FEBS Letters 567 (2004) 92–95 FEBS 28440

Minireview

Miniaturized proteomics and peptidomics using capillary liquid separation and high resolution mass spectrometry

Margareta Ramström, Jonas Bergquist*

Department of Chemistry, Analytical Chemistry, Biomedical Centre, Uppsala University, P.O. Box 599, SE-751 24 Uppsala, Sweden

Received 23 April 2004; accepted 28 April 2004

Available online 8 May 2004

Edited by Horst Feldmann

Abstract Knowledge of the protein and peptide content in a tissue or a body fluid is vital in many areas of medical and biomedical sciences. Information from proteomic and peptidomic studies may reveal alterations in expression due to, e.g., a disease and facilitate the understanding of the pathophysiology and the identification of biological markers. In this minireview, we discuss miniaturized proteomic and peptidomic approaches that have been applied in our laboratory in order to investigate the protein and peptide contents of body fluids (such as plasma, cerebrospinal and amniotic fluid), as well as extracted tissues. The methods involve miniaturized liquid separation, i.e., capillary liquid chromatography and capillary electrophoresis, combined with high resolution mass spectrometry (MS), i.e., Fourier transform ion cyclotron resonance MS. These approaches provide the opportunity to analyze samples of small volumes with high throughput, high sensitivity, good dynamic range and minimal sample handling. Also, the experiments are relatively easy to automate.

© 2004 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Proteomics; Peptidomics; Liquid separation; Mass spectrometry

1. Introduction

Proteomics and peptidomics, the studies of all proteins and peptides expressed in an organism, a tissue or a body fluid at a given times are fields of great importance in biological and medical research. Investigations of the protein and peptide composition of a body fluid, the sequences and structures of these analytes as well as possible post-translational modifications need to be performed before a description of a proteome or peptidome can be provided. It should be noted that proteomes and peptidomes are not constant; they are influenced by, e.g., metabolism, disease states, stress, and drug interactions. The possibility to monitor alterations in protein and peptide patterns, that can be correlated to a certain disease, is valuable for a deeper understanding of its etiology and will facilitate its diagnosis.

Several aspects have to be taken into account when developing or choosing a method for proteomic studies. The samples under study are often very complex and the components are present in a wide concentration range. Analytes of very different chemical properties, e.g., hydrophobicity, size and p*I*, should be detected in the same experiment. Hence, sensitive and non-discriminating methods need to be applied. The analysis time for running one sample should be short and the method easy to automate. Also, the experiments generate an enormous amount of data, which means that there are high demands on data-handling and bioinformatics.

The traditional proteomic approach starts with separation of the proteins by two-dimensional sodium dodecyl sulfate–polyacrylamide gel electrophoresis (2D-PAGE) [1,2], where the proteins are separated in two dimensions with respect to their isoelectrical points (p*I*-values) and sizes, respectively. The spots on the gels are then selected and the proteins are enzymatically digested before they are analyzed, e.g., by MALDI-TOF mass spectrometry (MS). This method has proved to be of great use for many applications. Alternative approaches that are becoming more popular are based on liquid separation in one or several dimensions followed by MS or tandem mass spectrometry (MS/MS) [3]. The experiments start with enzymatic digestion (often tryptic cleavage) of all proteins present in the sample, whereby the resulting peptides are separated and detected. The advantages of these methods are that they are generally much faster and easier to automate than the 2D-PAGE approach. In "shot-gun" proteomics, the identification of the proteins is based on sequences information on the tryptic peptides. Alternatively, accurate mass tags can be used to identify the components.

There is no clear cut-off definition between proteins and peptides, but most of those molecules referred to as peptides have masses of <20 kDa. The traditional 2D-approach is restricted to proteins in the mass range of around 10–200 kDa. Hence, this tool cannot be used in the field of peptidomics. Instead, the classical procedure for investigations of peptidomes involves purification of peptides from tissue extracts using several consecutive liquid chromatography (LC)-steps [4,5]. After purification, the peptides can be characterized using either Edman degradation or MS. Today, MS or preferentially MS/MS is the dominating technique for investigation of peptides. There is no need for complete purification of the components and the amount of sample required for the analysis is minor.

In this minireview, we will describe and discuss some non-gel-based methods that we utilize in our laboratory for the identification and characterization of proteins and peptides in

* Corresponding author. Fax: +46-18-471-3692.
E-mail address: jonas.bergquist@kemi.uu.se (J. Bergquist).

0014-5793/$22.00 © 2004 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2004.04.074

complex biological mixtures. These methods are based on packed capillary LC or capillary electrophoresis (CE) coupled on-line to high resolution MS, Fourier transform ion cyclotron resonance (FTICR) MS. LC and CE are based on two different separation mechanisms, in LC the analytes are separated with respect to their hydrophobicity, while size and charge are the important parameters in CE separation. The combination of both methods with electrospray (ESI) MS is today well-established tool for detection and identification of proteins and peptides [6,7]. Great advantages of the approaches are the low sample volume required for one experiment. While capillary LC has a higher peak capacity, CE benefits from being faster and has high separation efficiency. Hence, the methods should be regarded as alternatives and complements to each other. Previously, the combinations of LC and CE with FTICR MS have proved to be useful for the investigation of microbial and mouse cell proteomes [8,9]. The research in our laboratory concerns mainly studies of the protein and peptide composition of different mammalian (e.g., human) body fluids and tissues.

2. Materials and methods

A general description of the experimental procedures used for FTICR MS studies of body fluids and tissues is given below.

2.1. Sample handling

The samples discussed in this review are tryptic global digests of body fluids, e.g., plasma [10,11], CSF [6,11,12], salivary, urine [11] and extracted peptides from, e.g., pancreatic islets [13]. For a general description of sample preparation and handling, see the review by Bergquist et al. [11].

2.2. Liquid chromatography

The rather conventional HPLC-system used in these experiments consisted of two pumps (JASCO 1580, JASCO Japan) that delivered mobile phase gradients at a flow rate of ~1 μL/min. Typically, 10 μL of the sample volume was injected into a six-port injector valve (Valco Instruments Co. Inc., Schenkon, Switzerland). The peptides were separated on a 10-cm long in-house packed capillary C_{18}-column, packing material ∅5 μm ODS-AQ (YMC Europe, Schermbeck, Germany). They passed a UV-detector as a controlled step before they were electrosprayed on-line into the FTICR MS.

2.3. Capillary electrophoresis

The CE instrumentation used in our laboratory was designed and built in-house [6]. A Pyrex glass cylinder (diameter 5 cm, length 10 cm, volume 0.2 L) was fitted between two PVC plastic lids with gas-proof Teflon O-ring seals. Gas-proof tightened PEEK connections for the high voltage cable, inlet end of the capillary, and gas inlet were drilled into the top lid. A sample carousel of Plexiglas was mounted on a PVC piston running through the lid (also through a gastight seal), enabling fast manual sample switching. A vent valve inserted into the top lid adjusted the applied pressure. An outer safety shell of Plexiglas was mounted around the Pyrex glass cylinder. The high voltage power supply (Bertan, Hicksville, NY) operated in constant negative voltage mode. Fused silica capillary 25 μm i.d. 360 μm o.d. (Polymicro Technologies, Tucson, AZ) was cut into 100 cm lengths.

2.4. Mass spectrometry

The electrospray was maintained using the in-house constructed Black Dust (polyimide-graphite) sheathless electrospray emitters [14] in all experiments described in this review. The 9.4 T FTICR mass spectrometer (Bruker Daltonics, Billerica, MA) used in the studies is described in detail elsewhere [15]. In the LC experiments, a mass spectrum was typically collected during 10 s and 256 spectra were collected in each experiment. In the CE-runs, a mass spectrum was recorded during a few seconds.

3. Results

3.1. FTICR MS of body fluids

Many projects in our laboratory concern the development of methods to identify and characterize biomarkers in body fluids of clinical interest. The tool that we have used for detection is high resolution MS, e.g., FTICR MS. In addition to ultra-high resolution, FTICR MS provides very high mass accuracy and high sensitivity [16,17]. When using this approach, the mass measurement errors in a well-calibrated spectrum are on the ppm-level and hence the possible peptides corresponding to a certain detected mass are limited. Our approach starts with the simultaneous cleavage of all proteins in the samples, followed by MS detection of the peptides.

Direct infusion experiments of tryptic digests of plasma, CSF, salivary and urine, demonstrated the possibility to detect several interesting proteins in all body fluids under study [11], taking advantage of the high resolving power and the mass accuracy of the MS. This approach is fast, a mass spectrum can be obtained within seconds, and the experimental setup is rather straightforward. A direct infusion mass spectrum of a tryptic digested body fluid will hence give an overview of its main components and the complexity of the sample. This should be regarded as a fast screening method. To further improve the sensitivity of the analysis, LC [10,12] and CE [18] have been coupled prior to the mass spectrometer. When introducing a separation step prior to MS, ion suppression in the spray will be reduced and the spectra will be less complex. Hence, the possibility to observe peptides from low abundant components will increase and also the elution times of the peptides provide additional information. The information achieved from one experiment is very extensive. For example, LC-FTICR experiment of CSF typically results in 4000–7000 detected peptides, while around 1500 peptides are detected in a CE-experiment. Hence, the LC-approach provides more information and accepts higher sample load, while a CE experiment is faster to perform. Fig. 1 shows a typical mass chromatogram of tryptic digested CSF.

A great advantage of the presented method is that it is rather general and does not need to be optimized for different body fluids. LC-FTICR MS experiments have been performed on tryptic digests of CSF [12], plasma [10] and also on amniotic fluid (unpublished data), utilizing similar sample preparation

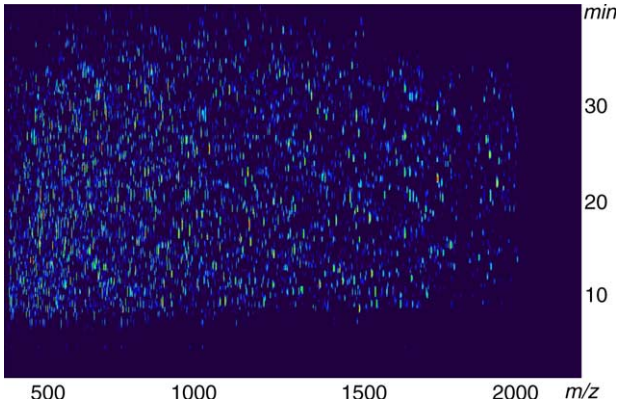

Fig. 1. An overview of an LC-FTICR mass chromatogram of a tryptic digest of CSF. Typically, 4000–6000 peptides are detected in one experiment. The most abundant peptides are those from albumin.

and experimental setup. The protein patterns from these biological mixtures are diverging but show some similarities. In all cases, the dominating peptides are those from the abundant protein HSA. These peptides are spread throughout the mass chromatogram and can be used as internal calibrants. In all studies, proteins of clinical interest were identified. The number of identified proteins were comparable to those achieved running a traditional 2D-gel experiment. In the plasma analysis, the LC-FTICR MS approach was demonstrated to be a good complement or an alternative to LC-iontrap MS/MS [10]. The results from these two studies also agreed well, several different groups of proteins were identified, e.g., transport proteins, immunoglobulins, glycoproteins, coagulation factors, enzymes and inhibitors.

3.2. Data handling

As described above, the analysis of a tryptic digest from a complex biological sample results in mass chromatograms of thousands of peaks. The mass measurement error using an FTICR could be at the ppm-level. It should, however, be mentioned that to achieve this, a proper calibration needs to be performed. It is a challenging task to calibrate a mass chromatogram from a CE- or a LC-run, since the total charge varies during the chromatographic run and local space charge effects will influence the calibration. The approach that we use is to calibrate internally with respect to known tryptic or endogenous peptides in the body fluid or tissue extract. To correctly identify proteins in the fluid, reliable data algorithm for identification needs to be constructed. When analyzing a peptide pattern using accurate mass measurement, a number of tryptic peptides from one protein need to be detected. In the algorithm that has been used for identification so far, the mass measurement error distributions and the distribution in the protein sequence of the detected peptides have been taken into account. Also, the separation step provides additional information that can be utilized to improve the identification [19]. When these algorithms were applied, around 40 proteins could be identified at a significance level >95% from the ~6000 peptides in the LC-run and 30 proteins from the 1500 peptides in the CE experiment of CSF proteins. The number of identified peptides is in the same order as from traditional 2D-gel electrophoresis [20–22]. However, the time required for one experiment is shorter. The concentration range of the identified proteins covers four orders of magnitudes.

3.3. On-line electron capture dissociation and nozzle-skimmer fragmentation

Tandem mass spectrometry is a powerful tool for the identification and characterization of proteins and peptides. A fragmentation method that is so far only applied in FTICR MS is called electron capture dissociation (ECD). This technique provides a "mild" fragmentation, allowing for fragmentation without the loss of labile post-translational modification. This is an attractive alternative to more established methods. In our laboratory we have shown, for the first time, the possibility to combine high rate ECD with on-line LC [23] and CE [24] for the analysis of tryptic digests of proteins as well as standard peptides. Sequence information from ECD was successfully provided in alternating spectra. We have also applied the LC-ECD FTICR MS approach when studying an extract of peptides from mouse pancreatic islets, and on-line nozzle-skimmer fragmentation was applied to the same sample [13]. Sequence information on components derived from the main components was achieved taking both approaches. Complementary information was achieved due to the different fragmentation mechanisms of the two methods. However, the sensitivity of ECD needs to be further improved before lower abundant components in the samples can be sequenced and studied. If this can be done, we believe in the combination of on-line separation ECD and nozzle-skimmer fragmentation as being a very powerful approach to achieve useful complementary sequence information on peptide extracts as well as tryptic digests.

3.4. Quantification of proteins and peptides

Investigation of a proteome does not only concern what proteins are present, but also in what concentration they are expressed. There are several ways to quantify proteins using MS. One way to perform relative quantification of proteins in, e.g., plasma would be to normalize them to coeluting peptides of albumin. HSA is known to vary by less than 7% between individuals. For better quantification labelling methods need to be applied.

Labelling markers can be divided into two subgroups; those that label only specific amino acid residues in the peptides and those that label chemical groups present in all peptides. The latter techniques are referred to as being global. A novel global labelling technique for comparative proteomics was tested for the analysis of tryptic digests of standard proteins. The markers used were the Quantification-Using-Enhanced-Signal-Tags (QUEST) markers, which are covalently linked to the lysines and N-terminal of all peptides [25]. These markers had previously been described to be useful in MALDI MS experiments, and our study is the first one using ESI and FTICR MS. Both direct infusion experiments and on-line LC-FTICR MS were tested on tryptic digests of standard proteins of known concentration ratios [26]. High sequence coverage of the QUEST-labelled proteins was observed and the results indicated that this method was useful for determining the relative quantities of the proteins. The advantage of using a global labelling technique is that all peptides in a mixture can be quantified. However, the complexity of the samples increases since two pools of peptides labelled with different markers are added. When choosing a method for quantification, these aspects need to be taken into account. We believe that the usage of these markers has a great potential. The next step will be to use the markers in comparative peptidomics and proteomics.

4. Concluding remarks

High resolution MS is a powerful tool in proteomics. Research in our laboratory relies on the combination of LC and CE with FTICR MS. The great advantages of these approaches are that small volumes are needed, that they are fast and rather general. Our results clearly demonstrate the possibility to identify proteins of clinical interest in several body fluids. An application for these novel techniques should be to search for biomarkers of, e.g., a disease. The results from the previous experiments have encouraged us to start a comparative proteomic study of CSF from patients suffering from a neurodegenerative disease and CSF from controls. Mass chromatograms of the two groups have been recorded and

compared using a pattern recognition approach. The results show promises that this approach, after further evaluation, can be used as a diagnostic tool (manuscript submitted to Proteomics). It should also be interesting to include the QUEST-markers in this study, to monitor possible disease-specific alterations. On-line fragmentation, nozzle-skimmer fragmentation and ECD can be applied to identify these proteins or peptides and to characterize post-translational modifications.

Future approaches on the sample side may include single cell proteomics, combining the high sensitivity of our methodology with laser capture microdissection technology, probing of biochemical functions following selected signalling pathways with cellular systems. On the separation side, we are investigating further miniaturization including on-chip LC and CE approaches, multidimensional LC or LC-CE systems as well as monolith columns and sol–gel beds. The challenge in proteomics and peptidomics is to combine the correct techniques with the proper samples and analytes. Many mass spectrometric approaches have proved to be outstanding in this research. The integration of proteomics and bioinformatics still needs to be taken to the next level, and is today the bottleneck for many applications.

Acknowledgements: We thank all our colleagues at Analytical Chemistry and the Division of Ion Physics, Uppsala University who have contributed to this work. Financial support from Knut and Alice Wallenberg, the Swedish Foundation for Strategic Research, the Swedish Society for Medical Research and the Swedish Research Council (Grant 621-2002-5261, 629-2002-6821) is gratefully acknowledged. J.B. has a senior research position at the Swedish Research Council (VR).

References

[1] Klose, J. (1975) Humangenetik 26, 231–243.
[2] O'Farrell, P.H. (1975) J. Biol. Chem. 250, 4007–4021.
[3] Domon, B. and Broder, S. (2003) J. Proteome Res., published on web.
[4] Clynen, E., De Loof, A. and Schoofs, L. (2003) Gen. Comp. Endocrinol. 132, 1–9.
[5] Schrader, M. and Schultz-Knappe, P. (2001) Trends Biotech. 19, S55-60.
[6] Wetterhall, M., Johnson, T. and Bergquist, J. (2002) in: Mass Spectrometry and Hyphenated Techniques in Neuropeptide Research (Silberring, J. and Ekman, R., Eds.), pp. 135–154, Wiley, New York.
[7] Gelpi, E. (1995) J. Chromatogr. A 703, 59–80.
[8] Shen, Y. et al. (2001) Anal. Chem. 73, 1766–1775.
[9] Smith, R.D. et al. (2001) Electrophoresis 22, 1652–1668.
[10] Wu, S.-L., Choudhary, G., Ramström, M., Bergquist, J. and Hancock, W.S. (2003) J. Proteome Res. 2, 383–393.
[11] Bergquist, J., Palmblad, M., Wetterhall, M., Håkansson, P. and Markides, K.E. (2002) J. Mass Spectrom. Rev. 21, 2–15.
[12] Ramström, M., Palmblad, M., Markides, K.E., Håkansson, P. and Bergquist, J. (2003) Proteomics 3, 184–190.
[13] Ramström, M. et al. (2003) Eur. J. Biochem. 270, 3146–3152.
[14] Nilsson, S., Wetterhall, M., Bergquist, J., Nyholm, L. and Markides, K.E. (2001) Rapid Commun. Mass Spectrom. 15, 1997–2000.
[15] Palmblad, M., Håkansson, K., Håkansson, P., Feng, X., Cooper, H.J., Giannakopulos, A.E., Green, P.S. and Derrick, P.D. (2000) Eur. J. Mass Spectrom. 6, 267–275.
[16] Amster, I.J. (1996) J. Mass Spectrom. 31, 1325–1337.
[17] Marshall, A.G., Hendrickson, C.L. and Jackson, G.S. (2000) In: Encyclopedia of Analytical Chemistry, Wiley, Chichester, pp. 11694–11728.
[18] Wetterhall, M., Palmblad, M., Håkansson, P., Markides, K.E. and Bergquist, J. (2002) J. Proteome Res. 1, 361–366.
[19] Palmblad, M., Ramström, M., Markides, K.E., Håkansson, P. and Bergquist, J. (2002) Anal. Chem. 74, 5826–5830.
[20] Sickmann, A., Dormeyer, W., Wortelkamp, S., Woitalla, D., Kuhn, W. and Meyer, H.E. (2000) Electrophoresis 21, 2721–2728.
[21] Sickmann, A., Dormeyer, W., Wortelkamp, S., Woitalla, D., Kuhn, W. and Meyer, H.E. (2002) J. Chromatogr. B 771, 167–196.
[22] Raymacker, J., Daniels, A., De Brabandere, V. and Missiaen, C. (2000) Electrophoresis 21, 2266–2283.
[23] Palmblad, M., Tsybin, Y.O., Ramström, M., Bergquist, J. and Håkansson, P. (2002) Rapid Commun. Mass Spectrom. 16, 988–992.
[24] Tsybin, Y.O., Håkansson, P., Wetterhall, M., Markides, K.E. and Bergquist, J. (2002) Eur. J. Mass Spectrom. 8, 389–395.
[25] Beardsley, R.L. and Reilly, J.P. (2003) J. Proteome Res. 2, 15–21.
[26] Hagman, C., Ramström, M., Håkansson, P. and Bergquist, J. (2004) J. Proteome Res., In press.