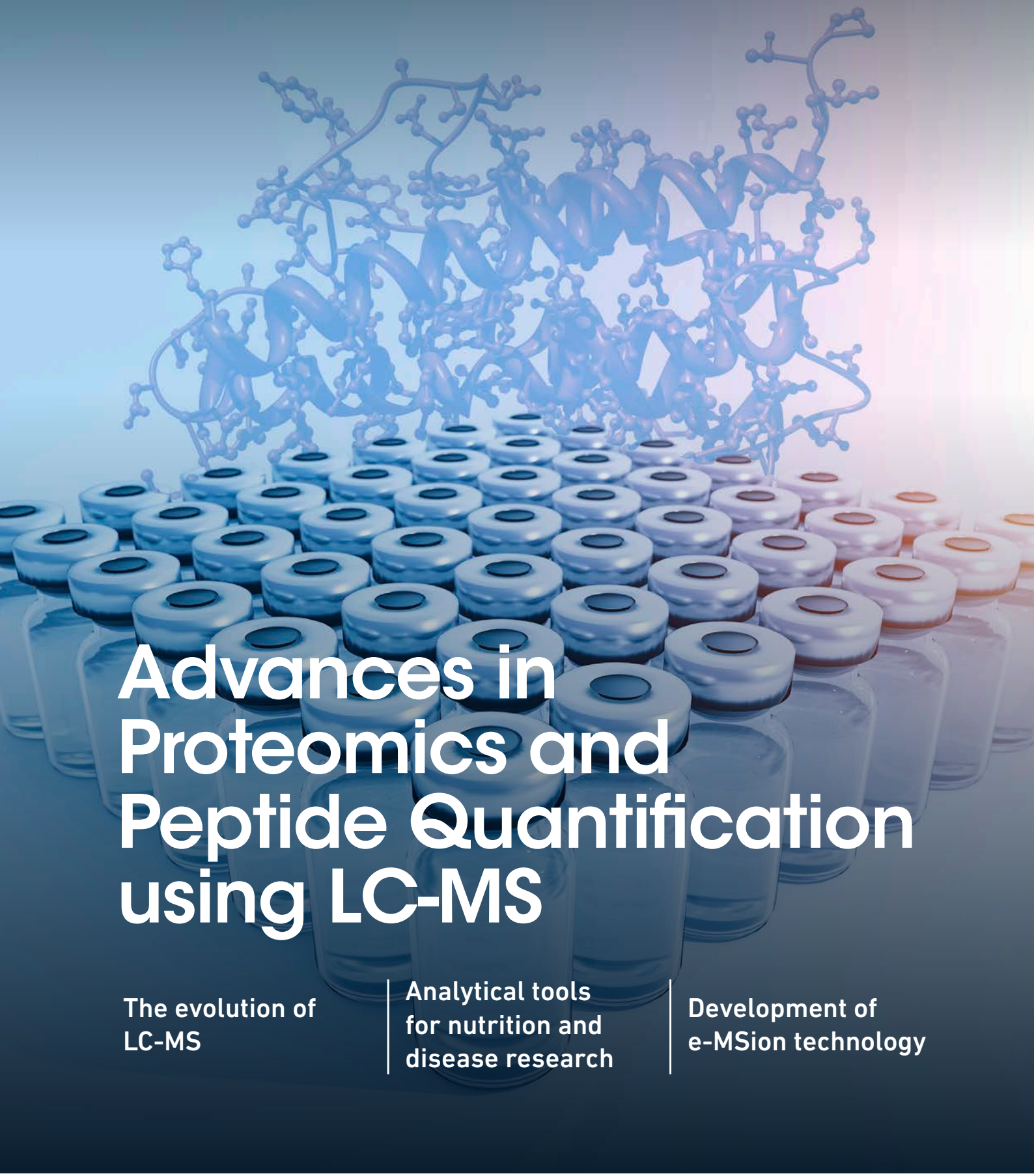
TECHNOLOGY NETWORKS



Advances in Proteomics and Peptide Quantification using LC-MS

The evolution of LC-MS

Analytical tools for nutrition and disease research

Development of e-MSion technology

SPONSORED BY

Agilent

Foreword

The field of proteomics has rapidly expanded to cover almost every aspect of protein research – from identification and characterization of post-translational modifications to biomarker discovery and quantitation. Similarly, peptide quantitation is a rapidly growing area that measures unique peptides from proteins in clinical research samples, based on prior knowledge. While protein and peptide quantitation present many distinct challenges, the development of comprehensive, optimized, and practical workflows is enabling more accurate and reproducible results.

Liquid chromatography mass spectrometry (LC-MS) is one such technique that has evolved to become an established bioanalytical platform for the quantitative determination of proteins, peptides, macromolecular drugs and biomarkers.

In this eBook, we explore how researchers are using LC-MS to advance their research as well as some of the latest solutions offered by Agilent to help these researchers overcome some of the common analytical challenges. In particular, we focus on the Agilent 6495 Triple Quadrupole LC/MS and 6560 Ion Mobility LC/Q-TOF demonstrating the benefits of incorporating these workflows for peptide and protein quantitation.

Contents

Quantification of Human Milk Proteins and their Glycoforms using LC/MS 4

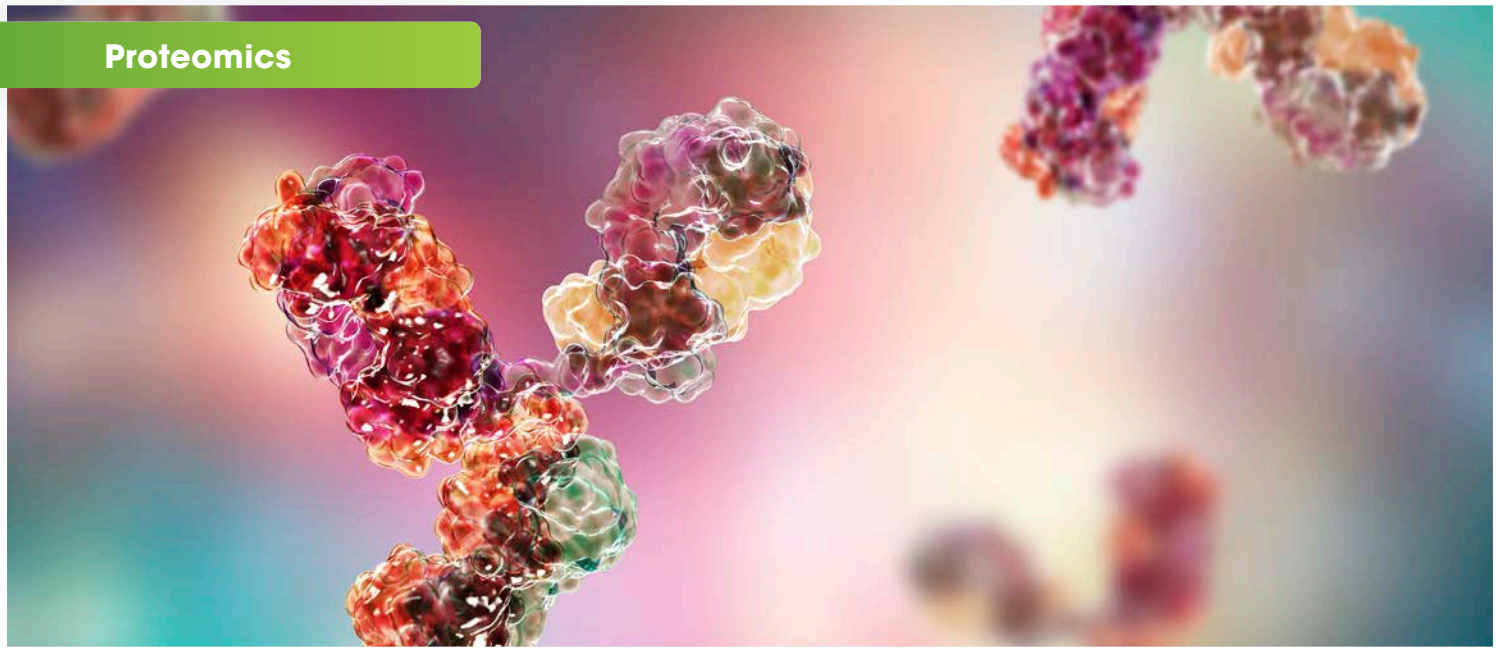
Robust and Reproducible Protein Quantification in Plasma using the Evosep One and the Agilent 6495 Triple Quadrupole LC/MS 7

From Nanoflow to Standard Flow LC/MS for Routine Quantitative Plasma Proteomics in Diabetic Kidney Disease Research 14

Peptide Quantification in Plasma Using the Agilent 6495 Triple Quadrupole LC/MS Coupled with the Agilent 1290 Infinity II LC System 23

Characterization of Proteins Using Electron Capture Dissociation Technology 30

A Novel, Automated, and Highly Selective Phosphopeptide Enrichment for Phosphopeptide Identification and Phosphosite Localization 34



Quantification of Human Milk Proteins and their Glycoforms using LC/MS

Human milk oligosaccharides are a highly abundant constituent in human milk that have attracted considerable attention. In addition to their protective and prebiotic properties, there is evidence that they serve as a beneficial food for gut bacteria, block pathogens, and aid brain development. In an effort to explore their utility as disease and nutrition biomarkers, rapid-throughput methods are being developed to investigate their biological activity in large cohort sets.

In this interview, we speak to Professor Carlito Lebrilla, a distinguished professor at the University of California, Davis. His research group is focused on analytical chemistry with specific emphasis in the area of nutrition, disease markers and mass spectrometric methods.

The team are developing mass spectrometry-based tools for the discovery of marks for cancer including ovarian, breast and prostate. They are pioneering the glycomic approach for the early diagnosis of cancer. In nutrition, they are examining human milk as the model for the perfect food and determining bioactive components in milk. His expertise in oligosaccharides and glycosylation of proteins and lipids makes his group particularly unique.

Here, Professor Lebrilla speaks about the applications of his work, his motivations and how he envisions the next few years will play out in the LCMS research space.

Q: The focus of your group is on developing analytical tools for nutrition and disease research. Can you discuss the applications of your work and then touch on recent study examples?

A: We have been studying human milk and looking at the bioactive components. We have discovered that there is a group of compounds called human milk oligosaccharides that are the third most abundant components, which are more abundant than proteins.

Initially, they seemed to have no nutritional value as there are no enzymes in the gut of the infant to break them down due to their structures. Therefore we began to question the role of these complicated and very abundant structures as they appeared to have no nutritional value.

We discovered that the human milk oligosaccharides are food for a microbe that has the enzymes to break down the compounds. The bacteria had been recruited by the mother to protect her infant. Interestingly, when this bacteria gets to the human milk oligosaccharides, it produces metabolites and products that the infant uses for protective and developmental purposes.

This bacteria, more recently, has been eliminated due to caesarean section birth and the use of antibiotics. It is rare to see this in infants in the US and other developed countries. Therefore we started a company to make this beneficial bacteria and you can buy this bacteria in hospitals in the US. The company itself is conducting a large clinical trial in Northern Europe to measure its potential translational beneficial effects on juvenile diabetes.

The other side of disease is that we've been looking for a way to diagnose cancers earlier, particularly ovarian cancer, through a blood biopsy or changes in the blood. We have discovered markers for these tests and have begun to commercialize these also by starting a company that are now developing these tests. I hope that within a few years we will have early tests for ovarian cancer, and other cancers, through this discovery.

Q: What tools and techniques are primarily used in your group? Are there any particular advances with certain analytical technologies that have enabled your research?

A: Our group uses a lot of liquid chromatography-mass spectrometer instruments, which allows for the reproducibility in the sensitivity of these methods. For example, in our human milk studies, we can do all of our analysis on 10 microlitres of human milk, which is great as we get samples from all over the world and may only have 100 microlitres to perform the analysis. Having sensitivity has been valuable.

We also have nanoflow technology, which means that the analysis itself takes a few hundred nanolitres of material to flow through the columns. The standard HPLC or LC methods would take milliliters of material. Standard HPLC, in the mobile phase, uses a large amount of sample. However, when you use nanoflow technology, you cut that sample size down significantly. This means you can use less material for your separation, but also you're using less of your sample because instead of having to move micrograms of compounds you can go down to nanograms and even picograms of material for your analysis.

The nanoflow has been valuable for us as it's a way that we've managed to increase the sensitivity. With that, and going back to human milk, the nanoflow lets us use minuscule amounts of material. This is also beneficial for blood analysis as the increased sensitivity allows us to sample around 10 to 20 microlitres of blood, which you can almost do on a pinprick rather than a vial.

Advances in separation science, advances in new methods for separating very complicated compounds, and most importantly the robustness of the methods is how we can reproducibly obtain the same results time and time again. This allows us to discover changes that are being made biologically rather than changes being made due to the instrument variation. In that regard, these methods have been very useful.

I'd like to mention that advances in informatics have also been beneficial, as you can generate hundreds of thousands of data points. Some of this analysis would produce hundreds of true compounds and maybe thousands of spurious signals, however, having the right informatics tools allows the team to figure out what are important peaks and what are the important signals. This can then be used to distinguish which compounds are changing.

Q: Could you talk to us about your recent study where you were to use LCMS and MS to determine immunoglobulins in mothers' milk and what your motivations were for conducting the work?

A: We had several reasons to look at immunoglobulins in milk but also in blood. We conducted the study in milk because when an infant is first born their immune system is not developed so the mother's immune system protects the infant. The immunoglobulins in the mothers' milk are important proteins. They complement the human milk oligosaccharides that I mentioned earlier. The proteins in the milk also have 'decorations' on them that affect the protein's function. The mothers' proteins are therefore also protecting the infant. Additionally, we also found that mothers were modifying the proteins in such a way that they interact with the microbes differently, which to us was fascinating.

Immunoglobulins are glycosylated, which are the short sugar chain decorations that interact with viruses, bacteria, and other pathogens by blocking them from binding with the gut of the infant. This is the mother protecting her infant. When an infant is born, the immune system develops gradually as the infant starts to produce their own immunoglobulins. The mother's immunoglobulin concentrations will decrease over time, whereas the infant's immunoglobulin will start to increase. This transition ensures that the infant is always protected.

Our team is looking at immunoglobulins to see precisely how the changes occur over time and looking closely at the modification of immunoglobulins early on in the lactation, and then later in lactation. The team then also wanted to understand how the same immune proteins from the mother differs when it is produced in her blood, milk, and saliva. We found that each of these immunoglobulins, even though they're the same IgA protein and is the most abundant antibody in mothers' milk, had altered glycosylation depending on which tissue it came from. This was directly related to the post-translational modification of glycosylation. We found that the immunoglobulin was

different because it had a different function in mothers' milk compared to blood and saliva.

The immunoglobulins also abundant in blood, and here we're interested in the blood as sources of biomarkers for diseases. In different diseases immunoglobulins change their glycosylation. We wanted to figure out how changing these immunoglobulins alter their function. One of the things that we found, for example, is that when a person has cancer, their immunoglobulins are also altered with cancer. For background, there have been several studies where scientists have looked at the breast tissue of women who died prematurely through other causes. They examined the breast tissue and found scarring. Scarring is typically an indication of cancer, however the frequency of scarring was much higher than that of breast cancer. Similar studies were performed on the testes of men who died prematurely, and here too, the frequency of scarring was much higher than the occurrence of testicular cancer.

The thought was that the immune system heals itself and it attacks cancer before the patient knows about it. They do this by having the immune system surveilling, and when they detect cancer, they kill it. When the immune system gets overwhelmed, the patient notices because the immune systems can no longer take care of the situation. We know that immunoglobulins and other proteins change with cancer, and they are one of the biomarkers that we have discovered as being able to provide earlier indications of cancer.

Glycan modifications are interesting and useful because they allow the protein to slightly change their function. Going back to immunoglobulins, when the mother is pregnant her immunoglobulin changes, and when the baby is born, these change again. Certainly gene regulation and protein production is one way to alter function, but now we know it is not just the protein abundances, but the proteins' structures are modified through post-translational modifications that also change functions.

Q: How do you envision the development of this workflow will influence other Omics research groups, and what advice would you give them if they were considering adopting the workflow?

A: The workflow that we have developed to monitor changes in glycosylation is relatively simple and employs a MS method called multiple reaction monitoring for quantitation. We do LC-MSMS to monitor the glycan modifications on the protein. I am confident that as we have shown this to be a straightforward method, others will try to replicate it and look at specific proteins and monitor these proteins throughout aging, disease states, and general changes in biological states of people. I think in this regard, people will be able to not just look at protein abundances, which is now common, but also look at how the protein modifications are changing to study other diseases.

If it goes back to the earlier question of why we use the method, it's because the method is robust and is highly reproducible. The reliability allows us look at things that are associated with the disease and not instrument variations from one day to the next. The reproducibility, for everyone, is really important. As for what advice I would give adapting the workflow, I think it's pretty robust and if you can follow the procedures that we've outlined, it's not that much more difficult to do. I think people are going to start using this a lot, that is if they're not doing so already.

Q: What will your next steps be in the LCMS research space?

A: One of the things that we're encountering is that the stationary phase in some of the columns that we have are not enough to separate the material that we need. We have taken one separation, then a second post-separation, maybe even a third post-separation so that we can get all the information because for the glycans there are many isomers. I think making multidimensional LC more reproducible is an area that needs more development.

Another area I predict we will see soon in the LCMS research space is surrounding sensitivity. Right now with LCMS, we are looking at femtomole to even attomole sensitivity. However, we need to see LCMS hit that lower-level more easily because the amount of material that we're getting is becoming less and less. Sometimes you get a biopsy and the amount of material you get maybe a few micrograms or a milligram, and you need to be able to do multiple analyses with that amount of samples.

I think achieving higher sensitivity is really what the community needs and I am hoping that's where a lot of the manufacturers will be going. If high sensitivity instruments were available, I think it would open up new avenues of research, but it would also help medical science tremendously.

Carlito B Lebrilla, PhD
UC Davis Department of Chemistry

Application Note

Proteomics

Agilent

Trusted Answers

Robust and Reproducible Protein Quantification in Plasma using the Evosep One and the Agilent 6495 Triple Quadrupole LC/MS

Authors

Linfeng Wu
Agilent Technologies, Inc.,
Santa Clara, CA, USA

Nicolai Bache and
Dorte B. Bekker-Jensen
Evosep Biosystems, Odense,
Denmark

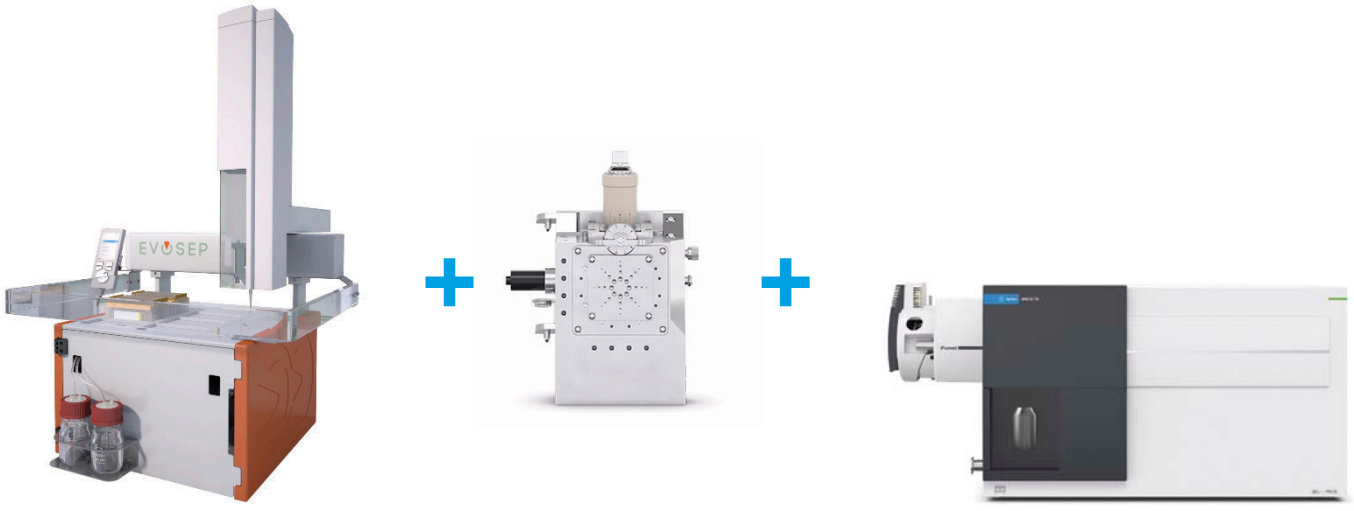
Abstract

This application note showcases the reproducibility, robustness, and analytical sensitivity of MRM-based LC/MS analysis of human plasma proteins using the Agilent 6495 triple quadrupole LC/MS with the Evosep One LC system. Results demonstrated superior retention time reproducibility over a 12-day analysis period, and equivalent sensitivity before and after robustness testing, underlining the suitability of the system for high-throughput protein quantification.

Introduction

As targeted workflows for protein biomarker verification using multiple reaction monitoring (MRM)-based methods are getting more popular, sample complexity and the low concentration of certain proteins are still the major challenges. Consequently, the development of MRM-based analytical methods using stable isotope-labeled standard (SIS) peptides for targeted quantitative proteomics in biological fluids is focused on improving method reproducibility, sensitivity, and robustness. Nanoflow liquid chromatography/mass spectrometry (LC/MS) uses the sample concentration effect of electrospray ionization (ESI) by producing the highest ionization efficiency using a small internal diameter column. However, nanoflow LC/MS typically requires a high degree of maintenance, and the limited nanocolumn capacity results in less robust chromatography for complex samples. Agilent Jet Stream technology coupled to the 6495 triple quadrupole LC/MS incorporating iFunnel technology has demonstrated outstanding analytical sensitivity and robustness for targeted protein quantification in complex matrices using standard flow LC/MS.[1,2] Nonetheless, there is still a need to address the prevalent challenges associated with throughput and robustness of nanoflow LC/MS for applications with limited amounts of samples. A conceptually novel chromatography system, called Evosep One, uses four low-pressure pumps in parallel to elute samples from a disposable and single-use trap column, the Evotip, while also using a pre-formed gradient, specifically designed to deliver the robustness and throughput required for clinical research applications.[3] This application note demonstrates the robustness, reproducibility, and analytical sensitivity performance of an Evosep One instrument coupled to a 6495 triple quadrupole LC/MS with an Agilent nanospray source (Figure 1).

To illustrate instrument robustness under challenging conditions, a balanced SIS peptide mix was spiked into human plasma digest and loaded directly onto Evotips. 574 replicate injections were analyzed on the LC/MS platform in a consecutive manner. To evaluate the precision and accuracy of the protein quantification, standard curve analyses were carried out both before and after the robustness test. A pre-formed gradient was used for separation, allowing 60 sample injections per day for both experiments. The instruments were not cleaned, adjusted, or tuned during the 12 days of continuous operation. The results show excellent robustness, reproducibility, and analytical sensitivity performance for targeted quantitative protein analysis in plasma.



Evosep One | Agilent nanospray source | Agilent 6495 triple quadrupole LC/MS

Figure 1. Evosep One coupled to an Agilent 6495 triple quadrupole LC/MS system with nanospray source.

Experimental

Instrumentation

- Evosep One
- Agilent nanospray source (G1992A)
- Agilent 6495 triple quadrupole LC/MS (G6495B)

Materials

Human plasma was purchased from Bioreclamation (catalog no. HMPLEDTA2). PeptiQuant Biomarker Assessment Kit (BAK-A6495-76) was purchased from Cambridge Isotope Laboratories.

Sample preparation

Human plasma was prepared by denaturation, reduction, alkylation, and trypsin digestion followed by lyophilization using a Speedvac. The plasma digest was reconstituted and spiked with the balanced SIS peptide mixture followed by a serial dilution for standard curve analysis. The final SIS peptide concentration ranged from 280 fmol/mL to 7 nmol/mL relative to the original plasma sample. A large stock of plasma sample spiked with 0.7 nmol/mL of the same SIS peptide mixture was also prepared for robustness testing. All samples for the 12-day continuous measurement were loaded on Evotips (~1 µg load) before starting the measurements and kept cold until analysis.

LC/MS analysis

All samples were separated by the Evosep One using a standardized pre-formed 21-minute gradient with a Pepsep column (100 µm × 8 cm) packed with 3 µM C18 beads, allowing 60 sample injections per day (Table 1). A stainless-steel emitter was incorporated into the needle holder (clamshell) for the Agilent nanospray source. LC/MS data was acquired using the Agilent 6495 triple quadrupole LC/MS in dMRM mode. A transition list for 33 pairs of heavy and endogenous peptides (198 transitions, 66 peptides matching to 31 protein biomarkers) was selected for the final dMRM method.

Table 1. LC/MS acquisition parameters.

Evosep One LC System	
Analytical Column (Length/ID/C18 Bead Size)	8 cm/100 µm/3 µm
Flow Rate	1 µL/min
Gradient Length	21 minutes
Cycle Time	24 minutes
Throughput	60 (samples/day)
Agilent 6495 Triple Quadrupole Mass Spectrometer	
Ion Mode	nanoESI, Positive
Gas Temperature	200 °C
Drying Gas Flow	11 L/min
Capillary Voltage	1750 V
High/Low Pressure RF Voltage	200/110 V
Delta EMV	200 V
Q1 And Q3 Resolution	Unit/Unit
Cycle Time	500 ms
Minimum/Maximum Dwell Time	5.90 ms/80.589 ms
Total MRMs	198

Data processing

Data analysis for targeted peptide quantification was carried out using the Agilent MassHunter workstation software (v10.0) and Skyline software (v19.1.0.193).

Results and discussion

Robustness test: Complex sample analysis

To assess the system robustness, 1 µg of human plasma digest spiked with the balanced SIS peptide mixture was injected 574 times in a consecutive manner, leading to 10 fmol of SIS peptide mixture on column per injection. The Evotip was used as a disposable trap column and the plasma sample was directly loaded onto this without further SPE cleanup. The sample was separated using a 21-minute LC-dMRM method. The retention time (RT) alignment of the targeted 33 pairs of heavy and endogenous peptides from 574 consecutive injections provided good and consistent reproducibility across all injections for the targeted peptides (Figure 2). This observation demonstrated the excellent reproducibility and robustness of the Evosep One for high-throughput studies.

The high reproducibility and robustness of the MS signal is illustrated by the distribution of the relative standard deviation (RSD) of peak area for all targeted peptides during the robustness test. The median RSD is 8.5%, with 62 out of the 66 peptides (93.9%) showing an RSD below 16%. Only two pairs of heavy and endogenous peptides show an RSD greater than 16%. One pair is hydrophilic peptides (RT = 2.6 minutes) and unstable in solution, as a signal degradation (>50%) was observed

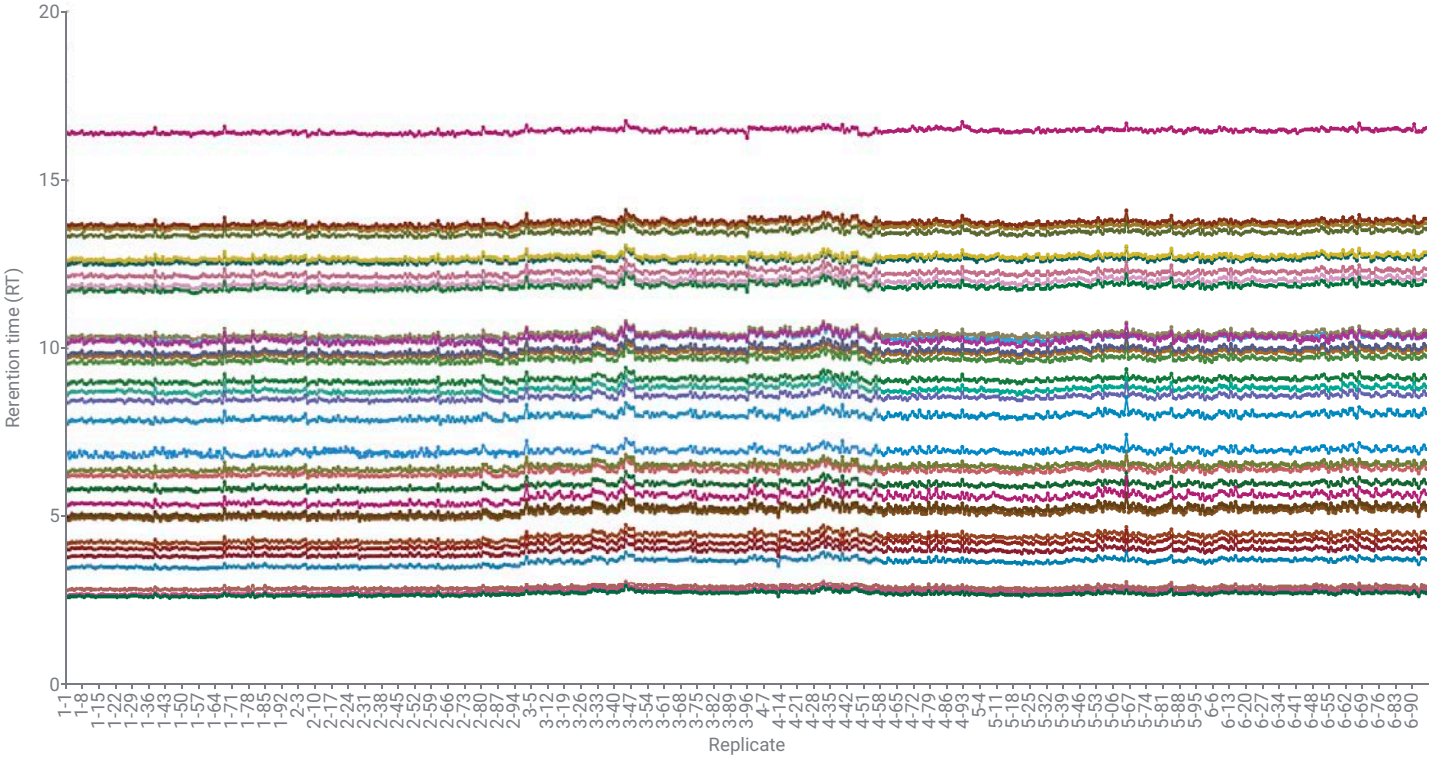


Figure 2. Retention time distribution of all targeted 33 pairs of peptides from 574 replicate injections during robustness test. The different peptides were color-coded.

4

after overnight storage in a refrigerator. The other pair suffered severe matrix interference, causing variation in peak integration. Therefore, the high RSD of these two pairs of peptides was not due to instrument variation (Figure 3).

The MRM peak area of four selected peptides matching four protein biomarkers demonstrated the outstanding robustness of the overall analytical platform (Figure 4):

- Very stable MS response without any adjustment on the LC/MS system (MRM peak area RSD = 6.5, 7.0, 7.9, and 6.0%, respectively, for n = 574)
- Good RT reproducibility (RSD = 0.69, 0.80, 1.04, and 0.59%, respectively, for n = 574)

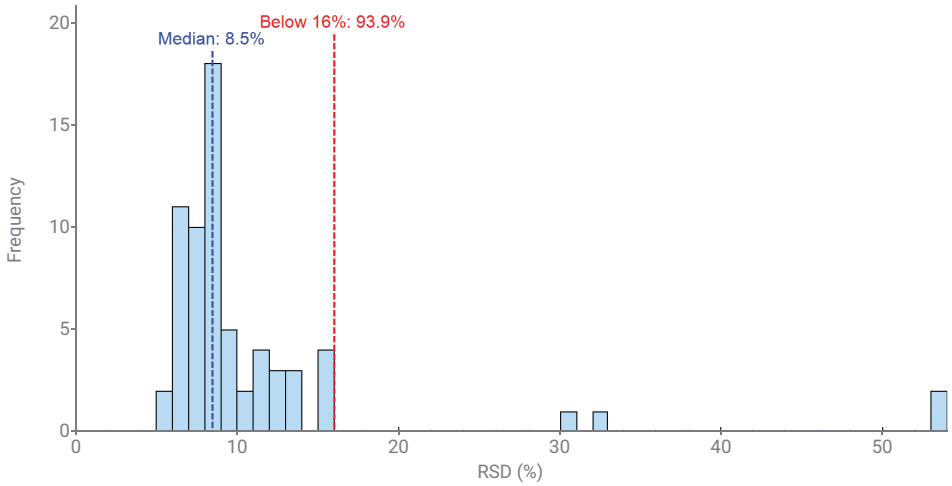
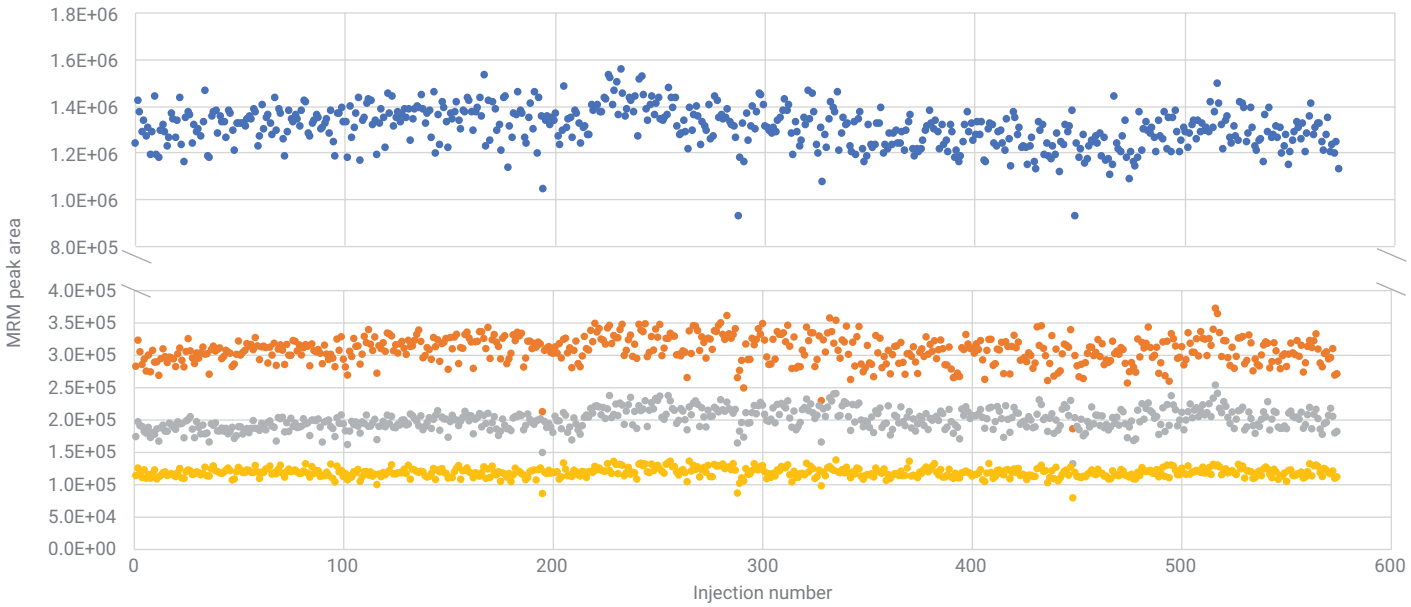


Figure 3. Distribution of peak area RSD for all targeted peptides from 574 replicate injections. The blue dash line represents the median RSD of 8.5%. The red dash line marks the 93.9% of the peptides having an RSD below 16%.



Complement C3
Hemopexin
Serotransferrin
Alpha-1B-glycoprotein

Protein	Peptide	Average MRM Peak Area	MRM Peak Area %RSD (n = 574)	Average RT (min)	RT %RSD (n = 574)
Complement C3	SGIPIVTSPYQIHFTK	1.31E+06	6.5	12.2	0.69
Hemopexin	NFPSPVDAAFR	3.08E+05	7.0	10.4	0.80
Serotransferrin	EGYYGYTGAFR	2.00E+05	7.9	8.8	1.04
Alpha-1B-glycoprotein	LETPDFQLFK	1.18E+05	6.0	13.4	0.59

Figure 4. MRM peak area of four selected SIS peptides from 574 replicate injections during the robustness test.

5

Quantitation of peptide standards in human plasma

To evaluate the analytical sensitivity for quantification of proteins in plasma, the SIS peptide mixture was spiked into human plasma digest at eight different concentrations ranging from 280 fmol/mL to 7 nmol/mL relative to the original plasma sample, then loaded directly on Evotips (~1 µg plasma digest). A standard curve was measured before and after the robustness test, with each injection being from Evotips, which were loaded in parallel before initiating the robustness test and stored cold until analysis. Five replicate injections of each dilution were measured before and after the robustness test to evaluate the precision and accuracy of the quantification up-front and with the impact of the robustness test.

The results from the two standard curves, measured before and after the robustness test, for the SIS peptide SGIPIVTSPYQIHFTK from Complement C3 are summarized below:

- Low amol-level sensitivity with a limit of detection (LOD) of 4 amol on column and a lower limit of quantification (LLOQ) of 10 amol on column in heavy matrices for both standard curves using a quantitative criterion of RSD <15% and accuracy of 80 to 120% (Figure 5 and Table 2)
- Excellent standard curve fitting for a dynamic range over 4 orders of magnitude (4 amol to 100 fmol on column) in heavy matrices with R^2 = 0.9987 and 0.9998 for the two standard curves, respectively (Figures 5A and 5B)
- Excellent precision and accuracy observed at all the tested levels including the LLOQ levels (Table 2)
- Good RT reproducibility (RSD = 0.43 and 0.32%, respectively, for n = 40)

The highly reproducible results between the two standard curves demonstrate the excellent instrument robustness and sample storage stability on Evotips even for the low-amol peptides. The samples for the standard curve measured after the robustness test were stored cold at 4 °C for about 10 days before being subjected to LC/MS. Despite this, the LOD of 4 amol and LLOQ of 10 amol remained the same as for the freshly prepared Evotips loaded just prior to analysis before the robustness test.

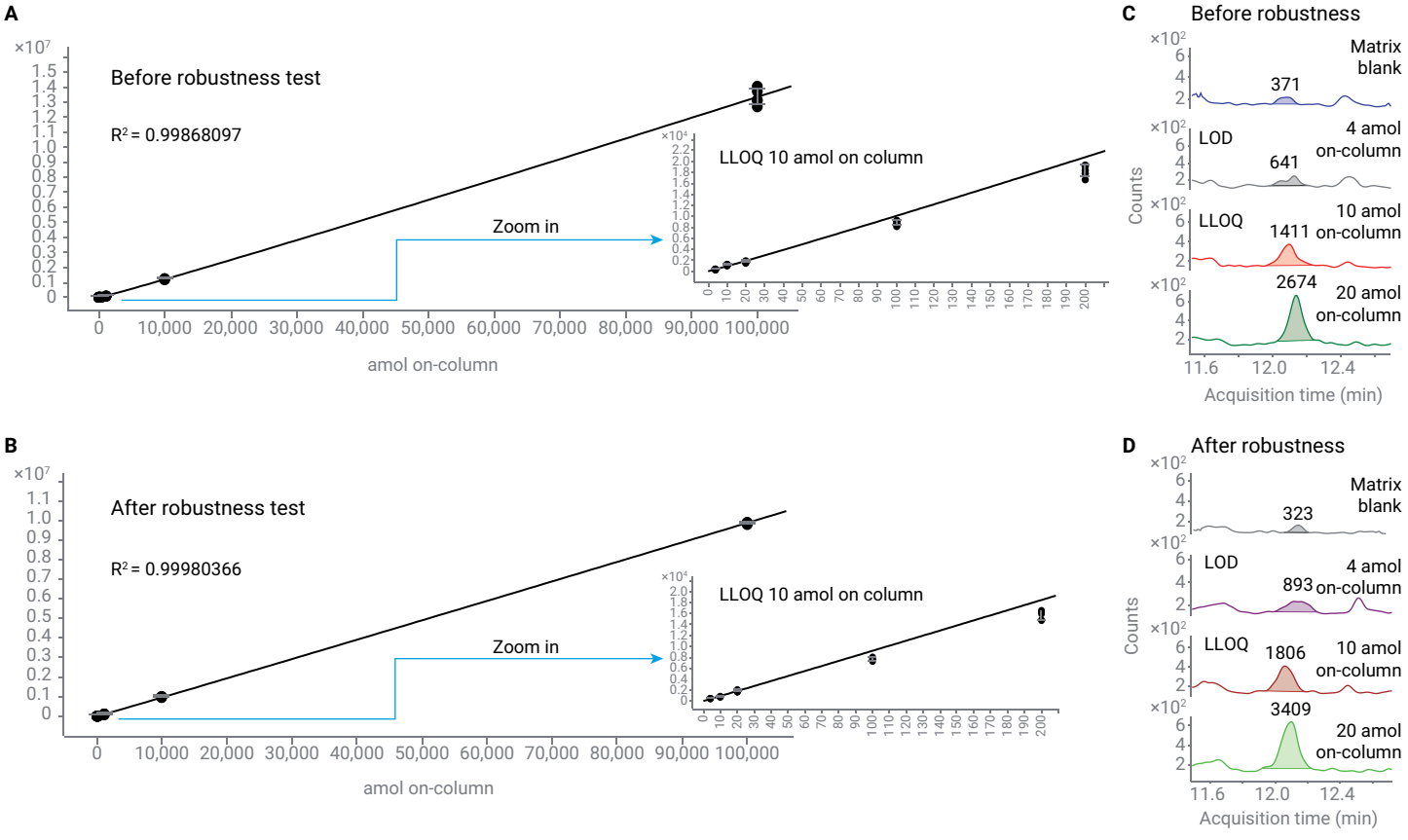


Figure 5. Standard curve analyses of SIS peptide SGIPIVTSPYQIHFTK from Complement C3 in plasma both before and after robustness test. A,B) standard curves before and after robustness test. C,D) Stacked extracted ion chromatograms showing the LOD and LLOQ.

All LC/MS analyses for the robustness test and standard curve evaluation were carried out on a single analytical column and was not replaced. The column was conjugated with the stainless-steel emitter inside the clamshell, which shields the column temperature from room-temperature fluctuation. Throughout the 12-day testing, the backpressure of the analytical column did not increase; the MS signals show excellent stability without spray needle changing, spray voltage adjustment, or mass spectrometer tuning. In summary, all these results demonstrated the outstanding analytical sensitivity, robustness, and reproducibility of this LC/MS platform for high-throughput studies under challenging conditions.

Table 2. Precision and accuracy for the standard curve analysis of the Complement C3 heavy peptide standard SGIPIVTSPYQIHFTK in trypsinized human plasma.

Amount On-column (amol)	Before Robustness Test		After Robustness Test	
	%RSD (n = 5)	%Accuracy	%RSD (n = 5)	%Accuracy
4	17.7	96.4	11.5	126.0
10	10.5	119.5	10.3	87.2
20	11.2	91.0	11.4	97.9
100	5.6	87.3	4.4	83.2
200	5.8	88.5	5.5	84.9
1,000	4.7	97.3	6.2	94.7
10,000	3.3	101.2	1.4	102.2
100,000	4.0	99.9	0.6	99.9

Conclusion

The reproducibility, robustness, and analytical sensitivity of MRM-based LC/MS methods for peptide quantification in biological fluids are important considerations for high-throughput protein biomarker quantification studies. This application note discusses the performance evaluation of the Evosep One LC system when coupled to the nanospray source and 6495 triple quadrupole LC/MS for the quantification of proteins in human plasma. For the analysis of a complex plasma digest matrix over an extended period of 12 days, the LC/MS system maintained excellent response stability and RT reproducibility without any cleaning, adjustment, or tuning. The analytical sensitivity was evaluated both before and after the robustness test and resulted in the same LOD and LLOQ. All these results demonstrate the suitability of this high-performance LC/MS system for high-throughput protein quantification in complex matrices.

References

1. The Agilent 6495 Triple Quadrupole LC/MS: Peptide Quantitation Performance. *Agilent Technologies technical overview*, publication number 5991-6898EN, **2016**.
2. Jet Stream Proteomics for Sensitive and Robust Standard Flow LC/MS. *Agilent Technologies technical overview*, publication number 5991-5687EN, **2015**.
3. Bache, N *et al.* A Novel LC System Embeds Analytes in Pre-formed Gradients for Rapid, Ultra-robust Proteomics. *Mol. Cell Proteomics* **2018**, *17(11)*, 2284–2296.

Printed in the USA, May 15, 2020
5994-1928EN

Application Note

Clinical Research

Agilent

Trusted Answers

From Nanoflow to Standard Flow LC/MS for Routine Quantitative Plasma Proteomics in Diabetic Kidney Disease Research

Authors

Orla Coleman, Angela McArdle, and Stephen Pennington
Atturos Ltd, UCD,
Dublin 4 Ireland

Scott Bringans,
Richard Lipscombe
Proteomics International,
Perth, Western Australia,
Australia

Linfeng Wu
Agilent Technologies, Inc.

Abstract

This application note showcases the capability, reproducibility, and analytical sensitivity of MRM-based LC/MS analysis of human plasma proteins using the Agilent 1290 Infinity II LC coupled with the Agilent 6495 triple quadrupole (TQ) mass spectrometer. The results demonstrate equivalent data from eight peptides across 12 samples on the Agilent standard flow system compared to a nanoflow counterpart system. Successful and reproducible analysis of the eight peptides in whole plasma comparable to depleted plasma analysis highlights the analytical sensitivity and robustness offered by the 6495 LC/TQ for routine protein measurement in biomarker research.

Introduction

Diabetic kidney disease (DKD) is a significant complication of diabetes with one in three adult diabetics having some degree of DKD. The current clinical tests to monitor DKD are the urinary albumin:creatinine ratio (ACR) and the estimated glomerular filtration rate (eGFR), which have minimal power to predict early DKD.[1] Therefore, there is a need to identify new protein biomarkers and develop tests for early DKD prediction.

PromarkerD is a MS-based clinical research test method developed to predict chronic kidney disease (CKD) in a cohort validation of people with type 2 diabetes.[2,3] It was developed using a proteomics biomarker discovery and validation workflow by Proteomics International (PI). Initial development of PromarkerD involved the measurement of 12 proteins in immunodepleted plasma samples using a nanoflow LC/MS platform.[4] During method development the number of protein markers was narrowed down to four proteins (APOA4, CD5L, FHR2, and IBP3) with a total of eight peptides. However, a nanoflow LC/MS system is not suitable for routine protein biomarker measurements. As a future option for clinical assay development, a standard flow LC/MS system provides an alternative platform with excellent performance for routine protein measurement.[5]

This application note explores the ease of transferring the nanoflow LC/MS-based PromarkerD method to a standard flow LC/MS platform using the Agilent 1290 Infinity II LC coupled to the Agilent 6495 LC/TQ. The extent to which nanoflow LC/MS can be matched by standard flow LC/MS measurements is also explored. It is demonstrated that the four PromarkerD protein markers could be comparably measured from nanoflow to standard flow LC/TQ systems, and in whole plasma and depleted plasma samples. Since the standard flow LC/MS platform is a more robust and reproducible platform than nanoflow LC/MS, our study indicates that the Agilent 1290 Infinity II LC coupled to the Agilent 6495 LC/TQ is more suitable for routine measurement of PromarkerD protein markers.

Materials and methods

Plasma samples

All subjects' plasma samples were provided by the Fremantle Diabetes Study (FDS), a longitudinal observational cohort.[6] EDTA (Ethylenediaminetetraacetic acid) plasma was collected from all subjects after an overnight fast and stored at −80 °C until required. The FDS protocol was approved by the South Metropolitan Area Health Service Human Research Ethics Committee (07/397), Western Australia. All subjects gave informed consent before participation. A standard reference plasma was collected from three healthy volunteers and combined before aliquoting and storage at −80 °C.

Method transfer

To establish the PromarkerD method on the Agilent 6495 LC/TQ (G6495B), a synthetic peptide mix was generated and spiked into Atturos reference serum (ARS). Synthetic peptides were purchased from Thermo Fisher Scientific and resuspended in 50/50 acetonitrile (ACN)/H_2O (ddH_2O, 18.2 Ω) and a stock solution was prepared to give a final concentration of 0.1 pmol/µL. The synthetic peptide mix (1 pmol) was spiked into 4 µg of ARS for analysis of PromarkerD. An MRM method, which targets the 12 proteins including 21 endogenous peptides and 17 isotopically labeled peptides was generated.

LC/MS analysis

All samples were separated using the Agilent ZORBAX Eclipse Plus Rapid Resolution C18 analytical column: 50 × 2.1 mm, 1.8 µm in size (part number 959757-902). LC/MS data was acquired for PromarkerD on the Agilent 6495 LC/TQ using the parameters shown in Table 1. Buffer A consisted of 99.9% H_2O (ddH_2O, 18.2 Ω) with 0.1% formic acid and buffer B consisted of 99.9% acetonitrile with 0.1% formic acid. Once the MRM assay was set up and the peptide retention times determined, the data was acquired in dynamic MRM mode. Data analysis for targeted peptide quantification was carried out using Skyline software.

Table 1. Overview of LC/TQ acquisition parameters.

Agilent 6495B Triple Quadrupole Mass Spectrometer			
Ion Mode	AJS ESI, positive mode		
Gas Temperature	150 °C		
Drying Gas Flow	15 L/min		
Capillary Voltage	4,000 V		
Nozzle Voltage	300 V		
High/Low Pressure RF Voltage	150/60 V		
Delta EMV	300 V		
MS1/MS2 Resolution	Unit/Unit		
Cycle Time	500 ms		
Agilent 1290 Infinity II LC			
Time (minutes)	**A (%)**	**B (%)**	**Flow (mL/min)**
2	92	8	0.400
18	70	30	0.400
22	10	90	0.400
25	10	90	0.400
25.10	97	3	0.400
26	97	3	0.400

Results and discussion

Nanoflow to standard flow comparison

To assess the measurement comparability of the PromarkerD biomarkers on nanoflow and standard flow LC/TQ platforms, 12 plasma samples (PI-4205 A-L) were processed by Proteomics International (PI) and then sent to Atturos. PI analyzed the 12 samples on a non-Agilent nanoflow LC/TQ platform. At Atturos, the samples were separated using a 26-minute LC-dMRM method, using the Agilent standard flow 6495 LC/TQ. The MRM peak area of eight selected peptides matching four protein biomarkers (APOA4, CD5L, FHR2, and IBP3) were extracted for relative peak area comparison to the average peak area of each peptide (n = 12, in % terms). This brought the two sets of MS data to scale with each other and allowed observation of any trends across the 12 samples for each peptide. Comparison results for the four proteins and eight peptides are shown in Figure 1. Overall, a consistent trend was observed between results obtained on the nanoflow and standard flow LC/TQ platforms. The data obtained on the Agilent system are largely equivalent to that of PI.

Occasionally deviation between the two data sets was observed, e.g. peptides from APOA4 in PI-4205D. Examination of the targeted MS data of the LEPYADQLR peptide from APOA4 on the 6495 LC/TQ proved good chromatographic peaks with high dot-product (dotp) values between the peak areas and the matching MS/MS peak intensities (Figure 2). In addition, both LEPYADQLR and ISASAEELR peptides of APOA4 showed the same trend in the plasma sample of PI-4205D. This indicates the high quality of targeted MS data on the 6495 LC/TQ for this peptide in PI-4205D. Therefore, the observed deviation might be due to some unknown reason in the sample handling.

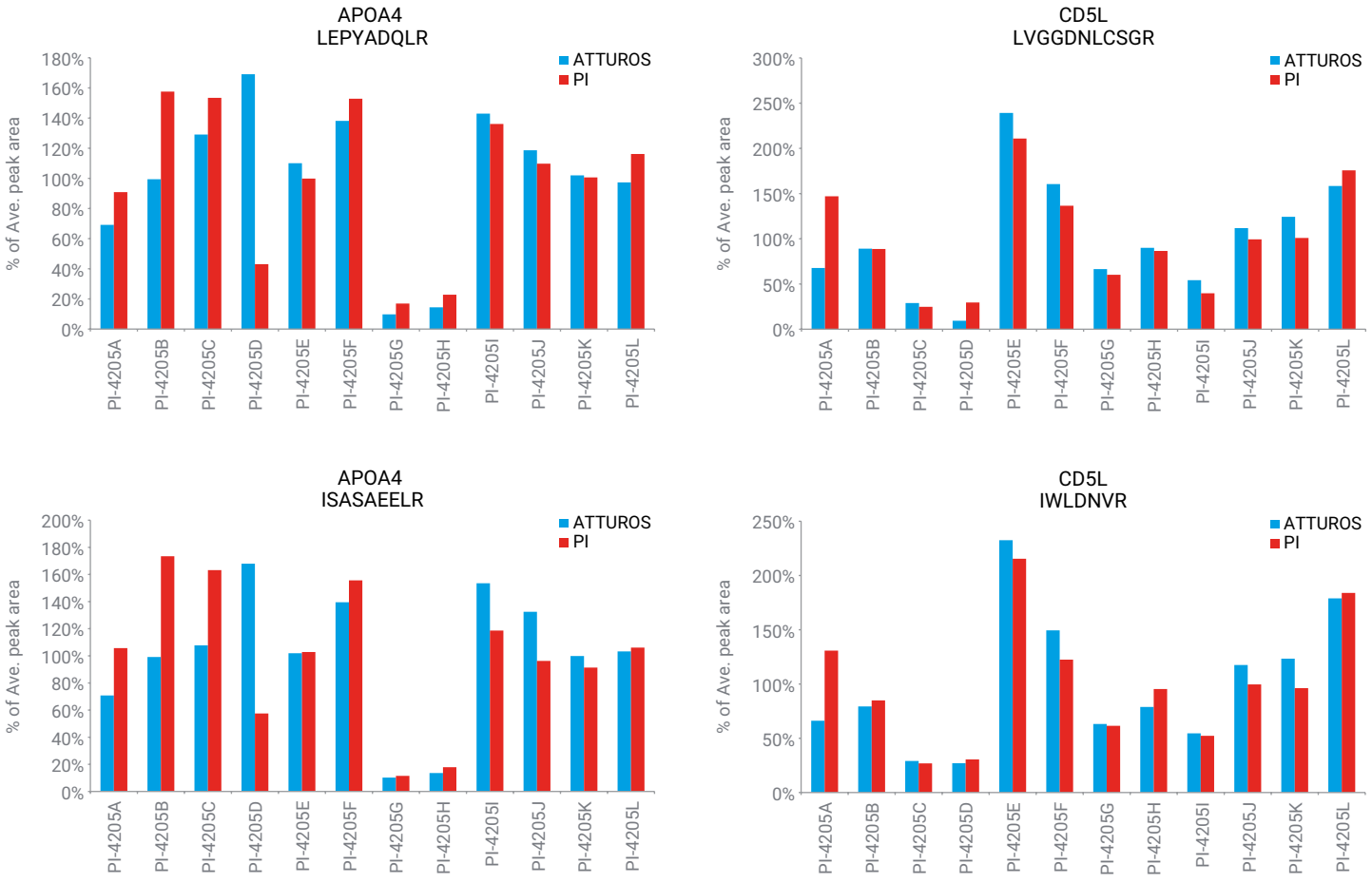
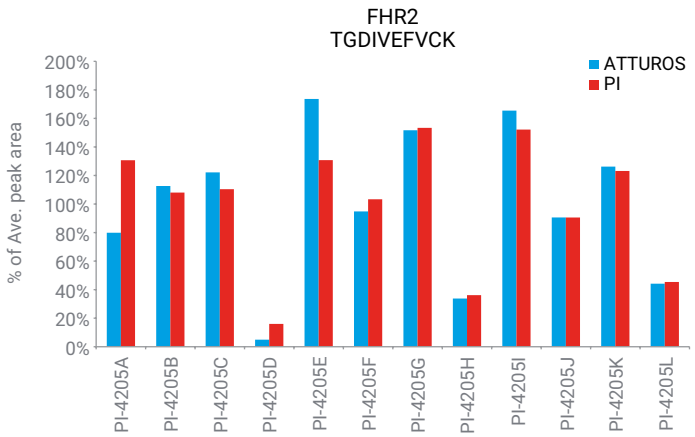
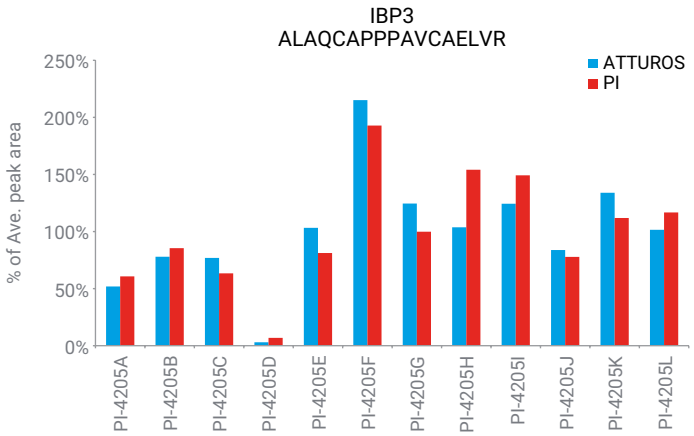
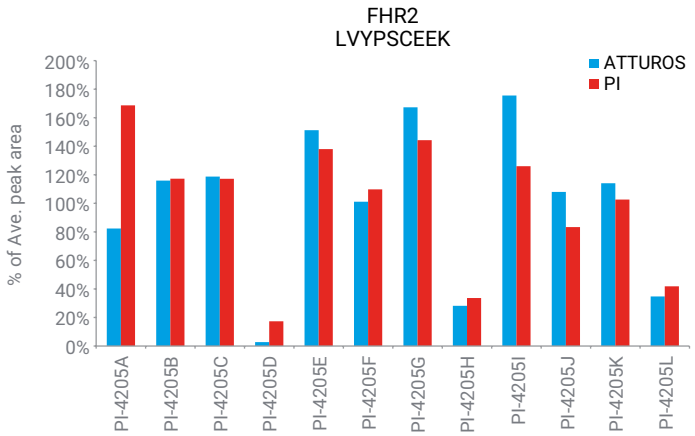
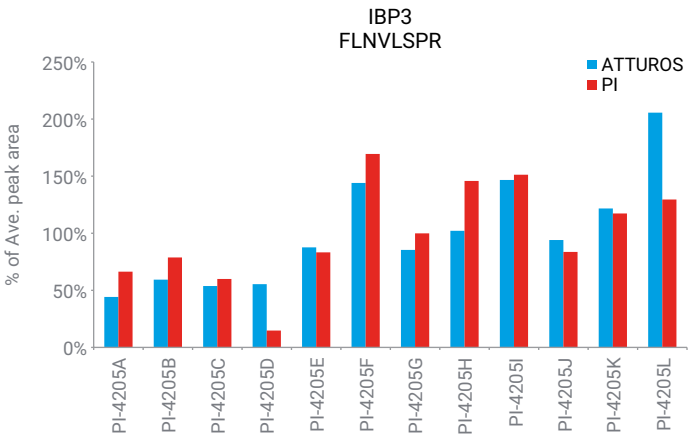


Figure 1. Comparison of eight peptides from four selected proteins (APOA4, CD5L, FHR2, and IBP3) across 12 plasma samples analyzed by nanoflow LC/TQ in PI (red bars) and standard flow LC/TQ in Atturos (blue bars). The individual subject's peak area for a given peptide is expressed as a percentage of the average peak area for that peptide across the 12 plasma samples (continued on next page).


FHR2
TGDIVEFVCK
ATTUROS
PI
% of Ave. peak area
200%
180%
160%
140%
120%
100%
80%
60%
40%
20%
0%
PI-4205A
PI-4205B
PI-4205C
PI-4205D
PI-4205E
PI-4205F
PI-4205G
PI-4205H
PI-4205I
PI-4205J
PI-4205K
PI-4205L


IBP3
ALAQCAPPPAVCAELVR
ATTUROS
PI
% of Ave. peak area
250%
200%
150%
100%
50%
0%
PI-4205A
PI-4205B
PI-4205C
PI-4205D
PI-4205E
PI-4205F
PI-4205G
PI-4205H
PI-4205I
PI-4205J
PI-4205K
PI-4205L


FHR2
LVYPSCEEK
ATTUROS
PI
% of Ave. peak area
200%
180%
160%
140%
120%
100%
80%
60%
40%
20%
0%
PI-4205A
PI-4205B
PI-4205C
PI-4205D
PI-4205E
PI-4205F
PI-4205G
PI-4205H
PI-4205I
PI-4205J
PI-4205K
PI-4205L


IBP3
FLNVLSPR
ATTUROS
PI
% of Ave. peak area
250%
200%
150%
100%
50%
0%
PI-4205A
PI-4205B
PI-4205C
PI-4205D
PI-4205E
PI-4205F
PI-4205G
PI-4205H
PI-4205I
PI-4205J
PI-4205K
PI-4205L

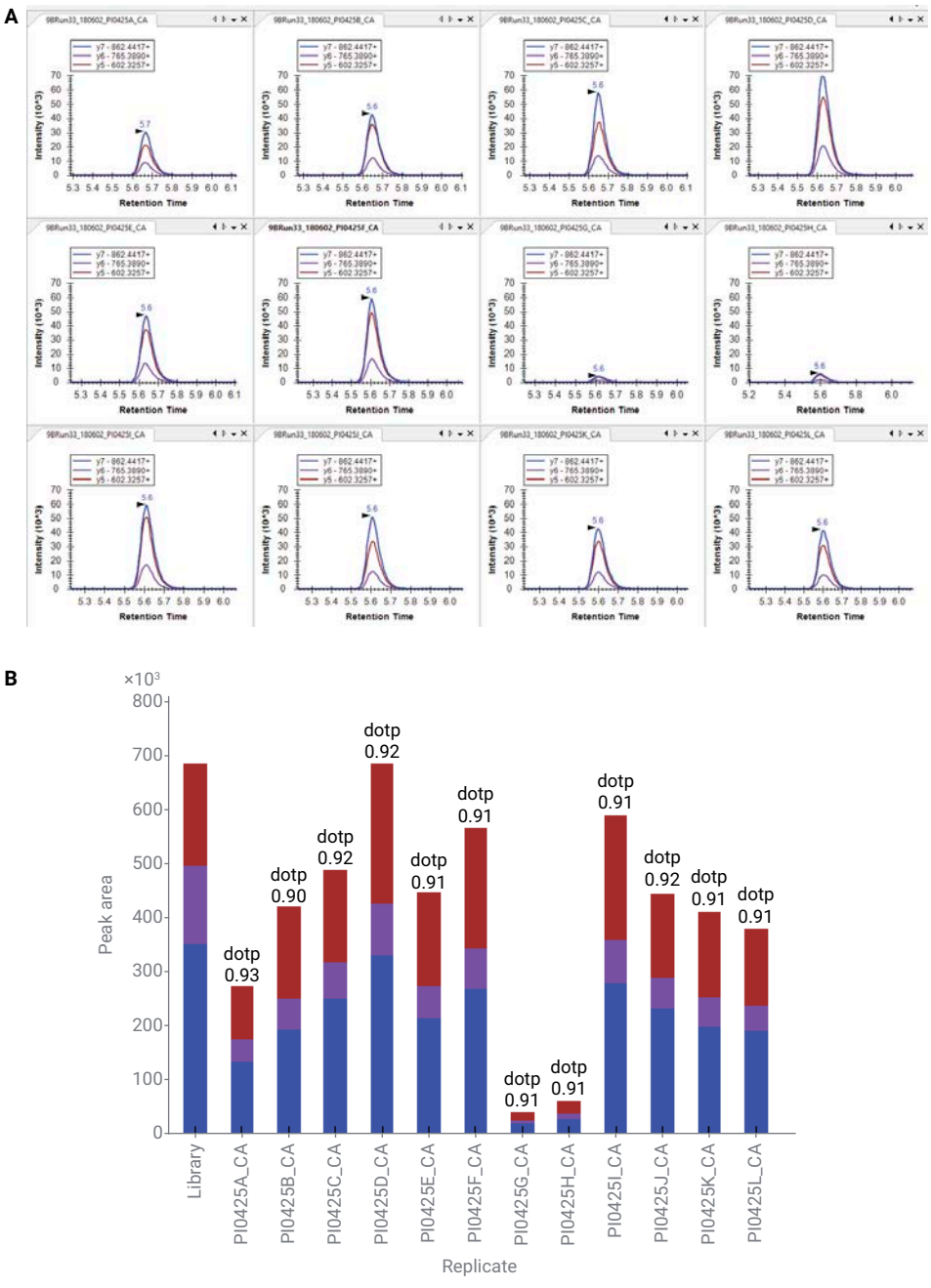


Figure 2. Peak area of the LEPYADQLR peptide from APOA4 across the 12 plasma samples. (A) MRM chromatograms of LEPYADQLR peptide in 12 plasma samples. (B) Stacked bar chart of peak areas of the targeted transitions displayed with dot-product (dotp) value between the peak areas and the matching MS/MS peak intensities.

Depleted plasma versus whole plasma on the 6495 LC/TQ

The PromarkerD biomarkers were measured on the Agilent 6495 LC/TQ in whole reference plasma and reference plasma depleted of the top 14 serum proteins using the Agilent Multiple Affinity Removal Column Human 14. The samples were analyzed using the PromarkerD method parameters described previously in technical replicate injections (n = 6). Replicates 1 to 3 were run at the beginning of the worklist, 40 injections were then run to analyze the 12 plasma samples described previously and finally replicates 4 to 6 of the depleted and whole reference plasma were analyzed; 17-hours after replicates 1 to 3. Figure 3 shows the comparison of MRM chromatograms of peptides from the eight peptides matching four selected proteins (APOA4, CD5L, FHR2, and IBP3) in whole and in depleted reference plasma. In whole reference plasma, the PromarkerD biomarkers produced good MRM chromatograms for all peptides from APOA4 and CD5L, and good MRM chromatograms for at least one peptide from each of FHR2 and IBP3. APOA4, CD5L, and IBP3 have since been independently validated on a separate, larger cohort of people with type 2 diabetes and proved alongside standard clinical features to be an accurate prognostic test for future renal decline.[3] The PromarkerD test may be useful for risk stratification in future clinical trials.

Similar reproducibility was observed in depleted and whole reference plasma with all peptides having a relative standard deviation (RSD) below 20% across the six replicates (Table 2). This result demonstrated that the standard flow LC/MS on the 6495 LC/TQ could be used to measure the PromarkerD biomarkers in both whole plasma and depleted plasma.

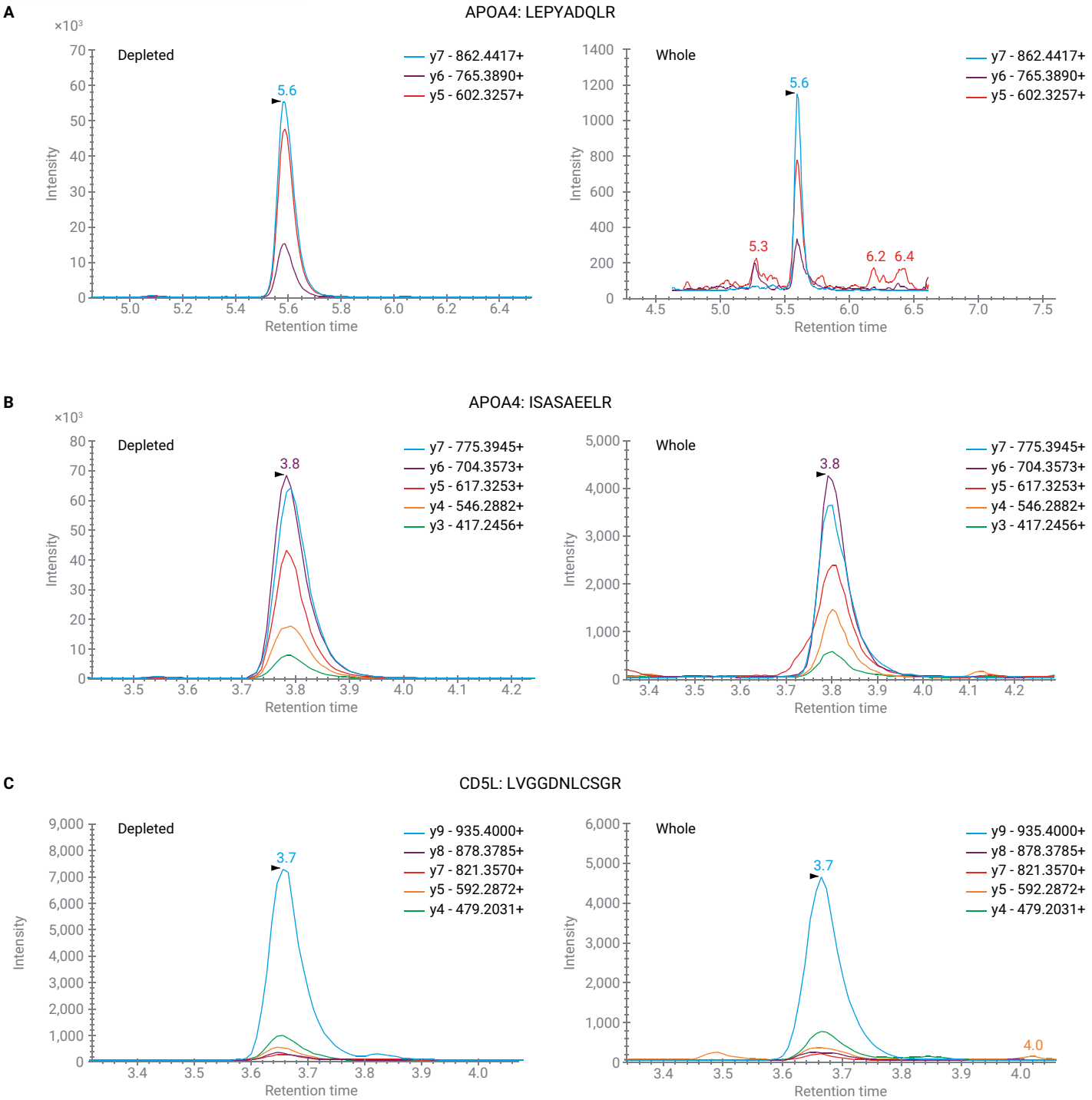
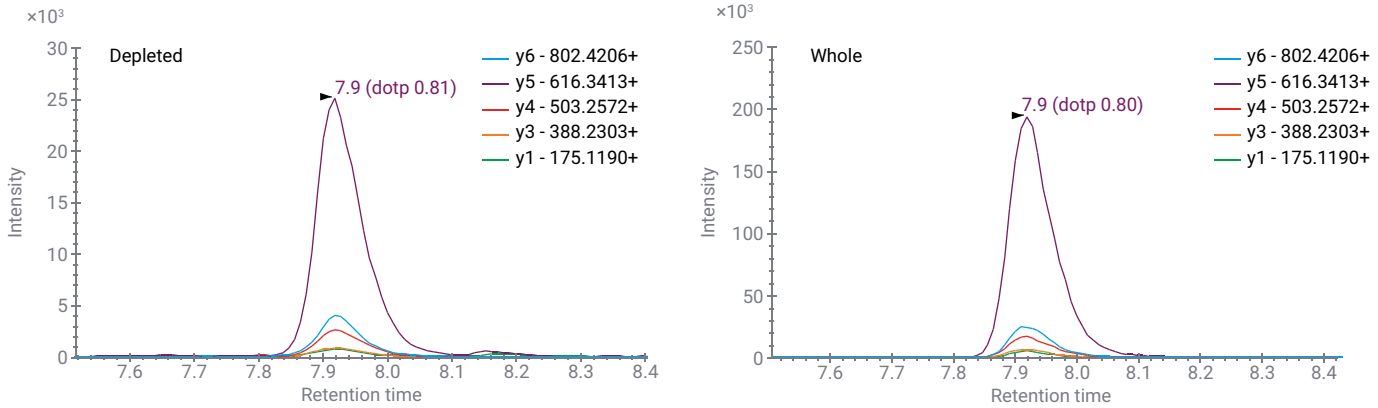


Figure 3. MRM chromatograms for all eight peptides (A-H) targeting four proteins (APOA4, CD5L, FHR2, and IBP3) in depleted and whole reference plasma (continued on next page).

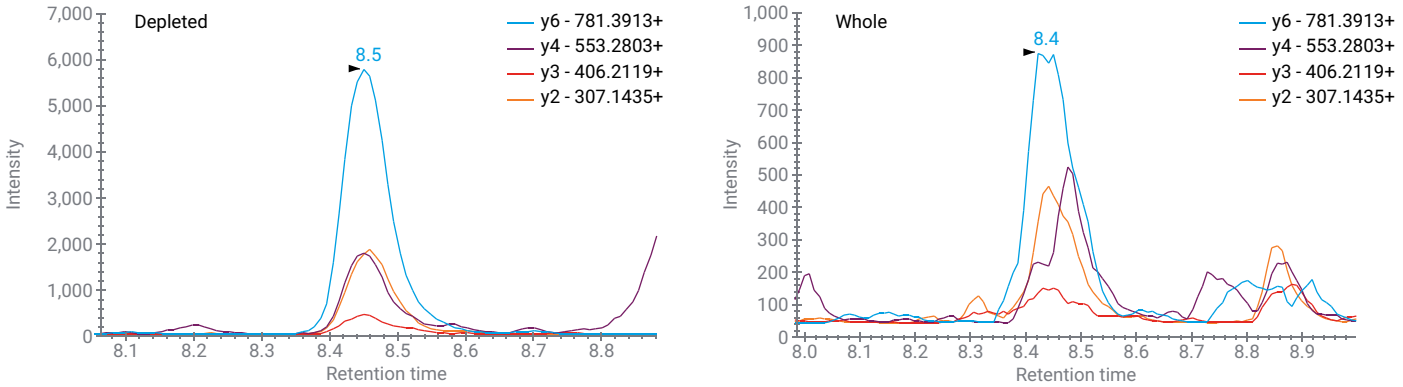
D

CD5L:IWLDNVR



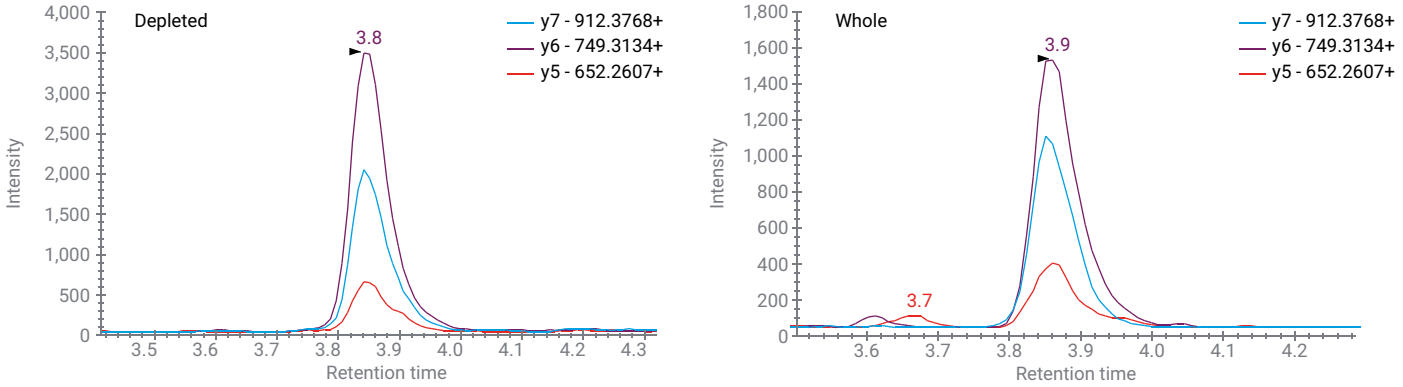
E

FHR2: TGDIVEFVCK

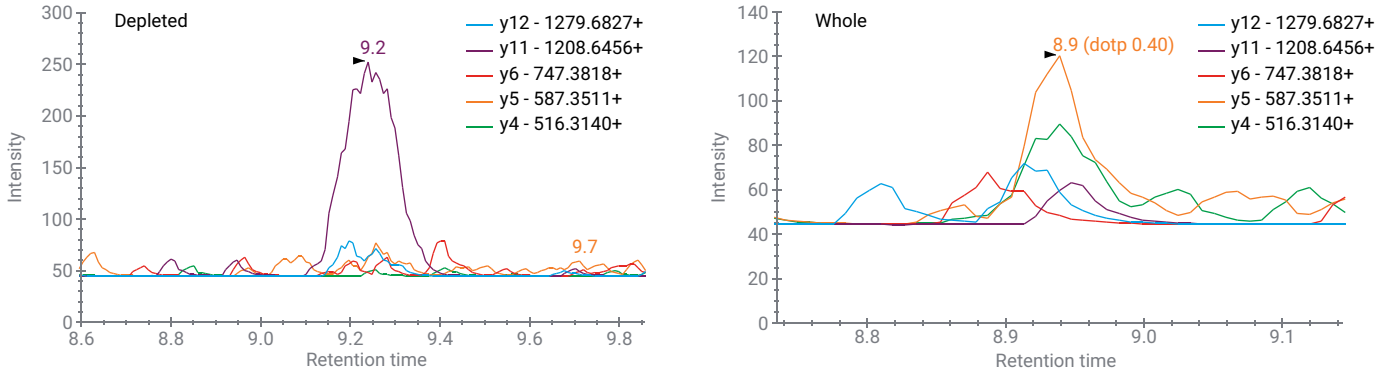


F

FHR2:LVYPSCEEK



G



H

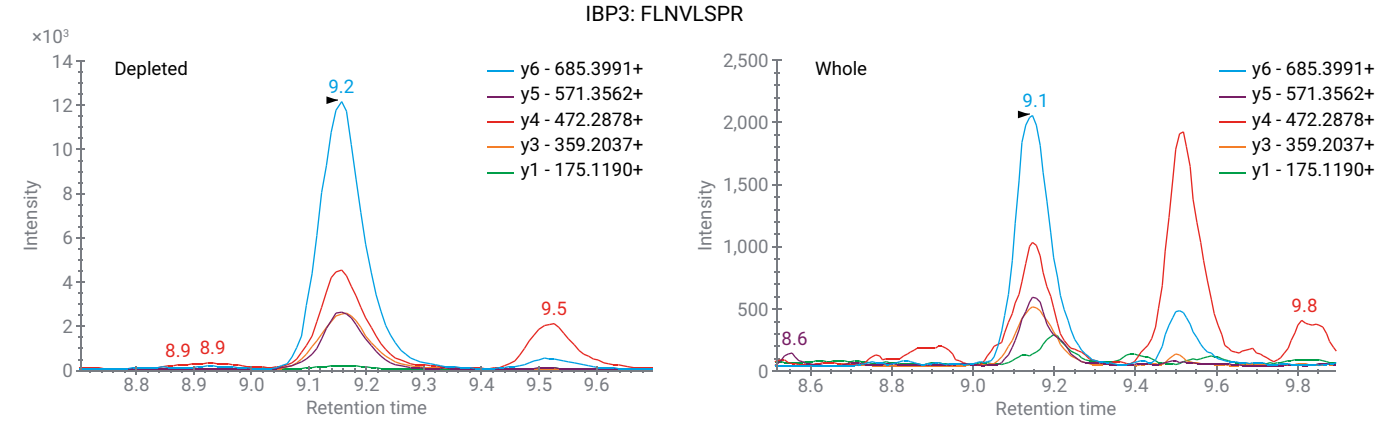


Table 2. Peak areas for all eight peptides in depleted and whole reference plasma. Reproducibility is shown in terms of percentage relative standard deviation (RSD) across six nonconsecutive- technical replicate injections of each sample measured over a 22-hour worklist.

			Depleted Reference Plasma (n = 6)		Whole Reference Plasma (n = 6)	
Accession	Protein	Peptide	Peak area	Peak area RSD	Peak area	Peak area RSD
P06727	APOA4	LEPYADQLR	257,570	9.97%	3,944	14.52%
P06727	APOA4	ISASAEELR	303,069	10.99%	18,759	14.11%
O43866	CD5L	LVGGDNLCSGR	31,643	10.50%	19,117	12.86%
O43866	CD5L	IWLDNVR	136,872	7.05%	947,395	6.12%
P36980	FHR2	TGDIVEFVCK	26,877	10.58%	7,392	17.39%
P36980	FHR2	LVYPSCEEK	15,943	6.91%	6,898	6.60%
P17936	IBP3	ALAQCAPPPAVCAELVR	1,894	16.67%	127	18.54%
P17936	IBP3	FLNVLSPR	68,437	10.00%	16,725	13.00%

Conclusion

The reproducibility, robustness, and analytical sensitivity of MRM-based LC/MS methods for peptide quantification in biological fluids are important considerations for high-throughput protein biomarker quantitation studies. This application note discusses the transfer of an MRM method from nanoflow to standard flow LC/TQ platforms and the analytical sensitivity of the Agilent 6495 LC/TQ in whole and depleted human plasma. For 12 subjects, equivalent results were obtained on the standard flow system compared to nanoflow. Standard flow LC/TQ measurement of protein biomarkers shown here does not negatively affect the data output. This application note also shows the analytical sensitivity of standard flow; when coupled with the 6495 LC/TQ it is capable of reproducibly detecting the four PromarkerD biomarker proteins in whole plasma without depletion. These results demonstrated the suitability of this high-performance LC/MS system on routine protein measurement in complex matrices for potential clinical research assay development in the future.

References

1. Dunkler, D. *et al.* Risk Prediction for Early CKD in Type 2 Diabetes. *Clin. J. Am. Soc. Nephrol.* **2015** Aug 7, *10(8)*, 1371–9.
2. Peters, K. E. *et al.* Identification of novel circulating biomarkers predicting rapid decline in renal function in type 2 diabetes: the Fremantle diabetes study phase II. *Diabetes Care*. **2017**, *40(11)*, 1548–55.
3. Peters, K. E. *et al.* Validation of a Protein Biomarker Test for Predicting Renal Decline in Type 2 Diabetes: The Fremantle Diabetes Study Phase II. *J. Diabetes Complications* **2019** Aug 27, 107406.
4. Bringans, S. D. *et al.* Comprehensive Mass Spectrometry Based Biomarker Discovery and Validation Platform as Applied to Diabetic Kidney Disease. *EuPA Open Proteomics* **2017** Mar 1, *14*, 1–10.
5. Percy, A. J. *et al.* Comparison of Standard- and Nano-Flow Liquid Chromatography Platforms for Mrm-Based Quantitation of Putative Plasma Biomarker Proteins. *Anal. Bioanal. Chem.* **2012** Sep, *404(4)*, 1089–101.
6. Davis, T. M.; Bruce, D. G.; Davis, W. A. Cohort Profile: The Fremantle Diabetes Study. *Int. J. Epidemiol*. **2013** Apr 1, *42(2)*, 412–21.

Printed in the USA, October 9, 2020
5994-2381EN

Application Note

Proteomics

Agilent

Trusted Answers

Peptide Quantification in Plasma Using the Agilent 6495 Triple Quadrupole LC/MS Coupled with the Agilent 1290 Infinity II LC System

Author

Linfeng Wu
Agilent Technologies, Inc.

Abstract

This application note showcases the quantitative performance of multiple reaction monitoring (MRM)-based LC/MS analysis of peptides derived from protein biomarkers in human plasma using the Agilent 1290 Infinity II LC system coupled to the Agilent 6495 triple quadrupole LC/MS with Jet Stream ionization source. Results demonstrated outstanding performance of peptide quantification in human plasma using this standard flow-based LC/TQ system.

Introduction

In translational research, increased throughput as well as increased multiplexing mass spectrometry using MRM-based LC/MS methods have become popular for targeted, bottom-up quantification of protein biomarkers. Researchers often target signature peptides from a limited number of proteins in large cohorts and monitor their expression levels during a specified time period, often leading to hundreds or thousands of biological samples. In such large-volume studies, high-throughput, robustness, and reproducibility are essential when deploying LC/MS methods. For this reason, standard-flow triple quadrupole LC/MS (LC/TQ) is superior to other low-flow LC/MS platforms. In addition to reproducibility and robustness, analytical sensitivity is another important consideration for peptide quantification. The 6495 triple quadrupole LC/MS system coupled with an Agilent Jet Stream (AJS) ionization source is a standard flow-based LC/TQ platform that provides outstanding performance for targeted peptide quantification due to its high sensitivity, reproducibility, and robustness as well as ease of handling and maintenance.[1,2]

This application note demonstrates the quantitative analytical sensitivity of 99 signature peptides from 75 protein biomarkers in plasma using the 6495 triple quadrupole LC/MS system and AJS ionization source. To assess overall quantitative performance for protein biomarker peptides, a commercially available mixture of stable isotope-labeled standard (SIS) peptides was spiked into human plasma digest at nine different concentrations. Standard curves of 99 SIS peptides were generated to determine their lower limit of quantification (LLOQ) in plasma matrix. The results show excellent analytical performance under standard flow conditions with this LC/TQ system for targeted peptide quantification in heavy matrix.

Experimental

Instrumentation

- Agilent 1290 Infinity II LC system including:
  - Agilent 1290 Infinity II High Speed Pump (G7120A)
  - Agilent 1290 Infinity II Multisampler (G7167B) with sample cooler option (option 100)
  - Agilent 1290 Infinity II Thermostatted Column Compartment (G7116B)
- Agilent Jet Stream source (G1958-65638)
- Agilent 6495C triple quadrupole LC/MS system (G6495C)

Materials

Raw human plasma was purchased from Bioreclamation (catalog no. HMPLEDTA2). PeptiQuant Biomarker Assessment Kit (BAK-A6495-76) was purchased from Cambridge Isotope Laboratories.

Sample preparation

Human plasma was diluted with 25 mM ammonium bicarbonate followed by denaturation with TFE, reduction with DTT, and alkylation with Iodoacetamide. The sample was then trypsin digested overnight at 37 °C. The digested plasma was dried down, then reconstituted at around 1 µg/µL and spiked with the balanced stable-isotope standard (SIS) peptide mixture at nine different concentrations ranging from 5 amol/µL to 100 fmol/µL (which is converted to 350 fmol/mL to 7 nmol/mL in the original raw human plasma).

LC/MS analysis

All samples were injected with replicates (n = 5) and analyzed using the LC/MS acquisition method provided by the PeptiQuant Biomarker Assessment Kit with minor modifications.[3] As briefly summarized in Table 1, 10 µL of samples were loaded onto an Agilent ZORBAX RRHD Eclipse Plus C18 column (2.1 × 150 mm, 1.8 µm, part number 959759-902) and separated using the same 43-minute LC gradient provided by the kit. On the current system, the peptide retention times were first updated using the *Dynamic MRM Update Options* in Agilent MassHunter software to adjust the minor retention time shift, then AJS source parameters were optimized using the Source Optimizer. The final MS parameters are listed in Table 1. All the other LC/MS parameters, such as MRM transitions and their collision energies, were directly copied from the original method. A total of 99 SIS peptides from 75 proteins with MRM responses in the original SIS peptide mixture were used in the final method for standards curve creation.

Data processing

Agilent MassHunter workstation Quantitative Analysis software (v10.1) was used for peptide quantification analysis.

Results and discussion

Chromatography of standard-flow LC/MS

The 1290 Infinity II LC system, the next generation in UHPLC, gives excellent chromatography resolution and higher retention precision. The human plasma digest spiked with SIS peptides at 10 fmol/µL was injected onto a ZORBAX RRHD Eclipse Plus C18, 2.1 × 150 mm, 1.8 µm column using a 43-minute LC/MS analysis performed in dynamic MRM mode. Excellent peak shapes and separation were achieved (Figure 1). All targeted peptides eluted between 2.4 to 36.5 minutes, showing a median full width at half maximum (FWHM) of 3.5 seconds. Figure 1B shows the excellent peak shapes of four selected SIS peptides

Table 1. LC/MS parameters.

Agilent 1290 Infinity II LC Parameters	
Analytical Column	Agilent ZORBAX RRHD Eclipse Plus C18, 2.1 × 150 mm, 1.8 µm (p/n 959759-902)
Mobile Phase A	H_2O, 0.1% formic acid
Mobile Phase B	Acetonitrile, 0.1% formic acid
Flow Rate	0.4 mL/min
Injection Volume	10 µL
LC Gradient Time	43 min
Column Temperature	50 °C

Agilent 6495 LC/TQ Parameters	
Ion Mode	Agilent Jet Stream, positive
Gas Temperature	150 °C
Drying Gas Flow	17 L/min
Nebulizer Gas	30 psi
Sheath Gas Temperature	250 °C
Sheath Gas Flow	12 L/min
Capillary Voltage	3500 V
Nozzle Voltage	0 V
High-/Low-Pressure RF Voltage	200/110 V
Delta EMV	200 V
Q1 and Q3 Resolution	Unit/Unit
Cycle Time	600 ms
Minimum Dwell Time	7.02 ms
Total MRM Transitions	606
HED in Tune File	–20 kV

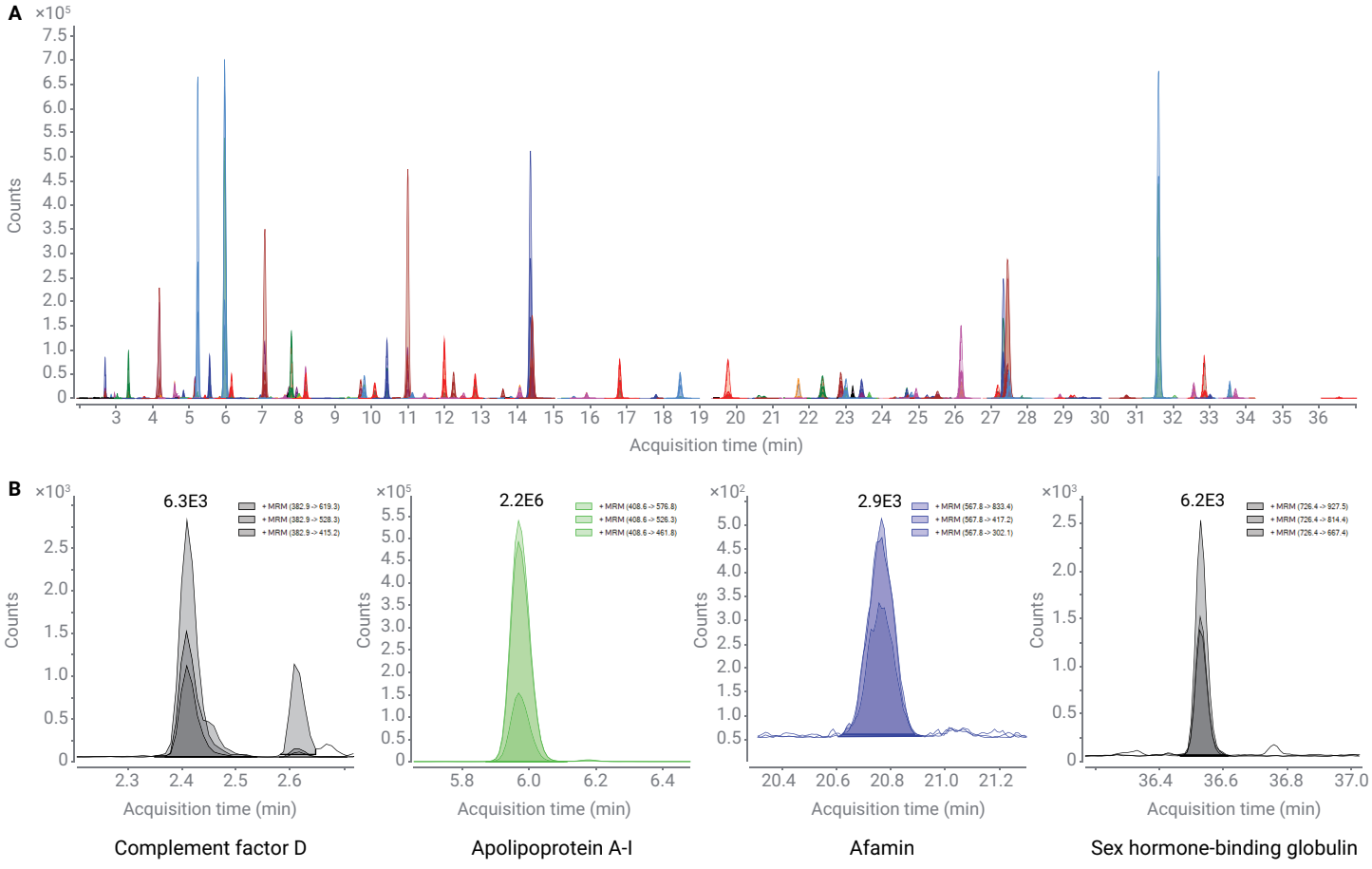


Figure 1. MRM chromatograms of heavy and endogenous peptides in human plasma digest spiked with SIS peptide mixture at 10 fmol/µL. (A) total dynamic MRM chromatograms of all targeted heavy and endogenous peptides. (B) MRM peak chromatograms of four selected SIS peptides spanning wide ranges of retention time (2.4 ~ 36.5 minutes) and signal response (peak area 2.9E3 ~ 2.2E6).

spanning a wide range of retention time (2.4 ~ 36.5 minutes) and signal response (peak area 2.9E3 ~ 2.2E6). The targeted SIS peptides differ widely in terms of signal response when loaded equally on column. This indicates broad differences in the analytical sensitivity of various peptides.

To assess the precision of peptide retention time (RT), the relative standard deviation (RSD) of all targeted endogenous peptides were calculated from all the 45 injections (9 levels × 5 replicates) since they have good and equal signal response across all the spiking samples. Distribution of peptide retention time versus its corresponding RSD shows the RSD ranges from 0.01 to 0.42%, with a median RSD of 0.13% (Figure 2). This result shows much better RT reproducibility compared to a low microflow LC system.[4] It demonstrated the excellent retention time precision of the standard flow 1290 Infinity II LC system.

Quantification of peptide standards in human plasma

To evaluate quantitative performance of the 6495 triple quadrupole LC/MS with AJS ionization source for peptide quantification, the balanced SIS peptide mixture was spiked at nine different concentrations ranging from 5 amol/µL to 100 fmol/µL in 1 µg/µL plasma digest. Replicate (n = 5) injections were made for all samples to make linear standard curves. Eighty-one percent (80 out of 99) of targeted SIS peptides show R^2 equal to or greater than 0.99. The LLOQ of each SIS peptide was determined as the lowest spiking level with MRM response RSD <20% and linear quantification accuracy between 80 to 120%. As expected, the LLOQs of all 99 SIS peptides show a wide range of difference (Figure 3). For well-responding peptides, LLOQs as low as 5 amol/µL in complex plasma matrix could be achieved. This indicates the importance of peptide and MRM transition selection to achieve the best quantitative sensitivity.

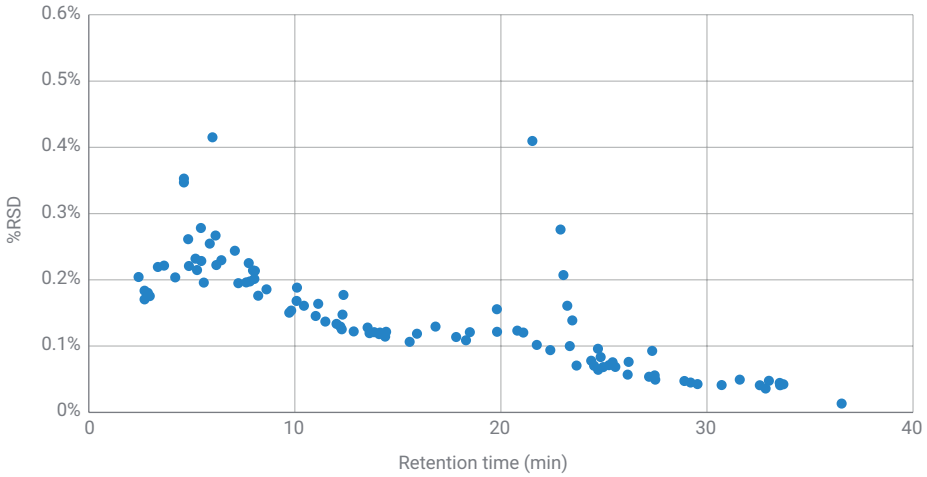


Figure 2. Distribution of retention time versus its corresponding RSD for 99 targeted endogenous peptides in human plasma digest (n = 45).

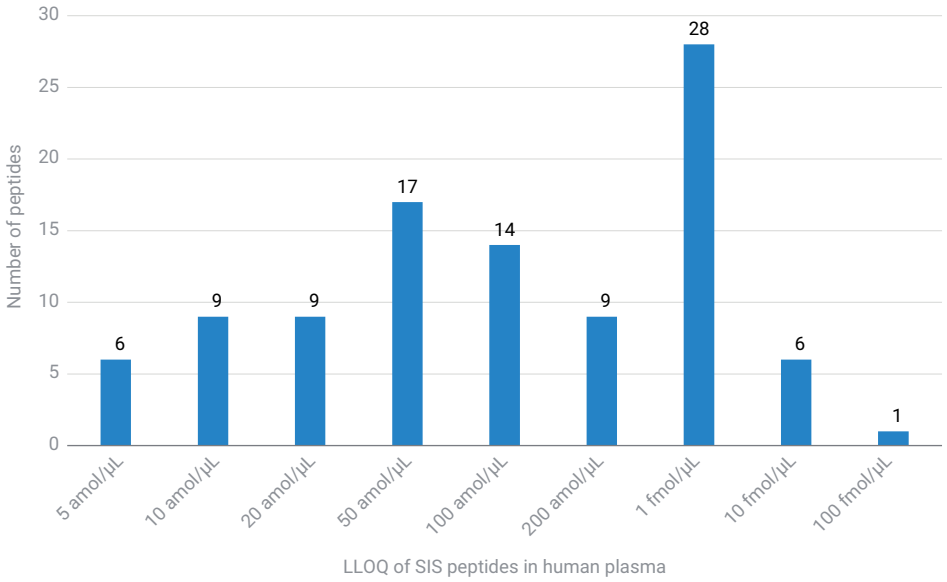


Figure 3. Distribution of LLOQ of SIS peptides spiked into human plasma. 10 µL of each standard sample was injected leading to 10 µg plasma digest matrix on column. The LLOQ of SIS peptide was determined as the lowest spiking level, with MRM response RSD<20% and linear quantification accuracy between 80 to 120%.

Figure 4 and Table 2 show the excellent quantitative performance of the SIS peptide AFLLTPR from apolipoprotein M in plasma including:

- Excellent linear dynamic range (5 amol/µL to 100 fmol/µL) with R^2 = 0.9998 (Figure 4A)
- Excellent analytical sensitivity with an LLOQ of 5 amol/µL in 1 µg/µL trypsinized plasma (Figure 4)
- Superior precision and accuracy observed at all levels, including the LLOQ level (Table 2)

Table 2. Precision and accuracy for the Agilent 6495 triple quadrupole LC/MS analysis of the SIS peptide standard AFLLTPR in 1 µg/µL plasma digest.

Level (amol/µL)	Average Response	RSD (%) (n = 5)	Accuracy (%) (n = 5)
Matrix blank	3	NA	NA
5	117	9.3	114.2
10	251	9.7	102.4
20	562	4.6	104.0
50	1,295	1.6	91.3
100	2,925	4.1	100.9
200	5,174	3.3	88.6
1,000	28,396	1.3	96.4
10,000	302,313	2.5	102. 5
100,000	2,944,790	1.1	99.8

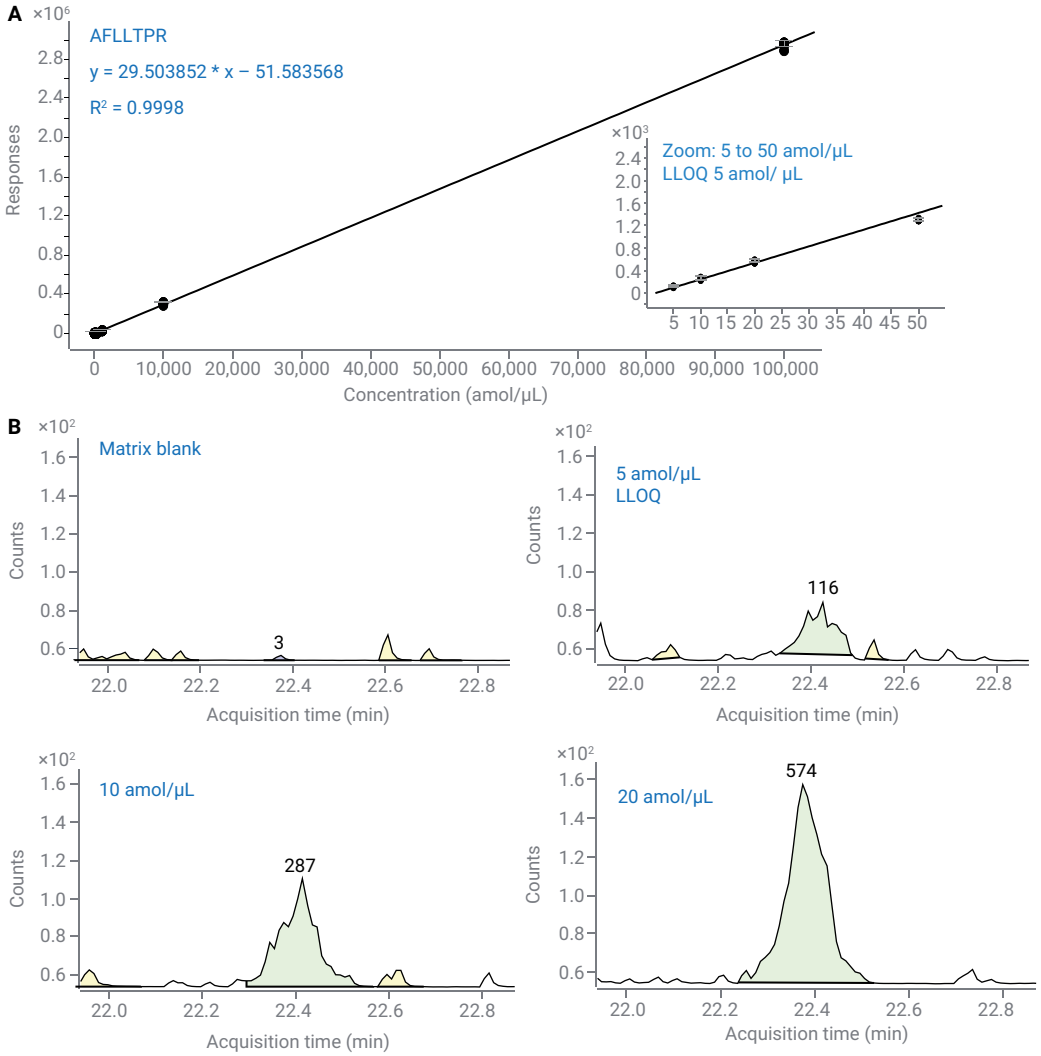


Figure 4. Quantitative performance of the SIS peptide AFLLTPR from apolipoprotein M in plasma. (A) standard curve of SIS peptide AFLLTPR for the range from 5 amol/µL to 100 fmol/µL in 1 µg/µL plasma digest with inset showing detail for the range of the curve from 5 to 50 amol/µL. (B) extracted ion chromatograms for peptide AFLLTPR showing the LLOQ.

The LLOQ results show that analytical sensitivity not only depends on the strength of monitored transition signal response, but may also be affected by matrix interference, response reproducibility, linearity, etc. For example, Figure 5 shows extracted ion chromatograms of SIS peptide SGIPIVTSPYQIHFTK from complement C3 at low levels labeled with the response ratio of qualifier ion transition relative to quantifier ion transition. The matrix blank contains interferences, showing the ion response ratios deviated from the expected ratios. These interferences affected the linear quantification accuracy at the low concentration of 5 amol/µL (RSD = 2.7%, accuracy = 157.5% for n = 5). Therefore, the LLOQ of peptide SGIPIVTSPYQIHFTK was determined to be 10 amol/µL (RSD = 2.4% and accuracy = 112.8% for n = 5), even though the linear standard curve R^2 of 0.9995 was achieved. Note that the quantitative sensitivity of SIS peptide SGIPIVTSPYQIHFTK is comparable to a previous result using a microflow LC system coupled to 6495 LC/TQ, which has a lower sample loading capacity.[4] This demonstrated that the 6495 triple quadrupole LC/MS, coupled with the 1290 Infinity II LC system and Jet Stream ionization source, can achieve similar quantitative sensitivity to the microflow LC/TQ system if not sample restricted.

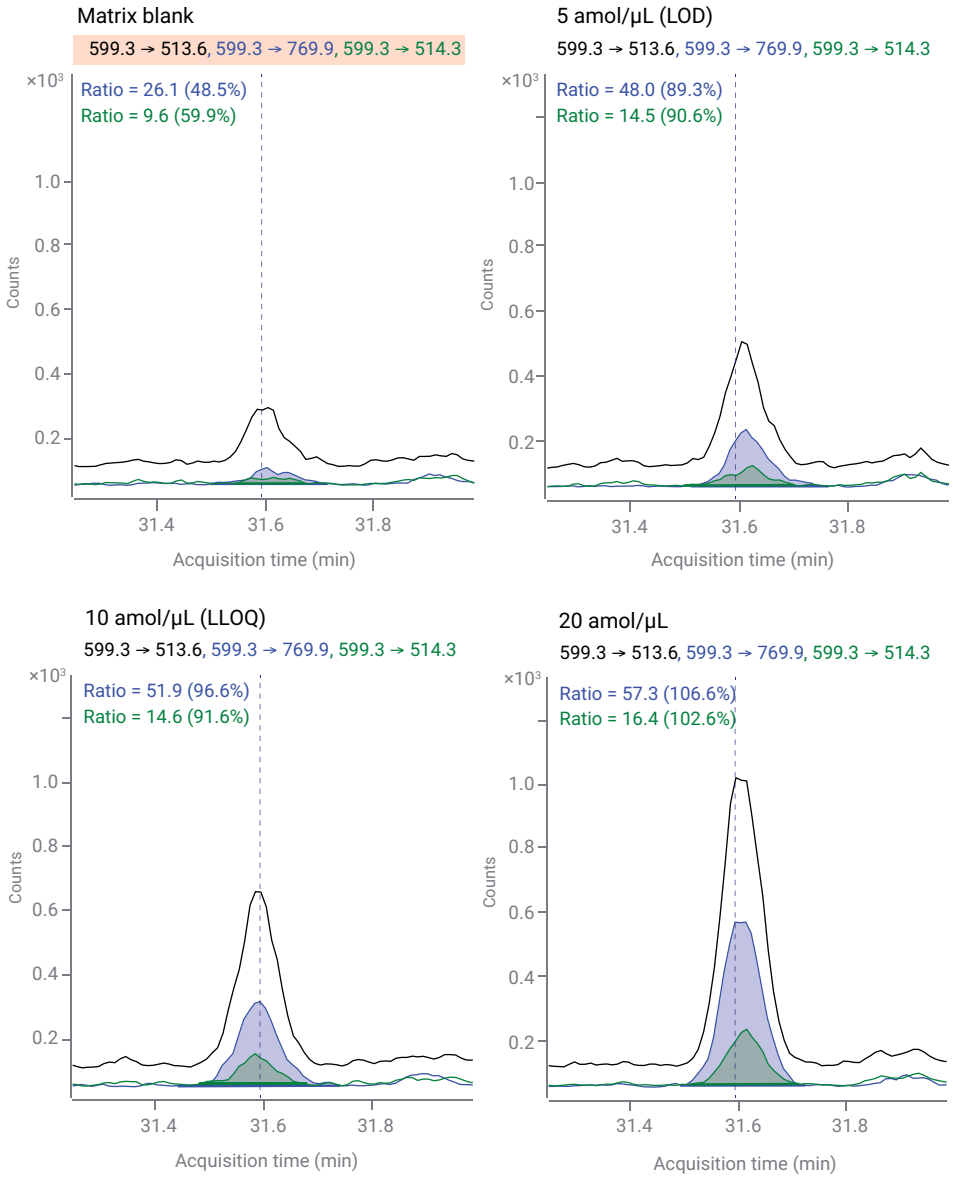


Figure 5. Extracted ion chromatograms of SIS peptide SGIPIVTSPYQIHFTK from complement C3 at low levels. The measured response ratios of qualifier relative to quantifier were labeled with percent of expected ratio in brackets. The title of chromatogram was highlighted in color when the percent of expected ratio deviated above 30% uncertainty in this case. The dashed line indicates the reference retention time of the targeted peptide.

Conclusion

The analytical sensitivity, precision, accuracy, and robustness of MRM-based LC/MS methods for peptide quantification in biological fluids are important considerations for high-throughput biomarker verification. This application note describes the overall analytical sensitivity of the 6495 triple quadrupole LC/MS for peptide quantification in human plasma. Using a commercially available biomarker assessment kit from a third-party vendor, the outstanding performance of the 6495 LC/TQ coupled with the 1290 Infinity II LC system and Jet Stream ionization source was demonstrated for peptide quantification. Based on the physicochemical property of the peptides at hand, various peptides have wide differences in MRM signal response, which leads to huge variation (>four orders of magnitude) on quantitative sensitivity. In this experiment, an LLOQ of 5 amol/µL was achieved in heavy plasma matrix for some peptides, which is comparable to the results using a low-microflow LC/TQ platform.[4] This result demonstrated the excellent analytical sensitivity of the 6495 triple quadrupole LC/MS coupled with the 1290 Infinity II LC system and Jet Stream ionization source.

References

1. The Agilent 6495 Triple Quadrupole LC/MS: Peptide Quantitation Performance. *Agilent Technologies technical overview*, publication number 5991-6898EN, **2016**.
2. Jet Stream Proteomics for Sensitive and Robust Standard Flow LC/MS. *Agilent Technologies technical overview*, publication number 5991-5687EN, **2015**.
3. Percy, A.J. *et al.* A Standardized Kit for Automated Quantitative Assessment of Candidate Protein Biomarkers in Human Plasma. *Bioanalysis* **2015**, *7(23)*, 2991–3004.
4. Robust and Reproducible Protein Quantification in Plasma using the Evosep One and the Agilent 6495 Triple Quadrupole LC/MS. *Agilent Technologies application note*, publication number 5994-1928EN, **2020**.

Printed in the USA, September 10, 2020
5994-2285EN

Agilent
Trusted Answers

Characterization of Proteins Using Electron Capture Dissociation Technology

Proteoforms – the different forms of proteins made from the genome with a variety of splice variations, potentially thousands of post translational modifications – are the critical elements in all biological systems. Recent technological developments have enabled comprehensive proteoform analyses in complex samples, and an increasing number of laboratories are adopting top-down proteomics workflows allowing for the highest molecular precision for analysing primary structures by examining proteins in their intact state.

In this interview, we speak with Dr Joe Beckman a professor at Oregon State University who has used mass spectrometry for two decades to probe the causes of amyotrophic lateral sclerosis (ALS).

Dr Beckman has recently been collaborating with Dr. Voinov and Prof. Doug Barofsky to found e-MSion to develop their inventions making electron capture dissociation (ECD) possible in common mass spectrometers. e-MSion aims to advance proteomics research by improving the accessibility of electron-based fragmentation in mass spectrometry (MS). Dr Beckman provides insights into proteoform analysis and how e-MSion can be used in combination with MS-based technologies to better characterize proteoforms.

Q: What are proteoforms and how are they created?

A: The term "proteoform" was first coined to describe all of the changes that happened to proteins after they were synthesized in the cell. After the human genome was sequenced, it was shocking to find that only approximately 20,000 genes make up the human body; this is not much different from a mouse or even a fruit fly. However, splicing of messenger RNA and decoration with hundreds of protein modifications result in the formation of at least one million different proteoforms from these 20,000 genes. If one considers antibodies, everyone can produce billions of distinct antibodies emanate from a basic set of a few hundred genes.

The real challenge we face as biochemists is to characterize the multiplicity of proteoforms, as we aren't able to get that information directly from DNA sequences. The biopharmaceutical industry faces further challenges during protein production in bioreactors; the cells can make unknown changes and even more modifications occur after therapeutic proteins are purified more changes occur in storage. Characterizing these subtle modifications really pushes the limits of current technology. The biopharmaceutical industry applies a battery of about 20 different biophysical techniques to understand what's happening to therapeutic proteins and how they are modified. Looking beyond the biopharmaceutics, gaining further insights into protein modifications is critical for understanding human biology and the processes

underlying disease.

Q: Why is it important to characterize proteoforms, and what are the traditional methods used for characterization?

A: For 50 years, the primary method to probe protein structure and modifications has been a combination of gel electrophoresis with specific antibodies selected to recognize specific forms of protein modifications. Monoclonal antibodies were a huge advance to have to understand biological processes. However, antibodies can take months to develop, are expensive, and can be imprecise and subject to many artifacts.

Mass spectrometry was an important advance in the early 1990s, as it provided much greater precision for characterizing proteins. However, at the time these instruments could only handle small fragments of a protein. Nowadays we are starting to work with much larger pieces frequently the entire protein at once. This enables us to characterize how these proteins are modified and what these modifications are by weighing the entire protein to the precision of a single neutron.

Q: Can you discuss the development of the e-MSion technology and the ExD cell?

A: In the late 1990s, thanks to the work by Roman Zubarev, electrons were recognized as being highly effective at breaking proteins into manageable pieces for analysis. However, the approach was rather impractical as it could only be done with very expensive high-end research instruments. By chance, I became involved in a project to make this technique practical on much more affordable instruments as a professor at Oregon State University. In the early 2000's, we invited a Russian scientist, Valery Voinov to work with us to develop better instrumentation.

During the first week after he arrived, he started to examine a complicated research apparatus that used resonant electrons to ionize molecules, which wasn't working. We came back from lunch to find he had taken apart the core of this complex instrument and held gold-plated electrical elements over a candle to coat the outside with soot. He was screwing it all back together and it started to work. He then very simply said "backscattered electrons". It was a problem we had struggled with for months. It became my end of the day to sit with Valery in the evening while he was taking apart and redesigning mass spectrometers and learn what he is doing.

From his knowledge of magnetic fields, Valery developed an interesting idea to confine sufficient low energy electrons to create a simpler way to fragment proteins. With his early results, we were able to get research grants that brought in significant funding to develop the technique. We started an academic partnership through George Stafford with Agilent and they provided us with a triple quad mass spectrometry instrument. They were concerned when they learnt that Valery had drilled a hole in the side of the vacuum chamber. But this soon overlooked after he started to produce some highly promising results. But it was hard to convince mass spec manufacturers that the technology would be of practical use and commercially viable. The university was encouraging professors to commercialize their inventions, so we filed for patents and launched e-MSion to develop the technology.

The name we chose started came from Valery, combining an e^- for an electron, the MS for mass spectrometry, and then the "ion". We then added "ECD for the Masses", which meant that this technology could work with almost any mass spectrometer and affordable for any laboratory.

We have been successful in developing the company by applying for a small business grants called SBIR's and are now making sales. Agilent has become a really important partner for e-MSion and allowed deep access inside the instruments whereby we are producing amazing new capabilities. It allowed us to solve far more complicated problems, using affordable instruments.

Q: How does ExD cell overcome challenges associated with current fragmentation methods?

A: The current methods that are common on instruments are basic; you need to add a lot of energy to peptides to collide into gas molecules, in effect sandblasting peptides to break them apart.

This collision induced dissociation (CID) knocks off many bits from all over the peptide. Although the method is efficient and simple to implement, it can give hundreds to thousands of peaks. The vast majority just aren't recognizable. This complicates the interpretation and increases the chances of false identifications.

This is why we developed the ExD cell, a 30mm long collection of DC electrodes, magnets and a hot filament that efficiently fragments proteins into smaller pieces without slowing the flow of ions through the mass spectrometer.

In a typical mass spectrometer, peptides and proteins are flying at the speed of a commercial jet airline through the instrument and the few microseconds needed to traverse the ExD cell, we've completed the fragmentation process and broken them into interpretable fragments. The advantages are the clean spectrum that breaks an entire protein as neatly as a peptide, but without losing the post translational modifications.

ECD gives several important advantages over CID. For example, leucine and isoleucine are two amino acids that appear to be identical in a standard MS analysis. After an electron is captured it migrates to the side chain and breaks off different sized pieces from these two amino acids. And this is what lets us distinguish these two different forms that you can't see in a traditional mass spec. As 16% of the human proteome contains Leucine or isoleucine, this creates a lot of ambiguity when you're trying to understand what is happening. Current instruments don't detect this, although they know it's a 'nice to have', and students aren't taught this either as its so uncommon to see such sidechain fragmentation. With the ExD cell, it's a pretty common feature. It lets us look for other modifications that happen with aging. This is a huge advantage as although they weigh the same amount we can break the molecule in ways that reveal their presence.

Q: How can ExD cell be used in combination with MS-based technologies to better characterize proteoforms?

A: This is the beauty of the cell. We can take an existing mass spectrometer that someone's already been using for years, and within an hour, add in our technology. The instrument still works as it did before, but adds the option to turn on electron-based fragmentation that makes the instrument able to work with more complex samples.

Another method that's become popular combined with mass spectrometry called ion mobility separations (IMS). This method separates proteins or peptides based on size. This is similar to doing size exclusion chromatography, which is a classic biochemistry method, but get higher resolution separations on the time scale of 30 milliseconds. Our technology adds the capability to fragment proteins after they are separated at this rapid time scale.

You still have to average together hundreds to thousands of ion mobility separations extending over a couple of minutes to get sufficient signal-to-noise with proteins, but it's a phenomenal method that can provide considerable information about a protein's structure. The combination of our ExD cells with IMS is exciting as we can work with much larger materials that have previously been impossible to work with.

Q: What are the applications of this? Can you describe research examples?

A: I have three applications that I'd like to talk about here. The first is related to Coronavirus that we have recently started. The nucleocapsid protein in SARS-CoV-2 has a serine-rich (SR) domain that controls RNA replication by keeping the viral RNA highly localized in a cell. This allows the virus RNA to replicate rapidly.

Once the virus has made many copies or RNA, a cellular kinase known as GSK is thought to bind to the nucleocapsid protein, decorating a key motif in the middle of with phosphate groups virus RNA. This may be a signal that causes viral RNA to start being packaged and exported. Hence, post translational modifications are important for controlling how fast the virus can replicate.

Being able to decipher the phosphorylation code controlling this small motif in the neucleocapsid can help identify existing drugs to interfere in viral replication and packaging.

However, characterizing where the phosphates are added in that region is tough because the classic localization methods don't work well. We're able to sequence the entire length of the motif and quantify where up to eight phosphates are located. We have known the value of ExD in general to localize phosphorylation in other proteins, however, it works brilliantly for these serine-arginine rich proteins from viruses as well as cellular proteins known to regulate RNA splicing, the process that generates many proteoforms.

My second application comes with protein characterization. One of the biggest challenges we face is from pharmaceutical companies, and again for context, they make therapeutic antibodies that can bind and activate the SARS-CoV-2 virus. However, when antibodies are produced in bioreactors, it is hard to know if the antibodies correctly folded and modified.

By combining ion mobility with the ECD cell, it may be possible to answer whether antibodies are folded correctly. This is made possible by an innovation known called collision-induced unfolding being developed by Brandon Routolo at the University of Michigan. This technique allows proteins to be progressively unfolded and determine what subtle modifications might have occurred. I believe this will be a huge advantage for the biopharmaceutical industry as well as basic research efforts to understand bigger proteins in complexes.

The final application I'd like to mention is how and why I use mass spectrometry in my academic research.

I study motor neuron disease, also known as ALS and Lou Gehrig's disease. This is a fatal, progressive paralysis due to motor neurons dying in the spinal cord. It kills patients within 2-5 years after diagnosis. In 1993, mutations to a small metalloprotein known as SOD were shown to cause ALS. The scientific community still does not understand how SOD causes motor neurons to die.

My research group has shown that mutations to SOD make it more likely for zinc to fall out, leaving while copper remains bound; we also have shown that the copper in SOD causes motor neurons to die. We have uncovered the

biochemical mechanisms that show how copper in SOD kills motor neurons in cells, but we could never find the murder weapon in the few thousand dying motor neurons located deep in the spinal cord.

My personal goal in developing the ExD technology has been to find the putative murder weapon – that is to prove that copper is bound to SOD without zinc directly from motor neurons. Using ExD, we have established that the SOD protein can be broken into pieces that retain copper and zinc on these fragments. Furthermore, we can do this from the few motor neurons in the spinal cord that are destined to die in ALS. I have not been able to solve this problem with any other technique in thirty years of experimentation and now that the pandemic is ending can return to this research.

Q: Can you discuss current challenges and future directions for the field?

A: The challenge that the field faces is we can take amazingly large groups of protein complexes in a mass spectrometer and get a good estimate of their size, shape, and weight with a precision that I never conceived could be possible.

Sir John Fenn, who won the Nobel Prize for figure out how to ionize proteins, would go to lectures and would start by talking about Dumbo, the flying elephant. He would jokingly say "Who thought a protein would fly through a mass spectrometer?" And in fact, it's not just elephants that fly anymore, it can be a pods of whales.

We now have the tools needed to understand all the ways that proteins become decorated in vivo and how these proteins interact with other proteins. Protein fragmentation by ECD allows far bigger proteins and other macromolecules to be analysed faster and more completely than ever possible.

Joe Beckman
President/Chief Executive Officer e-MSion

Application Note

Proteomics

Agilent

Trusted Answers

A Novel, Automated, and Highly Selective Phosphopeptide Enrichment for Phosphopeptide Identification and Phosphosite Localization

Authors

Valery G. Voinov and Joseph S. Beckman
e-MSion Inc.
Corvallis, OR, USA

Shuai Wu, Kenneth Newton, Linfeng Wu, and Jordy J. Hsiao
Agilent Technologies Inc.,
Santa Clara, CA, USA

Introduction

Phosphopeptide enrichment is one of the most challenging sample preparation steps for LC/MS analysis due to the variation in reproducibility and selectivity caused by manual sample preparation. The LC/MS analysis of the phosphopeptide samples can be challenging due to a variety of factors including:

- The metal ions on the surface of the HPLC system or within the solvents can form phosphopeptide-metal ion complexes, which will cause severe suppression of the phosphopeptide peaks. This effect can substantially impair phosphopeptide identification especially for multiphosphorylated peptides.
- The ionization efficiency of phosphopeptides is another frequently discussed issue. In complex mixtures with nonphosphorylated peptides, phosphopeptides show much less ionization efficiency.
- The phosphoryl group is a labile post-translational modification (PTM), and is generally eliminated as phosphoric acid upon fragmentation using CID. This neutral loss of phosphoric acid is usually the prominent fragmentation, and leads to missing site-specific information in the peptide backbone fragmentation, making it difficult to localize the phosphosite in phosphopeptide.

In the Phosphopeptide Challenge of the Human Proteome Project (HPP), each participating laboratory received two sample vials. The vial labeled "Phosphopeptide" contains a set of synthesized phosphopeptides of human sequence at various concentrations, mixed with their nonphosphorylated counterparts. For some peptides, there is more than one phosphorylated form. The second vial labeled "Phosphopeptide-Yeast" contains the same peptides in 6 µg of tryptic-digested yeast lysate. Each vial was provided dry. This study requests HUPO members to perform peptide sequence analysis using different methods. Each laboratory needs to identify the peptides in the sample, and determine the number and location of the phosphosite on each peptide. Each lab needs to determine the relative abundance of phosphorylation at each modified site by comparison with its nonphosphorylated counterpart. The third request includes phosphopeptide enrichment from the sample containing the yeast matrix and reanalysis by MS (Figure 1). In compliance with this study, HUPO provided a list of 89 human peptide sequences with a worksheet to record results for peptide identification, phosphosite localization, relative quantification, and enrichment.

The Agilent AssayMAP Bravo platform provided a fully automated and highly selective phosphopeptide enrichment workflow using high-capacity Fe(III)-NTA cartridges. Using microchromatography cartridges, this automation platform allows users to easily enrich phosphopeptides in a routine workflow even with a small sample amount. Sample analysis was conducted on an Agilent 1290 Infinity II LC coupled to an Agilent 6550 iFunnel Q-TOF LC/MS with an electromagnetostatic (EMS) ECD cell. To improve the peak shape and recovery of the phosphopeptides, Agilent InfinityLab deactivator additive was added to the HPLC mobile phase. Peptide identification was performed using data-dependent acquisition (DDA), and data were analyzed by Agilent Spectrum Mill B.06 software. After peptide identification using DDA, phosphosite localization was analyzed using targeted MS/MS with ECD, and data were processed with Byonic software (Protein Metrics, Cupertino, CA, USA) based on the HUPO sequence list. Phosphopeptide relative quantitation was performed using single MS1 scan, the spectra library generated by DDA, together with Skyline software (MacCoss Group, University of Washington, Seattle, WA, USA). The enrichment selectivity was evaluated based on both the overall distinct phosphopeptide number and the phosphopeptide number on the HUPO sequence list.

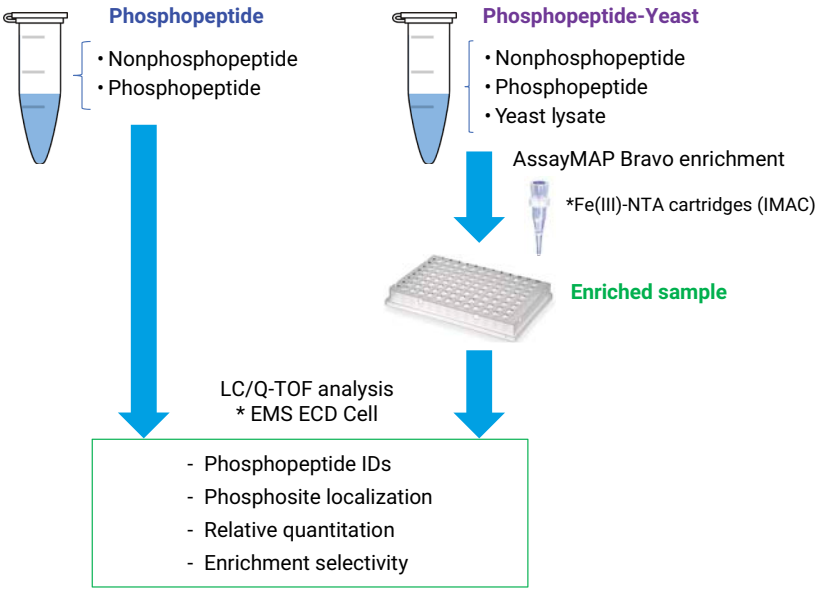


Figure 1. HUPO phosphopeptide challenge experimental design.

Experimental

Material

Two sample vials were received from HUPO:

- **"Phosphopeptide"** contained a set of synthesized phosphorylated (Ser, Thr, or Tyr) peptides of human sequence origin at various concentrations, mixed with their nonphosphorylated counterparts. For some peptides, there is more than one phosphorylated form. HUPO provided 89 peptide sequences.
- **"Phosphopeptide-Yeast"** contained the same peptides in a background matrix consisting of 6 µg of trypsin-digested yeast lysate.

AssayMAP Fe(III)-NTA cartridges were from Agilent Technologies Inc. (Santa Clara, CA, USA). All other chemicals were from Sigma-Aldrich (St. Louis, MO, USA).

Phosphopeptide enrichment and cleanup using AssayMAP Bravo

"Phosphopeptide-Yeast" sample vials were resuspended in 100 µL of 80% ACN, 0.1% TFA. The sample was sonicated for two minutes and transferred to a 96-well PCR plate on position A1.

The Agilent AssayMAP Phosphopeptide Enrichment v2.0 App was used for automated phosphopeptide enrichment using an Fe(III)-NTA cartridge. Figure 2 shows the user interface and application settings. One Fe(III)-NTA cartridge was pretransferred to deck location 2 at A1. Labware was chosen according to the labware table, except that LoBind 96 Eppendorf plates were used on deck locations 4 and 9. The cartridge was primed with 100 µL, 50% ACN, 0.1% TFA using a high flow rate of 300 µL/min. Following that, the cartridge was equilibrated using 80% ACN containing 0.1% TFA, the same solvent as the binding buffer (Table 1). Sample loading onto the cartridge was a critical step. In accordance with our previous study, the flow rate was set at 3.0 µL/min to give enough time for efficient binding of the phosphopeptides. An internal cartridge with 80% ACN containing 0.1% TFA wash step was used after loading. The enriched phosphopeptides were eluted with 20 µL of 1% ammonium hydroxide (approximately pH 11) directly into a LoBind PCR plate containing 80 µL of 2.5% formic acid (Table 1 and Figure 2).[1]

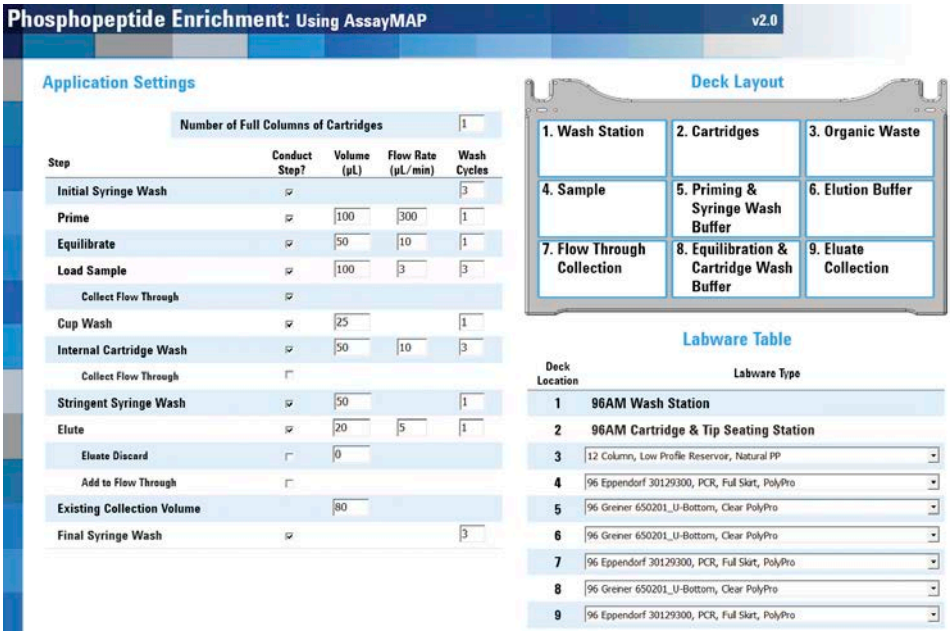


Figure 2. AssayMAP Bravo Phosphopeptide Enrichment v2.0 App.

Table 1. Phosphopeptide enrichment and cleanup protocol using AssayMAP Bravo.

	Phosphopeptide Enrichment	Peptide Cleanup
Affinity Medium	IMAC	Reversed phase
Resin Volume	5 µL	5 µL
Support	Fe(III)-NTA, 100 nmol Fe(III)	C18
Prime Buffer	50% ACN, 0.1% TFA in water	50% ACN, 0.1% TFA in water
Equilibration Buffer	80% ACN, 0.1% TFA in water	0.1% TFA in water
Loading Buffer	80% ACN, 0.1% TFA in water	Ammonium formate in water
Loading Volume	100 µL	100 µL
Loading Flow Rate	3 µL/min	3 µL/min
Washing Buffer 1	80% ACN, 0.1% TFA in water	0.1% TFA in water
Washing Volume 1	50 µL	50 µL
Number of Washes 1	1	1
Elution Buffer 1	1% Ammonium hydroxide	70% ACN, 0.1% TFA in water
Elution Volume 1	20 µL	20 µL
Existing Collection Volume	80 µL	0 µL

The Agilent AssayMAP Peptide Cleanup v2.0 App was used to desalt the enriched phosphopeptides with C18 cartridges (Figure 3). One C18 cartridge was pretransferred to deck location 2 at A1. Labware was chosen according to the labware table, except that LoBind 96 Eppendorf plates were used on deck location 9. The enriched phosphopeptide plate was set at deck location 4. The C18 cartridge was primed with 100 µL of 50% ACN, 0.1% TFA in water, equilibrated with 50 µL of 0.1% TFA, loaded with enriched phosphopeptide sample at a flow rate of 3 µL/min, washed with 50 µL of 0.1% TFA, and eluted with 20 µL of 70% ACN, 0.1% TFA at a flow rate of 5 µL/min.

The sample was dried at room temperature in a SpeedVac vacuum concentrator (Thermo Fisher Scientific, Inc., Waltham, MA, USA), resuspended in 25 µL of 10% ACN, 0.1% FA, and sonicated for two minutes. The sample was further diluted with 25 µL of 0.1% FA in water, and the final sample was in 50 µL of 5% ACN, 0.1% FA.

Peptide identification using data-dependent acquisition

"Phosphopeptide" samples were resuspended in 25 µL of 10% ACN, 0.1% FA. The sample vials were vortexed and sonicated for two minutes. Samples were further diluted in 25 µL of 0.1% FA, and the final samples were in 50 µL of 5% ACN, 0.1% FA.

The 1290 Infinity II LC system was converted to nanoflow LC by coupling with the Agilent Infinity UHPLC Nanodapter. This nanoflow LC was connected to the Agilent nanospray ESI source and coupled with the Agilent 6550 iFunnel Q-TOF LC/MS for peptide identification (Figure 4).

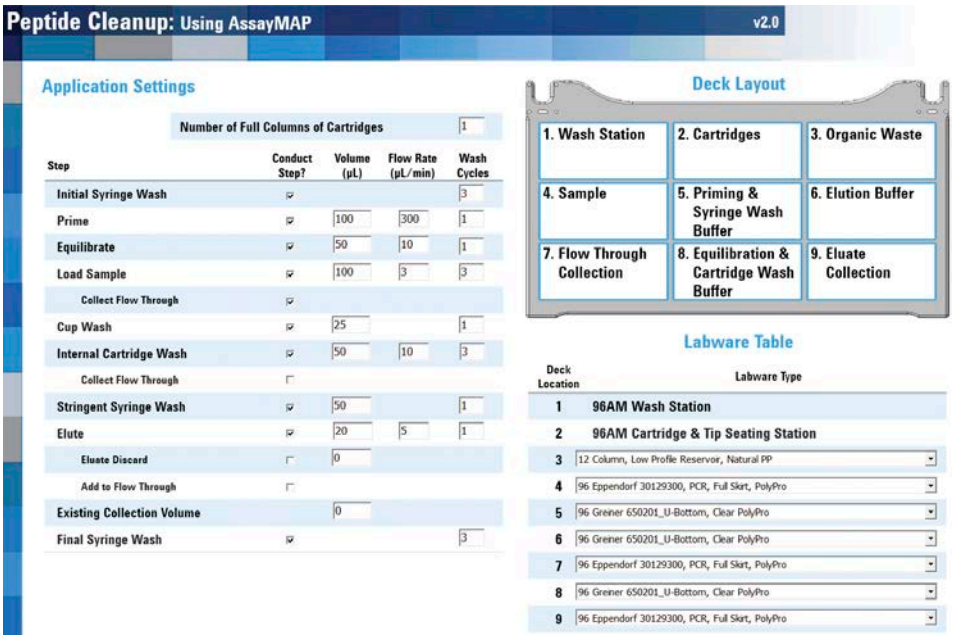


Figure 3. Agilent AssayMAP Bravo Peptide Cleanup v2.0 App.

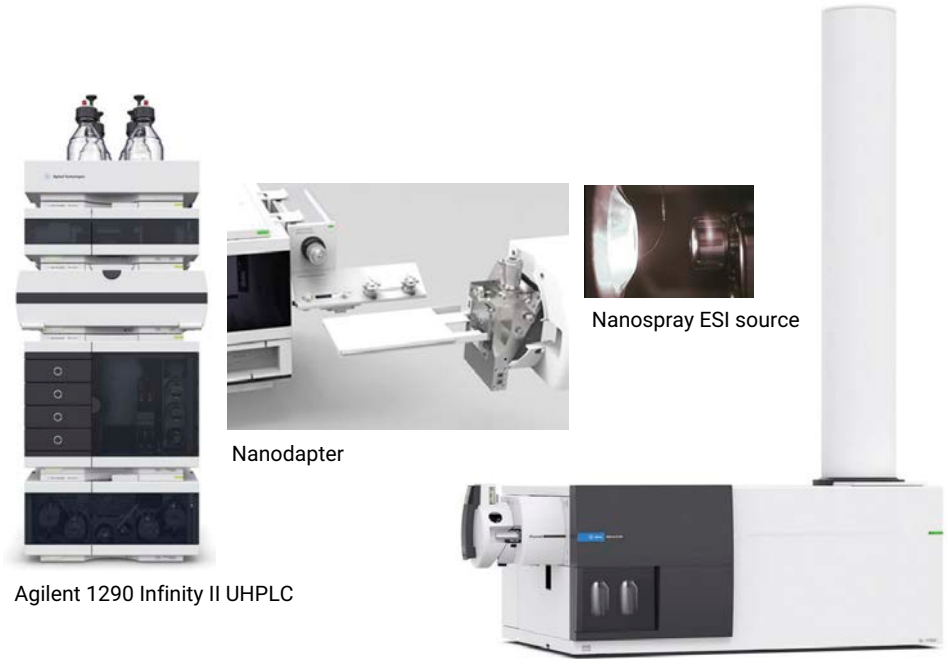


Figure 4. Nanodapter converts standard-flow LC to nanoflow LC coupled with an Agilent 6550 iFunnel Q-TOF LC.

Table 2 lists the LC parameters. The Nanodapter was configured in direct injection mode. A 75 µm × 25 cm C18 column was kept at 60 °C and used for peptide separation with a 90 minute gradient in a total 120 minute LC run time.[2] To enhance the chromatographic performance of the phosphopeptides with minimal ion suppression, 0.1% InfinityLab Deactivator additive was added to solvent A.[3] For peptide identification, 2 µL of "Phosphopeptide" and enriched "Phosphopeptide-Yeast" sample were injected with triplicate injections for each sample. Data-dependent acquisition was used with a selection of top 15 precursor ions. Table 3 lists the detailed setup for the 6550 iFunnel Q-TOF LC/MS for DDA using CID.

Table 2. Nano-LC parameters.

LC Conditions	
Nanodapter Configuration	Direct injection mode
Guard Column	PepMap C18, 75 µm × 2 cm
Analytical Column	PepMap C18, 75 µm × 25 cm
Column Temperature	60 °C
Solvent A	0.1% formic acid, 0.1% deactivator in water
Solvent B	0.1% formic acid in 90 % acetonitrile
Flow Rate	0.085 mL/min primary flow 300 nL/min on-column flow rate
Q-TOF Gradient	Time (min) B (%) 0 3 90 37 95 70 97 70 100 3
Stop Time	115
Post Time	5
Injection Volume	2 µL for CID, 4 µL for ECD, 1 µL for MS1

Table 3. Agilent 6550 iFunnel Q-TOF parameters.

Parameter	Value	
Spray Needle	New objective noncoated needle, 25 µm id, 10 µm tip id, 5 cm length, orthogonally positioned	
Gas Temperature	200 °C	
Drying Gas	11 L/min	
Acquisition Mode	Extended dynamic range (2 GHz) *m/z* 100 to 1700 High analytical sensitivity	
	MS	**MS/MS**
Mass Range	*m/z* 300 to 1700	*m/z* 50 to 1700
Acquisition Rate	3 spectra/s	>3 spectra/s
Isolation Width	Medium (~4 *m/z*)	
Collision Energy	(Slope)*(*m/z*)/100+Offset Charge Slope Offset 2 3.1 1 3 3.6 -4.8 >3 3.6 -4.8	
Maximum Precursor/Cycle	15	
Precursor Threshold	1,000 counts and 0.01 %	
Active Exclusion	Excluded after 1 spectra Released after 0.2 minutes	
Isotope Mode	Peptides	
Sort Precursors	By abundance only; +2, +3, >+3	
Scan Speed Varied Based on Precursor Abundance	Yes	
Target	25,000 counts/spectrum	
Use MS/MS Accumulation Time Limit	Yes	
Purity Stringency	100 %	
Purity Cutoff	30 %	

Spectrum Mill was used to search against the Swiss-Prot human protein database with a 1.5% false discovery rate. Trypsin was the digestion enzyme, and up to two missed cleavages were allowed. Carbamidomethylation was set as the fixed modification while N-term Acetyl, deamidation (N), and phosphorylation of serine (S), threonine (T), and tyrosine (Y) were set as the variable modifications. Table 4 lists the detailed search parameters.

Phosphorylation site localization using ECD

The ECD cell was designed to merge with a shorter hexapole collision cell to replace the original hexapole collision cell in Agilent Q-TOF systems. The length of the ECD cell combined with shorter collision cell is the same as the original collision cell (Figure 5). A peptide standard Substance, P, was used to tune the ECD signal to obtain an approximately 1% conversion efficiency for the +2-charge state. Although acquisitions were done with zero added collision energy, the settings used for the ECD experiments were equivalent to approximately 5 eV of applied collision energy.

A targeted list of precursor ions that contains *m/z*, charge state, retention time, delta retention time, isolation width, and acquisition time was generated using the phosphopeptide list identified from the DDA experiment. Collision energy was set to 0 eV, and the same precursor ions were fragmented four to five times in one run. Data were analyzed with Byonic. Phosphopeptides that were characterized in the first run were removed from the targeted MS/MS list, and further experiments were conducted for unidentified phosphopeptides with more injections.

Table 4. Spectrum Mill search parameters.

Spectrum Mill Search Parameters		
Database	Swiss-Prot human proteins	
Enzyme	Trypsin	
Maximum Missed Cleavage	2	
Modifications	**Fixed** Carbamidomethylation	**Variable** Acetyl (N-term) Deamidated (N) Phosphorylated (S) Phosphorylated (T) Phosphorylated (Y)
Mass Tolerance	**MS1** 15 ppm	**MS2** 30 ppm
Maximum Ambiguous Precursor Charge	5	
Calculate Reversed Database Scores	Yes	
Dynamic Peak Thresholding	Yes	
Peptide FDR	1.50%	
Precursor Charge Range	2 to 7	

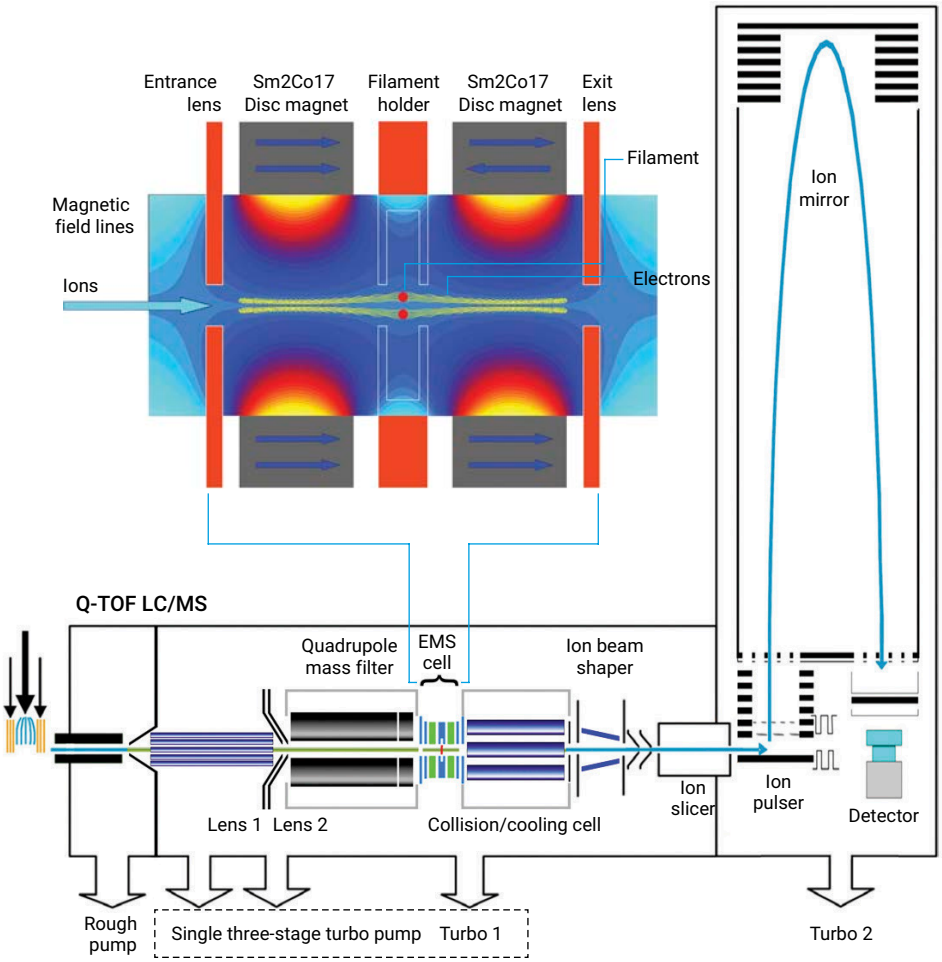


Figure 5. Schematic of an electromagnetostatic (EMS) ECD cell and its position in an Agilent Q-TOF LC/MS.

Byonic software was used to analyze ECD spectra. The same Swiss-Prot human protein database was used with trypsin as the enzyme and up to two missed cleavages. Carbamidomethylation was set as the fixed modification, while deamidation (N, Q) and phosphorylation of serine (S), threonine (T), and tyrosine (Y) were set as the variable modifications.

Relative quantitation of phosphorylation at each modified site

Triplicate MS1 experiments were also conducted for "Phosphopeptide" samples with 1 µL injections using the same LC gradient. The DDA data were imported into Skyline first, and a peptide CID spectra library was created. The full MS1 data were then imported into Skyline, and the identified peptide peaks were integrated based on MS1 scan. The peak areas (or ion intensity from both +2 and +3 precursor ions) were exported into Excel and added for both nonphosphopeptides and phosphopeptides including their deamidated forms. The ratio of phosphopeptide/nonphosphopeptide at each phosphorylation site were calculated based on the ratio of ion intensity.

Results and discussion

Spectrum Mill analysis of DDA data

Figure 6A shows the total ion chromatogram (TIC) of "Phosphopeptide" sample with a 90 minute gradient. With triplicate sample injections using DDA, Spectrum Mill identified 437 distinct peptides with 294 distinct phosphopeptides (Table 5). To evaluate the analytical results, HUPO provided a list of 89 peptide sequences. The 89 nonphosphopeptides were all identified from the "Phosphopeptide" sample. Their phosphorylated counterparts were also identified without confirmation of most phosphosite location using CID. The enriched "Phosphopeptide-Yeast" sample was also analyzed using the same method. Figure 6B shows the TIC of the enriched "Phosphopeptide-Yeast" sample using a 90-minute gradient. Table 5 lists the detailed search result for both the "Phosphopeptide" and the enriched "Phosphopeptide-Yeast" samples including each replicate and the combined search result. After enrichment, a total of 287 distinct peptides were identified with 264 distinct phosphopeptides. The selectivity of the enrichment (phosphopeptide/peptide) from the overall peptide ID numbers is approximately 92%.

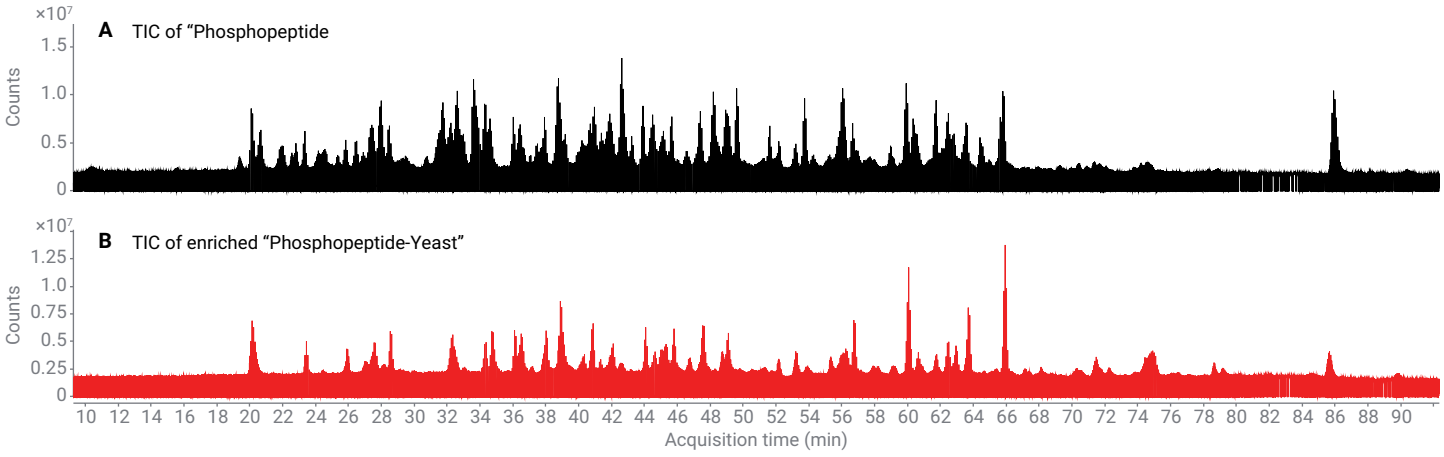


Figure 6. A) TIC of "Phosphopeptide" sample using 90 minute gradient. B) TIC of enriched "Phosphopeptide-Yeast" sample using 90 minute gradient.

Table 5. Spectrum Mill search results.

	"Phosphopeptide"				Enriched "Phosphopeptide-Yeast"			
	Replica 1	Replica 2	Replica 3	Combine	Replica 1	Replica 2	Replica 3	Combine
Total Distinct Peptides	316	297	308	437	193	203	203	287
Total Distinct Phosphopeptides	195	182	193	294	179	192	189	264
Mono-	123	125	127		117	127	123	
Di-	63	50	58		52	56	53	
Tri-	8	6	7		8	8	12	
Tetra-	1	1	1		2	1	1	
% Phosphopeptides/Total Peptides	61.7	61.3	62.7	67.3	92.7	94.6	93.1	92.0
Total Phosphosites	277	247	268		253	267	269	
Sites Assigned %	63	61.9	63.7	62.9	62.4	63.4	62.4	62.7
Sites Not Assigned	102	94	97		95	98	101	

ECD spectra analysis using Byonic

The result for peptide VVEAVNSDSDSEFGIPK analysis is used as an example to demonstrate how the phosphosites were localized. Figure 7A shows the CID spectrum of nonphosphorylated peptide VVEAVNSDSDSEFGIPK. Spectrum Mill identified the precursor ion *m/z* 896.93 (z = +2, mass error = 1.1 ppm) with no phosphosite on this peptide. The sequence matching result showed a complete sequence coverage for this peptide with b, y type ions.

Figures 7B, 7C, and 7D show the ECD spectra of peptides with the same sequence but with mono-, di-, and triphosphorylation. In Figure 7B, Byonic identified the precursor ion *m/z* 936.92 (z = +2, mass error = −0.07 ppm) with one phosphosite for this peptide. The diagnostic ions c6 (*m/z* = 629.36) and c7 (*m/z* = 796.37) confirmed the phosphosite at serine 7 on this peptide. In Figure 7C, Byonic identified the precursor ion *m/z* 651.60 (z = +3, mass error = −3.69 ppm) with two phosphosites for this peptide. The diagnostic ions c7 (*m/z* = 796.35) and z11 (*m/z* =1325.45), c9 (*m/z* =1078.37) and z9 (*m/z* =1043.40) confirmed the two phosphosites at serine 7 and 9 for this peptide. In Figure 7D, Byonic identified the precursor ion *m/z* = 678.25 (z = +3, mass error = −11.54 ppm) with three phosphosites for this peptide. The diagnostic ions c7 (*m/z* = 796.37), c9 (*m/z* = 918.42), and z9 (*m/z* = 1123.40) confirmed the three phosphosites at serine 7, 9, and 11. The CID spectra (not shown) for peptide VVEAVNSDSDSEFGIPK with different degrees of phosphorylation can only identify the number of phosphosites without confirming the location of the phosphosites.

Comparing Figure 7B with 7C, when the size of the phosphopeptide increases, the precursor ion with +3 charge state became more abundant, resulting in a better ECD spectrum with a higher sequence coverage (Figure 7C). Precursor ion with a +2 charge state and lower abundance generated fewer fragment ions, but we were still able to confirm the location of the phosphosite (Figure 7B). Comparing Figure 7C with 7D, peptides with multiple phosphosites (≥3) generally have less ionization efficiency in the positive ion mode. The more abundant precursor ion with the same charge state (+3) also results in a better ECD spectrum (Figure 7C). Comparing Figure 7B with 7D, the two precursor ions have nearly the same abundance. The one with a +3 charge state still generates more fragment ions than the one with a +2 charge state.

In this study, HUPO provided a list of 89 peptide sequences. The total number of distinct phosphopeptides based on this sequence list and confirmed by ECD is 94 in the "Phosphopeptide" sample. Moreover, 93 out of 94 phosphopeptides spiked into the yeast lysate were still identified from the enriched "Phosphopeptide-Yeast" sample. This result showed that a high number of phosphopeptides were recovered from the enrichment. Conversely, 9 out of 89 nonphosphopeptides were still detected after enrichment of the "Phosphopeptide-Yeast" sample. The selectivity of the enrichment (phosphopeptide/peptide) based on the peptide sequence list is approximately 91.3%.

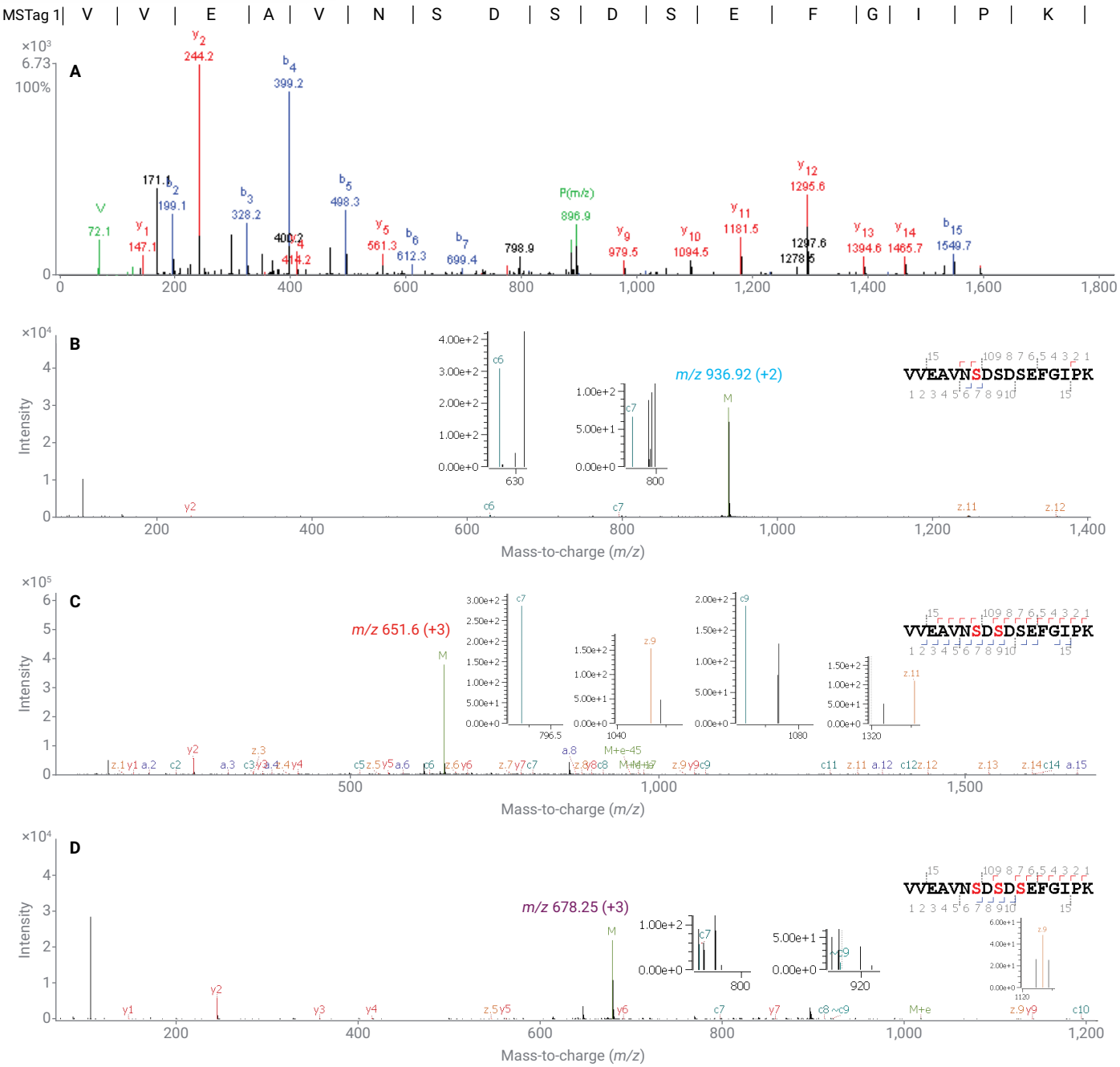


Figure 7. A) CID spectrum of peptide VVEAVNSDSDSEFGIPK showing a complete sequence coverage with b, y type ions. B), C), D) ECD spectra of peptides VVEAVNSDSDSEFGIPK with mono-, di- and tri phosphorylation localized by c, z type ions. Note the presence of most of the sequence ions in the ECD spectra even with the expected low efficiency for low charge state.

Relative quantitation of peptide phosphorylation

Figure 8 shows the extracted ion chromatograms (EICs) of peptide VVEAVNSDSDSEFGIPK with different degrees of phosphorylation. All the peptides were well separated using the C18 column with the 90-minute gradient. For di- and triphosphopeptides, the ion intensity of the deamidated phosphopeptides were added together with the original phosphopeptide. The relative abundance of phosphopeptide/nonphosphopeptide is calculated based on ion intensity. For peptide VVEAVNSDSDSEFGIPK, if the abundance of nonphosphopeptide is 1.00, the ratio of mono-, di- and triphosphopeptides are 0.03, 4.71, and 0.02 respectively, which are annotated in Figure 8. The relative abundance of phosphorylation at each modified site was calculated for all the 94 phosphopeptides in the "Phosphopeptide" sample based on the ratio of phosphopeptide/nonphosphopeptide. Note that this relative quantitation does not consider the different ionization efficiency from different peptide species. It is only a comparison from the ion intensity of different peptides.

Conclusion

An automated phosphopeptide enrichment with both qualitative and quantitative analysis using the Agilent AssayMAP Bravo platform and an LC/Q-TOF system was implemented for the HUPO Phosphopeptide Challenge.

A CID experiment was performed for peptide identification in which 437 distinct peptides with 294 phosphopeptides were identified in the "Phosphopeptide" sample. All 89 nonphosphopeptides from the HUPO sequence list were identified. The ECD experiment determined the location of 124 phosphosites from 94 phosphopeptides based on the 89 nonphosphopeptide sequences. The rest of the peptides not on the sequence list were also reported back to HUPO.

Within the enriched "Phosphopeptide-Yeast" sample, 287 distinct peptides were identified, of which 264 were distinct phosphopeptides. The overall selectivity of the enrichment was approximately 92.0 %.

Moreover, 93 out of 94 phosphopeptides spiked into the yeast were still identified from the enriched "Phosphopeptide-Yeast" sample. Agilent showed the highest number of phosphopeptides recovered from the enrichment compared to other labs in this study.

References

1. Russell, J. D.; Murphy, S. Agilent AssayMAP Bravo Technology Enables Reproducible Automated Phosphopeptide Enrichment from Complex Mixtures Using High-Capacity Fe (III)-NTA Cartridges, *Agilent Technologies Application Note*, publication number 5991-6073EN, **2016**.
2. Wu, S.; Wu, L. Human Breast Cancer Cell Line Phosphoproteome Revealed by an Automated and Highly Selective Enrichment Workflow, *Agilent Technologies Application Note*, publication number 5994-0315EN, **2018**.
3. Hsiao, J. J. *et al.*, Improved LC/MS Methods for the Analysis of Metal-Sensitive Analytes Using Medronic Acid as a Mobile Phase Additive. *Anal. Chem.* **2018**, *90(15)*, 9457–9464.

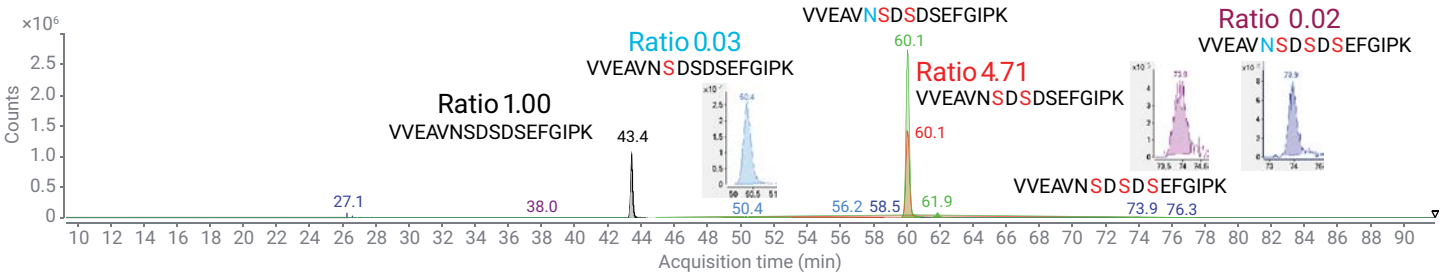


Figure 8. EIC of all VVEAVNSDSDSEFGIPK peptides with the ratio of phosphopeptide/nonphosphopeptide annotated.

Printed in the USA, November 2, 2020
5994-1235EN