Published in final edited form as:

Nat Methods. 2009 March ; 6(3): 184–185. doi:10.1038/nmeth0309-184.

A Computational Approach to Correct Arginine-to-Proline Conversion in Quantitative Proteomics

Sung Kyu Park¹, Lujian Liao¹, Jin Young Kim, and John R. Yates III^{*}

Department of Chemical Physiology, The Scripps Research Institute, La Jolla, CA 92014

To the editor

Stable isotope labeling of proteins using heavy amino acids is a widely used method to measure quantitative changes in mass spectrometry-based proteomics¹. When using a heavy form of arginine, however, heavy isotope labels can be inserted into proline through arginine catabolism. If the stable isotope incorporated into proline is not considered, ratios of proline-containing light and heavy peptides can be incorrectly calculated, leading to a reduction in intensity of the isotopic labeled heavy peptide (Fig. 1a). Proposed solutions have included decreasing the arginine concentration², or increasing the proline concentration³. Manipulating the amino acid concentration in culture media may result in sub-optimal culturing conditions for certain cell lines⁴. Van Hoof et al. proposed a method that replaces ¹²C¹⁴N arginine with ¹²C¹⁵N arginine in the light media, allowing the level of converted proline to be normalized by quantifying the monoisotopic peak in the MS spectra. Quantification solely based on monoisotopic peak can also compromise accuracy.

In a high-resolution mass spectrometer, isotope clusters are well defined, making it straightforward to distinguish the converted heavy proline clusters from the heavy arginine clusters (Fig. 1a). We extract individual isotope peaks rather than simply summing up all ion intensities within an *m/z* range of a predicted isotope distribution in which noisy peaks can be potentially included⁵ (Fig. 1b). Using an LTQ-Orbitrap mass spectrometer, we analyzed a 1:1 mixture of cultured cortical neurons grown in light or heavy media. Due to the nature of the primary cell culture (cells are not doubling) used, an incomplete heavy isotope labeling (ranging from 60 to 90 percent) is observed⁶. By using the ratio distribution of no-proline-containing peptides (Fig. 1c) as the target for the correction, the distribution plot with a single proline correction shows an obvious shift toward the expected ratio, and the distribution becomes more focused, indicating a more accurate overall ratio calculation. When considering the second heavy proline, the distribution plot became almost identical to the expected ratio. For PC12 cells that were completely labeled with heavy isotope, the proline correction resulted in a complete overlap of the ratio distribution profile between the proline-containing and no-proline-containing peptides and the distribution is centralized to zero (Fig. 1d). We also applied the same approach to low-resolution data where we found both single and double proline correction improved the accuracy of the ratio (Supplementary Methods, Fig. 1 and Table 1 online).

Depending on cell type and media conditions, heavy proline can contribute up to 30–40% of the heavy peptide⁴. By taking heavy proline into consideration, we achieve quantification that is more accurate, as demonstrated by the decreased standard deviation of ratios of five

^{*} Author to whom correspondence should be addressed. The Scripps Research Institute, 10550 North Torrey Pines Rd. SR11, Department of Chemical Physiology, La Jolla, CA 92037, USA, jyates@scripps.edu.

¹ These authors contributed equally to the paper.

representative proteins (Supplementary Table 2 online). The second satellite peak, a result of two proline residues in the peptide, contributes much less to the ratio correction than the first satellite peak (Figure 1a,c). We also examined the possibility of arginine-to-glutamate conversion as part of arginine catabolism, but the amount of heavy glutamate found in these samples was negligible (Supplementary Fig. 2 online).

In conclusion, our approach should allow the accurate calculation of expression ratios between light and heavy isotope labeled peptides without modifying the culture media used for labeling. The software is available at <http://fields.scripps.edu> (see Supplementary Methods for detailed instructions)

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors acknowledge NIH funding support UOM/DMID-BAA-03-38, SR21 AI072615-02, P41 RR011823, 5 R01 MH067880, and P30 NS057096.

References

1. Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* 2005;1:252–262. [PubMed: 16408053]
2. Blagoev B, Mann M. Quantitative proteomics to study mitogen-activated protein kinases. *Methods* (San Diego, Calif 2006;40:243–250.
3. Bendall SC, et al. Prevention of amino acid conversion in SILAC experiments with embryonic stem cells. *Mol Cell Proteomics* 2008;7:1587–1597. [PubMed: 18487603]
4. Van Hoof D, et al. An experimental correction for arginine-to-proline conversion artifacts in SILAC-based quantitative proteomics. *Nature methods* 2007;4:677–678. [PubMed: 17762871]
5. Park SK, Venable JD, Xu T, Yates JR 3rd. A quantitative analysis software tool for mass spectrometry-based proteomics. *Nature methods* 2008;5:319–322. [PubMed: 18345006]
6. Liao L, Park SK, Xu T, Vanderklish P, Yates JR. Quantitative proteomic analysis of primary neurons reveals diverse changes in synaptic protein content in *fmr1* knockout mice. *Proc Natl Acad Sci U S A* 2008;105:6.

Figure 1.

(a) Example of high resolution MS scan of two peptides identified as “AVFVDLEPTVIDEVR” (Upper panel) and “AEDGAAPSPSSETPK” (Lower panel). Heavy and light samples were mixed with an equal total protein amount. Note the presence of one or multiple proline(s) in the peptide sequence results in heavy proline clusters. (b) Theoretical spectra depicting the method used to calculate light/heavy peptide ratios. Ion current from multiple light isotope peaks are summed with a defined mass window to calculate the light peptide. While ion current from several clusters of heavy isotope labeled peptide are summed similar to that of light peptides, except that extra heavy isotope peaks derived from proline conversion were also added to get the heavy isotope ion current. (c) Histogram of peptide ratios (light versus heavy) with or without proline correction of incompletely labeled proteins quantified by LTQ-Orbitrap. A 1:1 protein mixture from light and heavy labeled neurons was digested and analyzed by MudPIT, with peptide ratios plotted here. Proline-containing peptides from high resolution data were plotted with no correction, with single proline correction, and with double proline correction. While single proline correction makes the distribution plot shift more than double proline correction, the double correction makes the plot closer to the expected ratio. (d) Similar to (c) except that the SILAC experiment was performed on PC12 cells such

that the heavy isotope labeling was complete. Because proline conversion is not prominent, only single proline correction was applied.