

# **HHS Public Access**

Mol Biosyst. Author manuscript; available in PMC 2016 November 02.

Published in final edited form as:

Author manuscript

Mol Biosyst. 2016 July 19; 12(8): 2385–2388. doi:10.1039/c6mb00385k.

## **Differential quantification of isobaric phosphopeptides using data-independent acquisition mass spectrometry**

**Simone Sidoli**a,\* , **Rina Fujiwara**a,\* , **Katarzyna Kulej**a,b, and **Benjamin A. Garcia**a,†

aEpigenetics Program, Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>b</sup>Division of Cancer Pathobiology, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

### **Abstract**

Phosphorylation is a post-translational modification (PTM) fundamental for processes such as signal transduction and enzyme activity. We propose to apply data-independent acquisition (DIA) using mass spectrometry (MS) to determine unexplored phosphorylation events on isobarically modified peptides. Such peptides are commonly not quantitatively discriminated in phosphoproteomics due to their identical mass.

> Phosphorylation is a reversible modification that regulates many biological events in a cell, including cell cycle, growth, apoptosis and signal transduction pathways<sup>1</sup>. Phosphorylation is currently the most studied and characterized post-translational modification (PTM) in biological systems, and likely one of the most abundant in nature. Its misregulation can be detrimental, giving rise to abnormal phenotype and many diseases such as cancer<sup>2, 3</sup>. Epidermal growth factor (EGF) binds to the EGF receptor on external cell membrane and stimulates cell signalling, modulating the activity of enzymes involved in specific pathways including transcriptional regulators, small GTPases and kinases by stimulating phosphorylation events<sup>4–6</sup>. Since EGF stimulation influences a complex cell signalling pathways, there are numerous phosphorylation and de-phosphorylation events in response to EGF<sup>4</sup>. Antibody based approaches for characterization of phosphorylations are labor intensive and time consuming to study a large number of protein phosphorylation events.

> Mass spectrometry (MS), especially when coupled with nano-liquid chromatography (nLC), has become the method of choice to characterize PTMs in large scale and in an unbiased manner<sup>7</sup>. In MS numerous acquisition methods have been developed to address different needs in protein and peptide characterization. For instance, selected reaction monitoring (SRM) gives the highest sensitivity and selectivity due to its ability to select predetermined precursor and product ions from complex mixtures. On the other hand, methods like data dependent acquisition (DDA) are preferred for discovery proteomics, since the instrument

<sup>†</sup>To whom correspondence should be addressed: Benjamin A. Garcia, Epigenetics Program, Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Room 9–124, 3400 Civic Center Blvd, Bldg 421, Philadelphia, PA 19104, USA. bgarci@mail.med.upenn.edu; Phone: 1-215-573-9423; Fax: 215-573-4764. \*Contributed equally to this work

Electronic Supplementary Information (ESI) available: Materials and Methods (document); Supplementary tables 1–3 (Excel tables).

automatically determines which signals should undergo fragmentation, generating a list of MS/MS spectra suitable for database searching and peptide identification<sup>8</sup>. More recently, data independent acquisition (DIA) methods have emerged, in order to provide a compromise between SRM and  $DDA<sup>9, 10</sup>$ . These methods use a first mass analyzer to generate wide isolation windows (10–25  $m/z$ ) stepping across a mass range, collecting sequential MS/MS at every instrument duty cycle. This leads to the production of an ion map of fragments from all detectable precursor masses<sup>11, 12</sup>. This type of raw file is suitable for the so-called "virtual SRM", where the abundance of selected peptide candidates is integrated using MS and MS/MS extracted ion chromatography  $(XIC)^{12}$ . However, such complex spectra are generally not ideal for peptide identification. Thus, the most common workflow currently is to perform a DDA run for peptide identification followed by DIA runs for accurate quantification. For this purpose, bioinformatics tools as  $(e.g.)$  Skyline<sup>13</sup>, OpenSWATH<sup>™ 14</sup> or other commercial software have been developed to integrate such datasets and facilitate the analysis.

Due to its nature, DIA has the potential to differentially quantify isobaric peptides, i.e. peptides with same sequence and PTMs, but with PTMs on different amino acid residues. We recently proved its efficiency in the analysis of histone samples, as histone peptides are highly enriched in isobaric modifications<sup>15, 16</sup>. However, histone protein samples are relatively simple mixtures. Thus far, no attempt was made for proteome-wide entire cell lysates. In this study, we performed DDA followed by DIA on an Orbitrap Fusion™ (Thermo Scientific) for the analysis of the entire phosphoproteome of EGF stimulated HeLa cells. After serum starvation, we treated HeLa cells with EGF for 0, 5, and 20 min. Then we extracted total proteins, performed traditional protein digestion using trypsin and phosphopeptide enrichment using titanium dioxide  $(TiO<sub>2</sub>)$  chromatography.

First, we analyzed the phosphoproteome using DDA, where we identified approximately 12,000 phosphopeptides (Figure 1 and Table S1). This number includes phosphopeptides with ambiguous localization site, as shown by the localization confidence score in Table S1. After using DDA to generate a spectral library we performed DIA runs. From the elution profile, most of the peaks width was at least 1 min (calculated by manual estimation of the peak baseline); given that instrument duty cycle was about 2 seconds this indicated that each chromatographic profile was accurately defined by about 30 data points. Peptide identifications from DDA and DIA raw files were uploaded in Skyline<sup>13</sup>. The software integrated the identification list with peak area integration; this list was filtered using the mProphet algorithm available in Skyline itself. This led to the quantification of more than 7,000 phosphopeptides (Figure 1 and Table S2). Of these, about 1,500 had at least one isobaric form also quantified (Figure 1 and Table S3).

We selected this last table and we attempted to perform differential quantification of isobaric phosphopeptide species using unique fragment ions, i.e. same fragment ion type (e.g. y5) with a different mass between the two species due to different position of the phosphoryl group. We focused only on phosphopeptides present in only two isobaric forms to avoid differential quantification from spectra with three or more mixed species. This manual approach led to a dramatic decrease in usable peptides; in fact, only 27 positional isomer pairs were obtained (Figure 1). The reasons of this relative small number as compared to the

Mol Biosyst. Author manuscript; available in PMC 2016 November 02.

total list of identified phosphopeptides are multiple: (i) only part of isobaric peptides have sufficiently intense unique fragment ions for differential quantification; (ii) it is required that both species are identified with comparable scores during database searching, which might not happen as one of the two species could be of low abundance and thus considered as not a confident identification; (iii) there is presently no automated data analysis software, implying that all calculations of relative ratios was tediously manually performed. In addition, 28 isobaric phosphopeptides (14 pairs) we considered had different retention times showing distinct chromatographic profile (Figure 2A). These do not require MS/MS based quantification, as their abundance can be easily discriminated using traditional MS based XIC.

One example of co-eluting and discriminated isobaric phosphopeptides is shown in Figure 2B. The peptide AASPSPQSVR (747, 756) was identified as belonging to the translated cDNA FLJ61739 with the phosphorylation pattern Ser 749 and 751, but also Ser 749 and 754. At the precursor mass level these two species completely overlap and thus would have the same abundance if quantified by MS XIC. By using the y5 fragment, which has a unique mass for each of the two species, we could determine the relative abundance of these peptides in all three analyzed conditions. It is worthwhile to note that phosphorylation on serine 751 decreases 5 min after EGF stimulation while phosphorylation on serine 754 does not decrease until 20 min after EGF treatment. More examples are illustrated in the Skyline file included as supporting information (for review, available at [https://](https://www.dropbox.com/sh/1tcsvanalqg6cd6/AADhClPRYcCtaTBoKH4T-jUaa?dl=0) [www.dropbox.com/sh/1tcsvanalqg6cd6/AADhClPRYcCtaTBoKH4T-jUaa?dl=0](https://www.dropbox.com/sh/1tcsvanalqg6cd6/AADhClPRYcCtaTBoKH4T-jUaa?dl=0)).

Finally, in order to evaluate the accuracy of MS/MS based quantification we considered the variance between technical replicates for those analyzed isobaric species (Figure 2D). The precursor mass ion signal provided the most reproducible observation (average CV 9.66%) as expected, being the most intense detectable ions. We also estimated the variance of the ions used to estimate the relative ratio between isobaric phosphopeptides performing three different comparisons. At first, we just estimated the variance of the intensity of the fragment ions between replicates (product, CV 17.70%); then, we calculated the ratio between the fragment ion of the isobaric species A and isobaric species B, and verified the reproducibility of the measurement between replicates (ratio product, CV 21.14%); finally, we calculated when possible the ratio of two isobaric peptides A and B using different fragment ions, and verified the reproducibility of the resulting ratio (Multiple ratio products, CV 72.46%). Results indicated that precursor and product ion signals had excellent reproducibility in their measurements. However, when calculating the ratio between two isobaric peptides the measurement was more variable depending on which fragment ion was used. This was expected, as very low abundance fragment ions, even if detectable, might have compressed dynamic range and therefore create a bias in the ratio calculation. In this study, if more than one unique fragment was detected, we considered the average ratio between the two isobaric isoforms.

#### **Conclusions**

We perform proof of principle experiments to introduce DIA as valuable tool to perform differential quantification of isobaric phosphopeptides in proteomics analyses. The

Mol Biosyst. Author manuscript; available in PMC 2016 November 02.

application of this method is still limited due to the challenge of (i) identifying multiple species from the same MS/MS spectrum, (ii) the sensitivity in detecting unique fragment ions of the same type for the two isobaric isoforms (e.g. y5 and y5′) and (iii) the mathematical limitations in calculating the relative ratio between more than two isoforms. However, we speculate that the rapidly increasing interest of DIA methods would lead to the development of suitable software for data processing, identification and quantification of positionally isobaric modified peptides, as much is currently missed due to our inability of discriminating the abundance of co-eluting peptides with same precursor mass.

#### **Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

#### **Acknowledgments**

We gratefully acknowledge funding from NIH grants R01GM110174 and R01AI118891, and a Leukemia & Lymphoma Society Dr. Robert Arceci Scholar Award.

#### **References**

- 1. Cohen P. Trends in biochemical sciences. 2000; 25:596–601. [PubMed: 11116185]
- 2. Chi P, Allis CD, Wang GG. Nature Reviews Cancer. 2010; 10:457–469. [PubMed: 20574448]
- 3. Portela A, Esteller M. Nat Biotechnol. 2010; 28:1057–1068. [PubMed: 20944598]
- 4. Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M. Cell. 2006; 127:635–648. [PubMed: 17081983]
- 5. Tong JF, Li LJ, Ballermann B, Wang ZX. Plos One. 2016:11.
- 6. Zhao YT, Feresin RG, Falcon-Perez JM, Salazar G. Traffic. 2016; 17:267–288. [PubMed: 26728129]
- 7. Karch KR, Denizio JE, Black BE, Garcia BA. Frontiers in genetics. 2013; 4:264. [PubMed: 24391660]
- 8. Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR 3rd. Chemical reviews. 2013; 113:2343–2394. [PubMed: 23438204]
- 9. Ma D, Chan MK, Lockstone HE, Pietsch SR, Jones DN, Cilia J, Hill MD, Robbins MJ, Benzel IM, Umrania Y, Guest PC, Levin Y, Maycox PR, Bahn S. Journal of proteome research. 2009; 8:3284– 3297. [PubMed: 19400588]
- 10. Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, Bonner R, Aebersold R. Molecular & cellular proteomics : MCP. 2012; 11:O111 016717.
- 11. Hopfgartner G, Tonoli D, Varesio E. Analytical and bioanalytical chemistry. 2012; 402:2587–2596. [PubMed: 22203371]
- 12. Reiter L, Rinner O, Picotti P, Huttenhain R, Beck M, Brusniak MY, Hengartner MO, Aebersold R. Nature methods. 2011; 8:430–435. [PubMed: 21423193]
- 13. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. Bioinformatics. 2010; 26:966–968. [PubMed: 20147306]
- 14. Rost HL, Rosenberger G, Navarro P, Gillet L, Miladinovic SM, Schubert OT, Wolski W, Collins BC, Malmstrom J, Malmstrom L, Aebersold R. Nat Biotechnol. 2014; 32:219–223. [PubMed: 24727770]
- 15. Sidoli S, Lin S, Xiong L, Bhanu NV, Karch KR, Johansen E, Hunter C, Mollah S, Garcia BA. Molecular & cellular proteomics : MCP. 2015; 14:2420–2428. [PubMed: 25636311]
- 16. Sidoli S, Simithy J, Karch KR, Kulej K, Garcia BA. Analytical chemistry. 2015; 87:11448–11454. [PubMed: 26505526]

Mol Biosyst. Author manuscript; available in PMC 2016 November 02.



#### **Figure 1. Numbers of characterized phosphopeptides**

Large circle, DDA identified phosphopeptides using Proteome Discoverer (Thermo). Second circle, number of quantified using Skyline confidence threshold. Third circle, number of isobaric phosphopeptides present in the Skyline result table. Small circle, differentially quantified isobaric phosphopeptides in this study.



#### **Figure 2. Quantification of isobaric phosphopeptides**

(A) LC-MS chromatogram displaying an example of two isobaric phosphopeptides, HNSpASVENVSLR and HNSASpVENVSLR, baseline separated by chromatography. (B) Quantification of two co-eluting isobaric forms of the phosphopeptide AASPSPQSVR, modified as S749/S751 (left) or S749/S754 (right). Bar plot and XIC represent the precursor signal of the given peptide. (C) The two co-eluting peptides were distinguished and quantified using the chromatogram of the unique fragment ion y5. The first showed decreased abundance after 5min of EGF stimulation, whereas the abundance was unchanged as compared to the control for the S754. (D) Distribution of the coefficient of variation between LC-MS replicates for the analyzed isobaric phosphopeptides. We considered precursor mass signals, product mass signals, calculated ratios and differential ratios calculated using different fragment ions (where possible).