pI strips is an attractive alternative to in-gel protein digestion for proteome analysis. Proteomics 2008;8:4862-72

43. Cargile BJ, Sevinsky JR, Essader AS, Stephenson JL Jr, Bundy JL. Immobilized pH gradient isolelectric focusing as a first-dimension separation in shotgun proteomics. J Biomol Tech 2005;16(3): 181-9

44. Branca RM, Orre LM, Johansson HJ, et al. HiRIEF LC/MS enables deep proteome coverage and unbiased proteogenomics. Nat Methods 2014;11(1):59-62

45. PeptideAtlas. Available from: www.peptideatlas.org

46. Atanassov I, Urlaub H. Increased proteome coverage by combining PAGE and peptide isoelectric focusing: comparative study of gel-based separation approaches. Proteomics 2013;13(20):2947-55

47. Wiśniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Methods 2009;6(5):359-63

48. Wiśniewski JR, Ostasiewicz P, Mann M. High recovery FASP applied to the proteomic analysis of microdissected formalin fixed paraffin embedded cancer tissues retrieves known colon cancer markers. J Proteome Res 2011;10:3040-9

49. Wiśniewski JR, Duś K, Mann M. Proteomic workflow for the analysis of archival formalin-fixed and paraffin-embedded clinical samples to a depth of 10000 proteins. Proteomics Clin Appl 2013;7(3-4):225-33

50. Gillet LC, Navarro P, Tate S, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics 2012;11(6):O111.016717

• **A key publication on data-independent acquisition, a potential paradigm change for in-depth proteome analysis.**

51. Chapman JD, Goodlett DR, Masselon CD. Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. Mass Spectrom Rev 2013. [Epub ahead of print]