Protein Analysis using Proteotypic Peptides and LC-MS:

Choosing the Chromatographic Separation for Optimal Coverage or Throughput

***Chris Barton and Richard Kay** - Senior Scientists, Drug Development Services, Quotient BioResearch Ltd.*

The science of protein analysis encompasses a range of analytical technologies including long established approaches such as gel electrophoresis and enzyme-linked immunosorbent assays [1, 2] Over the past ten years, mass spectrometry, linked with soft ionisation techniques such as electrospray ionisation (ESI), nanospray and matrix assisted laser desorption/ionisation (MALDI), has developed into an excellent tool for studying proteins in complex matrices. The combination of nanospray and liquid chromatography has since become the technique of choice for the majority of protein mass spectrometrists.

Mass spectrometric analysis of intact proteins by LC-MS is a challenging and often insensitive approach, mainly due to the limited m/z range of current instruments, the formation of a number of ions with differing mass to charge ratios and the low resolution of intact protein separation techniques. For high sensitivity analysis using LC-MS proteins are more usually enzymatically cleaved into surrogate, mass spectrometer friendly, proteotypic peptides [3, 4]. These are significantly easier to separate chromatographically than their undigested precursor (Figure 1)

A major drawback of the enzymatic cleavage approach is the significant increase in complexity of the protein sample. Dependant on its size, a single protein can result in tens or even hundreds of surrogate peptides after digestion. Separating, detecting and quantifying these specific peptides in the mileu can be a very challenging task. However, improvements in mass spectrometers, chromatographic systems and columns, along with the ready availability of labelled standards, are transforming this situation.

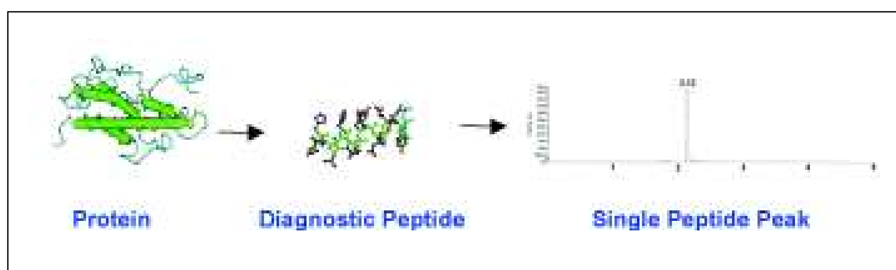

Figure 1. Proteotypic Peptide Analysis of Proteins.
A protein is reduced, alkylated and then digested using a specific enzyme such as trypsin to produce sequence-specific diagnostic peptides (proteotypic peptides). These act as quantitative surrogates of the parent protein, and can be separated and detected using LC-MS/MS.

New techniques are transforming the potential for using LC-MS as a quantitative technique for protein analysis. At Quotient Bioresearch, HPLC and triple quadrupole MS systems have been used to perform qualitative and quantitative analysis of proteins, in a range of areas including sports testing and analytical support for pharmaceutical development. This has provided experience in the application of liquid chromatography to the separation of peptides and how the available chromatographic techniques impact upon the detection technology. Different LC-MS based approaches can be tailored to specific applications in order to provide a balance between throughput and the number of analytes covered in a single multiplexed analytical run.

Large Scale Protein Identification and Quantification

The analysis of proteins by LC-MS/MS has rapidly evolved in the field of proteomics [5]. The most common feature of a proteomics experiment is the separation of a complex mixture of proteolytically derived peptides by reversed phase LC, and the detection and

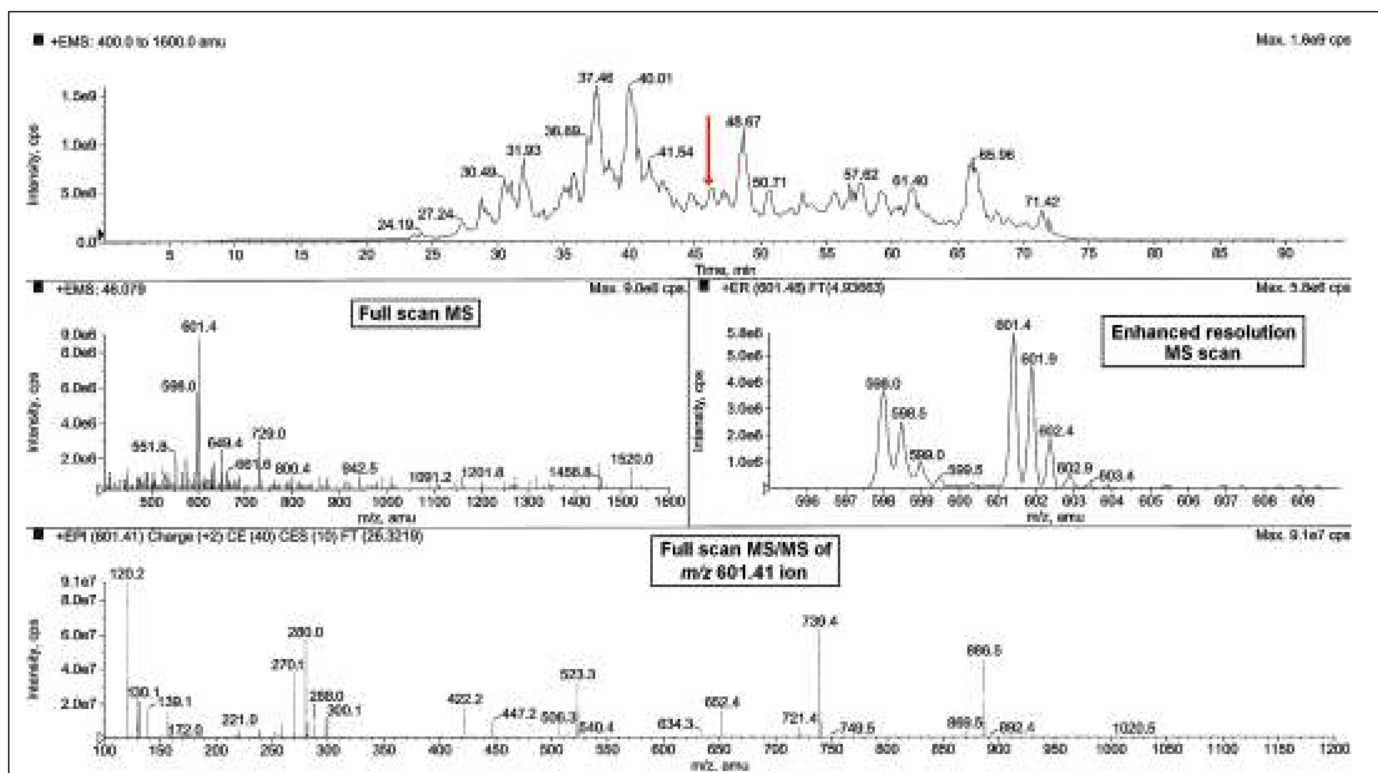

Figure 2. Identification of a protein by LC-MS/MS. Peptides are separated by reverse phase nano-LC column (Pepmap, Dionex, 0.075 x 150 mm x 3µm particle nanobore columns, flow rate 300 nL a minute) on a Dionex Ultimate 3000, linked to a Sciex 4000 QTrap system. MS and MS/MS provides information on a peptides mass and fragmentation behaviour, which can be used to identify the peptide's sequence.

identification of these peptides by tandem mass spectrometry (Figure 2). A peptides precursor mass and fragmentation pattern gives information on its specific amino acid sequence, and this can be combined with data from the organism's genome to identify the protein of origin.

A major advantage of this approach is that the technique used to identify proteolytic peptides can be rapidly modified to provide a selective technique for detection of a specific peptide set; for example with a triple quadrupole instrument through the use of multiple reaction monitoring (MRM).

The use of nano-scale chromatography combined with electrospray ionisation and increasingly accurate tandem mass spectrometry (such high resolution TOF, Ion-Trap and FT-ICR instrumentation) gives unparalleled sensitivity. Combining these techniques with sample fractionation, e.g. multi-dimensional chromatography, increases the number of proteins that can be seen in a single experiment into the thousands [6]. These techniques have been made semi-quantitative, for example with the iTRAQ reagent series which tags peptides from up to 8 samples with distinct heavy-isotope labels prior to analysis in a single run [7]. However, the tools used for this type of proteomic analysis (nano-LC and high resolution mass spectrometers) are expensive and rarely found in bioanalytical laboratories.

Targeted peptide analysis using Multiple Reaction Monitoring

The application of multiple reaction monitoring (MRM) to peptide analysis has finally brought proteomics into the reach of bioanalytical laboratories. The MRM analysis technique involves the pre-selection of a specific precursor ion m/z for fragmentation, and the monitoring of a specific product ion. This scan function can be performed on a number of instruments; however the triple quadrupole is the most efficient, and commonly used system. MRM analysis can result in a large increase in sensitivity for the peptide of interest, and by cycling the MS system through a series of transitions, multiplexed peptide analysis can be performed successfully on highly complicated peptide mixtures [8].

LC-MS using MRM is a standard technique in bioanalytical laboratories, and has become the method of choice for LC-MS analysis of peptides at Quotient Bioresearch. In small molecule bioanalysis, a target molecule is analyzed by comparing its response to an isotopically labelled internal standard, and quantified with respect to an external standard curve. This workflow can be applied to protein analysis, with synthetic isotopically labelled tryptic peptides [9], or even whole proteins [10], being used as internal standards, and recombinant proteins being used for external standard curves.

Impact of Chromatography on Targeted Peptide Analysis

The use of nano-LC for separation of peptides prior to MRM detection gives unparalleled sensitivity, but is time consuming and lacks robustness for the analysis of large numbers of samples. A typical nano-LC peptide separation using 0.075 x 150 mm nanobore columns packed with 3 µm particles is operated at a flow

rate of approximately 300 nL a minute. This technology is perfect for peptide identification, with the approximately 60 s peak widths giving ample opportunity for a number of full MS/MS scans to be obtained for a given peptide. When combined with MRM detection, analysis of a large number of peptides (100+) can be multiplexed in a single 60 minute run (Figure 3).

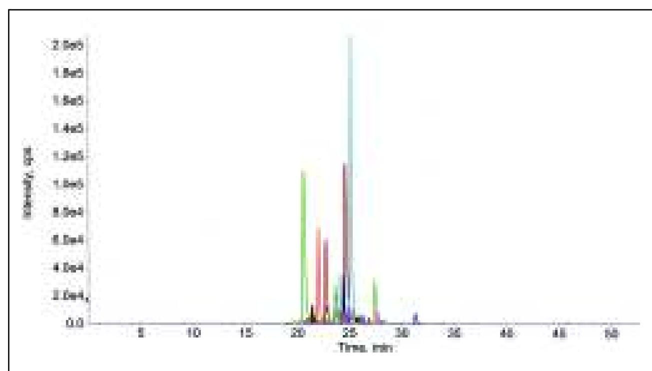

Figure 3. Nano-LC-MS/MS of serum protein digest. Nano-LC monitoring of 80 peptides from an enzyme digest of undepleted equine plasma in 55 minutes. Data was acquired on the same LC-MS system as in Figure 2.

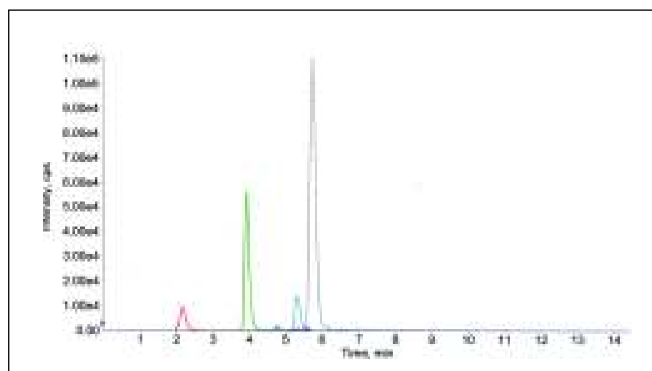

Figure 4. Capillary LC-MS/MS using a monolithic column. A 15 minute run Capillary-LC-MS/MS analysis of 15 peptides in acetonitrile depleted plasma using a polystyrene divinylbenzene 0.2 x 150 mm monolith (Dionex). Data acquired on the same LC-MS system as in Figure 1.

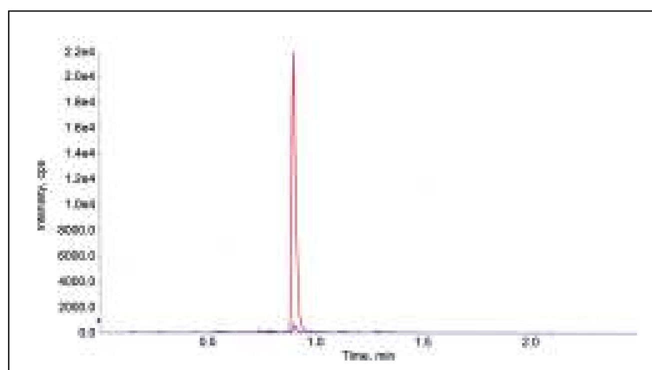

Figure 5. UHPLC-MS of GH-abuse biomarkers in ACN depleted plasma. UHPLC monitoring of 4 peptides (two shown) in a 2.5 minute run. Data acquired on a Waters Acquity (2.1 x 100 mm BEH C18 column with 1.7µm, 300 Å particles at 0.7 mL/min) linked to a Sciex API4000 MS system.

Nano-LC provides high sensitivity analysis, because of low solvent volume analyte concentration. In order to fully exploit the increased analyte concentration, a specialised nano-spray source designed to handle low mobile phase flow rates with minimal peak broadening is required. These nano-spray ionization sources then provide very high ionisation efficiency and sensitivity. However, increasing sample throughput using nano-LC is difficult, due to the very low flow rate and therefore long column washing and re-equilibration times. Further, nano-LC is less reliable than standard HPLC, and more prone to column blockages and resultant batch failures.

Monolithic columns, which consist of a single 'rod' of a macroporous polymer, have been trialled in an attempt to increase sample throughput. The macroporous structure reduces the diffusion path length and increases the efficiency of mass transfer, giving highly efficient chromatographic separations without the increases in back pressure associated with smaller particles sizes. These columns also allow a more rapid separation than standard nano-LC systems (due in part to a higher flow rate of 3 µL/min), and 10-15 minute runs are possible (Figure 4). The use of nanospray sources with capillary LC still retains the high ionisation efficiency of nano-LC-MS/MS, but because of the faster run times and decreased peak widths, fewer analytes can be analysed in a single run whilst retaining a suitable MS duty cycle. Capillary LC-MS/MS is therefore an alternative to nano-LC for faster, more targeted analysis. The one disadvantage of the system is that, like nano-LC, it suffers from similar technical problems inherent in low flow rate chromatography coupled to nano-spray ionisation.

The application of UHPLC (Ultra high pressure liquid chromatography) has brought great gains in bioanalytical separation in recent years. The use of LC pumps capable of sustaining higher back-pressures, combined with more robust column housing technologies has facilitated the use of sub 2 µm particle size chromatography packings. Smaller particle sizes bring benefits in efficiency of mass transfer, giving increased resolution, and perhaps more importantly, the ability to retain the same chromatographic resolution at increased flow rates. The benefit of UHPLC is thus faster, better resolved runs and ultimately higher sample throughput.

These advances in chromatographic resolution and speed impact upon the required performance of the detection system. In order to fully utilise the increased flow rates the ion source must be capable of handling upwards of 1 mL per minute of high aqueous content mobile phase whilst avoiding the introduction of chromatographic dead volume. In addition, the instrument data acquisition rate must be sufficiently rapid to allow a high duty cycle and minimal cross talk between MRM channels. Current triple quadrupole instruments have been designed with these factors in mind and bioanalytical laboratories have had a great deal of success applying this technology to small molecule analysis [11]. Based upon this experience, the application of UHPLC separation for the analysis of complex peptide mixtures derived from digested plasma and serum was investigated.

The combination of UHPLC separation and MRM was found to enable rapid, high throughput quantification of peptide surrogates of proteins. This was initially demonstrated with a 5 minute method for the separation and quantification of 10

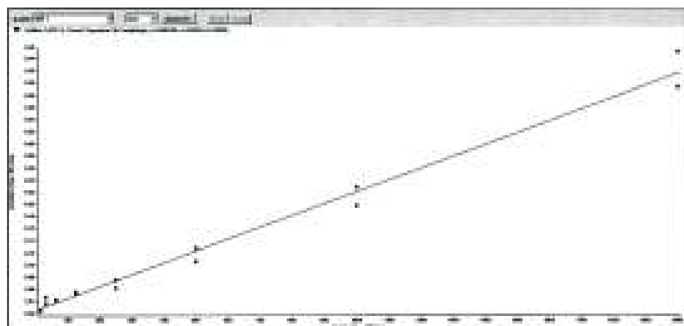

Figure 6. Standard Curve obtained using targeted UHPLC-MS/MS for analysis of IGF-1 in human serum.

	Nano-LC	Nano-LC	UHPLC
Flow rate	100 - 600 nL/min	2.5 - 25 L/min	0.3 - 1 mL/min
Particles	3 m, 100Å	Monolithic Polymer	1.7 m, 300Å
Ionisation Source	Nano-spray	Nano-spray	Turbo-ion spray
Backpressure	100 - 200 bar	100 - 200 bar	Up to 1000 bar
Run Time	60 - 240 min	10 - 20 min	2.5 - 5 min
No. of Analytes per run	100+	10-30	2-10
Sensitivity	Gold Standard	2-Fold less than nano-LC	10-20 fold less than nano-LC
Peak Width (Approx)	60 s	30 s	6 s
Robustness	Low; low flow rates and column blockage	Medium; low load-ability	Good, >2000 injections/column

Table 1: Comparison of chromatographic techniques used for peptide analysis

apolipoproteins and related proteins in an undepleted serum digest [12]. This was improved upon recently with a 2.5 minute method for the simultaneous quantification of two protein biomarkers of growth hormone (GH) abuse (see Figure 5). An initial concern was with the use of heated electrospray source to handle the high flow rates associated with UHPLC. This has an inherently lower ionisation efficiency than a nanospray source and could potentially have invalidated this as an analytical approach. However, the increased concentration of the analyte in the narrower peaks obtained with UHPLC (6 second widths) partially compensated for this loss of sensitivity, and it is thus possible to transfer assays from nano to UHPLC systems whilst retaining assay lower limits of quantification.

For the application of targeted protein quantification, UHPLC provides robustness and reproducibility benefits compared to the use of nano-LC or capillary LC. In our experience, the separation and analysis of peptides by UHPLC was not only far faster than nano-LC, but more reliable, with currently over 2000 injections being performed on a single column. Further, batch failure rates due to column blockage were much lower. A further benefit arises from the high accuracy and precision of the UHPLC-MS system in protein quantification (Figure 6), with co-efficient of variance values (CVs) of 5-10%. Even in assays where proteins must be initially extracted from a matrix, digested and then quantified, the technique was found to be highly reproducible, with CVs less than 15%. The disadvantage of UHPLC is that, because of the short run times and narrow peaks, only a small number of analytes can be studied in a single run whilst retaining MS duty cycle.

Conclusions and Future Directions

In conclusion, the application of multiple reaction monitoring to the analysis of peptide surrogates of proteins makes the routine quantification of proteins by LC-MS a viable alternative to other technology platforms. The impact this technology has on the types of chromatography that can be used, and the various advantages and disadvantages of each, is summarised in Table 1. This highlights that different chromatographic platforms are suitable for solving different problems, depending on the number of samples and the number of analytes to be studied.

The technologies supporting this platform are advancing rapidly. One current strand of research in the proteomics community is aimed at increasing the sensitivity of the approach by enriching the target from the matrix prior to analysis using different extraction conditions, including immuno-magnetic peptide capture [13]. Further, the recent improvements in chromatographic separation have made the MS system a limiting factor in multiplexed peptide analysis. It is hoped that future improvements in detector technologies will be made to drive this field forward.

Acknowledgements

Research discussed in this article was funded by the World Anti-Doping Agency and Horse racing and Betting Levy Board, with support from the British Horse-racing Authority and Waters.

References

1. *Lequin R (2005) Enzyme Immunoassay (EIA)/Enzyme-Linked Immunosorbent Assay (ELISA) Clinical Chemistry 51: 2415-2418*
2. *Lilley KS, Razzaq A and Dupree P (2002) Two-dimensional gel electrophoresis: recent advances in sample preparation, detection and quantitation. Current Opinions in Chemical Biology 6: 46-50*
3. *Domon B and Aebersold R (2006) Mass spectrometry and protein analysis. Science 312:212-7*
4. *de Hoog CL, Mann M (2004) Proteomics. Annuual Review of Genomics and Human Genetics 5: 267-93*
5. *Shi Y, Xiang R, Horváth C and Wilkins JA (2004) The role of liquid chromatography in proteomics. Journal of Chromatography A 1053: 27-36*
6. *Anderson NL, Polanski M, Pieper R, Gatlin T, Tirumalai RS, Conrads TP, Veenstra TD, Adkins JN, Pounds JG, Fagan R, Lobley A. (2004) The human plasma proteome: a nonredundant list developed by combination of four separate sources. Molecular and Cellular Proteomics 3: 311-26*
7. *Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Molecular and Cellular Proteomics.3: 1154-69.*
8. *Anderson L, Hunter CL (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Molecular and Cellular Proteomics 5: 573-88.*
9. *Kirkpatrick DS, Gerber SA, Gygi SP (2005) The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. Methods 35: 265-73*
10. *Rivers J, Simpson DM, Robertson DH, Gaskell SJ and Beynon RJ. (2007) Absolute multiplexed quantitative analysis of protein expression during muscle development using QconCAT. Molecular & Cellular Proteomics 6: 1416-27*
11. *Houghton R and Grace P (2008) UPLC – why all the Hype? Chromatography Today 1(2): 5-7*
12. *Kay RG, Gregory B, Grace PB and Pleasance S. (2007) The application of ultra-performance liquid chromatography/tandem mass spectrometry to the detection and quantitation of apolipoproteins in human serum. Rapid Communications in Mass Spectrometry 21: 2585–2593*
13. *Anderson NL, Anderson NG, Haines LR, Hardie DB, Olafson RW, Pearson TW. (2004) Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). Journal of Proteome Research 3: 235-44.*