A Proteomics Approach to the Protein Normalization Problem: Selection of Unvarying Proteins for MS-Based Proteomics and Western Blotting

Jacek R. Wiśniewski*†‡ and Matthias Mann‡

Biochemical Proteomics Group, ‡Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany

Supporting Information

ABSTRACT: Proteomics and other protein-based analysis methods such as Western blotting all face the challenge of discriminating changes in the levels of proteins of interest from inadvertent changes in the amount loaded for analysis. Mass-spectrometry-based proteomics can now estimate the relative and absolute amounts of thousands of proteins across diverse biological systems. We reasoned that this new technology could prove useful for selection of very stably expressed proteins that could serve as better loading controls than those traditionally employed. Large-scale proteomic analyses of SDS lysates of cultured cells and tissues revealed deglycase DJ-1 as the protein with the lowest variability in abundance among different cell types in human, mouse, and amphibian cells. The protein constitutes 0.069 ± 0.017% of total cellular protein and occurs at a specific concentration of 34.6 ± 8.7 pmol/mg of total protein. Since DJ-1 is ubiquitous and therefore easily detectable with several peptides, it can be helpful in normalization of proteomic data sets. In addition, DJ-1 appears to be an advantageous loading control for Western blot that is superior to those used commonly used, allowing comparisons between tissues and cells originating from evolutionarily distant vertebrate species. Notably, this is not possible by the detection and quantitation of housekeeping proteins, which are often used in the Western blot technique. The approach introduced here can be applied to select the most appropriate loading controls for MS-based proteomics or Western blotting in any biological system.

KEYWORDS: DJ-1, PARK7, "Total Protein Approach", quantitative proteomics, proteomic data normalization, Xenopus, Western blot, loading control

INTRODUCTION

Western blotting (WB) is one of the most widely used techniques in the life sciences. Employing specific antibodies, the method allows the identification and relative quantification of proteins separated by polyacrylamide gel electrophoresis proteins.1 Ubiquitously expressed proteins, such as β-actin, β-tubulin, and glyceraldehyde phosphate dehydrogenase (GAPDH), whose levels are thought to remain unchanged, serve as loading controls in WB. However, there are many pitfalls to the technique, which make it semiquantitative at best.2 Several studies have found that normalization approaches using integration of total protein blotted appear to be more accurate.3−5 In particular, the use of sensitive fluorescent labels allows precise normalization and protein quantitation at a higher dynamic range compared with standard "enhanced chemiluminescence"-based detection.6 Nevertheless, these methods are not commonly employed because of the requirement of specific scanning equipment. In any case, the WB technique remains challenging because of the variability of antibodies,6 which mainly reflects their specificity and affinity to antigens. In addition, quantitation with antibodies is usually limited to 1 order of magnitude of antigen concentration.

Mass spectrometry (MS)-based proteomics has become increasingly powerful and now allows the identification and quantification of thousands of proteins in a streamlined and robust manner. In particular, we have recently demonstrated that label-free proteomics using the "total protein approach" allows the determination of concentrations of thousands of proteins.7,8 We hypothesized that comparing quantitative proteomic data across diverse cell types and tissues could pinpoint ubiquitously expressed proteins, which could serve as more accurate normalization controls. We report the identification of deglycase DJ-1, a protein that is ubiquitously expressed and occurs with the lowest titer variability in proteomes across tissues and cells of vertebrate organisms.

Received: May 4, 2016
Published: June 13, 2016
EXPERIMENTAL PROCEDURES

Tissue Isolation and Lysis

Mouse (10-week-old female C57BL/6 mice were bred on a 12 h/12 h light/dark cycle and had free access to standard chow diet) and Xenopus laevis organs and cultured human cells were lysed in a buffer containing 0.1 M Tris-HCl (pH 7.4), 2% SDS, and 0.05 M DTT as described previously.9 The animals were bred at the Institute’s Animal Facility. Whole organs were homogenized with Ultra Turrax T8 homogenizer (IKA Labotechnik). For muscle lysate, the entire gastrocnemius was used. The cells were directly lysed in the buffer. Total protein content in the lysates was determined by means of a tryptophan fluorescence (WF) assay.10

Protein Digestion

Protein lysates were processed using the MED-FASP protocol11 with modifications described recently.12 Briefly, aliquots containing 50 μg of total protein were mixed with 200 μL of 8 M urea in 0.1 M Tris/HCl (pH 8.5) (UA) in the Microcon 30K ultrafiltration unit and then centrifuged at 10000g at 18 °C for 15 min. The eluates were discarded, 100 μL of UA was pipetted into the filters, and the units were centrifuged again. Then 50 μL of 0.05 M iodoacetamide in UA was added to the filters, and the samples were incubated in darkness for 20 min. The filters were washed twice with 100 μL of UA followed by two washes with 100 μL of 0.05 M Tris/HCl (pH 8.5) (digestion buffer, DB) and were digested in 40 μL of the buffer used for washing at 37 °C for 18 h using endoproteinase LysC (MED FASP) at an enzyme to protein ratio of 1:50. The released peptides were collected by centrifugation at 10000g for 10 min followed by two washes with 100 μL of DB. The material remaining on the filters was digested with trypsin using the above conditions, except that the cleavage reaction was performed for only 2 h. The peptide eluates containing 10 μg of total peptide were concentrated to a volume of 4−5 μL and were stored frozen at −20 C until mass spectrometric analysis. Peptides were quantified by the WF assay.10

LC−MS/MS Analysis

Analysis of the peptide mixtures was performed in an LTQ Orbitrap instrument (Thermo Fisher Scientific) as described previously.9 Briefly, aliquots containing 6−8 μg of peptides were injected and separated on a reversed-phase column (20 cm × 75 μm inner diameter) packed with 1.8 μm C18 particles (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) using a 4 h acetonitrile gradient in 0.1% formic acid at a flow rate of 250 nL/min. The LC was coupled to the mass spectrometer via a nanoelectrospray source (Thermo Fisher Scientific). The
The LTQ Orbitrap instrument was operated in data-dependent mode with survey scans acquired at a resolution of 60 000 at m/z 400. For collision-induced dissociation (CID) fragmentation, up to the 10 most abundant precursor ions from the survey scan with charge ≥+2 within the m/z 300–1700 range were selected. The normalized collision energy was 35. The dynamic exclusion parameters were 90 s and 5 ppm. The MS/MS spectra were acquired in the linear ion trap.

Data Analysis

The MS data from mouse and human material were analyzed using the MaxQuant software version 1.5.3.14. Proteins were identified by searching MS and MS/MS data for peptides (54 658 peptides in mouse tissues, 49 726 in human cell lines, and 25 502 in X. laevis) against a decoy version of UniProtKB (May 2013). Carbamidomethylation of cysteines was set as a fixed modification, and N-terminal acetylation and oxidation of methionine were set as variable modifications. Up to two missed cleavages were allowed. The initial allowed mass deviation of the precursor ion was up to 6 ppm, and for the fragment masses it was 0.5 Da. The “match between runs” option enabled transfer of identifications across samples within a time window of 2 min of the aligned retention times. The maximum false peptide discovery rate was specified as 0.01. Protein titers were calculated on the basis of the raw spectral protein intensity (without normalization) and N-terminal acetylation and oxidation of methionine were set as variable modifications. Up to two missed cleavages were allowed. The initial allowed mass deviation of the precursor ion was up to 6 ppm, and for the fragment masses it was 0.5 Da. The “match between runs” option enabled transfer of identifications across samples within a time window of 2 min of the aligned retention times. The maximum false peptide discovery rate was specified as 0.01. Protein titers were calculated on the basis of the raw spectral protein intensity (without normalization) and N-terminal acetylation and oxidation of methionine were set as variable modifications. Up to two missed cleavages were allowed. The initial allowed mass deviation of the precursor ion was up to 6 ppm, and for the fragment masses it was 0.5 Da. The “match between runs” option enabled transfer of identifications across samples within a time window of 2 min of the aligned retention times.

Western Blot Analysis

Mouse organ and cell lysates were electrophoresed, transferred to nitrocellulose, and probed as described previously. Briefly, aliquots of whole-tissue lysates containing 10 μg of total protein were separated on SDS gels and blotted onto nitrocellulose. The completeness of protein transfer was checked by Coomassie staining of the gel after blotting. After protein fixation with 0.5% glutaraldehyde, the blots were stained with Poceau S to test the quality of the protein transfer. The fixation enabled reuse of the same blot several times, thus saving labor and increasing the comparability of the blots generated with various antibodies. Antibodies against DJ-1 (PARK 7) were purchased from Pierce/Thermo-Fisher (PAS-13404, Darmstadt, Germany), ABCAM (ab18257, Cambridge, U.K.), Biozol (LS-C93055, Eching, Germany), and Sigma-Aldrich (SAB4500249, Taufkirchen, Germany). Antibodies against β-actin, β-tubulin, and GAPDH were obtained from Cell Signaling Technology (Leiden, The Netherlands) with catalog numbers 4967L, 2125S, and 2118, respectively. The blots were stripped with a solution containing 2% SDS and 15 mM 2-mercaptoethanol in water at 55 °C for 1 h. Following several washes with excess water, the blots were reused. Primary antibody binding was visualized using secondary antibodies and an enhanced chemiluminescence (ECL) detection system and scanned using an ImageQuant LAS 4000 instrument (GE Healthcare).

RESULTS AND DISCUSSION

Proteomic Screens To Identify Proteins with the Most Constant Expression Levels

SDS lysates of six mouse organs and six human cell lines were analyzed by label-free proteomics and in parallel by WB (Figure 1A–E and Table S1). Proteins exhibited a wide range of variabilities in measured expression values in these diverse systems. β-Actin and β-tubulin were in the middle of this distribution (Figure 1A) and showed considerable variation as expected widely used loading controls to be among the most stably expressed under different conditions. Instead, the protein with the most constant expression levels in mouse tissue and cultured human cells was an ubiquitously expressed enzyme, the deglycase DJ-1 (Figure 1A). This was surprising because one would have expected widely used loading controls to be among the most stably expressed under different conditions. Instead, the protein with the most constant expression levels in mouse tissue and cultured human cells was an ubiquitously expressed enzyme, the deglycase DJ-1 (Figure 1A). This was surprising because one would have expected widely used loading controls to be among the most stably expressed under different conditions. Instead, the protein with the most constant expression levels in mouse tissue and cultured human cells was an ubiquitously expressed enzyme, the deglycase DJ-1 (Figure 1A). This was surprising because one would have expected widely used loading controls to be among the most stably expressed under different conditions. Instead, the protein with the most constant expression levels in mouse tissue and cultured human cells was an ubiquitously expressed enzyme, the deglycase DJ-1 (Figure 1A). This was surprising because one would have expected widely used loading controls to be among the most stably expressed under different conditions. Instead, the protein with the most constant expression levels in mouse tissue and cultured human cells was an ubiquitously expressed enzyme, the deglycase DJ-1 (Figure 1A).
were often based on loading controls, which varied much more than DJ-1 in our data sets.

In contrast to relative analyses, the total protein approach (TPA)\textsuperscript{7,8} allows absolute quantitation based on normalized intensities, thereby enabling the comparison of protein titers among different archival data sets generated using different peptide fractionation and mass spectrometric platforms. We took advantage of this feature to compare DJ-1 concentrations that were assessed in several previously published studies from our laboratory. The levels of DJ-1 turned out to be similarly stable within these data sets and between mouse and human (Table 1 and Figure 1F). The protein constitutes 0.069 ± 0.017% of total cellular protein and occurs at a specific concentration of 34.6 ± 8.7 pmol/mg of total protein. The assessed DJ-1 titer values vary up to 2-fold among different tissues and cell lines. Since DJ-1 is structurally conserved down to prokaryotes,\textsuperscript{18} we asked whether its abundance is also evolutionarily conserved. We prepared and analyzed whole-tissue lysates of liver, heart, skeletal muscle, and oocytes from the African frog \textit{Xenopus}. TPA analysis revealed that the concentrations of DJ-1 in the amphibian tissues and cells are similar to those in mammals (Figure 1G and Figure 2C). This finding suggests that the cellular concentration of DJ-1 is well-maintained in diverse vertebrate species.

**DJ-1 as a Loading Control in Western Blot Analysis**

DJ-1 is a 20 kDa cytoplasmic protein, and therefore, it is easily extractable as well as quantitatively transferable from polyacrylamide gels to blotting membranes. Since small proteins often dissociate from the blotting membrane,\textsuperscript{19} we routinely fix proteins after the blotting procedure using diluted glutaraldehyde over a short time.\textsuperscript{15,20,21} Blots with fixed proteins can be stripped and reused for probing with different antibodies. This procedure allows more accurate comparison of different antigens on a single blot compared with several blots. Care should be taken to avoid “overfixation”, which can affect the antibody–antigen interaction. The expression levels of DJ-1 are 10–100-fold lower than those of β-actin, β-tubulin, and GAPDH, and thus, problems with band saturation\textsuperscript{7} can be avoided. Currently there are several different antibodies against

---

**Figure 2.** Performance of antibodies against DJ-1 (PARK7) protein in Western blot analysis. Antibodies from four different commercial sources (indicated in parentheses) were tested on blots with whole-cell/tissue lysates of (A) cultured human cells, (B) mouse tissues and HeLa cells, and (C) \textit{Xenopus laevis} with mouse liver and human cells. Each of the antibodies used was declared by the manufacturer to stain human and mouse DJ-1, but no information on the reactivities these antibodies against DJ-1 in \textit{X. laevis} was available. (D) Total protein dilution series using whole tissue and cell lysate. DJ-1 protein was visualized using antibodies B (Abcam). (E) Quantification of blots shown in (D). Each data point shows an average of two independent measurements. The dashed lines show linear regression of the data in the range of 1–10 μg total protein.
DJ-1 on the market. We tested the performance of four different products that have been declared as reactive against human and mouse DJ-1 on WB (Figure 2). All four antibodies detected the antigen in human cell line samples in a uniform way (Figure 2A). Two antibodies stained with a small variation observed in the proteomics analyses of human and mouse proteins (Figure 2B). These two antibodies were also useful for detection of Xenopus DJ1 protein, but the intensity of staining was weaker than for mouse and human controls (Figure 2C). These results also highlight one key limitation of the WB technique, namely, the properties of antibodies that often lack desired specificity and affinity to the intended antigen.

Next we analyzed the intensity of antibody staining across a series of sample dilutions (Figure 2D,E). This demonstrated the utility of the anti-DJ1 antibodies (antibodies B) as a quantification reference for sample loads from about 1 to about 20 μg of total protein. In this range, staining with β-actin showed saturation, which disqualifies this protein from use as a loading control. Western blotting is ubiquitously used and also frequently demanded by reviewers for validation of proteomic results. In view of the work presented here, it is our opinion that WB can at best play only a supportive role in the evaluation of proteomic results. WB can provide unambiguous results only when thoroughly tested antibodies are used, which is rarely the case, as has been pointed out recently. In this context, it should be considered whether another proteomic analysis using a different proteomic platform would be a better validation of proteomic data. For instance, this is readily achieved by varying the sample preparation method, digestion strategy, or peptide prefractionation or alternatively by employing different mass spectrometers. It is worth emphasizing that the DJ-1 concentrations presented in this work were achieved by varying the sample preparation method, digestion rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.

CONCLUSIONS

Here we have established that the quantitative and unbiased nature of MS-based proteomics can be used to advantage when selecting proteins with stable expression patterns across diverse biological systems. Specifically, we have identified a protein marker for proteomic normalization in MS-based experiments as well as in Western blot analyses in human, mouse, and amphibian tissues and cells. The approach suggested here can be used more widely to determine the suitability of proteins or sets of proteins as loading and normalization controls for any biological system. As we have shown here, historically selected “household proteins” may not be the most suitable for this task for technical and biological reasons. Conversely, expression data across diverse systems can inform the selection of proteins to serve as biomarkers.

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.6b00403.

Protein content and concentration of identified proteins, including the DJ-1 protein (Table S1) (XLSX)

AUTHOR INFORMATION

Corresponding Author

*Address: Biochemical Proteomics Group, Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany. E-mail: jwisniew@biochem.mpg.de. Tel.: +49 89 8578 2205.

Notes

The authors declare no competing financial interest.

RAW files and the MaxQuant search results have been deposited in the PRIDE repository database with the data set identifier PXD003422.

ACKNOWLEDGMENTS

We thank Mrs. Katharina Zettl for brilliant technical assistance, Prof. Hermann Koepell and Dr. Thorsten Keller (University of Würzburg) for kindly providing the oocyte lysate, and Dr. Marcus Rösche (Max Planck Institute of Biochemistry, Martinsried) for providing tissues from Xenopus laevis. This work was supported by the German Research Foundation (DFG/Gottfried Wilhelm Leibniz Prize).

REFERENCES

(3) Welinder, C.; Ekblad, L. Coomassie staining as loading control in Western blot analysis. J. Proteome Res. 2011, 10 (3), 1416–9.