AUTOMATED SPECTRAL EXTRACTION FROM PARTIAL METABOLIC LABELING IN PLANTA EXPERIMENTS ENABLING PROTEIN TURNOVER CALCULATIONS

David Lyon, M[a] Angeles Castillejo, Christiana Staudinger, Vlora Mehmeti, Volker Egelhofer and Stefanie Wienkoop

david.lyon@univie.ac.at

Introduction

In shotgun proteomics the FCP (Fold Change in Protein) is widely used to compare protein levels of various samples, but neither resolves the dynamics of the proteome in the different biological states that are being compared nor the mechanisms whereby the system changes from one state to the other [1–3]. Metabolic (in situ) labeling is characterized by the incorporation of stable isotopes into the proteins of organisms via growth on media or food [4]. Mass spectrometric measurement of the ratio between light (naturally occurring isotopic distribution) and heavy (enriched with ^{15}N) isotopes and their respective degrees of enrichment provide a means to measure the synthesis and degradation rates of individual proteins [1,5].

We hereby present an approach enabling the automated extraction and quantification of ^{15}N partial metabolic LC/MS shotgun proteomics data enabling the calculation of ^{15}N incorporation rates (such as q-values q = I (^{15}N) / (I (^{14}N) + I (^{15}N)) [6] (Fig. 1 and 4). The major advantages are automated high throughput data analysis and that the approach is not restricted to any organism or tissue.

Methods

Workflow

MS raw data and a SelPEX*-list (Selective Peptide EXtraction) [8] (consisting of amino acid sequence, charge state, and retention time) serve as input. The extracted MS[1] spectra of the peptides (m/z-values and their corresponding intensities) are the output. If a protein fasta file is provided, the peptides can be affiliated with proteins (Fig. 2).

Biological application: monitoring 5 days of plant drought stress-recovery

Our experimental setup consisted of seven week old plants (grown in perlite/vermiculite substrate) that were randomly separated into two groups: drought and control. The drought treated plants endured ten days of withholding water, while the control group was continuously watered. After ten days of drought, each group was again separated into two subsets, ^{14}N and ^{15}N respectively. A re-watering experiment was carried out where ^{15}N subsets were fertilized with ammonium nitrate enriched with ^{15}N (98%), while the ^{14}N subsets were supplied with unaltered ammonium nitrate (SILIP Stable Isotope Labeling In Planta [7]). The plants were initially harvested after 10 days of drought and subsequently with the onset of rewatering on a daily basis for five days.

MS setup

Time resolved LC/MS shotgun proteomics data from ^{15}N partial metabolic labeling In Planta of the model plant Medicago truncatula were analysed. The samples were subjected to LC/MS/MS shotgun proteomics using an LTQ-Orbitrap-XL and GC/MS/MS metabolomics. A SelPEX approach was applied, resulting from database dependent identification of peptides, to generate a list of 1384 peptide sequences, their charge state and retention time. The following constraints were applied to the list of identifications: proteins identified in all six replicates (three biological and two technical replicates), at least two peptides per protein, and no modifications. From this list the 100 most abundant peptides (amino acid sequence, charge state and retention time) were selected as a training set for the program. When inspecting the MS[1] spectra of the selected peptides, a time dependent formation of the isotopic envelope (Fig. 2 and 3) could automatically be extracted within minutes. Supplemental GC/MS data also showed a time dependent increase of ^{15}N labeled metabolites.

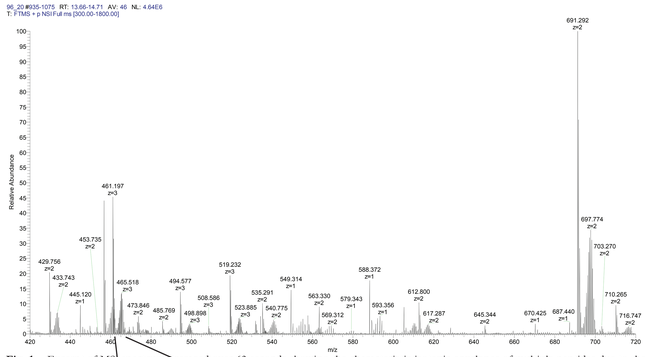

Fig. 1. Excerpt of MS[1] spectrum averaged over 63 seconds showing the characteristic isotopic envelopes of multiple peptides due to the incorporation of ^{15}N.

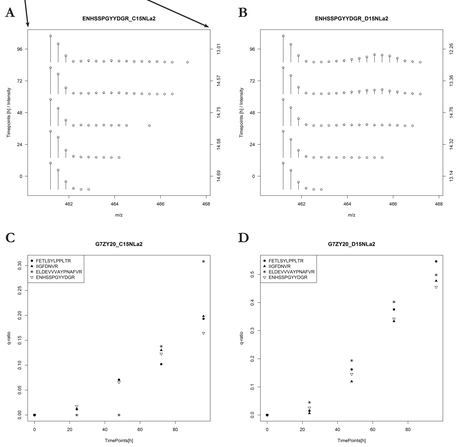

Fig. 2. Preliminary output of the program: automatically extracted spectra of control (**A**) and drought stressed (**B**) peptide, as well as the q-ratios of their corresponding protein (**C** and **D**). The individual spectra are normalized to the base peak. The drought stressed spectra show increased ^{15}N incorporation in comparison to the control. This is also evident in the q-ratios (**C** and **D**) of the corresponding protein (G7ZY20 is RUBISCO small chain).

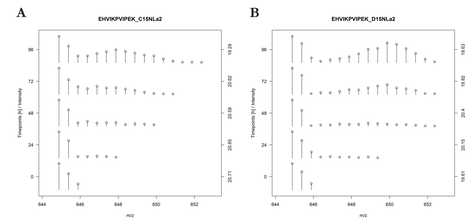

Fig. 3. Example of the automatically extracted isotopic envelope of the SelPEX peptide EHVIKPVIPEK (control **A**, drought **B**). Five measurements of increasing incorporation of stable isotope labeling are plotted above each other (from bottom to top, maximally labeled on top). Each data point in a spectrum is normalized to the basepeak of the given spectrum. The control is displayed on the left side and the corresponding drought treatment on the right. A shift in the skewness is apparent when comparing the spectra, indicating a varying degree of fractional labeling.

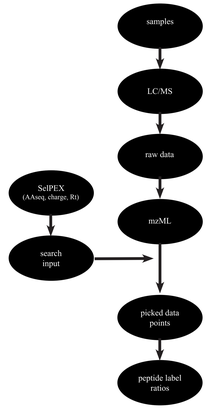

Fig. 4. Overview of the workflow: Only the SelPEX list is organism specific, the workflow and the program are not. The program calculates the possible isotopic envelope of a given sequence and charge state and extracts the corresponding spectrum from the raw data. The ratio (q-value) of ^{15}N (all ions of the isotopic envelope due to ^{15}N enrichment) to ^{14}N (all ions of the naturally occuring isotopic envelope) reflects protein dynamics.

Motivation

The output of the program (the extracted spectral envelope of individual peptides) can be used to calculate q-ratios, synthesis and degradation rates of proteins, as well as the degree of fractional labeling [9]. The q-ratio reflects the incorporation of newly synthesized protein species enriched with ^{15}N, in conjunction with the degradation of existing protein species. Thus large q-ratios indicate high protein turnover. While the synthesis and degradation rate show distinct processes. Short half-life proteins are often involved in signal transduction, whereas long half-life proteins are frequently involved in housekeeping roles, cytoskeletal in nature or are nuclear localized [10]. Regardless of the source of control, differentially regulated protein half-lives can be attributed to amino acid composition, length, localization within the organism and, probabyl most important, biological functions. Correlative analysis of q-values, synthesis and degradation rates between treatments, may reveal groups of functionally interacting proteins. These groups may be associated with specific degradation pathways or prove to be part of unknown regulatory mechanisms. We expect to elucidate proteins and or peptides responsible for the physiological adaptation occurring during drought stress and recovery and find the underlying mechanisms reflected by the differential regulation of protein levels, relevant for this adaptation.

Discussion

Our preliminary results demonstrate that an automated analysis of protein synthesis and degradation for previously identified SelPEX targets from complex samples is now possible. This approach is not restricted to any organism or tissue. Compared to previously published data on cell culture levels, our data show, that in Planta ^{15}N incorporation rates are slower (biological realtime), demonstarting the relevance of in Planta experiments for the interpretation of biological processes.

The initial data show that with this approach we are able to find differentially regulated q-ratios between treatments (drought vs. control) alluding to physiological mechanisms characteristic for drought (Fig. 2 and 3). By combining data from GC/MS metabolomics and physiological parameters with the presented LC/MS proteomics data, we aim for a holistic understanding of the underlying processes. Due to sample complexity and the occurrence of overlapping spectral envelopes, one or two dimensional gel electrophoresis precedes LC/MS analysis. Our program distinguished these signals if the mass-resolving-power is feasably high. To the best of the author's knowledge, this is the first SILIP gel-free shotgun proteomics study of a complex sample without any protein pre-fractionation.

(1) Pratt, J. M.; Petty, J.; Riba-Garcia, I.; Robertson, D. H. L.; Gaskell, S. J.; Oliver, S. G.; Beynon, R. J. Molecular & Cellular Proteomics 2002, 1, 579–591.
(2) Hoehenwarter, W.; Wienkoop, S. Rapid Communications in Mass Spectrometry 2010, 3609–3614.
(3) Li, L.; Nelson, C. J.; Solheim, C.; Whelan, J.; Millar, a H. Molecular & cellular proteomics : MCP 2012, 11, M111.010025.
(4) Kline, K. G.; Sussman, M. R. Annual review of biophysics 2010, 39, 291–308.
(5) Cargile, B. J.; Bundy, J. L.; Grunden, A. M.; Stephenson, J. L. Analytical chemistry 2004, 76, 86–97.
(6) Gustavsson, N.; Greber, B.; Kreitler, T.; Himmelbauer, H.; Lehrach, H.; Gobom, J. Proteomics 2005, 5, 3563–70.
(7) Schaff, J. E.; Mbeunkui, F.; Blackburn, K.; Bird, D. M.; Goshe, M. B. The Plant journal : for cell and molecular biology 2008, 56, 840–54.
(8) Castillejo, Ma. A. C.; Staudinger, C.; Egelhofer, V. and Wienkoop, S. Chapter 15 In: Plant Proteomics Methods and Protocols. Ed. J.V. Jorrin Novo, S. Komatsu, S. Wienkoop, W. Weckwerth: Springer New York. (in press)
(9) Sperling, E.; Bunner, A. E.; Sykes, M. T.; Williamson, J. R. Anal. Chem. 2008, 80, 4906–4917.
(10) Lei, Li. Application of 15N labeling to measure protein turnover rate in Arabidopsis thaliana. PhD thesis, 2013.

SelPEX is based on database dependent protein identification. Peptides with good ionisation properties (and ideally proteotypic), are selected for a robust ^{15}N analysis. This approach allows for the best accuracy for quantification and identification.

FWF Project Number: P23441-B20

FWF Der Wissenschaftsfonds.
Austrian Science Fund

Department
Molecular Systems Biology
universität wien
Universität of Vienna | Faculty of Life Sciences | Althanstrasse 14 | 1090 Vienna | Austria