Reproducible and Consistent Quantification of the Saccharomyces cerevisiae Proteome by SWATH-mass spectrometry*§

Nathalie Selevsek‡, Ching-Yun Chang§, Ludovic C. Gillet‡, Pedro Navarro¶¶, Oliver M. Bernhardt¶¶, Lukas Reitter¶¶, Lin-Yang Cheng§, Olga Vitek|||, and Ruedi Aebersold‡‡‡§§

Targeted mass spectrometry by selected reaction monitoring (S/MRM) has proven to be a suitable technique for the consistent and reproducible quantification of proteins across multiple biological samples and a wide dynamic range. This performance profile is an important prerequisite for systems biology and biomedical research. However, the method is limited to the measurements of a few hundred peptides per LC-MS analysis. Recently, we introduced SWATH-MS, a combination of data independent acquisition and targeted data analysis that vastly extends the number of peptides/proteins quantified per sample, while maintaining the favorable performance profile of S/MRM. Here we applied the SWATH-MS technique to quantify changes over time in a large fraction of the proteome expressed in Saccharomyces cerevisiae in response to osmotic stress.

We sampled cell cultures in biological triplicates at six time points following the application of osmotic stress and acquired single injection data independent acquisition data sets on a high-resolution 5600 tripleTOF instrument operated in SWATH mode. Proteins were quantified by the targeted extraction and integration of transition signal groups from the SWATH-MS datasets for peptides that are proteotypic for specific yeast proteins. We consistently identified and quantified more than 15,000 peptides and 2500 proteins across the 18 samples. We demonstrate high reproducibility between technical and biological replicates across all time points and protein abundances. In addition, we show that the abundance of hundreds of proteins was significantly regulated upon osmotic shock, and pathway enrichment analysis revealed that the proteins reacting to osmotic shock are mainly involved in the carbohydrate and amino acid metabolism. Overall, this study demonstrates the ability of SWATH-MS to efficiently generate reproducible, consistent, and quantitatively accurate measurements of a large fraction of a proteome across multiple samples. Molecular & Cellular Proteomics 14: 10.1074/mcp.M113.035550, 739–749, 2015.

In systems biology and biomedical studies targeted mass spectrometry via selected reaction monitoring (SRM)† (also known as multiple reaction monitoring, MRM) has emerged as a powerful technique for the consistent and reproducible quantification of proteins across numerous complex samples (1–6). Optimal sets of precursor/fragment ion pairs, called transitions, uniquely represent a specific peptide. They constitute a definitive mass spectrometric assay for the detection of targeted peptides, and thus the proteins from which they derive, in the complex matrix of trypsinized biological samples (1, 7). Protein quantification is then performed by relating the intensity of the acquired transition signals to suitable reference signals. Most quantification strategies commonly used in proteomics are compatible with this method (8). Recently, the high-throughput development of S/MRM assays has been achieved via the generation of MS/MS spectral libraries from the measurements of thousands of synthetic peptides representing proteotypic peptides (9). Moreover, many experimental and bioinformatics workflows have been developed for assay generation, assay optimization, data evaluation, and the dissemination of optimized S/MRM assays (10–16). In combination, these developments have supported the creation of mass-spectrometric maps of entire proteomes of selected species including Streptococcus pyogenes, Myco-

---

* Indicates equal contribution.

† The abbreviations used are: SRM, selective reaction monitoring; MRM, multiple reaction monitoring; DIA, data independent acquisition; FDR, false discovery rate.
bacterium tuberculosis, and Saccharomyces cerevisiae (5, 17–19) and the robust use of these resources to quantify specific protein sets across multiple biological samples.

Currently, targeted proteomics by S/MRM can be multiplexed to a maximum set of ~100 proteins that can be measured in a single LC-S/MRM run at optimal quantitative accuracy, limit of detection and dynamic range. The quantification of higher numbers of proteins per run compromises some of the performance parameters of the method because of well understood tradeoffs (8). Attempts have been made to further increase the degree of multiplexing of S/MRM, either by automated adjustment of the scheduled detection windows (20) or by acquiring, in a data-dependent manner, the complete set of precursor-fragment ion pairs of a given assay (21). Alternatively, parallel reaction monitoring (PRM) approach operated on quadrupole-orbitrap mass spectrometer has shown detection and quantification performances similar or better than those obtained in SRM, because of the increased selectivity of the mass analyzer (22–24). These approaches are promising, but their application relies on prior knowledge of the precursor ions that need to be targeted during the data acquisition, and they still are subject of the above-mentioned tradeoffs.

Recently, we developed a novel MS strategy that combines data independent acquisition (DIA) of trypsinized protein samples with S/MRM-like, in silico targeted analysis of the acquired complete fragment ion maps (25). We termed the method SWATH-MS, and applied the sequential isolation window acquisition principle (26) to repeatedly cycle, in a single injection, through 32 consecutive 25-Da precursor isolation windows (swaths). The process acquires fragment ion spectra of all precursors in a space defined by the 400–1200 m/z precursor range and a user-specified retention time window. We used the prior information in MS/MS spectral libraries to extract groups of signals that uniquely identify a specific peptide, and to demonstrate that peptides could be identified and quantified over a dynamic range of four orders of magnitude, even when the precursors were not detectable in a survey MS scan. For the 45 proteins involved in the central carbon metabolism of yeast, we demonstrated that the accuracy of quantification was equivalent to that of S/MRM (25). However, because of the lack of adequate software tools at that time, the extensive high-throughput targeted data analysis of the SWATH-MS maps could not be fully demonstrated in that first study.

Here we demonstrate the multiplexing capabilities of SWATH-MS for the detection and quantification of significantly larger fractions of a proteome as compared with S/MRM, without compromising reproducibility, consistency, and quantitative accuracy. We describe the large scale deployment of fragment ion spectral libraries and the use of S/MRM-like analysis tools specifically adapted to SWATH-MS data for the detection and quantification of temporal changes of the S. cerevisiae proteome in response to osmotic stress.

**MATERIALS AND METHODS**

**Yeast Culture, Protein Isolation, and Digestion—** Three series of six cultures each from the yeast strain BY4741 MATa his3Δ leu2Δ met15Δ ura3Δ were grown in S.D. medium until they reached an A<sub>600</sub> of 0.8. To apply the osmotic shock, 0.4 m NaCl was added to each 50 ml culture and after 0 min (T0), 15 min (T1), 30 min (T2), 60 min (T3), 90 (T4), and 120 min (T5) the cells were harvested. At the respective time points, the culture media were quenched by addition of trichloroacetic acid (TCA) to a final concentration of 6.25% and the cells were harvested by centrifugation at 1500 × g for 5 min at 4 °C. The supernatants were discarded and the cell pellets were washed three times by centrifugation with cold (−20 °C) acetone to remove interfering compounds. The final cell pellets were resolubilized in lysis buffer containing 8 m urea, 0.1 m NH₄HCO₃, and 5 mM EDTA and cells were disrupted by glass bead beating (five times 5 min at 4 °C). The total protein amount from the pooled supernatants was determined by BCA Protein Assay Kit (Thermo). Yeast proteins were reduced with 12 mM dithiothreitol at 37 °C for 30 min and alkylated with 40 mM iodoacetamide at room temperature in the dark for 30 min. Samples were diluted with 0.1 m NH₄HCO₃ to a final concentration of 1.5 m urea and the proteins were digested with sequencing grade porcine trypsin (Promega, Madison, WI) at a final enzyme:substrate ratio of 1:100. Digestion was stopped by adding formic acid to a final concentration of 1%. Peptide mixtures were desalted using reverse phase cartridges Sep-Pak TC18 (Waters, Milford, MA) according to the following procedure: wet cartridge with one volume of 100% methanol, wash with one volume of 80% acetonitrile, equilibrate with four volumes of 0.1% formic acid, load acidified digest, wash six volumes of 0.1% formic acid, and elute with one volume of 50% acetonitrile in 0.1% formic acid. Peptides were dried using a vacuum centrifuge and resolubilized in 100 µl of 0.1% formic acid and frozen at −20 °C. For in depth-fractionation experiments of NaCl-untreated yeast cells, the peptide mixtures were separated by off gel electrophoresis (OGE) using a pH 3–7 IPG strip (Amersham Biosciences) and a 3100 OFFGEL Fractionator (Agilent Technologies, Santa Clara, CA) with collection in 12 wells and then submitted for C18 clean-up. All samples were spiked with the retention time standard peptides iRT-Kit (Biognosys, Schlieren, Switzerland).

**LC-MS Set Up—** Yeast tryptic peptides were analyzed on a 5600 TripleTOF™ mass spectrometer (ABSciex, Concord, Ontario). Chromatographic separation of peptides was performed on a NanoLC-2DPlus HPLC system (Eksigent, Dublin, CA) coupled to a 75 μm diameter fused silica emitter packed with 20 cm of Magic C18 AQ 3 μm resin (Michrom BioResources, Auburn, CA). Peptide samples (2 µl injections containing 2 µg of total peptide amount) were loaded on the column from a cooled (4 °C) Eksigent autosampler using a flow of 300 nL/min of an isotropic 98% Buffer A (Buffer A: 2% acetonitrile, 0.1% formic acid) and 2% Buffer B (Buffer B: 98% acetonitrile, 0.1% formic acid). Peptides were then separated at a flow rate of 300 nL/min with a 180 min linear gradient of 2% to 35% Buffer B. For shotgun experiments, the mass spectrometer was operated with a "top 20" method: Initially, a 500 ms survey scan (TOF-MS) was collected from which the top 20 precursor ions were automatically selected for fragmentation, whereby each MS/MS event consisted of a 150 ms fragment ion scan. The main selection criterion for parent ion selection was precursor ion intensity. Ions with an intensity exceeding 20 counts per second and a charge state greater than 2+ were preferentially selected. The selected precursors were then added to a dynamic exclusion list for 20s. Ions were isolated using a quadrupole resolution of 0.7 amu and fragmented in the collision cell using the collision energy equation (0.0625 × m/z - 3.5) with an additional
collision energy spread of 15 eV within the 150 ms accumulation time to mimic SWATH like fragmentation conditions. If ≤ 20 precursor ions meeting the selection criteria were detected per survey scan, the detected precursors were subjected to extended MS/MS accumulation times to maintain a constant total cycle time of 3.5 s. For SWATH-MS based experiments the mass spectrometer was operated in a looped product ion mode. In this mode the instrument was specifically tuned to allow a quadrupole resolution of 25 amu per mass selection. The stability of the mass selection was maintained by the operation of the RF and DC voltages on the isolation quadrupole in an independent manner. Using an isolation width of 25 amu, a set of 32 overlapping windows (1 amu overlap) was constructed covering the mass range 400 to 1200 amu. The collision energy for each window was determined based on the collision energy for a putative 2+ ion centered in the respective window (equation: 0.0625 × m/z - 3.5) with a collision energy spread of 15 eV. An accumulation time of 100 ms was used for each fragment ion scan and for the survey scans acquired at the beginning of each cycle, resulting in a total cycle time of 3.3 s. The sequential precursor isolation window set-up was as follows: 499–425, 424–450, 449–475,..., 1174–1200, with an effective (100%) transmission of ~25 Da and ~0.3 Da attrition on either side of the isolation window.

**Generation of Spectral Library and Database Searching**—The shotgun spectral library was generated using a total of 46 shotgun injections on an ABSciex TripleTOF 5600 instrument: two series (technical replicate injections) of 12 off-gel fractions from the BY 4741 strain (in exponential phase, OD600 = 0.8), 18 injections (six time points in biological triplicates) from the complete BY 4741 osmotic shock study, and four technical replicate injections of a FY4 strain pooled in TPP 4.5.2. The transition MS/MS coordinates for the peptides used to build a consensus spectra library using SpectraST (included in TPP 4.5.2) (27). The transition MS/MS coordinates for the peptides in the experiment. However, multiple methods allow us to use the p values to control the FDR in the list of the identifications. We used the method of Storey et al. (29), which uses a two-group mixture model of the p values to estimate the q-values, and to control the FDR. In this study, the estimates based on this method turned out to be very accurate and generally slightly conservative (supplemental Fig. S2A). This process of FDR estimation in targeted proteomics experiments has also been validated in a similar manner for S/MMRM. For a detailed description please see the supplementary of (15).

For the technical replicate measurements, the features confidently identified below 1% FDR threshold were used to estimate the elution IRT value of the feature. Then we recovered the features with q-values over 1% FDR but below 10% FDR, for which an elution IRT value was close (i.e. less than 0.35 IRT units) to the estimated elution IRT. To evaluate the FDR after this identification recovery, we considered that the IRT tolerance window only covers 8.33% of the total extraction window, and therefore only 8.33% of the recovered false identifications meet the recovery criterion, assuming that false identifications are spread uniformly in the retention time space. This keeps the actual FDR below 1%.

**Statistical Analysis**—The investigation aimed at identifying proteins that significantly changed their abundance across the different time points. All the identified and quantified peak intensities were first transformed by the logarithm with base 2. To further filter out incorrectly identified or quantified transitions, we calculated the Pearson
correlation between the intensity of each transition and the average intensity of the protein’s transitions across all MS runs, and transitions with correlation below 0.5 were removed. The remaining data were subjected to constant normalization to equalize the median peak intensities of transitions between runs (30).

Protein-level quantification and testing for differential abundance were performed using MSstats (31) based on a linear mixed-effects model (32). The model decomposes log-intensities into the effects of times, biological replicates, SWATH features (i.e. a combination of peptides and transitions), MS runs, and statistical interactions. The model specified the reduced scope of biological replication. For each investigated comparison, MSstats provided model-based estimates of fold changes, as well as \( p \) values that were adjusted to control the FDR at the cutoff 0.05 (33). All the input and output files from the MSstats analysis together with the R scripts performed for the manuscript are provided in supplemental Data S1.

RESULTS

Detection of Yeast Proteins by SWATH-MS—We first determined the fraction of the S. cerevisiae proteome that could be detected in a trypsinized, unfraccionated yeast cell lysate by SWATH-MS. For that purpose, we selected MS coordinates consisting of precursor/fragment ion pairs, the relative fragment ion intensities and peptide elution time for each targeted peptide, and extracted signal groups corresponding to these coordinates from SWATH-MS datasets. The MS coordinates were generated as prior information via in-depth fractionation and MS sequencing of yeast cell digests using the same model 5600 QqTOF mass spectrometer that was also used to acquire the SWATH-MS datasets (Fig. 1). For the DDA measurements we generated protein extracts from cells in different perturbed conditions, including stages of a diauxic shift experiment and cells subjected to salt stress. Tryptic digests of these samples were fractionated by isoelectric focusing via off-gel electrophoresis and analyzed by DDA MS.

The resulting MS/MS spectra were searched against the Saccharomyces Genome Database (SGD) and the identified peptides were statistically filtered at 1% FDR at the PSM level. We compiled the resulting MS/MS spectra into a spectral library using SpectraST (27) and reported the final number of spectra, peptides and proteins in supplementary Results. Next, we used this spectral and an in-house python script to extract the coordinates of MS/MS transitions to target in SWATH-MS. In total, we targeted 331 4110 fragment ion traces (transition signals) corresponding to 66 007 proteotypic peptides and 4 596 proteins in each SWATH-MS. The fragment ion traces were extracted using the Spectronaut software, and the results were filtered at 1% FDR at the assay level using the error rate model originally described for mProphet (15). The results indicated that 16,178 peptides corresponding to 2578 proteins could be confidently detected with a FDR of 1% estimated by mProphet in an unfraccionated digest of a yeast cell lysate (supplemental Table S3). Overall, 71% of the proteins were identified with at least two peptides per protein (Fig. 2A). According to the comprehensive Western blot analysis performed by tandem affinity purification tagged yeast ORFs (34) the detected proteins spanned a concentration range of four orders of magnitude down to \( 1 \times 10^6 \) to 100 copies/cell (Fig. 2B). Furthermore, the data showed that the detected proteins were not biased against low abundant proteins, and that SWATH-MS could confidently detect \( > 300 \) proteins that could not be detected by the quantitative Western blot analysis (Fig. 2C). The fraction of the spectral library that was not detected by SWATH-MS may be explained by the fact that SWATH-MS is still not sensitive at the level of the SRM technology (35). Overall, these results demonstrate that a single-
Proteins are spanning a concentration range of four orders of magnitude from $10^6$ to 100 copies/cell according to quantitative Western blot analysis. Multiple injections at a high degree of reproducibility, and to detect a significant fraction of the proteome across different states, and used the Spectronaut/mProphet software tools for targeted data extraction. We integrated the fragment ions traces from assays that were confidently identified in all six runs and in both isotopic channels, i.e., a total of 3354 peptides from 780 proteins. In the case of the 1:1 mixture, an estimated fold change of 0.92 ± 0.14 S.E. was obtained whereas for the 10:1 mixture, 10.47 ± 4 S.E. by MSstats (Fig. 4), thus, demonstrating that SWATH-MS achieves accurate relative quantification at the proteome level. The quantification results of the six SWATH-MS runs are represented in supplemental Table S5.

Quantification of Changes in the Yeast Proteome upon Osmotic Shock—We next deployed the technique to quantify the changes in the S. cerevisiae proteome induced by osmolarity stress. The osmolarity stress response has been extensively studied in yeast and it occurs frequently in the yeast cell’s natural environment. Specifically, we investigated the time-resolved response of the yeast proteome to NaCl, a salt that is commonly used for inducing an osmotic stress response (36). We added NaCl to a concentration of 0.4 M to cells in exponential growth phase and sampled the cultures at different time-points (0 min (T0), 15 min (T1), 30 min (T2), 60 min (T3), 90 min (T4), and 120 min (T5)) after salt addition in

**Fig. 2.** SWATH-MS analysis of an unfraccionated yeast cell lysate. A, Number of peptides per protein identified in a yeast cell digest. B, Proteins are spanning a concentration range of four orders of magnitude from $10^6$ to 100 copies/cell according to quantitative Western blot analysis. C, Distribution of the cellular abundances (copies/cell) of the detected proteins according to quantitative Western blot analysis (34).
biological triplicates. The proteins isolated from the 18 yeast cultures were digested and subjected to SWATH-MS analysis. Targeted data extraction of the queried proteins was performed using the spectral library described above. The results of the targeted data extraction are presented in supplemental Table S6. Identified and quantified peak intensities were log-transformed and normalized using MSstats as described in material and methods. For the downstream statistical analysis, we only considered the peptides that were detected in at least three replicates and at least one time point. As the result of this filtering, 29% of transitions were removed before the statistical analysis, but after the identification step. In other words, the protein identification step by Spectronaut used four transitions for every protein, whereas the MSstats analysis step used between one and four transitions per protein. After all the filtering steps, the input to MSstats was a set of 2589 target proteins, where each protein was represented by 1–129 peptides, and each peptide was represented by 1–4 transitions and no missing values. Out of the 2589 proteins, only 75 proteins had one transition as input to MSstats.

Next, protein-level quantification and testing for differential abundance were performed based on a linear mixed-effects model (32). Several comparisons with respect to the baseline at time point 0 min (T0) (specifically, comparing protein abundances at time 0 min (T0) to the abundances at 15 min (T1), 30
min (T2), 60 min (T3), 90 min (T4), and 120 min (T5) after addition of salt) were tested for each protein.

Next, we performed unsupervised clustering of proteins according to their patterns of change in abundance over time. We enumerated all the possible $3^5$ such patterns (i.e. significantly up-regulated, significantly down-regulated and no statistically significant change, at times 15 min (T1), 30 min (T2), 60 min (T3), 90 min (T4), and 120 min (T5) as compared with time 0 min (T0)). We retained four clusters with more than 50 proteins each for further examination (Fig. 5A and B). Proteins in cluster 1 ($n = 266$) and cluster 2 ($n = 67$) are up-regulated along the complete time course, with a $> 20$ min delay in the response for proteins belonging to cluster 2. Pathway enrichment analysis tool DAVID (http://david.abcc.ncifcrf.gov/) revealed an over-representation of proteins involved in carbohydrate metabolism such as the glycolysis-gluconeogenesis pathway ($p = 4.4 \times 10^{-6}$), the starch and sucrose metabolism ($p = 7.2 \times 10^{-6}$) and the pentose phosphate pathway ($p = 8.3 \times 10^{-5}$). These pathways are directly linked to the glycerol, threhalose and glycogen metabolism, which are known to be induced by osmotic shock to trigger the production of glycerol, an essential osmolyte for osmoadaptation (36). In contrast, down-regulated protein profiles were obtained for cluster 3 ($n = 219$) and cluster 4 ($n = 567$) upon addition of salt.
Proteome Quantification by SWATH-MS

with a > 20 min delay in the response for proteins belonging to cluster 3. Mainly enzymes involved in amino acid biosynthesis (i.e. the glycine, serine, threonine metabolism pathway; \( p = 3.4\times10^{-4} \) and the phenylalanine, tyrosine, and tryptophan pathway, \( p = 2.3\times10^{-2} \)) were found to be repressed as previously suggested (37).

Validation of Protein Fold Changes by S/MRM—To validate the protein fold changes obtained by SWATH-MS, we quantified a subset of 100 proteins in the 18 osmotic shock time course experiment using S/MRM. Among these proteins, 24 were up-regulated along the complete time course according to SWATH-MS, 22 were down-regulated along the complete time course, and five were down-regulated along the complete time course after a > 20 min delay. The remaining 49 proteins were not regulated between any time points according to SWATH-MS. For each of the 100 proteins, we chose the highest-responding peptide for each protein that was confidently identified in SWATH-MS, together with their corresponding fragment ions (i.e. four transitions in total), for S/MRM quantification. The fragment ion traces of these peptides detected with S/MRM were extracted and integrated using Skyline (13) together with the corresponding peptides detected with SWATH-MS (supplemental Data S2). This allowed us to consistently process both data sets with the same integration/quantification parameters using the same software. Because the optimal peptides and transitions were chosen for the S/MRM quantification, all the quantified transitions were of a relatively good quality, and no downstream filtering was necessary. The quantitative values were used as input to MSstats to estimate log fold changes and their associated confidence intervals, and to test proteins for differential abundance, between 15 min (T1), 30 min (T2), 60 min (T3), 90 min (T4), 120 min (T5), and the initial time point 0 min (T0).

Supplemental Data S3 shows log fold change profiles and their associated confidence intervals of the 51 regulated proteins quantified by S/MRM and SWATH-MS. To be consistent, for the results for SWATH-MS we only used the subset of the peptides and of the fragments that were also quantified by S/MRM, even though for SWATH-MS additional transition signals for these peptides and additional peptides were concurrently acquired. The data show that the majority of the confidence intervals overlap, indicating good agreement between SRM and SWATH-MS quantification. To further formalize the comparison between the methods, we compared the outcomes of tests for differential abundance. We classified the outcome of each test for differential abundance as significant up-regulation (denoted as +1), significant down-regulation (denoted as −1), and absence of significant regulation (denoted as 0) (supplemental Table S7). These tests for differential abundance were applied to the five time points of the study resulting in five values per protein across the time course. Table I shows that for 64.7% of the proteins the conclusions from the two datasets agree in at least four time points. Conversely, only for 17.6% of the proteins conclusions from the two datasets agree in less than three time points, whereas seven proteins had no agreement because of interferences or low quality peak shape (see supplemental Data S4). The remaining discrepancies can also be caused by other reasons. First, the S/MRM and SWATH-MS acquisitions occurred several weeks apart. Second, external factors such differences in chromatographic conditions (3 h in SWATH versus 30 min in SRM) may generate differences in ionization or ion suppression between both measurements, in a way that impacts peptide quantification accuracy of one or the other method. Finally, the nature of the mass analysis between the two methods may result in differences in sensitivity and detection. Overall, the results demonstrated that SWATH-MS targeted analysis of complex samples can provide biological conclusions that are in a high agreement with S/MRM, but at a much higher throughput.

Correlation of Protein Fold Changes with Transcriptomics—We next correlated the protein abundance profiles obtained by SWATH-MS with their corresponding transcript profiles across the four most significantly regulated pathways (i.e. the glycolysis-gluconeogenesis pathway, the pentose phosphate pathway, the glycine, serine, threonine metabolism pathway and the phenylalanine, tyrosine, and tryptophan pathway). For this purpose, we used a transcriptomic data set that was previously generated for yeast treated under similar experimental conditions (37). Fig. 6 represents the quantitative patterns (as heat maps) for each transcript and protein along the osmotic shock time course together with a frequency plot assessing of their similarity in terms of Pearson correlation coefficients for each of the four pathways. The results revealed that the correlation between transcripts and protein abundance greatly differs depending on the biological pathway. In the case of the glycolysis-gluconeogenesis pathway, 50% of the detected proteins are reflecting good levels of correlation (i.e. Pearson correlation coefficient > 0.5) with their corresponding transcripts, whereas 50% of the detected proteins involved in the phenylalanine, tyrosine, and tryptophan pathway are showing consistent anticorrelation (i.e. Pearson correlation coefficient < 0.5) between protein and transcript abundance changes. However, 50% of the protein profiles measured for the pentose phosphate pathway and the glycine, serine, and threonine metabolism pathway showed no clear correlation (i.e. −0.5 < Pearson correlation coefficient < 0.5) between the protein abundance and their

| Number of agreements | 0 | 1 | 2 | 3 | 4 | 5 |
|----------------------|---|---|---|---|---|---|
| Protein (%)          | 3.9| 5.9| 7.8| 17.6| 19.6| 45.1 |

Table I

Distribution of the number of agreements when conducting the test of differential abundance of T1-T0, T2-T0, T3-T0, T4-T0, T5-T0 for SRM and SWATH-MS, respectively. 51 regulated proteins are in our discussion.
corresponding RNA profiles, which may be mainly because of a slight delay observed for the protein response compared with the RNA response. Protein and transcript profiles for all genes can be found in the supplemental Data S5.

**DISCUSSION**

Reliable and precise quantitative data sets can be acquired through the consistent recording of the MS signals of a peptide across all the desired conditions or perturbations in a biological system. Currently, two MS strategies are fulfilling such criteria, namely S/MRM and “data independent acquisition” (DIA). In S/MRM, the instrument acts as a mass filter and selectively target only the precursor and fragment ions of interest. The resulting chromatographic peak groups are then analyzed by using the information (i.e. fragmentation pattern or chromatographic properties) available in MS/MS spectral libraries. In the case of DIA-MS, MS/MS spectra are triggered within a user defined ion window independently of precursor intensities and then analyzed with traditional database search algorithms.

Recently we developed the SWATH-MS technology that combines the DIA acquisition approach with targeted data analysis of S/MRM (35) and could thus demonstrate on a set of 60 peptides higher sensitivity than analyzing the data with regular database searching. In the present study, we demonstrated the multiplexing capabilities of SWATH-MS for large scale quantitative proteomics studies. We were able to confidently detect around 2500 proteins spanning the dynamic range of protein expression in yeast in a 3-hours single sample injection. In contrast, S/MRM would have required 48 h of

---

**Fig. 6. Comparison between proteomics and transcripts data sets.** Heat maps for each transcript (top raw) and the corresponding protein log₂-fold changes (bottom raw) between 30 min (T2), 60 min (T3), 90 min (T4), 120 min (T5), and the initial time point 0 min (T0) together with the distribution of Pearson correlation coefficients calculated for all detected protein-transcript pairs in A, Glycolysis-gluconeogenesis pathway, B, Pentose phosphate pathway, C, Glycine, serine and threonine metabolism pathway, and D, Phenylalanine, tyrosine, and tryptophan pathway. Significant up-regulation is shown in red, significant down-regulation in blue, and no statistically significant change in abundance is shown in yellow. Color intensity reflects the corresponding log₂ fold change.
Proteome Quantification by SWATH-MS

instrument time to detect the same number of proteins in one single sample because of its lower multiplexing capacity (i.e. ~100 proteins/run). Thus, the SWATH-MS measurements for the 18 osmotic shock time course samples were completed in ≤ 3 days. Because the data structure of SWATH-MS data is equivalent to that of S/MRM, a similar/extended bioinformatics workflow was implemented. S/MRM, several chromatographic peak groups extracted for the same targeted peptide, evaluated using automated and objective probabilistic scoring model. To confidently identify the targeted peptides by SWATH-MS, we applied Spectronaut, a bioinformatics tool the scoring strategy mProphet was developed for the automatic evaluation of S/MRM signals (15). In addition to the chromatographic (S/MRM-like) scores (i.e. coelution, peak shape similarity, intensity, and correlation of fragmentation pattern between peak groups and assays), SWATH-MS specific scores were added such as mass accuracy and isotopic distribution of fragment ions. A combined score (i.e. discriminated score) was then calculated for each detected peak group and finally used for FDR estimation. Thus, for the automated analysis of the 18 SWATH-MS runs by Spectronaut less ≤ 2 days were required and allowed the detection of ≥2500 yeast proteins with high confidence along the time course study. To further pinpoint the proteins that were significantly changing in abundance between the different time points, we applied MSstats, a statistical modeling framework for protein significance analysis previously designed for S/MRM experiments (32). MSstats uses an intensity-based approach decomposing the MS signals obtained for each protein across isotopic labels, peptides, charge states, transitions, samples, and conditions. It has been shown that MSstats performed better in terms of sensitivity and accuracy than simple statistical methods like t tests. By applying MSstats to our SWATH-MS data sets, we could show that out of the 2589 quantified proteins, 333 and 786 were found significantly up-and down-regulated, respectively, along the complete time course study or with a delay of 20 min. Many of these proteins were shown to be involved in metabolic pathways and were known to be active upon osmotic shock (37). Besides, we could show on a set of 51 yeast proteins that SWATH-MS delivered similar quantitative protein profiles than S/MRM along the 18 osmotic shock time course samples. Furthermore, the reproducibility of SWATH-MS runs was demonstrated with low within-run coefficients of variation (CVs) of ≤ 10% for the majority of the targeted peptides, with only minimal dependence of peptide abundance, results that are similar to those obtained by S/MRM (38). In conclusion, the results demonstrate that SWATH-MS allows for the quantification of large set of proteins across multiple samples with a precision, reproducibility, and accuracy that is comparable to that obtainable by S/MRM. With the adaptation of S/MRM-based software tools for the SWATH-MS targeted data analysis, the technique can be rapidly applied to any type of system biology or biomedical investigation, as it was successfully demonstrated in the last years with S/MRM but with a higher throughput and higher degree of multiplexing.

Accession Codes—The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org/) via the PRIDE partner repository with the data set identifier PXD001010.

The SRM data have been deposited to the PASSEL (11) via the peptide atlas website with the accession number PASS00504.

Acknowledgments—We thank Lucia Espona for data deposition; Stefania Vaga and Jonas Grossmann for helpful comments on the manuscript.

* This work was supported by EU FP7 Unicellsys (Grant No. 201142); ERC Proteomics v3.0 (Grant No. 233226) and ETH Zurich.

[This article contains supplemental Figs. S1 to S3, Data S1 to S5, and Tables S1 to S7.]

§§ To whom correspondence should be addressed: Institute of Molecular Systems Biology, Wolfgang-Pauli-Str. 16, ETH Zürich, CH-8093 Switzerland. Tel.: +41 44 633 3170; Fax: +41 44 633 1051; E-mail: aebersold@imsb.bioc.ethz.ch.

REFERENCES

1. Picotti, P., Bodenmiller, B., Mueller, L. N., Domon, B., and Aebersold, R. (2009) Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. Cell 138, 795–806
2. Costenoble, R., Picotti, P., Reiter, L., Stallmach, R., Heinemann, M., Sauer, U., and Aebersold, R. (2011) Comprehensive quantitative analysis of central carbon and amino-acid metabolism in Saccharomyces cerevisiae under multiple conditions by targeted proteomics. Mol. Syst. Biol. 7, 464
3. Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. Mol. Syst. Biol. 4, 222
4. Addona, T. A., Shi, X., Keshishian, H., Mani, D. R., Burgess, M., Gillette, M. A., Clauser, K. R., Shen, D., Lewis, G. D., Farrell, L. A., Fifer, M. A., Sabatine, M. S., Gersten, R. E., and Carr, S. A. (2011) A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease. Nat. Biotechnol. 29, 635–643
5. Picotti, P., Clement-Ziza, M., Lam, H., Campbell, D. S., Schmidt, A., Deutsch, E. W., Rost, H., Sun, Z., Rinner, O., Reiter, L., Shen, Q., Michaelson, J. J., Frei, A., Alberti, S., Kusebauch, U., Wollscheid, B., Moritz, R. L., Beyer, A., and Aebersold, R. (2013) A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. Nature 494, 266–270
6. Picotti, P., and Aebersold, R. (2012) Selected reaction monitoring-based proteomics: workflows, potential, pitfalls, and future directions. Nat. Methods 9, 555–566
7. Stahl-Zeng, J., Lange, V., Ossola, R., Eckhardt, K., Krek, W., Aebersold, R., and Domon, B. (2007) High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. Mol. Cell. Proteomics 6, 1809–1817
8. Domon, B., and Aebersold, R. (2010) Options and considerations when selecting a quantitative proteomics strategy. Nat. Biotechnol. 28, 710–721
9. Picotti, P., Rinner, O., Stallmach, R., Daulet, F., Farrah, T., Domon, B., Wenschuh, H., and Aebersold, R. (2010) High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. Nat. Methods 7, 43–46
10. Deutsch, E. W., Chambers, M., Neumann, S., Levander, F., Binz, P. A., Shofstah, J., Campbell, D. S., Mendoza, L., Ovelleiro, D., Helsens, K., Martens, L., Aebersold, R., Moritz, R. L., and Brusniak, M. Y. (2012) TraML—a standard format for exchange of selected reaction monitoring transition lists. Mol. Cell. Proteomics 11, R111 015040
11. Farrah, T., Deutsch, E. W., Kreisberg, R., Sun, Z., Campbell, D. S., Men-
Proteome Quantification by SWATH-MS

dozza, L., Kusebauch, U., Brusniak, M. Y., Huttenhain, R., Schiess, R., Selevsek, N., Aebersold, R., and Moritz, R. L. (2012) PASSEL: the PeptideAtlas SRMxperiment library. Proteomics 12, 1170–1175

12. Brusniak, M. Y., Kwok, S. T., Christiansen, M., Campbell, D., Reiter, L., Picotti, P., Kusebauch, U., Ramos, H., Deutsch, E. W., Chen, J., Moritz, R. L., and Aebersold, R. (2011) ATAGS: A computational software tool for high throughput transition optimization and validation for selected reaction monitoring mass spectrometry. BMC Bioinformatics 12, 78

13. MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., and MacCoss, M. J. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26, 966–968

14. Abbatiello, S. E., Mani, D. R., Kesharian, H., and Carr, S. A. (2010) Automated detection of inaccurate and imprecise transitions in peptide quantification by multiple reaction monitoring mass spectrometry. Clin. Chem. 56, 291–305

15. Reiter, L., Rinner, O., Picotti, P., Huttenhain, R., Beck, M., Brusniak, M. Y., Hengartner, M. O., and Aebersold, R. (2011) mProphet: automated data processing and statistical validation for large-scale SRM experiments. Nat. Methods 8, 430–435

16. Sherman, J., McKay, M. J., Ashman, K., and Molloy, M. P. (2009) Unique ion signature mass spectrometry, a deterministic method to assign peptide identity. Mol. Cell. Proteomics, 8, 2051–2062

17. Picotti, P., Lam, H., Deutsch, E. W., Mirzaei, H., Ranish, J., Domon, B., and Aebersold, R. (2008) A database of mass spectrometric assays for the yeast proteome. Nat. Methods 5, 913–914

18. Karlsson, C., Malmstrom, L., Aebersold, R., and Malmstrom, J. (2012) Proteome-wide selected reaction monitoring assays for the human pathogen Streptococcus pyogenes. Nat. Commun. 3, 1301

19. Schubert, O. T., Moursitsen, J., Ludwig, C., Rost, H. L., Rosenberger, G., Arthur, P. K., Claassen, M., Campbell, D. S., Sun, Z., Farrah, T., Genzgenbacher, M., Maiolica, A., Kaufmann, S. H., Moritz, R. L., and Aebersold, R. (2013) The Mtb proteome library: a resource of assays to quantify the complete proteome of Mycobacterium tuberculosis. Cell Host Microbe 13, 602–612

20. Gallien, S., Peterman, S., Kiyonari, R., Souady, J., Duriez, E., Schoen, A., and Domon, B. (2012) Highly multiplexed targeted proteomics using precise control of peptide retention time. Proteomics 12, 1122–1133

21. Kiyonari, R., Schoen, A., Prakash, A., Peterman, S., Zabrouskov, V., Picotti, P., Aebersold, R., Huhmer, A., and Domon, B. (2011) Increased selectivity, analytical precision, and throughput in targeted proteomics. Mol. Cell. Proteomics 10, M110 002931

22. Gallien, S., Duriez, E., Demeure, K., and Domon, B. (2013) Selectivity of LC-MS/MS analysis: implication for proteomics experiments. J. Proteomics 81, 148–158

23. Gallien, S., Duriez, E., Crone, C., Kellmann, M., Moehring, T., and Domon, B. (2012) Targeted proteomic quantification on quadrupoleorbitrap mass spectrometer. Mol. Cell. Proteomics 11, 1709–1723

24. Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S., and Coon, J. J. (2012) Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. Mol. Cell. Proteomics 11, 1475–1488

25. Gillet, L. C., Navarro, P., Tate, S., Roest, H., Selevsek, N., Reiter, L., Bonner, R., and Aebersold, R. (2012) 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 11, 1–17

26. Venable, J. D., Dong, M. Q., Wohlschlegel, J., Dillin, A., and Yates, J. R. (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. Nat. Methods 1, 39–45

27. Lam, H., Deutsch, E. W., Eddes, J. S., Eng, J. K., Stein, S. E., and Aebersold, R. (2008) Building consensus spectral libraries for peptide identification in proteomics. Nat. Methods 5, 873–875

28. Escher, C., Reiter, L., MacLean, B., Ossola, R., Herzog, F., Chilton, J., MacCoss, M. J., and Rinner, O. (2012) Using iRT, a normalized retention time for more targeted measurement of peptides. Proteomics 12, 1111–1121

29. Storey, J. D., and Tibshirani, R. (2003) Statistical significance for genome-wide studies. Proc. Natl. Acad. Sci. U.S.A. 100, 9440–9445

30. Zien, A., Aigner, T., Zimmer, R., and Lengauer, T. (2001) Centralization: a new method for the normalization of gene expression data. Bioinformatics 17, S323–S331

31. Choi, M., Chang, C. Y., Clough, T., Broudy, D., Killeen, T., Maclean, B., and Vitek, O. (2014) MSstats: an R package for statistical analysis of quantitative mass spectrum-based proteomic experiments. Bioinformatics 30, 2524–2526

32. Chang, C. Y., Picotti, P., Huttenhain, R., Heinzelmayer-Schwarz, V., Jo- Vanovic, M., Aebersold, R., and Vitek, O. (2012) Protein significance analysis in selected reaction monitoring (SRM) measurements. Mol. Cell. Proteomics 11, M111 014662

33. Benjamin, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. Roy. Stat. Soc. 57, 289–300

34. Ghaemmaghami, S., Huh, W. K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O’Shea, E. K., and Weissman, J. S. (2003) Global analysis of protein expression in yeast. Nature 425, 737–741

35. Gillet, L. C., Navarro, P., Tate, S., Roest, H., Selevsek, N., Reiter, L., Bonner, R., and Aebersold, R. (2012) 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 11, 0111 016717

36. Rep, M., Krantz, M., Thevelein, J. M., and Hohmann, S. (2000) The transcriptional response of Saccharomyces cerevisiae to osmotic shock. Hot1p and Msn2p/Msn4p are required for the induction of subsets of high osmolarity glycerol pathway-dependent genes. J. Biol. Chem. 275, 8290–8300

37. Lee, M. V., Topper, S. E., Hubler, S. L., Hose, J., Wenger, C. D., Coon, J. J., and Gasch, A. P. (2011) A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. Mol. Syst. Biol. 7, 514

38. Addona, T. A., Abbatiello, S. E., Schilling, B., Skates, S. J., Mani, D. R., Bunk, D. M., Spiegelman, C. H., Zimmerman, L. J., Ham, A. J., Keshishian, H., Hall, S. C., Allen, S., Blackman, R. K., Borchers, C. H., Bucz, C., Cardenas, H. L., Cusack, M. P., Dedder, N. G., Gibson, B. W., Held, J. M., Hilte, T., Jackson, A., Johansen, E. B., Kinsinger, C. R., Li, J., Mesli, M., Neubert, T. A., Niles, R. K., Pulsipher, T. C., Ransohoff, D., Rodriguez, H., Rudnick, P. A., Smith, D., Tabb, D. L., Teget, T. J., Varjathy, A. M., Vega-Montoto, L. J., Wahlander, A., Waldemarson, S., Wang, M., Whiteaker, J. R., Zhao, L., Anderson, N. L., Fisher, S. J., Liebler, D. C., Paulovich, A. G., Regnier, F. E., Tempst, P., and Carr, S. A. (2009) Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. Nat. Biotechnol. 27, 633–641