Published in final edited form as:

Rapid Commun Mass Spectrom. 2016 May 15; 30(9): 1101–1107. doi:10.1002/rcm.7544.

Use of captive spray ionization to increase throughput of the data-independent acquisition technique PAcIFIC

John D Chapman¹, J. Scott Edgar¹, David R. Goodlett², and Young Ah Goo²

¹Department of Medicinal Chemistry, School of Pharmacy, University of Washington, Seattle, WA

²Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD

Abstract

Rationale—The Precursor Acquisition Independent From Ion Count (PAcIFIC) method is a data-independent acquisition technique capable of identifying proteins over eight orders of magnitude in a single analysis in human plasma. Widespread application of this technique in proteomic studies is hindered by its time intensive nature. There exists a need to explore strategies to increase the throughput of the PAcIFIC method.

Methods—The PAcIFIC acquisition technique was optimized for use with an Orbitrap mass spectrometer fitted with a captive spray ionization (CSI) source. Chromatographic methods, PAcIFIC acquisition parameters, for example channels interrogated per chromatographic gradient, time span of chromatographic gradient, and sample loading amount were investigated to achieve a maximum number of peptide and protein identifications on a yeast proteome where protein copy number had been previously determined.

Results—A 24-hour CSI PAcIFIC method was developed with minimal reduction of peptide and protein identifications from the 4.2-day nano-ESI (nESI) PAcIFIC method. Analysis of a yeast cell lysate with the 4.2-day nESI PAcIFIC method resulted in 13,468 peptide and 2,120 protein identifications. A 24-hour CSI PAcIFIC method resulted in 11,277 peptide and 1,753 protein identifications. Increased sample loading of the CSI system allowed for an 8% increase in peptide and protein identifications.

Conclusion—A dramatic decrease in the overall analysis time of the PAcIFIC method (24 hours with CSI versus 100.8 hours with nESI) was achieved with minimal reduction of peptide and protein identifications. Furthermore, the CSI PAcIFIC method demonstrated the high degree of sensitivity and capability of identifying proteins across a large dynamic range observed with the nESI PAcIFIC method (CSI PAcIFIC identified proteins as low as 46 molecules per cell).

Keywords

Data-independent acquisition; PAcIFIC; Captive spray ionization (CSI); Proteomics

*Corresponding authors: Young Ah Goo, Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, 20 N. Pine Street, Room N719, Baltimore, MD 21201, ygoo@rx.umaryland.edu, Phone 410.706.3380, Fax 410.706.0886; David R. Goodlett, Department of Pharmaceutical Sciences, School of Pharmacy University of Maryland, Baltimore, MD 21201 USA, Phone: 001.410.706.1490, Fax: 001.410.706.0886, dgoodlett@rx.umaryland.edu.

Introduction

Distinctive from the relatively static genome, proteomes are living, breathing, and dying networks that control the outcomes of cellular response.[1] The ability to identify dynamic forces within a proteome is vital for our understanding of cellular functions at the molecular level. However, the massive dynamic range of protein expression levels within a sample makes identification of low expression level proteins extremely challenging. Therefore, the means by which we generate detailed lists of proteins and their relative abundance levels with respect to environmental or patho-physiological conditions are critical in solving challenging questions in biology and medicine.[2, 3]

Previously, our laboratory developed a data-independent acquisition (DIA) technique termed Precursor Acquisition Independent From Ion Count (PacIFIC) for use with nano-electrospray ionization (nESI).[4] As noted by Gillet *et al.* in their publication of the SWATH method,[5] the nESI PacIFIC acquisition method is capable of identifying proteins over the largest dynamic range (eight orders of magnitude at an FDR of 0.5%) as compared to other currently available DDA and DIA schemes.[4] PacIFIC operates by successively interrogating predefined 2.5 m/z isolation widths, or channels, by tandem-MS across a user-defined m/z range (typically 1000 m/z). As opposed to an intensity based precursor ion selection used in data-dependent methods, PacIFIC uses predefined 2.5 m/z isolation width and fragment all ions present regardless of the intensity of ions, which in turn, greatly increases the measureable dynamic range without a need for additional data processing software that is often required when a wider isolation window is used as is the case with methods like SWATH [5] The nESI PacIFIC method has proven an invaluable tool for researchers looking to increase proteome coverage of complex biological matrices such as serum, plasma, cell culture and urine samples.[6-10] For example, Hengel *et al.* illustrated effective application of the nESI PacIFIC method in comparison to a data-dependent 4-window gas phase fractionation (GPF) mass spectrometry technique. The study determined that the pro-metastatic protein anterior gradient 2 (AGR2) was significantly up regulated following tamoxifen treatment and increased expression of AGR2 was implicated in tamoxifen drug resistance. This protein was identified in nESI PacIFIC method but not the data-dependent GPF technique or a transcriptomic screen, highlighting the necessity of being able to more thoroughly catalog biological and clinical samples.[9] Additionally, it has been cited that the nESI PacIFIC method provides a quantitative advantage over other data-dependent acquisition techniques due to the repetitive sampling of each m/z channel during a single chromatographic peak that can provide additional benefit especially for label-free quantification.[6, 11]

However, a current drawback of the nESI PacIFIC method is the time intensive nature required for each sample to be analyzed. Current optimized methods have been established for the LTQ Orbitrap and LTQ Velos (Thermo Scientific, San Jose, CA, USA): 44 sample injections each covering 15 2.5 m/z channels (about 4.2 days), and 27 sample injections each covering 25 2.5 m/z channels (about 2.5 days), respectively.

Here, we explored use of a captive spray ionization (CSI) source (Bruker, Billerica, MA, USA) to reduce the overall analysis time of the PacIFIC method. The CSI source is a high-

voltage non-tapered capillary emitter situated precisely in front of the mass spectrometer inlet to allow for the vacuum of the mass spectrometer to pull ambient air into the source, directing a high percentage of sample ions into the mass spectrometer. The CSI source provides (1) axial gas flow that keeps the capillary emitter clean for prolonged spray stability; (2) gas flow directed at the spray tip to aid in desolvation and to nebulize the solvent into a Taylor cone; and (3) gas flow to create a vortex around the Taylor cone to funnel all of the sample into the mass spectrometer inlet.[12] In turn, it is possible to achieve a higher solvent flow rate with comparable sensitivity to nESI.[13] The higher flow rate allows for the efficient use of larger bore nano-columns, increased gradient control, as well as faster trapping, peptide elution, and column equilibration times. The CSI source has been successfully implemented in small molecule and data-dependent proteomic applications. [12, 14, 15]

In this report, we demonstrated an enhanced PACIFIC method to analyze complex proteomics samples in a fraction of the time it takes for the original PACIFIC analysis with a minimal change in the number of peptide and protein identifications and the median copy number of the detected proteins in the yeast proteome.

Experimental

Chemicals and solvents

Unless otherwise noted, all chemicals and solvents were purchased from Sigma Aldrich (St. Louis, MO, USA).

Saccharomyces cerevisiae growth

A single colony of strain BY4742 was grown in yeast extract peptone dextrose (YEPD) media overnight. The following day, the media was diluted and the yeast were grown to mid-log phase ($OD_{600} = 0.56$) at conditions matching those of Ghaemmaghami *et al.*[16] *S. cerevisiae* were harvested by centrifugation at $5,000 \times g$ for 10 minutes at 4 °C. The supernatant was removed and the cells were washed twice with ice cold PBS, each time quickly repelleting the cells at $5,000 \times g$ for 2 minutes. After the final wash, the PBS was removed and the resulting *S. cerevisiae* cell pellets were stored at -80 °C until protein extraction and mass spectrometry sample preparation.

Protein extraction and quantification

S. cerevisiae pellets were lysed in 1mL buffer (0.74% β -mercaptoethanol and 0.815 M NaOH) for 10 min at 4 °C followed by a probe sonication. In short, cell pellet was sonicated for roughly 5 seconds and then placed on ice to cool the suspension. This process was repeated twice more for each aliquot. Subsequently, lysed *S. cerevisiae* suspensions were centrifuged at 5,000 rpm at 4 °C for 10 minutes to pellet cellular debris. The supernatant, containing cellular proteins, was removed and placed in a clean eppendorf on ice. Protein quality was checked on a protein gel to confirm no degradation of protein.

Each aliquot of *S. cerevisiae* protein suspension was quantified using a BCA protein assay according to the manufacturer's instructions (Thermo Scientific/Pierce, Rockford, IL, USA).

Sample protein concentrations were normalized by making aliquots of 200 µg/100 µL *S. cerevisiae* protein with the dilution of a necessary amount of 100 mM ammonium bicarbonate.

Protein digestion and peptide preparation

Each 200 µg/100 µL *S. cerevisiae* protein sample was denatured with the addition of urea to 6M. Subsequently, samples were buffered with the addition of 7 µL 1.5 M Tris pH 8.8 (this raises the buffer pH since TCEP is so acidic), reduced with 2.5 µL of 200 mM tris(2-carboxyethyl)phosphine (TCEP) for 1 hour at 37 °C, alkylated with 20 µL of 200 mM iodoacetamide for 1 hour at room temperature in the dark, and then quenched with 20 µL of 200 mM dithiothreitol at room temperature. Prior to addition of sequencing-grade porcine trypsin (Promega, Madison, WI, USA) at a protein to enzyme ratio of 50:1, samples were diluted with 900 µL of 50 mM ammonium bicarbonate and 200 µL of MeOH. After an overnight incubation, peptides were desalted on a Vydac C18 macrospin column (The Nest Group, Southborough, MA, USA) according to the manufacturer's protocol. Resulting eluent was concentrated on a SPD 111V SpeedVac (Thermo Savant, San Jose, CA, USA) and either stored at -80 °C for future use or reconstituted with a solution of 95% water/5% acetonitrile/0.1% formic acid peptide for immediate mass spectrometry analysis.

Nano-electrospray ionization (nESI) mass spectrometry

Nano-HPLC was performed using a Waters NanoAquity (Milford, MA, USA). A homemade trapping column was made from 100 µm inner diameter (ID) capillary (Polymicro Technologies, Phoenix, AZ, USA) packed with 2 cm of 200 Å, 5 µm Magic C18AQ particles (Michrom, Auburn, CA, USA). Successive analytical separation was performed on a homemade 75 µm ID × 180 mm long analytical column pulled using a Sutter Instruments P-2000 CO₂ laser puller (Sutter Instrument Company, Novato, CA) and packed with 15 cm of 100 Å, 5 µm Magic C18AQ particles (Michrom, Auburn, CA, USA). For each sample injection, approximately 1 µg, unless otherwise noted, of the peptide sample was loaded on the trapping column at 4 µL/min with 95% water/5% acetonitrile/0.1% formic acid for 5 minutes. Trapped peptides were then eluted from the trapping column onto the analytical column using a variable gradient with a flow rate of 0.25 µL/minutes. The gradient utilized two mobile phase solutions: A, water/0.1% formic acid; and B, acetonitrile/0.1% formic acid. The variable gradient used is as follows: 0 minutes, A (95%), B (5%); 55 minutes, A (65%), B (35%); 65 minutes, A (15%), B (85%); 75 minutes, A (15%), B (85%); 80 minutes, A (95%), B (5%); 80-100 minutes, A (95%), B (5%). Peptide digests were analyzed on both a LTQ Orbitrap and a LTQ Velos (Thermo Fisher, San Jose, CA, USA) by nESI in positive ion mode. Ion source conditions were optimized using the tuning and calibration solution suggested by the instrument provider. For full range data-dependent acquisition, precursor ion scans were performed from 400-2000 m/z selecting top 5 or top 10 most intense ions for the LTQ Orbitrap and the LTQ Velos, respectively, and data-dependent ion selection for MS/MS analysis operated within that m/z range. For gas-phase fractionation (GPF), precursor ion scans and data-dependent ion selection for MS/MS analysis were performed in the following four restricted m/z ranges: 400-520 m/z; 515-690 m/z; 685-970 m/z; or 965-2000 m/z. For tandem MS analysis, the isolation window was set to 2 m/z and a normalized collision energy of 35% was used for all fragmentation events.

Singly charged ions were excluded from analysis and dynamic exclusion was enabled with a repeat count of 1, a repeat duration of 30 seconds, an exclusion duration of 45 seconds and an exclusion list of 250. For PACIFIC acquisition, the mass spectrometer parameters were set as defined by Panchaud *et al.*[4, 11]

Captive spray ionization mass spectrometry

Captive spray ionization was performed using a Waters NanoAquity (Milford, MA, USA). A manufactured fused silica 200 μm ID trapping column purchased from New Objective (Woburn, MA, USA) and packed with 2 cm of 200 $\text{\AA}$, 5 μm Magic C18AQ particles (Michrom, Auburn, CA, USA). Successive analytical separation was performed on a manufactured fused silica 200 μm ID column purchased from New Objective (Woburn, MA, USA) and packed with 15 cm of 100 $\text{\AA}$, 5 μm Magic C18AQ particles (Michrom, Auburn, CA, USA). For each sample injection, approximately 1 μg , unless otherwise noted, of the peptide sample was loaded on the trapping column at 6 $\mu\text{L}/\text{minutes}$ with 95% water/5% acetonitrile/0.1% formic acid for 2.5 minutes. Trapped peptides were then eluted from the trapping column onto the analytical column using a variable gradient with a flow rate of 2.5 $\mu\text{L}/\text{minutes}$, unless otherwise noted. The gradient utilized two mobile phase solutions: A, water/0.1% formic acid; and B, acetonitrile/0.1% formic acid. The variable gradient used for data-dependent acquisition methods is as follows: 0 minutes, A (95%), B (5%); 62 minutes, A (58%), B (42%); 65 minutes, A (15%), B (85%); 65.7 minutes, A (15%), B (85%); 65.75 minutes, A (95%), B (5%); 65.75-68 minutes, A (95%), B (5%). Peptide digests were analyzed on a LTQ Orbitrap (Thermo Fisher, San Jose, CA, USA) by captive spray ionization in positive ion mode. Ion source conditions were optimized using the tuning and calibration solution suggested by the instrument provider and under conditions recommended by the CSI manufacturer (Bruker, Billerica, MA, USA). For full range data-dependent acquisition, precursor ion scans were performed from 400-2000 m/z selecting top 5 most intense ions and data-dependent ion selection for MS/MS analysis operated within that m/z range. For gas-phase fractionation (GPF), precursor ion scans and data-dependent ion selection for MS/MS analysis were performed in the following four restricted m/z ranges: 400-520 m/z ; 515-690 m/z ; 685-970 m/z ; or 965-2000 m/z . For tandem MS analysis, the isolation window was set to 2 Th and a normalized collision energy of 35% was used for all fragmentation events. Singly charged ions were excluded from analysis and dynamic exclusion was enabled with a repeat count of 1, a repeat duration of 30 seconds, an exclusion duration of 45 seconds and an exclusion list of 250.

Captive spray ionization PACIFIC mass spectrometry

For CSI PACIFIC acquisition, the mass spectrometry parameters employed are as defined by Panchaud *et al.*[4, 11] To optimize the number of injections per total analysis, the following conditions were tested on the LTQ Orbitrap: 1) 44 sample injections using the variable gradient as follows: 0 minutes, A (95%), B (5%); 26 minutes, A (58%), B (42%); 29 minutes, A (15%), B (85%); 29.7 minutes, A (15%), B (85%); 29.75 minutes, A (95%), B (5%); 29.75-32 minutes, A (95%), B (5%); 2) 33 sample injections using the variable gradient as follows: 0 minutes, A (95%), B (5%); 34 minutes, A (58%), B (42%); 37 minutes, A (15%), B (85%); 37.7 minutes, A (15%), B (85%); 37.75 minutes, A (95%), B (5%); 37.75-40 minutes, A (95%), B (5%); and 3) 22 sample injections using the variable

gradient as follows: 0 minutes, A (95%), B (5%); 56 minutes, A (58%), B (42%); 59 minutes, A (15%), B (85%); 59.7 minutes, A (15%), B (85%); 59.75 minutes, A (95%), B (5%); 59.75-62 minutes, A (95%), B (5%).

Database search and protein quantification

Data acquired on the LTQ Orbitrap was converted from Thermo's RAW format to the universal mzXML format and searched against a database containing all known *S. cerevisiae* proteins using SEQUEST.[17] For the data-dependent acquisition techniques (400-2000 m/z and 4-window GPF), search parameters for SEQUEST included a precursor ion tolerance of 2.1 Da, fragment binning tolerance for MS/MS fragment ions at 1.0005, trypsin enzyme specificity (fully-tryptic), two allowed missed cleavage, cysteines modified with iodoacetamide and the variable option for methionines in reduced or oxidized form. The same search parameters were applied to all with an exception of the PACIFIC search parameter for SEQUEST was modified to increase the precursor ion tolerance to 3.75 Da. Results were analyzed with Peptide Prophet[18, 19] ensuring that peptide hits with a probability of >0.99 and unique peptide number ≥ 2 were accepted and linked to protein entries. Protein identifications were linked to the protein molecule per cell value as quantified by Ghaemmaghami *et al.*[16] in order to assess a quantification value for each identified protein.

Results and Discussion

Captive spray ionization gradient development

As discussed above, the directed ambient gas flow established by the CSI source allows for sensitivity comparable to nESI at increased solvent flow rates.[20] Here, to capitalize on this feature of the CSI source, we aimed to use the higher solvent flow rate to increase sample throughput while maintaining the same mobile phase gradient slope and duration used in our standard nESI applications.

A standard nESI chromatography column system used in our laboratory is a 100 μm inner diameter (ID) trapping column packed with 2 cm of Magic C18 packing material, linked through a four-way fitting to a 15 cm 75 μm ID analytical column situated in proximity to the mass spectrometer inlet. The trapping flow rate is 4 $\mu\text{L}/\text{minute}$ and maintains a pressure of approximately 1800 psi and the analytical flow rate is 0.25 $\mu\text{L}/\text{minute}$ and reaches a maximum pressure of approximately 900 psi. Following a trapping period of 5-minutes, peptides are eluted from the analytical column and enter the mass spectrometer using a 100-minute sample analysis method that contains a 65-minute gradient optimized peptide elution (Figure 1A). As indicated by the red triangles in Figure 1C, peptides from 1 μg of a trypsin-digested yeast cell lysate elute from the column after 10-minutes and continue to elute until approximately 67 minutes into the total chromatographic gradient. Due to the low flow rate of nESI, it takes approximately 17 minutes for a complete change in solvent to fully equilibrate the HPLC and chromatography column system, from HPLC pump to analytical column tip, for the next sample analysis. Therefore, following the 85% ACN wash step, an equilibration time of 20 minutes is allowed to ensure that before the start of the next sample

injection the solvent in the HPLC and chromatography system has returned to trapping conditions of 5% ACN/0.1% FA (Figure 1A).

The mobile phase gradient slope and duration used in the nESI system above was mirrored for use with the CSI source. To modulate the increase of pressure associated with an increase in solvent flow rate, we used 200 μm ID trapping and analytical columns, both packed with the same length of Magic C18 packing material described in our nESI system. We tested a spectrum of solvent flow rates and observed that at 2.5 $\mu\text{L}/\text{minute}$ a maximum level of sensitivity was achieved, based on measured intensities of a standard peptide, angiotensin II (data not shown). The signal intensity from a 1 pmole sample injection of angiotensin II on the CSI system was 0.1 to 0.5 orders of magnitude lower than that recorded with nESI (data not shown).

The ten-fold increase in solvent flow rate used with the CSI source allows for an equilibration of the entire HPLC and chromatography column system, from pump to analytical column tip, in approximately 2.5 minutes. Using the higher solvent flow rate of 2.5 $\mu\text{L}/\text{minute}$, it is possible to drastically shorten both the lapse time for peptides to elute from the analytical column at the beginning of the chromatographic gradient and the equilibration period at the end of each sample analysis. The result is a 68-minute chromatographic gradient with a similar gradient slope and time as used in the nESI sample analysis – an overall reduction of 32 minutes (Figure 1B). As shown in Figure 1D, using the CSI chromatographic gradient, peptides start to elute at approximately 3 minutes and continue to elute till 60 minutes into the sample analysis. Additionally, the larger ID trapping column allows for an increase in the solvent flow rate during the trapping period to 6 $\mu\text{L}/\text{minute}$; effectively reducing the needed trapping period to 2.5 minutes while maintaining a similar back pressure to that observed in the nESI analysis.

The CSI source, using the CSI chromatography method above, was used to analyze a complex proteomic sample, namely, a trypsin-digested yeast whole cell lysate, using data-dependent acquisition. As illustrated in Figure 1, the total ion current (TIC) is depicted for a sample injection of 1 μg trypsin-digested yeast cell lysate for both nESI (Figure 1C) and CSI (Figure 1D). While all steps of the gradient are essential, for this comparison we equate the amount of time that peptides elute from the analytical column as the MS time utilization for the sample analysis. MS time utilization is the period of time during the sample analysis where the mass spectrometer can attain tandem mass spectra for positive peptide identification. As a baseline, the 100-minute gradient of the nESI analysis has a percent time utilization of 57.1%. Using the established CSI conditions above, the percent time utilization is improved to 82.4%. While the overall analysis time of the CSI system is less than that of the nESI system, the increased percent time utilization allows for a similar number of peptide identifications as compared to nESI during data-dependent acquisition, specifically, 2702 peptides were identified using the nESI configuration and 2620 peptides were identified using the CSI system.

Additionally, comparison of a four-window GPF data-dependent analysis was completed. The total gradient time for the nESI GPF method is 400 minutes and results in 4459 peptide

identifications. The total gradient time for the CSI GPF method is 272 minutes (a 32% reduction in time) and results in 4408 peptide identifications.

As discussed above, the CSI system appears to have a small decrease in signal sensitivity in comparison to the nESI system. The yeast protein molecules per cell (copy number) value, as measured by Ghaemmaghami *et al.*, [16] was used to evaluate differences in the sensitivity of proteins identified from the two ionization sources, nESI and CSI (Figure 2). This was done for both a standard data-dependent (400-2000 m/z) and a four-window GPF acquisition technique. As illustrated in Figure 2, the frequency distribution of peptide identifications for nESI and CSI follow a similar pattern. Although, it should be noted that the median value of the assigned protein molecule per cell, denoted by a black triangle (▼), is lower in both of the CSI analyzes.

Coupling CSI to the PAcIFIC acquisition technique

As designed by Panchaud *et al.*, to maximize peptide identifications and benefit from improved quantitative measurements, the PAcIFIC duty cycle (i.e., interrogating each designated PAcIFIC m/z channel for a sample injection) should occur multiple times during elution of a single chromatographic peak. [4, 11] To analyze the m/z region of 400-1400 m/z, there are a total of 667, 2.5 m/z PAcIFIC channels to interrogate. Panchaud *et al.* arrived at an optimized nESI PAcIFIC method for the LTQ Orbitrap using 44 sample injections each interrogating 15 2.5 m/z channels.

By virtue of altering the source and chromatographic system, it was necessary to review and optimize the PAcIFIC acquisition method for use with CSI. For the CSI PAcIFIC method on the LTQ Orbitrap we fixed the overall sample analysis time to 24 hours (in comparison to 4.2 days for nESI PAcIFIC). With a fixed total analysis time and total number of 2.5 m/z channels to interrogate, there exists a relationship between the number of channels to interrogate for each chromatographic analysis and the time of each chromatographic analysis. We evaluated the following three methods: 1) 44 sample injections each interrogating 15 2.5 m/z channels with a single sample analysis time of 32.7-minutes per sample injection; 2) 33 sample injections each interrogating 20 2.5 m/z channels with a single sample analysis time of 43.6 minutes per sample injection; and 3) 22 sample injections each interrogating 30 2.5 m/z channels with a single sample analysis time of 65.5-minutes per sample injection. Empirically, the time of the PAcIFIC duty cycle on the LTQ Orbitrap when interrogating 15, 20 and 30 channels is approximately 3.9-5-seconds, 5.3-6.7-seconds, and 7.9-10.0-seconds, respectively (data not shown).

As a baseline reference, analysis of a trypsin-digested yeast cell lysate with the 4.2-day nESI PAcIFIC method identified 13,468 peptides originating from 2,120 proteins. As depicted by the frequency distribution of assigned protein molecules per cell, the median copy number value of the identified proteins is 5,040 (Figure 2). The results from the three CSI PAcIFIC methods of the same yeast cell lysate sample are also illustrated in Figure 2. It was observed that the conditions associated with 44 sample injections allow for a maximum number of peptide and protein identifications (11,277 peptides, 1,753 proteins), lowest assigned copy number median value (5941 protein molecules per cell), and the shortest PAcIFIC duty cycle of the settings tested.

Increased sample loading and application to CSI PACIFIC

As discussed above, to modulate the increase in pressure associated with higher solvent flow rates in our CSI system, the bore size of the trapping column was increased from 100 μm to 200 μm . Doubling the ID of the capillary quadruples the volume of Magic C18 material that can be packed in a fixed 2 cm trapping column length. This, in turn, increases the binding capacity of the trapping column. We recognized the potential of loading more sample per injection to effectively increase the amount of each peptide species available for analysis and explored such effects on the number of peptides identified.

We first compared the peptide and protein identifications from a series of analyses completed with the 44 injection CSI PACIFIC method to assess the effect of increasing the per sample injection amount from 1 μg to 2 μg . With no effect on the overall sample analysis time, this technique resulted in 7.2% more peptide identifications, and a 7.6% boost in protein identifications. Notably, the mean copy number value for the 2 μg sample load analysis decreased by 408.9 protein molecules per cell. By identifying proteins unique to the 2 μg analysis, it was found that these identifications originate from lower copy number per cell proteins with a mean and median value of 4062.6 and 2537.2 protein molecules per cell, respectively.

We further increased the sample amount from 2 μg to 3 μg per sample injection. Here, it was observed that the number of peptide and protein identifications decreased by approximately 8%. It is possible that at this sample load level the capacity of the trapping column has been exceeded and/or there exists an increased number of chimeric spectra[21] that are not appropriately handled with our current proteomics analysis pipeline.

Therefore, as seen from the data, it is possible to see great benefit by increasing the presence of peptide species by increased sample loading, but this must be done with caution as to not overload the trapping column.

12-hour CSI PACIFIC analysis

To further investigate decreasing the time per total sample analysis, we analyzed the data from the CSI PACIFIC analysis with 44 sample injections to select the m/z range with the highest rate of peptide identifications. We found that a bulk of peptide identifications originate from a limited m/z range, the 285 2.5 m/z channels located between 670 and 1096 m/z . With the trapping period and pauses in sample analysis (e.g., waiting for contact closure signals from the liquid chromatographer), analysis can be done in less than twelve hours. To validate this approach the original data set was reanalyzed to only include the above mentioned 19 sample injections. With a reduction in more than 50% of the time required to complete the PACIFIC analysis, 65.3% of the peptide identifications and an unexpected 90.5% of the number of protein identifications were still identified. Additionally, we observed no effect on the dynamic range and only a modest increase in the median value of protein molecules per cell (a rise from 5491.2 to 6210 protein molecules per cell).

Conclusion

Since the first report, the PACIFIC method has shown expanded detectable dynamic range, lower limit of detection, and improved overall confidence of peptide identifications and relative protein quantification measurements with no need for prior sample fractionation or additional data processing and special bioinformatics tools for DIA data analysis.[6-11] As demonstrated in this proof of principle study, the CSI source can be used to significantly decrease the overall analysis time of the PACIFIC method. Specifically, use of the CSI on the Orbitrap enabled us to reduce the overall PACIFIC analysis time from 4.2 days to 24 hours (interrogating the m/z range of 400-1400 m/z). Furthermore, this 24-hour method yielded comparable identification results as acquired using a nESI source on the same instrument (Figure 2). Moreover, we demonstrated that selection of high-yielding PACIFIC channels can further reduce the overall analysis time - in less than 12 hours, interrogation of 285 2.5 m/z channels located between 670 and 1096 m/z identified 65.3% of the peptide identifications and a surprising 90.5% of the number of protein identifications as identified using the 400-1400 m/z range. We note that channel selection may be further optimized to account for variables such as sample organism, digestion methods and proteases used, and targeting of peptides/proteins of interest.

Designing mass spectrometry application studies require a thoughtful balance of time/cost and rationale of acquisition technique. The PACIFIC method has demonstrated gains in identifying proteins over other data-dependent acquisition techniques, but the time/cost investment of using the nESI PACIFIC method is significant. As demonstrated here, use of CSI to decrease the overall analysis time of the PACIFIC method can make this sensitive acquisition technique a more practical and throughput research tool. Additionally, we note that this technique is highly dependent on the scanning speed of the mass spectrometer used. Further decreases in analysis time of the CSI PACIFIC method are achievable as scanning speeds of mass spectrometers continue to improve.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant number(s): This work was supported in part by the NIAID U54 AI057141-06, and University of Maryland Baltimore, School of Pharmacy Mass Spectrometry Center (SOP1841-IQB2014).

References

1. Wilkins MR, Appel RD, Van Eyk JE, Chung MC, Gorg A, Hecker M, Huber LA, Langen H, Link AJ, Paik YK, Patterson SD, Pennington SR, Rabilloud T, Simpson RJ, Weiss W, Dunn MJ. Guidelines for the next 10 years of proteomics. *Proteomics*. 2006; 6:4. [PubMed: 16400714]
2. Nilsson T, Mann M, Aebersold R, Yates JR 3rd, Bairoch A, Bergeron JJ. Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat Methods*. 2010; 7:681. [PubMed: 20805795]
3. Cravatt BF, Simon GM, Yates JR 3rd. The biological impact of mass-spectrometry-based proteomics. *Nature*. 2007; 450:991. [PubMed: 18075578]

4. Panchaud A, Scherl A, Shaffer SA, von Haller PD, Kulasekara HD, Miller SI, Goodlett DR. Precursor acquisition independent from ion count: how to dive deeper into the proteomics ocean. *Anal Chem.* 2009; 81:6481. [PubMed: 19572557]
5. Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, Bonner R, Aebersold R. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics.* 2012; 11:O111. [PubMed: 22261725]
6. Pak H, Pasquarello C, Scherl A. Label-free protein quantification on tandem mass spectra acquired in a data-independent mode provides accurate measurements over five orders of concentration magnitude in complex matrices. *J Integrated Omics.* 2011; 1:211.
7. Voss J, Goo YA, Cain K, Woods N, Jarrett M, Smith L, Shulman R, Heitkemper M. Searching for the noninvasive biomarker holy grail: are urine proteomics the answer? *Biol Res Nurs.* 2011; 13:235. [PubMed: 21586496]
8. Acosta-Martin AE, Panchaud A, Chwastyniak M, Dupont A, Juthier F, Gautier C, Jude B, Amouyel P, Goodlett DR, Pinet F. Quantitative mass spectrometry analysis using PACIFIC for the identification of plasma diagnostic biomarkers for abdominal aortic aneurysm. *PLoS One.* 2011; 6:e28698. [PubMed: 22163325]
9. Hengel SM, Murray E, Langdon S, Hayward L, O'Donoghue J, Panchaud A, Hupp T, Goodlett DR. Data-independent proteomic screen identifies novel tamoxifen agonist that mediates drug resistance. *J Proteome Res.* 2011; 10:4567. [PubMed: 21936522]
10. Goo YA, Cain K, Jarrett M, Smith L, Voss J, Tolentino E, Tsuji J, Tsai YS, Panchaud A, Goodlett DR, Shulman RJ, Heitkemper M. Urinary proteome analysis of irritable bowel syndrome (IBS) symptom subgroups. *J Proteome Res.* 2012; 11:5650. [PubMed: 22998556]
11. Panchaud A, Jung S, Shaffer SA, Aitchison JD, Goodlett DR. Faster, quantitative, and accurate precursor acquisition independent from ion count. *Anal Chem.* 2011; 83:2250. [PubMed: 21341720]
12. Ramanathan R, Raghavan N, Comezoglu SN, Humphreys WG. A low flow ionization technique to integrate quantitative and qualitative small molecule bioanalysis. *Int J Mass Spectrom.* 2011; 301:127.
13. Dahal UP, Jones JP, Davis JA, Rock DA. Small molecule quantification by liquid chromatography-mass spectrometry for metabolites of drugs and drug candidates. *Drug Metab Dispos.* 2011; 39:2355. [PubMed: 21937735]
14. Donegan M, Browning M. A review of recent developments in sample ionization interfaces used in mass spectrometry. *J Liq Chromatogr Related Technol.* 2012; 35:2345.
15. Taylor SW, Nikoulina SE, Andon NL, Lowe C. Peptidomic profiling of secrete products for pancreatic islet culture results in a higher yield of full-length peptide hormones than found using cell lysis procedures. *J Prot Res.* 2013; 12:3610.
16. Ghaemmaghami S, Huh W, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. Global analysis of protein expression in yeast. *Nature.* 2003; 425:737. [PubMed: 14562106]
17. Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. *J Am Soc Mass Spec.* 1994; 5:976.
18. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem.* 2003; 75:4646. [PubMed: 14632076]
19. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem.* 2002; 74:5383. [PubMed: 12403597]
20. Katz E, Fon M, Eigenheer RA, Phinney BS, Fass JN, Lin D, Sadka A, Blumwald E. A label-free differential quantitative mass spectrometry method for the characterization and identification of protein changes during citrus fruit development. *Proteome Sci.* 2010; 8:68. [PubMed: 21162737]
21. Houel S, Abernathy R, Renganathan K, Meyer-Arendt K, Ahn NG, Old WM. Quantifying the impact of chimera MS/MS spectra on peptide identification in large-scale proteomics studies. *J Proteome Res.* 2010; 9:4152. [PubMed: 20578722]

Chromatographic gradients used for (A) nESI and (B) CSI methods, with the time indicated before 0-minutes as the period allowed for solvent flow through the trapping column only. Total ion chromatograms from the analysis of 1 μ g of trypsin-digested yeast whole cell lysate using the above chromatographic gradients for both (C) nESI and (D) CSI analysis. The end time of each analysis is denoted by a red x-axis tick mark. Additionally, the beginning and end time of peptide elution is denoted by red triangles.

Figure 2.

Frequency distribution based on protein molecules per cell (copy number), as assigned from values published in Ghaemmaghami *et al.*, [16] of trypsin-digested yeast cell lysate mass spectrometry derived protein identifications from a variety of ionization source and mass spectrometer acquisition methods. Each pixel represents a spread of 200 protein molecules per cell, starting from zero. Listed for each technique is the: total number of peptide and protein identifications, a minimum and maximum value assigned from the protein molecules per cell (an average of three replicate experiments), and the median value for protein molecules per cell assignment, as denoted by the black triangle (▼).