

IntechOpen

Recent Advances in Proteomics Research

Edited by Sameh Magdeldin



RECENT ADVANCES IN PROTEOMICS RESEARCH

Edited by **Sameh Magdeldin**

Recent Advances in Proteomics Research

<http://dx.doi.org/10.5772/59739>

Edited by Sameh Magdeldin

Contributors

Raffaella Balestrini, Marco Chiapello, Silvia Perotto, Takhar Kasumov, Sheng Pan, Jeroen A.A. A.A. Demmers, Lennart Van Der Wal, Bruno Manadas, Cátia Santa, Sandra Anjo, Vera Mendes

© The Editor(s) and the Author(s) 2015

The moral rights of the and the author(s) have been asserted.

All rights to the book as a whole are reserved by INTECH. The book as a whole (compilation) cannot be reproduced, distributed or used for commercial or non-commercial purposes without INTECH's written permission.

Enquiries concerning the use of the book should be directed to INTECH rights and permissions department (permissions@intechopen.com).

Violations are liable to prosecution under the governing Copyright Law.



Individual chapters of this publication are distributed under the terms of the Creative Commons Attribution 3.0 Unported License which permits commercial use, distribution and reproduction of the individual chapters, provided the original author(s) and source publication are appropriately acknowledged. If so indicated, certain images may not be included under the Creative Commons license. In such cases users will need to obtain permission from the license holder to reproduce the material. More details and guidelines concerning content reuse and adaptation can be found at <http://www.intechopen.com/copyright-policy.html>.

Notice

Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

First published in Croatia, 2015 by INTECH d.o.o.

eBook (PDF) Published by IN TECH d.o.o.

Place and year of publication of eBook (PDF): Rijeka, 2019.

IntechOpen is the global imprint of IN TECH d.o.o.

Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

Additional hard and PDF copies can be obtained from orders@intechopen.com

Recent Advances in Proteomics Research

Edited by Sameh Magdeldin

p. cm.

ISBN 978-953-51-2201-2

eBook (PDF) ISBN 978-953-51-5412-9

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

3,800+

Open access books available

116,000+

International authors and editors

120M+

Downloads

151

Countries delivered to

Our authors are among the
Top 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Meet the editor



Sameh Magdeldin is senior researcher in the Medical School, Niigata University, Japan, and academic associate professor in the Physiology Department, Suez Canal University (SCU), Egypt. He received his M.V.Sc. and Ph.D. in Physiology and his second Ph.D. in Proteomics in July 2012. He has expertise in shotgun proteomics analysis, reversed-phase chromatography and label-free comparative proteomics approaches. Dr. Magdeldin has published outstanding articles on aquaporin research using proteomics technology. He also created the outstanding “All and None” methodology for analyzing large-throughput proteomics data published in a highly respected proteomics journal. He currently serves as a guest editor, associate editor and peer reviewer for several international journals. Dr. Magdeldin received several grants and awards, such as the national encouraging prize, 8th HUPO congress young investigator award, JSN award, grant-in-aid for young scientists and young researcher overseas grant from the Japan Society for the Promotion of Science (JSPS).

Contents

Preface XI

- Chapter 1 **Quantitative Mass Spectrometry-based Proteomics 1**
Lennart van der Wal and Jeroen A. A. Demmers
- Chapter 2 **Proteome Dynamics with Heavy Water – Instrumentations, Data Analysis, and Biological Applications 21**
T. Kasumov, B. Willard, L. Li, R.G. Sadygov and S. Previs
- Chapter 3 **Neuroproteomics – LC-MS Quantitative Approaches 57**
Cátia Santa, Sandra I. Anjo, Vera M. Mendes and Bruno Manadas
- Chapter 4 **Symbiotic Proteomics – State of the Art in Plant–Mycorrhizal Fungi Interactions 113**
Marco Chiapello, Silvia Perotto and Raffaella Balestrini
- Chapter 5 **Targeted Proteomics in Translational and Clinical Studies 135**
Eslam Nouri-Nigjeh, Ru Chen and Sheng Pan

Preface

When scientists and researchers talk about proteins, particularly their function and structure, proteomics should be mentioned. In fact, the term proteomics refers to the entire complement of proteins, including modification. This promising discipline has enabled us to study proteins from a massive and comprehensive point of view, empowering us to accurately understand the molecular basis for disease initiation, progression and efficacious treatment based on the discovery of unique biomarkers.

The book *Recent Advances in Proteomics Research* describes in five sections some of the applications of proteomics. This fine research has been written by leading experts worldwide.

This book is aimed mainly at those interested in proteins and in the field of proteins, particularly biochemists, biologists, pharmacists, advanced graduate students and postgraduate researchers.

Finally, I am grateful to all the experts who participated in this book and shared their valuable experiences. Indeed, without their participation, this book would not have come to light.

Sameh Magdeldin, M.V.Sc., Ph.D. (Physiology), Ph.D. (Proteomics)

Senior Researcher and Proteomics Team Leader,

Medical School, Niigata University, Japan

Associate Professor, Physiology Department,

Suez Canal University, Egypt

Quantitative Mass Spectrometry-based Proteomics

Lennart van der Wal and Jeroen A. A. Demmers

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/61756>

Abstract

Mass spectrometry-based proteomics, the large-scale analysis of proteins by mass spectrometry, has emerged as a powerful technology over the past decade and has become an indispensable tool in many biomedical laboratories. Many strategies for differential proteomics have been developed in recent years, which involve either the incorporation of heavy stable isotopes or are based on label-free comparisons and their statistical assessment, and each of these has specific strengths and limitations. This chapter gives an overview of the current state-of-the-art in quantitative or differential proteomics and will be illustrated by several examples.

Keywords: Mass spectrometry, quantitation, SILAC, heavy isotope labelling, chemical tagging, ^{18}O labelling

1. Introduction

Analysis of the proteome using mass spectrometry has proven to be an indispensable tool in biomedical research over the past 15 years or so. Originally, because of technical limitations, only qualitative measurements were performed for the identification of proteins in a sample. However, the need to put a quantitative label on proteomics analyses became evident rapidly. For this reason, several different technologies were developed for their use, in combination with mass spectrometry, to supply researchers with more quantitative data to investigate, e.g. the dynamics of a particular proteome. In this chapter, a brief overview of quantitative approaches in mass spectrometry-based proteomics will be given. Current protocols for quantitative analysis and software solutions for data analysis will be discussed and examples from the field (including our own laboratory) will be given to illustrate the power of these methods.

2. Mass spectrometry-based proteomics

Before the application of mass spectrometry, protein analysis was mostly based on the purification of single proteins or protein complexes, followed by the performance of experiments on these purified proteins or complexes. Usually, such biochemical experiments are quite laborious and mostly reliant on the extent to which the protein can be purified. Although mass spectrometry as a technique to study small molecules dates back to the beginning of the 20th century, the use of mass spectrometry in peptide and protein analysis is more recent. The development of both matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) in the 1980s was key to this development, as these techniques allowed the ionization of biomolecules such as peptides, proteins and nucleotides, which made their detection by mass spectrometry possible. The 2002 Nobel Prize in Chemistry was awarded to John B. Fenn and Koichi Tanaka for the development of these ionisation techniques [1]. With the possibility to analyse biomolecules, in particular peptides and proteins, using mass spectrometry, the key step towards proteomics was made. Equally important was the advent of the genomic age, supplying the databases which are instrumental for the analysis and identification of proteins, as well as the technical advances of both mass spectrometers and the (bio)informatic infrastructures that are essential for large data handling.

Mass spectrometry in itself is merely a qualitative analytical technique. The biochemical and biophysical properties of proteins and peptides are quite variable, which leads to large differences in properties such as 'sprayability' and, thus, in resulting ion intensities between different peptides, even though these may be present in equimolar amounts in the sample. In order for mass spectrometry to be useful not only for the qualitative analyses but also for the quantitative analysis, these caveats and problems need to be addressed and solved. Concerning the different types of mass spectrometers, there are several physical principles to choose from. While it goes beyond the scope of this chapter to discuss all of these in details, it is quite useful to be aware of the different possibilities available as these may influence the performance of the quantitative analysis. The type of mass spectrometers that are most widely used in proteomics are (1) time-of-flight (ToF), (2) quadrupole, (3) (Paul) ion trap, (4) FTICR, or (5) orbitrap [2,3]. In ToF analysis, the velocity of an ion is measured in order to determine the size of the particle. The quadrupole analyses the movement of an ion through an electric field, while the Paul ion trap is a type of quadrupole that uses static direct current and radio frequency oscillating electric fields to trap ions. In an FTICR mass spectrometer, ions are trapped in a strong magnetic field and the periodic movement of the ions is translated back to m/z ratios. In an orbitrap, ions are trapped in an orbital motion around a spindle, while the image current from the trapped ions is detected and converted to a high-resolution mass spectrum using Fourier transformation (for a review on orbitrap mass spectrometry, see [3]). Current mass spectrometers are usually hybrid instruments, which combine two or more of the above mentioned principles for the analysis of peptides and proteins.

3. Overview of quantitation methods

The first (semi-)quantitative approach to proteomics was achieved using '2D difference in gel electrophoresis' (2D-DIGE). With this technique, proteins are separated according to size and charge, and it includes the incorporation of fluorescent labels (CyDye) to allow the comparison of two conditions versus an internal standard [4,5]. After separation of the proteins, the spots are analysed using specialised software, measuring the relative fluorescence intensities. Spots that appear to be differentially regulated can then be excised from the gel and identified using mass spectrometry [6]. The usage of an internal standard, usually a mix of the two measured conditions, supplies this method with quantitative properties. However, the limitations of 2D-DIGE, and of 2D gel electrophoresis for complex samples in particular, have led to decreased usage of the technique. Because of limitations in the number of samples to be compared, in studying membrane-bound proteins, as well as in the relative low proteome coverage, alternative technologies have now superseded the use of 2D-DIGE as a quantitative proteomics method. Techniques currently used for quantitation are summarized in Figure 1 and discussed further below.

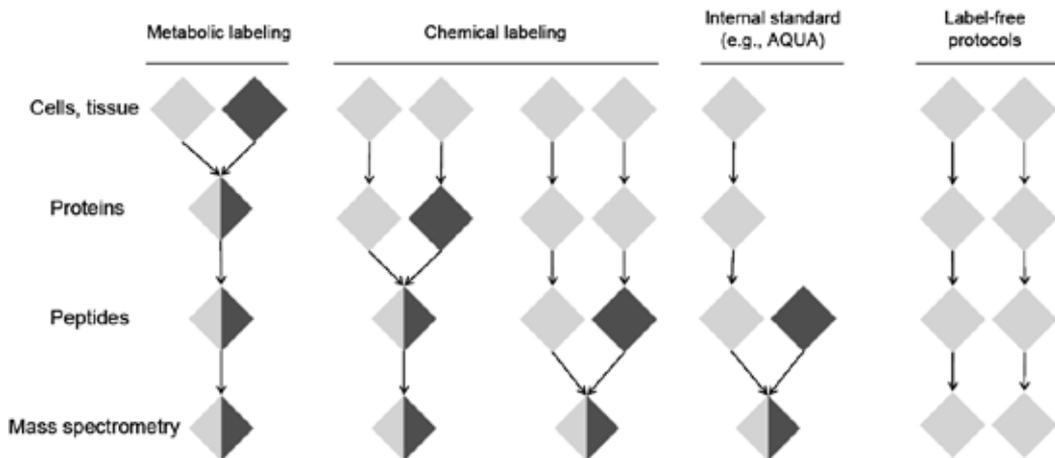


Figure 1. Overview of the stage in which incorporation of the stable isotope labels occurs using different labelling methods in quantitative proteomics. The colour of the diamonds represents the two proteins samples which are differentially labelled and compared in the workflow. (Figure adapted from [7]).

Nowadays, the techniques most frequently used to quantify proteins using mass spectrometry involve labelling proteins with isotopically labelled tags, which can be distinguished in the mass spectrometer because they differ in mass. Differential mass tags result in a (usually only small) mass difference between the 'light' and 'heavy' sample, while proteins and/or peptide properties such as the retention time on a chromatography column are not affected. This allows for the simultaneous analysis of the tagged proteins in a single mass spectrum or LC-MS run. Several methods based on the addition of labelled tags are used in modern proteomics, each with their strong and weak points. Furthermore, with the development of more sensitive and

faster mass spectrometers, methods that allow quantitation of proteins in a label-free manner have been developed, including spectral counting and the comparison of ion intensities. These techniques have the distinct advantage of requiring no (chemical) labelling of the sample, but the trade-off is the lower accuracy of the quantitation. All of these techniques will be described in this chapter, including several examples of how they are used to answer biomedical questions currently posed in the field.

4. Metabolic labelling

The use of amino acids with either of light or heavy stable N and/or C isotopes in growth medium is an approach that was introduced by the Mann lab [8]. Because the labelling takes place at the very beginning of the proteomics workflow, samples can be mixed at the earliest possible time point. Consequently, the occurrence of systematic errors that may be introduced during sample handling is reduced [8,9]. Although this method has shown to be a powerful way to perform quantitation in proteomics in many different applications, there are also several disadvantages to using metabolic labelling, most importantly the inability for application in human tissue samples. Because the samples need to be metabolically active in order to incorporate the label, this automatically precludes, e.g. blood and biopsy samples. This makes it impossible to use metabolic labelling in a diagnostic setting. Furthermore, for some metabolic labelling approaches, reliable software for data analysis is still lacking. The labelling itself is quite time-consuming, as it takes some time for cell cultures to become completely labelled. Finally, the costs of metabolic labelling approaches may be substantial due to the amount of expensive labelled reagents [10]. Below, several types of metabolic labelling will be discussed in more detail.

4.1. ^{15}N labelling

The use of heavy nitrogen (^{15}N) to label whole model organisms dates back to the 1960s, when it was applied to plants for the first time (see [10] for a review on the matter). In the late 1990s, this strategy took off for other organisms such as *E. coli*, yeast and *Drosophila* [9, 11, 12]. Protein labelling is usually achieved by adding salts containing labelled nitrogen into the medium (in the form of, e.g. NH_4Cl), which will then be metabolised by the organism and finally incorporated into proteins. The advantage of this labelling approach, apart from incorporation at the earliest possible moment, is that labelling of all peptides is guaranteed. However, it also provides increased complexity of the sample, most importantly because the mass difference between the light and the heavy counterpart peptides is basically variable for each peptide pair. This presents a major challenge for suitable bioinformatics solutions for qualitative and quantitative data analysis. ^{15}N labelling has successfully been used in the study of prokaryotes, for instance, the proteome of *S. aureus* was analysed extensively using ^{15}N labelled growth medium [13]. By tying the identified proteins to the completely sequenced genome of *S. aureus*, 80% of all expressed proteins were identified using this approach. A study by Kohlmann and coworkers [14] displayed the power of ^{15}N labelling in shotgun proteomics in analysing *R. eutropha*, a prokaryote that can grow on $^{13}\text{CO}_2$ as the sole carbon source and is

used for the industrial production of stable isotope-labelled biomolecules. *R. eutropha* has the ability to switch to lithoautotrophy, i.e. to switch to a source of reduced minerals to satisfy its energy needs. By studying the proteome of *R. eutropha* under both normal and lithoautotrophic conditions, a large upregulation of specific proteins was observed when the prokaryote had to switch its energy source, including chemotaxis-related proteins [14,15].

4.2. ^{13}C labelling

Another prime candidate for metabolic labelling is ^{13}C , as carbon is a key player in protein chemistry. ^{13}C labelling has been successfully used in the determination of protein turnover rates. For instance, by feeding *E. coli* on ^{13}C -labelled glucose, protein turnover rates using only a single culture could be measured by mass spectrometry [16]. Moreover, the method is applicable to shotgun proteomics, which allows for a broader overview of proteins and their turnover rates. In order to reduce the costs of metabolic labelling, a technique called 'subtle modification of isotope ratio proteomics' (SMIRP) was developed [17]. In SMIRP, an increase of only ~1% in isotope ratio can be used to relatively quantify proteins by calculating the ratio of isotopes and comparing it to the variability occurring in nature.

4.3. SILAC

In cultured cells, the metabolic labelling method of choice is stable isotope labelling using amino acids in culture (SILAC), which uses isotopically labelled amino acids (See Figure 2 for a typical SILAC workflow). In order for the amino acids to be incorporated into proteins, it is necessary to determine whether the studied organism is an auxotroph for said amino acid. If a cell or organism is an auxotroph for an amino acid, it cannot synthesize this amino acid itself and, therefore, the amino acid should be supplied in the food or in the growth medium [8]. Usually in SILAC, labelled lysine and arginine are used, which are particularly useful for proteins that are processed with trypsin. Since trypsin cleaves after lysine and arginine, in principle all peptides except for the C-terminal peptide are labelled. If the cells are auxotroph for the selected amino acids, all proteins in a cells are generally completely labelled after several doublings [8]. Conversely, this means that the cells must be dividing, which precludes the use of this technique on primary tissue samples. A complication that has been described in the literature that could potentially interfere with quantitation of SILAC labelled proteins is the natural occurrence of arginine-to-proline conversion. While lysine and arginine are relatively stable in the cell, it is possible for the cell to produce proline from spare arginine, which can then lead to heavy labelled proline. Obviously, this is undesirable and should be accounted for either experimentally or during data analysis (see e.g. [18]).

Labelling using SILAC can also be used to examine post-translational protein modifications such as phosphorylation and ubiquitination in a quantitative manner. An example of this is a phosphoproteomic study in yeast after the knockout of a kinase that plays a role in growth and division [19]. SILAC can in principle be used for any cultured cell type. A recent study from our lab into hormonal signalling in *Drosophila* combined SILAC mass spectrometry with transcriptome analysis [20]. *Drosophila* Kc cells were stimulated with the key insect hormone ecdysone and both mRNA expression and protein expression were studied during a time

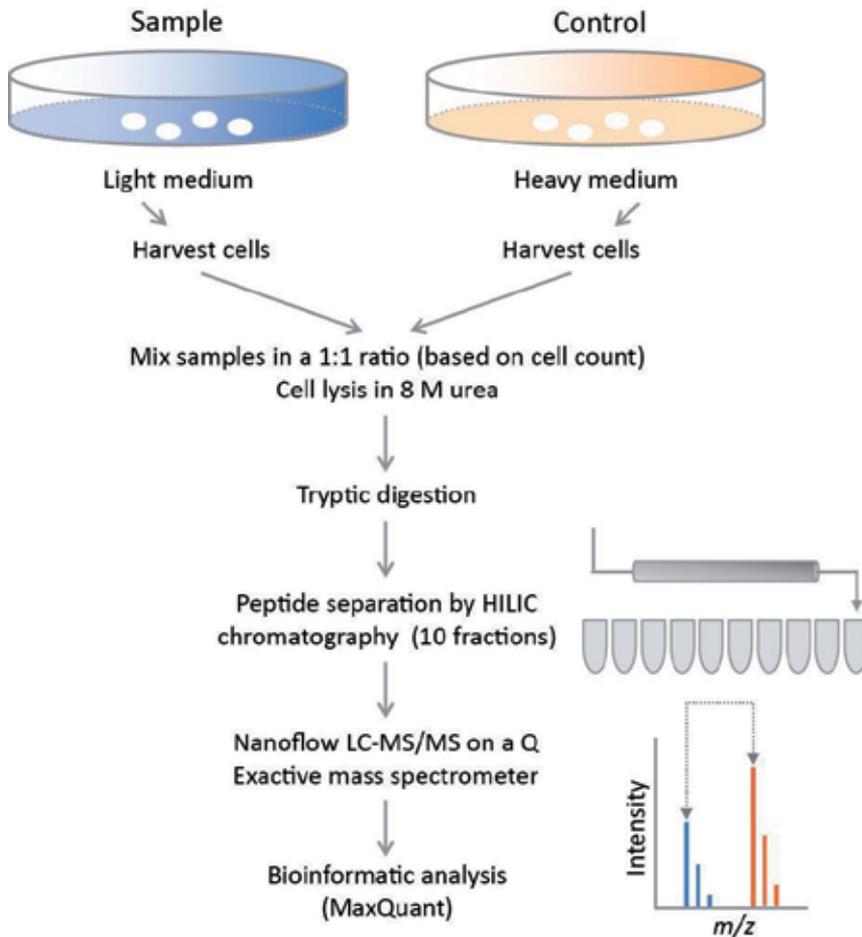


Figure 2. Typical SILAC workflow: cells representing two different biological conditions are grown in either light or heavy medium containing amino acid with stable heavy isotopes. Cells are then harvested and mixed in equal amounts and all sample preparation is performed on the mixed cell populations. In the final mass spectrum, a tryptic peptide will be observed as a peak pair, which represents the two sample conditions. By calculating the peak intensity ratio, the conditions can be compared in a quantitative fashion.

course. The results showed a correlation in the changing levels of mRNA and protein over time, although it became evident that in general there is a time delay between mRNA and protein expression. Not all mRNA–protein pairs showed this delay though, which could be attributed to post-transcriptional regulation events of mRNAs and to variable stability of proteins. Several interesting proteins linked to signalling pathways such as target of rapamycin (TOR) and Notch were identified as being regulated by ecdysone signalling, giving an indication of the scope of the ecdysone system. This study shows the applicability of SILAC in studies where a significant number of proteins are changed, and the correlation between mRNA and protein levels show the quantitative power of SILAC technology, as well as the power of this method to identify signalling networks in cellular systems. In general, there is

also a correlation, albeit weak, between steady-state levels of mRNA and protein (Figure 3). This is mainly true for products that show relatively high expression, which has also been reported in other studies. From this plot, it becomes also clear that for many mRNA products no corresponding protein was identified, illustrating the technical limitations in proteomics that still prevent very low abundant proteins to be detected. In addition, there were several protein products that could not be matched to mRNAs, indicating that, since the intensities of these proteins are generally similar to those with a matched mRNA, this could be attributed to the incomplete annotation of the *Drosophila* database.

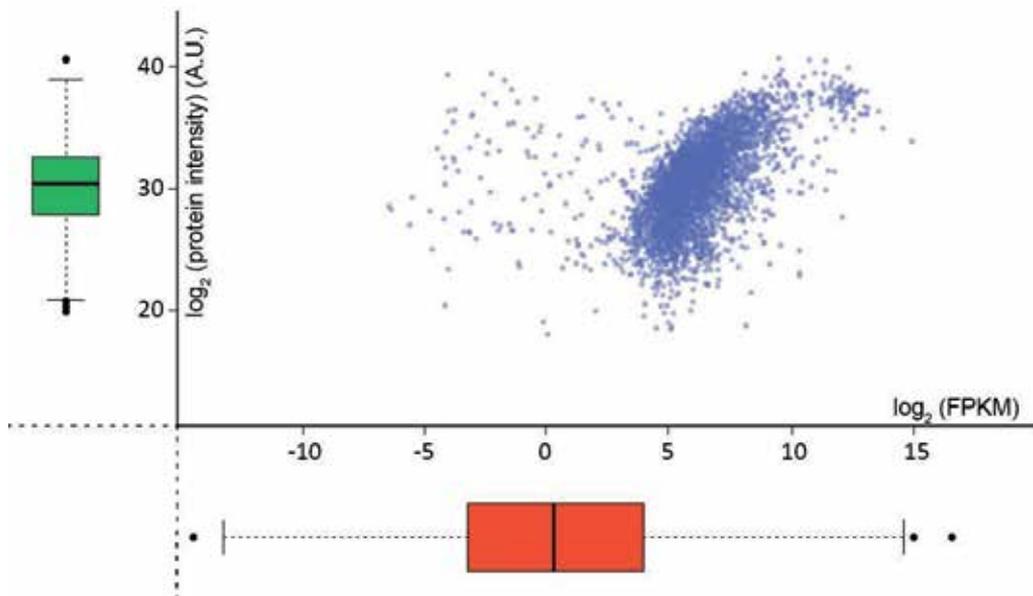


Figure 3. A scatter plot of absolute protein intensities (based on iBAQ values) versus absolute mRNA intensities (based on FPKM values) shows that steady-state levels of protein and mRNA show a weak correlation ($R^2=0.366$). The intensity distribution of proteins for which no corresponding hit in the transcriptome analysis was found is represented by the green box plot. This distribution is very similar to the distribution of overlapping hits (blue data points).

An interesting technological progression in the recent years has been the emergence of fully labelled SILAC organisms, such as fruit flies, mice and rats, which allows for *in vivo* quantitative protein analysis [12,21,22]. This allows scientists to study alterations of protein levels in lab mice with as little variation possible, which in turn makes it possible to study the dynamic proteome in tissue. Currently, the generation of SILAC labelled mice is limited by cost considerations due to the expenses required to raise the mice on a diet of labelled food and this has prevented large-scale usage thus far.

Finally, the so called 'super-SILAC' standard is a pool of multiple cell lines that have been labelled using SILAC, which is then spiked into experimental samples. By spiking all the samples with this standard, quantitation becomes possible without the necessity to label the samples themselves using SILAC. This allows the application of SILAC quantitation in patient

tissue, which can evidently not be labelled using traditional SILAC. It should be noted that it is recommended to have a representative sample for the tissue to be studied in the SILAC standard, which limits the usage of this technique to tissues with a representative cell line. For a more in-depth review on this topic, see [23].

5. Chemical labelling strategies

The use of chemical labelling strategies for relative quantitation in proteomics dates back to the late 1990s [24]. The major advantages of using chemical techniques rather than metabolic labelling are the reduced cost and the higher speed of sample processing and analysis. Where labelling cells with SILAC may take up to several days [8], chemical labelling protocols are usually performed in less than an hour [25]. Chemical labelling can be applied to any protein sample, not just metabolically active samples, and some of the techniques allow for a high number of samples to be analysed simultaneously [26]. However, since chemical labelling is done either at the protein level or at the peptide level and at a relatively late stage in the sample preparation protocol, systematic errors are introduced more readily. Also, labelling at the protein level requires specific proteins such as cysteine or lysine, which makes peptides without these amino acids not quantifiable [10,24].

5.1. Labelling with an Isotope-Coded Affinity Tag (ICAT)

The first chemical labelling technique that was described for quantitative mass spectrometry was the isotope-coded affinity tag (ICAT). In ICAT, a thiol reactive group is used to conjugate the tag to cysteine residues in the protein. Apart from the reactive group, the tag has a linker and a biotin moiety. The linker has either eight hydrogen atoms for the light version or eight deuterium atoms for the heavy version, which are used to distinguish two differentially labelled conditions by the 8 Da shift in the mass spectrum [26]. The biotin moiety of the tag can be used to affinity purify the tagged peptides after trypsinisation. The weakness of ICAT lies in the requirement of cysteine residues to be present in the peptide, which leads to a limitation in the amount of peptides tagged. Furthermore, the presence of deuterium causes a shift in elution times when peptides are fractionated using HPLC, which hampers subsequent data analysis [27]. This elution time shift problem was later solved by introducing ^{13}C instead of D into the linker moiety. ICAT labelling has, for instance, been used to investigate the redox state of proteins in a study to the formation of reactive oxygen species and the way this is dealt with by the cell [28]. The ability to use ICAT in human samples has been exploited in screening cerebrospinal fluid samples of Alzheimer patients to find novel prognostic biomarkers [29].

5.2. ICPL

Labelling using isotope-coded protein labels (ICPL) is based on a similar principle as ICAT. In ICPL, lysine residues in intact proteins are labelled, which are more common than cysteine residues. The mass difference between isotope pairs of the labelled and unlabelled peptides

depends on the amount of labelled lysine residues in the peptide and can be determined fairly simply, which provides strong constraints for database searches [30]. A disadvantage of labelling lysine residues is that modifying the residue side chain makes it impossible for trypsin to cleave at this particular lysine residue. As such, this results in much longer peptides after trypsin digestion, as cleavage will only occur after arginine residues, which may lead to proteolytic peptides that cannot be detected. It is therefore recommended to either use another or an additional protease for protein digestion, or to perform the labelling at the peptide level after proteolytic cleavage. A study on tumour cell senescence in which ICPL was successfully used is a good indicator for the power of quantitative proteomics in general. Here, an effect of tumour cell senescence on several important tumourigenesis proteins such as cMYC and key metabolic enzymes such as ATP synthetases were found [31].

5.3. Isobaric tagging

Tandem mass tags (TMT) and isobaric tag for relative and absolute quantitation (iTRAQ) are based on labelling peptides with isobaric tags. Here, the label is conjugated to the N-termini and lysine residues of peptides, so that in principle every peptide is labelled (Figure 4). The various isobaric tags themselves have different masses, but are balanced by a linker moiety that ensures identical intact masses for all possible combinations of tag plus linker. As a consequence, differentially labelled peptides end up in the same precursor peak in the mass spectrum. Only when this peak is subsequently selected for fragmentation, the linkers will be cleaved first, which leads to the appearance of peaks corresponding to the different tags ('reporter ions') in the low m/z region of the spectrum. The relative peak intensities of the tags are then used for quantitation [26]. Since identical peptides end up in the same peak, the complexity of the MS spectrum is not altered as a result of the labelling procedure. Furthermore, there are commercial kits available with up to 10 different tags, providing the possibility to run and compare 10 samples simultaneously. The most prominent disadvantage of this method is that the tag, just like most other chemical tags, is incorporated at the peptide level. Also, due to the low m/z values of the reporter ions, not all mass spectrometer types are suitable for detection.

Due to the high number of samples that can be measured in one run, its applicability to human tissue samples and the availability of high-resolution mass spectrometers capable of ion detection in the low m/z region, isobaric tagging has quickly become a popular method for the relative quantitation of proteins. For example, by using iTRAQ labelling for quantitation, differentially regulated phosphorylation sites could be detected that were phosphorylated by ATM/ATR, which are highly conserved kinases key in DNA damage repair [32]. iTRAQ labelling has recently been used to compare the proteome profiles of healthy brains to several prion diseased brains such as Creutzfeldt–Jakob disease [33]. This study showed that the changes in protein expression of different prion diseases are markedly similar, while most changes at the protein level were found in the cerebellum. This study provides an excellent example of biomarker research using mass spectrometry and could be a step towards defining biomarkers for different prion diseases, which are otherwise difficult to classify.

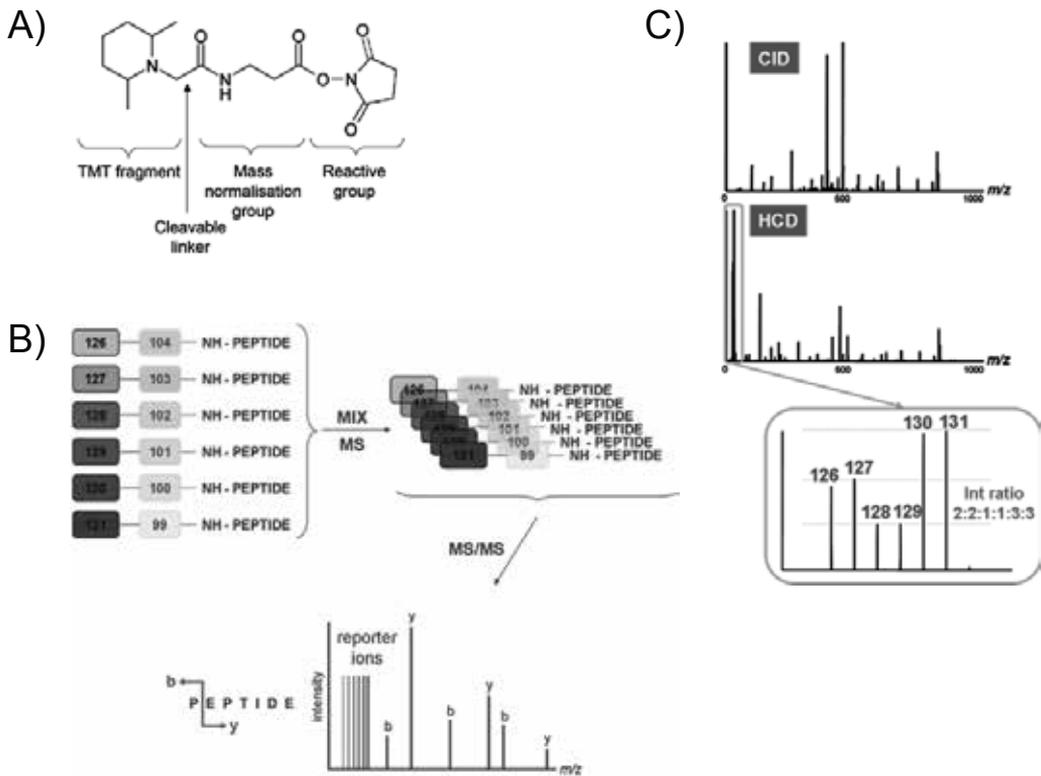


Figure 4. Principle of isobaric tagging. Peptides are tagged with chemical labels that have identical masses due to a delicate balance between individual tag and linker masses. Labeled peptides are then mixed and measured by mass spectrometry. In the mass spectrum, labeled peptides will appear as one peak, but only when these peptides are selected and fragmented by MS/MS, the mass tags will be released from the linker and will show up in the mass spectrum as differential reporter ions in the low m/z region. This allows for the relative abundance determination.

5.4. Dimethyl labelling

A simple method of labelling compounds at the peptide level for relative quantitation is dimethylation. Either light-labelled (with H) or heavy-labelled (with D) dimethyl groups are conjugated to the N-terminus of the peptides and to free lysine residue side chains. The advantages of dimethyl labelling include low cost, high speed and possibilities for automated sample preparation. However, since labelling occurs at the peptide level, variation between runs is still inherent to the process [10,34]. The first incarnation of dimethyl labelling was limited to only two different flavours. However, using isotopic isomers ('isotopomers') of formaldehyde with either only D or a combination of ^{13}C and D, up to three different samples can now be compared in a single run [35] (Figure 5). Although this may still be lower than the amount of different labels that can be achieved using isobaric tagging, it is significantly cheaper. Dimethyl labelling can be used for a variety of quantitative measurements, for instance, after a pulldown or immunoprecipitation enrichment protocol. Using an antibody to probe for phosphopeptides in combination with labelling allows one to quantitatively monitor

phosphorylation events [36]. Another possibility that was recently introduced is using dimethylation to study DNA–protein interactions, e.g. by using an oligonucleotide to pull down the proteins and performing the dimethylation labelling on the proteins enriched for [37]. These widely different applications show the power of dimethylation as a quantitative proteomics tool.

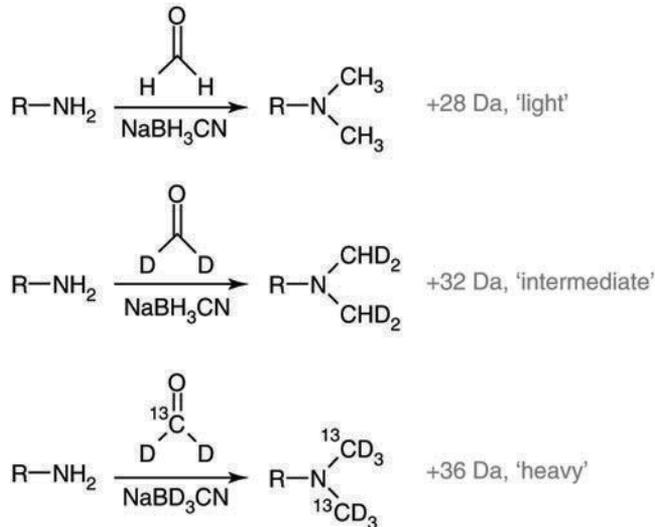


Figure 5. Labelling schemes of triplex stable isotope dimethyl labelling. R = remainder of the peptide. Figure adapted from [35].

5.5. ¹⁸O labelling

Another way to differentially label samples for quantitative purposes is the use of heavy oxygen. This labelling method is different from other labelling protocols in that the label incorporation is achieved during the digestion of proteins into peptides. By performing the digestion in water that contains ¹⁸O instead of ¹⁶O, the carboxyl terminus of every peptide will incorporate two ¹⁸O atoms. This method can be incredibly fast, with reports of labelling being achieved in 15 min [25]. A potential pitfall is that the labelling may be incomplete when not performed in a correct manner, leading to multiple peaks in the MS spectrum and therefore resulting into difficulties in quantitation [25,38]. Our lab has described a protocol to avoid incomplete labelling and to assure full incorporation of the heavy oxygen label [39]. By using immobilized trypsin under acidic conditions, all proteolytic peptides could be fully labelled with heavy oxygen with no traces of back-exchange. The labelling protocol was implemented into a protein–protein interaction analysis pipeline to differentiate between *bona fide* interaction partners of the low-level expressing cell cycle regulator cyclin-dependent kinase 9 (Cdk9) and non-specifically binding or background proteins (Figure 6). Previously known, as well as novel, interaction partners of Cdk9 were characterized, among which most notable are the Mediator complex and several other proteins involved in transcriptional regulation. It was

shown that a differential proteomics approach based on ^{18}O labelling provides a valuable method for high-confidence determination of protein interaction partners and is easily implemented in protein network analysis workflows.

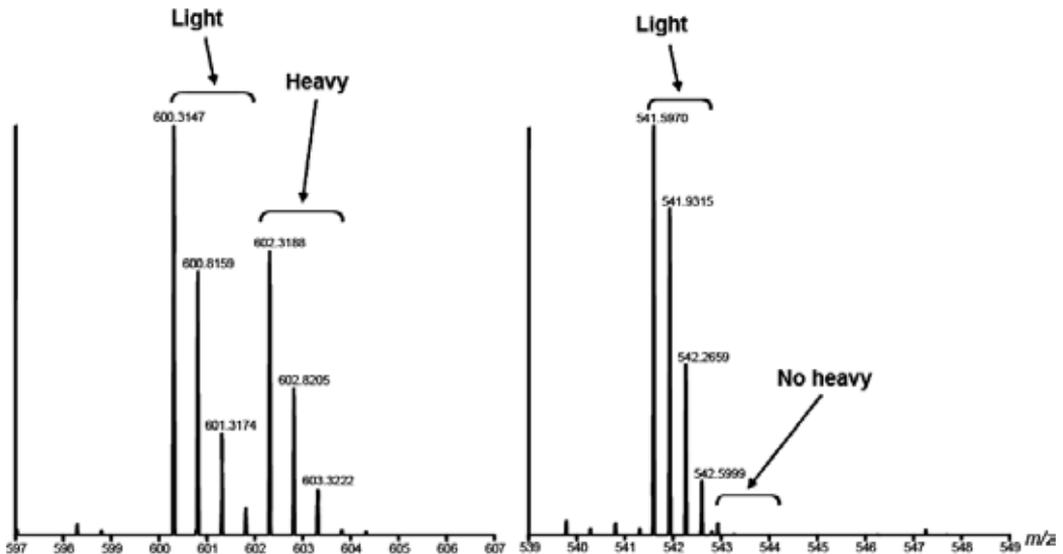


Figure 6. Example MS spectra of tryptic peptides from a 1:1 mixture of a Cdk9 co-IP experiment (light) and a control IP sample (heavy). (A) The tryptic peptide LGTPELSPTER, which originates from the contaminant acetyl-CoA carboxylase shows both the light and heavy forms of the peptide, and as such, is a non-specific protein. (B) The peptide GPPEETGAAVFDHPAK, of cyclin T1, can only be detected in the light sample and is therefore an interactor with Cdk9 [39].

Another method to achieve consistent labelling is to use alternative proteases besides trypsin, e.g. β -lactamase [40], which eliminates the incorporation of two heavy oxygen atoms and limits it to one atom consistently.

6. Absolute Quantitation (AQUA)

All label-based approaches described above are geared towards generating relative quantitative measurements. In many cases though, it would be interesting to measure absolute quantities of proteins instead. In order to gain absolute quantitation results, synthesized peptides or proteins containing heavy isotope labels that correspond to the target peptide or protein of interest can be spiked into the sample at a known concentration, after which the intensities of target and standard can be compared to one another. Obviously, the standard peptide can be modified with one or multiple post-translational modifications if needed [41]. Due to the fact that this spiked standard provides absolute rather than relative quantitation,

this technique has been dubbed absolute quantitation (AQUA). Spike-in components that can be used for AQUA include peptides with stable isotopes incorporated into one or several amino acids [41], a construct in which several peptides are strung together (which has the added advantage of being able to quantify multiple peptides in one run [42]), or an entirely labelled protein to quantify the amount of protein [43]. As with other quantitation techniques, the stage at which the label is incorporated largely determines the extent of the systematic quantitation error that is introduced into the sample. In studying hormonal influence on blood pressure, and more specifically angiotensin II, spiking in the synthesized heavy labelled angiotensin has been used to absolutely quantify protein levels in plasma. As such, it was shown that chronic kidney disease patients had strongly increased levels of angiotensin II [44]. These results show that AQUA can be useful in the field of biomarker research, although it has many more applications, such as in assessing the levels of enzymes in prokaryotes [45].

7. Label-free quantitation

With the development of better and faster mass spectrometers with higher sensitivity and heavier duty cycles, the number of studies that use label-free quantitation (LFQ) methods has increased over the past few years. The obvious advantage of LFQ is that no sample processing other than the standard LC-MS procedures is needed. Furthermore, there is no need for often expensive labelling kits. There are two major approaches employed in label-free quantitation: spectral counting and intensity-based quantitation. Quantitation by spectral counting is based on the observation that peptides that are more abundant will be detected and fragmented more often by the mass spectrometer, and as such the MS/MS count gives information about the abundance of the protein. However, there are several issues that should be taken into account here. In general, larger proteins generate more proteolytic peptides, which increases the chance that multiple peptides for one such protein are detected. Furthermore, in principle every peptide has different physicochemical properties, which influence the ionizability and, therefore, the detectability in the mass spectrometer. To address this, several modifications of spectral counting have been developed, which incorporate mathematical corrections, such as introducing a normalised spectral abundance factor into the equation to account for protein length variability (e.g. emPAI [46]). In intensity-based quantitation, on the other hand, the quantitation is based on the total amount of peptide that is detected in a specific retention time window for which the area under the curve in the chromatogram is accurately determined (extracted ion currents (or XICs) of peptides). LFQ has benefited greatly from recent developments in mass spectrometer hardware as it increases the number of quantifiable features present in a given LC-MS run and allows averaging over more peptides for protein quantitation [47]. In order for the ion intensity quantitation to be reproducible, normalization steps are required as differences in the total amount of protein loaded onto the LC-MS system and instrument variances need to be accounted for. Because of this, powerful software is required and has been developed to perform this type of peptide and protein quantitation (see [48] for an in-depth review). An interesting label-free quantitation technique has been described that combines peptide counting, spectral counting and ion intensities into the so-called normalized

spectral index [49]. Using this method, the variance between multiple LC-MS runs was largely eliminated. This method shows great promise in achieving reproducible label-free quantitation.

8. Software applications for quantitative mass spectrometry

Quantitative proteomic data are typically very complex and the data analysis requires specialized software. The main challenge concerns incomplete data, as even modern advanced mass spectrometers cannot sample and fragment every peptide ion present in a complex sample. As a consequence, only a subset of peptides and proteins present in a sample can be identified. Over the past years, several strategies for mass spectrometry-based quantitative proteomics and corresponding computational methodology for the processing of quantitative data sets have been developed (reviewed in ([50,51])), as different quantitative LC-MS methods require different software solutions for data analysis. Quantitation can be achieved by comparing peak intensities in differential stable isotopic labelling, via spectral counting, or by using the ion current in label-free LC-MS measurements. Many software solutions have been published and can be used freely, with specific instrument compatibility and processing functionality which can deal with these basically different quantitation methods. The researcher has to choose the appropriate software solution for his quantitative proteomic experiments based on the experimental and analytical requirements. Since it goes beyond the scope of this chapter to discuss all of the available software tools separately, we refer the reader to an extensive and up-to-date overview of software solutions including links to websites for downloads at <http://www.ms-utils.org>.

9. Concluding remarks

In summary, all of the mass spectrometry-based quantitation methods have their particular strengths and weaknesses and the researcher has to choose the best method from the multitude of methods that have emerged for the analysis of simple and complex (sub-) proteomes using quantitative mass spectrometry for his specific research. This choice depends on the availability of high-resolution mass spectrometer and LC equipment, the available expertise present in the lab and the financial aspects involved. Quantitative proteomics methods have become mature and can now be applied at a large scale to the study of proteomes and their dynamics. Using the labelling methods described in this chapter, thousands of proteins can be identified and quantified in a single experiment. However, there is still room for improvements to both the experimental strategies for the quantitative analysis of very complex mixtures and of their post-translational modifications and to appropriate bioinformatics and statistical approaches in order to obtain meaningful interpretations of the results. The ultimate goal is to generate quantitative proteomic data at a scale that would allow the comprehensive investigation of a biological phenomenon.

Author details

Lennart van der Wal and Jeroen A. A. Demmers*

*Address all correspondence to: j.demmers@erasmusmc.nl

Erasmus University Medical Center, Rotterdam, The Netherlands

References

- [1] "The Nobel Prize in Chemistry 2002". Nobelprize.org. Nobel Media AB 2014. Web. 3 Feb 2015. <http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2002/>.
- [2] Glish GL, Burinsky DJ. Hybrid mass spectrometers for tandem mass spectrometry. *J Am Soc Mass Spectrom.* 2008;19:161–72.
- [3] Scigelova M, Hornshaw M, Giannakopoulos A, Makarov A. Fourier transform mass spectrometry. *Mol Cell Proteomics* 2011;10:M111.009431.
- [4] Lucitt MB, Price TS, Pizarro A, Wu W, Yocum AK, Seiler C, et al. Analysis of the zebrafish proteome during embryonic development. *Mol Cell Proteomics* 2008;7:981–94.
- [5] Byrne JC, Downes MR, O'Donoghue N, O'Keane C, O'Neill A, Fan Y, et al. 2D-DIGE as a strategy to identify serum markers for the progression of prostate cancer. *J Proteome Res.* 2009;8:942–57.
- [6] Chakravarty B, Gallagher SR, Chakravarty DN. Difference gel electrophoresis (DIGE). *Curr Protoc Mol Biol* 2005;10(23):1–8.
- [7] Ong S-E, Mann M. Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* [Internet]. Nature Publishing Group; 2005 Oct;1(5):252–62. Available from: <http://dx.doi.org/10.1038/nchembio736>
- [8] Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002;1:376–86.
- [9] Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. Accurate quantitation of protein expression and site-specific phosphorylation. *Proc Natl Acad Sci U S A* 1999;96(June): 6591–6.
- [10] Gouw JW, Krijgsveld J, Heck AJR. Quantitative proteomics by metabolic labeling of model organisms. *Mol Cell Proteomics* 2010;9:11–24.

- [11] Paša-Tolic L, Jensen P. High throughput proteome-wide precision measurements of protein expression using mass spectrometry. *J Am Chem Soc* [Internet]. 1999;121:7949–50. Available from: <http://pubs.acs.org/doi/pdf/10.1021/ja991063o>
- [12] Sury MD, Chen J-X, Selbach M. The SILAC fly allows for accurate protein quantification in vivo. *Mol Cell Proteomics* 2010;9:2173–83.
- [13] Becher D, Hempel K, Sievers S, Zühlke D, Pané-Farré J, Otto A, et al. A proteomic view of an important human pathogen: towards the quantification of the entire staphylococcus aureus proteome. *PLoS One* 2009;4(12): e8176.
- [14] Kohlmann Y, Pohlmann A, Otto A, Becher D, Cramm R, Lütte S, et al. Analyses of soluble and membrane proteomes of *Ralstonia eutropha* H16 reveal major changes in the protein complement in adaptation to lithoautotrophy. *J Proteome Res* 2011;10:2767–76.
- [15] Pohlmann A, Fricke WF, Reinecke F, Kusian B, Liesegang H, Cramm R, et al. Genome sequence of the bioplastic-producing “Knallgas” bacterium *Ralstonia eutropha* H16. *Nat Biotechnol* 2006;24(September 2006):1257–62.
- [16] Cargile BJ, Bundy JL, Grunden AM, Stephenson JL. Synthesis / degradation ratio mass spectrometry for measuring relative dynamic protein turnover. *Anal Chem* 2004;76(1):86–97.
- [17] Whitelegge JP, Katz JE, Pihakari KA, Hale R, Aguilera R, Gómez SM, et al. Subtle modification of isotope ratio proteomics; an integrated strategy for expression proteomics. *Phytochemistry* 2004;65:1507–15.
- [18] Van Hoof D, Pinkse MWH, Oostwaard DW-V, Mummery CL, Heck AJR, Krijgsveld J. An experimental correction for arginine-to-proline conversion artifacts in SILAC-based quantitative proteomics. *Nat Methods* 2007;4(9):677–8.
- [19] Kettenbach AN, Deng L, Wu Y, Baldissard S, Adamo ME, Gerber SA, et al. Quantitative phosphoproteomics reveals pathways for coordination of cell growth and division by the fission yeast DYRK kinase Pom1. MCP, Press. 2015.
- [20] Sap KA, Bezstarosti K, Dekkers DHW, Van den Hout M, Van IJcken W, Rijkers E, et al. Global quantitative proteomics reveals novel factors in the ecdysone signaling pathway in *Drosophila melanogaster*. *Proteomics* 2014;(00):1–14.
- [21] Krüger M, Moser M, Ussar S, Thievensen I, Lubner CA, Forner F, et al. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell* 2008;134:353–64.
- [22] McClatchy DB, Dong MQ, Wu C, Venable JD, Yates III JR. ¹⁵N metabolic labelling of mammalian tissue with slow protein turnover. *J Proteome Res* 2007;6:2005–10.

- [23] Shenoy A, Geiger T. Super-SILAC: current trends and future perspectives. *Expert Rev Proteomics* [Internet]. 2015;12(1):13–9. Available from: <http://informahealthcare.com/doi/abs/10.1586/14789450.2015.982538>
- [24] Gygi SP, Rist B, Gerber S a, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 1999;17(October):994–9.
- [25] Mirza SP, Greene AS, Olivier M. ¹⁸O labeling over a coffee break: a rapid strategy for quantitative proteomics. *J Proteome Res* 2008;7:3042–8.
- [26] Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 2004;3:1154–69.
- [27] Coombs KM. Quantitative proteomics of complex mixtures. *Expert Rev Proteomics* 2011;8(5):659–77.
- [28] Leichert LI, Gehrke F, Gudiseva H V, Blackwell T, Ilbert M, Walker AK, et al. Quantifying changes in the thiol redox proteome upon oxidative stress in vivo. *Proc Natl Acad Sci U S A* 2008;105:8197–202.
- [29] Zhang J, Goodlett DR, Quinn JF, Peskind E, Kaye J a, Zhou Y, et al. Quantitative proteomics of cerebrospinal fluid from patients with Alzheimer disease. *J Alzheimers Dis* 2005;7:125–33; discussion 173–80.
- [30] Schmidt A, Kellermann J, Lottspeich F. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* 2005;5:4–15.
- [31] Binossek ML, Lechel A, Rudolph KL, Martens UM, Zimmermann S. Quantitative proteomic profiling of tumor cell response to telomere dysfunction using isotope-coded protein labeling (ICPL) reveals interaction network of candidate senescence markers. *J Proteomics* [Internet]. Elsevier B.V.; 2013;91:515–35. Available from: <http://dx.doi.org/10.1016/j.jprot.2013.08.007>
- [32] Roitinger E, Hofer M, Köcher T, Pichler P, Novatchkova M, Yang J, et al. Quantitative phosphoproteomics of the ATM and ATR dependent DNA damage response in *Arabidopsis thaliana*. *Mol Cell Proteomics* 2015;Preliminai.
- [33] Shi Q, Chen L-N, Zhang B-Y, Xiao K, Zhou W, Chen C, et al. Proteomics analyses for the global proteins in the brain tissues of different human prion diseases. *Mol Cell Proteomics* 2015;In press.
- [34] Hsu JL, Huang SY, Chow NH, Chen SH. Stable-isotope dimethyl labeling for quantitative proteomics. *Anal Chem* 2003;75(24):6843–52.
- [35] Boersema PJ, Aye TT, Van Veen TAB, Heck AJR, Mohammed S. Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates. *Proteomics* 2008;8:4624–32.

- [36] Giansanti P, Stokes MP, Silva JC, Scholten A, Heck AJR. Interrogating cAMP-dependent kinase signaling in Jurkat T cells via a protein kinase A targeted immune-precipitation phosphoproteomics approach. *Mol Cell Proteomics* [Internet] 2013;12(18):3350–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23882029>
- [37] Hubner NC, Nguyen LN, Hornig NC, Stunnenberg HG. A quantitative proteomics tool to identify DNA–protein interactions in primary cells or blood. *J Proteome Res* 2015;14:1315–29.
- [38] Zhao Y, Jia W, Sun W, Jin W, Guo L, Wei J, et al. Combination of improved 18O incorporation and multiple reaction monitoring: a universal strategy for absolute quantitative verification of serum candidate biomarkers of liver cancer. *J Proteome Res* 2010;9:3319–27.
- [39] Bezstarosti K, Ghamari A, Grosveld FG, Demmers JAA. Differential proteomics based on 18O labeling to determine the cyclin dependent kinase 9 interactome. *J Proteome Res* 2010;9:4464–75.
- [40] Wang M, Shen Y, Turko I V, Nelson DC, Li S. Determining carbapenemase activity with 18O labeling and targeted mass spectrometry. *Anal Chem* 2013;85:11014–9.
- [41] Gerber S a, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A* 2003;100:6940–5.
- [42] Rivers J, Simpson DM, Robertson DHL, Gaskell SJ, Beynon RJ. Absolute multiplexed quantitative analysis of protein expression during muscle development using QconCAT. *Mol Cell Proteomics* 2007;6:1416–27.
- [43] Brun V, Dupuis A, Adrait A, Marcellin M, Thomas D, Court M, et al. Isotope-labeled protein standards: toward absolute quantitative proteomics. *Mol Cell Proteomics* 2007;6:2139–49.
- [44] Schulz A, Jankowski J, Zidek W, Jankowski V. Absolute quantification of endogenous angiotensin II levels in human plasma using ESI-LC-MS/MS. *Clin Proteomics* [Internet]. 2014;11(1):37. Available from: <http://www.clinicalproteomicsjournal.com/content/11/1/37>
- [45] Voges R, Corsten S, Wiechert W, Noack S. Absolute quantification of *Corynebacterium glutamicum* glycolytic and anaplerotic enzymes by QconCAT. *J Proteomics* [Internet]. Elsevier B.V.; 2015;113:366–77. Available from: <http://dx.doi.org/10.1016/j.jprot.2014.10.008>
- [46] Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, et al. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 2005;4:1265–72.
- [47] Cox J, Hein MY, Lubner CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extrac-

- tion, Termed MaxLFQ. *Mol Cell Proteomics* [Internet]. 2014;13:2513–26. Available from: <http://www.mcponline.org/cgi/doi/10.1074/mcp.M113.031591>
- [48] Sandin M, Teleman J, Malmström J, Levander F. Data processing methods and quality control strategies for label-free LC-MS protein quantification. *Biochim Biophys Acta* [Internet]. Elsevier B.V.; 2014;1844(1):29–41. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23567904>
- [49] Griffin NM, Yu J, Long F, Oh P, Shore S, Li Y, et al. Label-free, normalized quantification of complex mass spectrometry data for proteomics analysis. *Nat Biotech* 2010;28(1):83–9.
- [50] Matthiesen R, Azevedo L, Amorim A, Carvalho AS. Discussion on common data analysis strategies used in MS-based proteomics. *Proteomics* 2011;11:604–19.
- [51] Mueller LN, Brusniak M-Y, Mani DR, Aebersold R. An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *J Proteome Res* 2008;7:51–61.

Proteome Dynamics with Heavy Water – Instrumentations, Data Analysis, and Biological Applications

T. Kasumov, B. Willard, L. Li, R.G. Sadygov and S. Previs

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/61776>

Abstract

The quantitative assessment of the synthesis of individual proteins has been greatly hindered by the lack of a high-throughput nonradioactive method. We recently developed a method that we call “proteome dynamics” and software that enables high-throughput kinetic analyses of peptides on a proteome-wide scale. Previous studies established that oral administration of heavy water ($^2\text{H}_2\text{O}$ or deuterium oxide, D_2O) is safe and well tolerated in humans. Briefly, a loading dose of $^2\text{H}_2\text{O}$, a nonradioactive isotope, is administered in drinking water. $^2\text{H}_2\text{O}$ rapidly labels body water and transfers ^2H from $^2\text{H}_2\text{O}$ to ^2H -labeled amino acids, which incorporates into proteins dependent upon the rate of synthesis of the specific protein. Proteins are analyzed by high-resolution mass spectrometry and protein synthesis is calculated using specialized software. We have established the effectiveness of this method for plasma and mitochondrial proteins. We demonstrated that fasting has a differential effect on the synthesis rates of proteins. We also applied this method to assess the effect of heart failure on the stability of mitochondrial proteins. In this review, we describe the study design, instrumentation, data analysis, and biological application of heavy water-based proteome turnover studies. We summarize this chapter with the challenges in the field and future directions.

Keywords: Heavy water, proteome dynamics, protein synthesis, modeling, isotopomers, mass spectrometry

1. Introduction

Prior to isotope studies, it was believed that the protein pool in the body was in a static state without any dynamic changes [1–3]. The pioneering work of Schoenheimer and his colleagues investigated the metabolic activities of body proteins using amino acid tracers and therein

established the dynamic nature of the protein pool [4, 5]. Subsequent experiments thoroughly studied the protein balance in the body and revealed that the diet only provides 60–80 g of proteins (per day) as a source of amino acids building blocks for protein synthesis, while the human body synthesizes 300–500 g of protein every day [6]. This discrepancy between dietary protein supply and synthesis suggests that the majority of the newly made proteins are synthesized from amino acids which are derived from degradation of preexisting proteins [7]. In addition, the *de novo* synthesis of nonessential amino acids from ammonia and intermediary metabolites derived from the glycolytic pathway, the pentose pathway, and the citric acid cycle also contributes to protein synthesis [8] (Fig. 1). It is now well recognized that protein turnover—synthesis and degradation—is critical for the maintenance of all cellular processes [9].

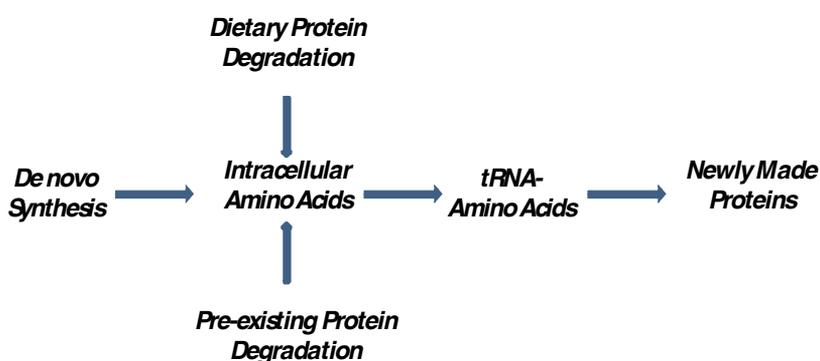


Figure 1. Sources of intracellular amino acids for protein synthesis.

The total protein synthesis rates in whole body and different organs have been measured using radioactive (^{14}C , ^{35}S , and ^3H) and stable (^{13}C , ^2H , and ^{15}N) isotope labeled amino acids in a tissue using the labeling ratio between the precursor amino acids and the protein products [10]. Because of the simplicity, radioactive isotopes dominated early protein turnover studies until gas chromatography-mass spectrometry (GC-MS) became commonly available for stable isotope-based tracer studies [11]. Radioactive amino acids were widely used in pulse-chase experiments that enabled quantification of both protein synthesis and degradation. However, due to safety concerns, radioactive isotopes found limited application in human studies. With the advancement in mass spectrometry instrumentation, the stable isotope-based amino acids found widespread use in clinical research. Similar to radioactive isotopes, two major designs, i.e., flooding dose or primed infusion of the stable isotope labeled amino acids are utilized to study protein turnover in human studies. Multiple studies investigated advantages and disadvantages of both methods [12, 13]. With a different degree of success, both methods enhanced our understanding of total protein dynamics in different tissues and circulation. However, both methods have been associated with several problems related to the assessment of true precursor enrichment and its impact on data interpretation; in addition, experimental design typically requires inpatient tracer administration. As discussed below, this is particularly critical for the short-term labeling protocol that is based on a precursor and product

relationship. The “true precursor” for protein synthesis is the intracellular tRNA-bound amino acids which are usually not accessible, particularly in human studies. Therefore, several extracellular surrogate markers of the “true precursor” have been used for calculation of the kinetic parameters with varying success. Finally, these methods generally require a large amount of expensive tracers, and in the case of stable isotopes, infusion of labeled amino acids elevates amino acid levels and perturbs normal protein metabolism. Until recently, all of these methods were only applicable in studies of total protein kinetics (i.e., consisting of a mixture of proteins) without giving any knowledge about the turnover rates of individual proteins. This shortcoming has particular relevance to health and disease, since it is recognized that proteins differentially respond to stress and the averaging of individual protein fluxes may result in a cancellation of changes in their kinetics. This point can be easily illustrated in the case of acute-phase response (APR) proteins. Due to the distinct dynamics of positive and negative APRs, they are differently affected in conditions associated with inflammation [14] or fasting [15]. Although advancement in methods surrounding protein isolation and sample preparation allowed the analysis of purified (individual) proteins, these methods are in general cumbersome, labor-intensive and, in many cases, it is difficult to purify proteins (specifically low abundant ones) from other contaminants.

Over the last 25 years, the development of novel analytical proteomics methods has provided a major advancement in medical research by allowing investigators to quickly identify and measure the relative amount of a large number of proteins in a plasma or tissue sample. On the other hand, like Western blots, these methods only provide static data on protein levels, and no information on the temporal changes on a given protein. By contrast, coupling of static proteomics with stable isotope-based metabolic labeling approaches enables the study of temporal protein dynamics on a proteome scale. Stable isotope labeled amino acids in cell culture (SILAC) [16] and ^{15}N -labeled algae feeding [17] were successfully applied to study protein turnover in cell culture and then *in vivo* in rodents. Although these methods enable quantification of virtually all identified proteins, the study of protein dynamics *in vivo* in humans is challenging. Since all amino acids have nitrogen, ^{15}N -labeled algae feeding enables tracing all proteins and label amplification in a newly synthesized peptide results in a mass shift relative to unlabeled peptides that simplifies the data interpretation. While ^{15}N -labeled algae provide a valuable tool for *in vitro* cell and *in vivo* rodent experiments, it is not practical in human studies because this would require the consumption of a fully ^{15}N -labeled diet. Although the SILAC method has been used in *in vivo* studies [18], the dietary administration of the SILAC tracers, e.g., [$^{13}\text{C}_6$]-lysine [19], [$^2\text{H}_8$]-valine [18], [$^2\text{H}_3$]-leucine [20], or [$^{13}\text{C}_6$]-arginine [19, 21], limits their application only to fed state which prevents comparisons of proteome dynamics in fed vs. fasted state [22]. In addition, the dietary tracer administration of ^{15}N -labeled algae and SILAC also prevents the modification of the diet as an experimental variable which limits the application of these methods to metabolic diseases that require the assessment of the role of multiple physiological parameters including glucose, insulin, and ketone body on protein synthesis in fasted state. Finally, the dietary administration of tracers in both methods does not allow to readily achieve a steady-state labeling in the precursor pool, a critical assumption made in protein turnover calculations based on precursor and product relationships. Deviation from a steady-state labeling in the precursor pool results in underes-

timination of protein synthesis using these methods and/or leads to complications in the mathematical modeling that is required to interpret the data.

Among all other tracers, $^2\text{H}_2\text{O}$ and H_2^{18}O have been used to study the protein turnover [22, 23]. The ubiquitous presence of H and O atoms in amino acids allowed investigators to consider both $^2\text{H}_2\text{O}$ and H_2^{18}O as unique tracers for the synthesis of virtually all proteins [2, 7, 24]. Since ^{18}O (M+2) isotope adds at least 2Da to each amino acid, the utilization of H_2^{18}O results in a larger mass shift that improves the sensitivity of the assay as compared to $^2\text{H}_2\text{O}$. However, H_2^{18}O is a relatively expensive tracer and is not necessarily affordable for use in humans.

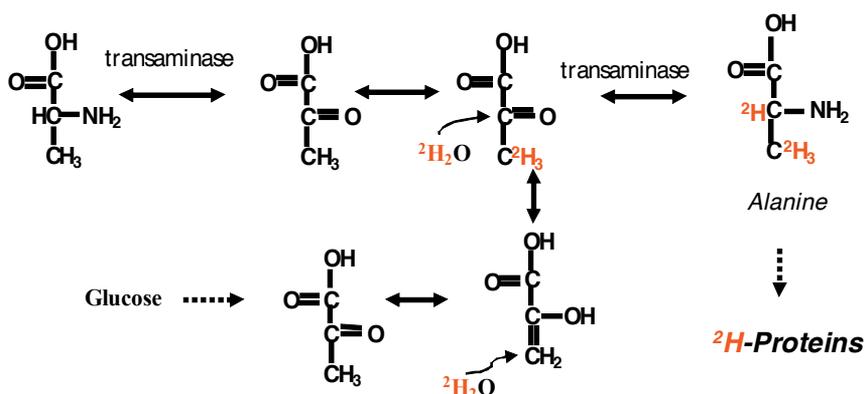


Figure 2. A simplified scheme of ^2H -labeling of alanine and proteins.

By contrast, $^2\text{H}_2\text{O}$ is a low-cost tracer which makes it practical for human application [25]. Similar to H_2^{18}O , $^2\text{H}_2\text{O}$ is safe and it easily equilibrates with total body water (TBW) and $^2\text{H}_2\text{O}$ also rapidly labels all amino acids (e.g., ~10–20 min in rodents and 1 h in humans) [15, 26]. Thus, the quick steady-state labeling of non-exchangeable H atoms in free amino acids after $^2\text{H}_2\text{O}$ administration demonstrates that the rate limiting step of ^2H incorporation into proteins is protein synthesis from amino acids (Fig. 2). Although the use of $^2\text{H}_2\text{O}$ in metabolic studies has a long history [24, 27], recently the $^2\text{H}_2\text{O}$ -metabolic labeling experienced a renaissance, for assessing DNA synthesis [28], gluconeogenesis [29], and lipid turnover [30, 31]. Previously, the $^2\text{H}_2\text{O}$ -metabolic labeling approach has been used by us and others to measure the average synthesis rate of mixed tissue proteins [32–35]. We, and others, recently pioneered $^2\text{H}_2\text{O}$ to study the synthesis rates of individual proteins using advanced mass spectrometry-assisted proteomics *in vivo* [15, 22, 36, 37]. We have continued to refine this approach by combining advanced high-resolution LC-MS (liquid chromatography-mass spectrometry)/MS proteomics with *in vivo* $^2\text{H}_2\text{O}$ -metabolic labeling to create a new method called “proteome dynamics,” which enables quantification of the rate of synthesis of individual proteins.

By giving $^2\text{H}_2\text{O}$ in the drinking water, one can enrich the precursor amino acid pool with ^2H and sustain it indefinitely without affecting the total concentration of precursor amino acids. The rationale is based on the observation that in the presence of $^2\text{H}_2\text{O}$, cells generate ^2H -labeled amino acids via transamination and/or *de novo* synthesis (Fig. 2). All amino acids, including

essential amino acids, can exchange at least one H atom as a consequence of a transamination reaction. However, since the equilibrium of ^2H incorporation from total body water into C–H sites of amino acids is not complete, therefore lower values of deuterium incorporation were observed for the essential amino acids [15]. For the nonessential amino acids, the asymptotic number of exchangeable hydrogen atoms varies depending on their structure and their metabolic origin. For example, *de novo* synthesized alanine and glutamine may incorporate up to four and five ^2H atoms, respectively. Although N–H, O–H, and S–H sites of amino acids also spontaneously exchange H with $^2\text{H}_2\text{O}$, these labile hydrogen atoms back-exchange with H_2O during the extensive sample preparation process. We have demonstrated that there is no back-exchange of C-bound ^2H atoms to ^1H from water after proteins have been synthesized and secreted, and therefore only C–H sites contribute to metabolic labeling during protein synthesis [31]. For the same reason, the *in vivo* $^2\text{H}_2\text{O}$ -metabolic labeling differs from *in vitro* H/D (^2H) exchange methodology that is widely used for protein structure analysis. In contrast to reversible H/D exchange of labile hydrogen atoms in preexisting proteins, the $^2\text{H}_2\text{O}$ -metabolic labeling irreversibly transfers ^2H to the carbon backbone of newly synthesized protein.

The incorporation of multiple copies of ^2H atoms into nonessential amino acids increases tryptic peptides ^2H labeling and improves the assay sensitivity. As a safe, nonradioactive tracer, $^2\text{H}_2\text{O}$ can be administered in the drinking water to free living organisms without interfering with their lifestyle routines. These valuable characteristics of $^2\text{H}_2\text{O}$ -metabolic labeling make it a unique tracer to study the synthesis rates of all proteins in different species, including humans.

2. The study design for heavy water-based proteome turnover studies

Essentially, all tracer-based protein turnover studies rely on establishing precursor (amino acid) and product (protein) relationships. When using a pre-labeled amino acid, one of the major challenges in protein turnover studies is determination of intracellular true precursor enrichment for the kinetic calculations. The true precursor in protein synthesis is an intracellular tRNA-bound amino acid which is in low quantities, and it is not accessible in extracellular fluids [38]. Therefore, the intracellular labeling of free amino acids has been used as the substitute for true precursor enrichment. Although this can be easily done in animal studies, the invasive tissue analysis is not suitable for human studies. In many experiments, only extracellular amino acids are accessible for the precursor enrichment measurements. Since amino acid movement through the cell membrane is a tightly regulated transporter-mediated process, there is an enrichment and concentration gradient of amino acids across the extracellular and intracellular space. To circumvent this issue, several approaches have been proposed to assess true precursor enrichment. For instance, the labeling of an extracellular α -ketoisocaproate (KIC), a metabolite of leucine, was used as a surrogate of intracellular leucine enrichment [39], while intracellular glycine enrichment was assessed based on urinary hippurate metabolite of glycine [40]. In other studies, intracellular amino acids labeling was assessed based on the analysis of protein-bound amino acid in a fast turnover protein like apoB100 [41]. Several studies have demonstrated that different surrogate precursors result in

substantially different kinetic calculations and therefore defining the true precursor and data interpretation are key issues in protein turnover studies [42–44].

In contrast to amino acids, $^2\text{H}_2\text{O}$ freely and rapidly equilibrates with the total body water in all organs and cell compartments and transfers ^2H to intracellular amino acids [15, 36]. This underlying assumption has been validated in multiple studies through analysis of total body water and intracellular amino acids labeling at different time points [15, 26, 45]. For the kinetic calculations, we assume that protein levels do not change during the $^2\text{H}_2\text{O}$ -metabolic labeling study period, and that there is steady-state flux of all proteins. We have validated this assumption through quantification of plasma proteins abundance using synthetic stable isotope-labeled peptides [31]. In addition, other investigators have performed a direct comparison of the heavy water method with a primed infusion of [$^2\text{H}_3$] leucine [45] and/or a flooding dose of [$^2\text{H}_5$]-phenylalanine [46]; these efforts suggest the validity and the reliability of the $^2\text{H}_2\text{O}$ -metabolic labeling approach.

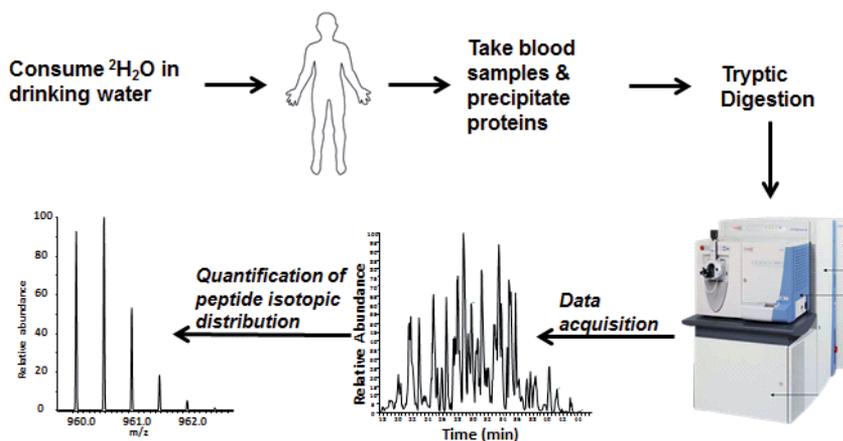


Figure 3. Flow scheme for experimental design and analysis of proteome dynamics with $^2\text{H}_2\text{O}$. After bolus load of $^2\text{H}_2\text{O}$ (0.3 ml/kg body weight), human subjects consume 0.5% in drinking water for 1 week and blood samples are collected at different time points.

These experimental results allow investigators to consider $^2\text{H}_2\text{O}$ as a precursor of ^2H tracer for proteins synthesis. Recently, we developed an algorithm (details discussed below) for calculating the enrichment of intracellular amino acid based on body water enrichment analysis (from accessible body fluids by simple headspace GC-MS analysis) [37]. This overcomes the issue related to true precursor enrichment. Furthermore, oral administration of heavy water after a bolus load easily maintains a steady-state labeling of total body water and amino acids that result in a substantial enrichment of analyzed proteins. When applied to plasma or serum proteins, the experimental design for $^2\text{H}_2\text{O}$ -metabolic labeling is as follows:

- $^2\text{H}_2\text{O}$ is given in a bolus dose followed by low intake in the drinking water to maintain a constant steady-state enrichment of $^2\text{H}_2\text{O}$ in body water (Fig. 3).

- Body water and tissue amino acids are rapidly labeled and attain a steady state without perturbing the normal concentration of amino acids. The ^2H labeling of free amino acids is directly proportional to the number of carbon-bound exchangeable hydrogen atoms and the $^2\text{H}_2\text{O}$ enrichment of body water, and is independent of the rate of protein synthesis.
- Blood is obtained and proteins are isolated at preselected time points.
- Tryptic peptides are analyzed by LC-MS/MS, which identifies several peptides from each protein of interest, both the endogenous mass and heavier peptides that are enriched with ^2H -labeled amino acids. Proteins are identified using online MASCOT or alternative software packages.
- The high-resolution full scan data analysis is performed using in-house-developed software allows quantification of mass isotopomer distribution of peptides.
- The time-course labeling of unique peptides represents the rate of synthesis of a protein that they originate from.

Protein life spans (or half-lives) range from minutes to more than 1 month. Although the heavy water-based metabolic labeling approach may not be suitable for the kinetic studies of very short-lived regulatory proteins such as glucagon, insulin, leptin, and adiponectin, it can capture the kinetics of thousands of proteins with the half-lives that are longer than the distribution and equilibration of $^2\text{H}_2\text{O}$ with amino acids.

This method has major advantages over other stable isotope methods that utilized amino acids pre-labeled with ^2H , ^{13}C , or ^{15}N , namely: (1) it enriches all proteogenic amino acids and thus increases the enrichment of newly synthesized proteins to a far greater extent than that can be achieved by infusion or feeding labeled amino acids or proteins, (2) it can be given to humans by multiple oral doses over the course of a day in drinking water and does not require IV infusion, and (3) it is relatively inexpensive (~\$350/person) compared to traditional amino acid tracers (\$1,000-\$4,000/person).

For the most accurate calculation of protein kinetics, two different short-term and long-term experimental designs with heavy water have been employed.

2.1. The short-term heavy water protocol for protein synthesis

The short-term protocol requires the bolus load of heavy water and the measurement of peptide enrichment during the semilinear increase segment of ^2H -labeling time-course curve [15, 22]. The optimal design for the short-term heavy water protocol requires multiple time points in the early period of protein synthesis, although a single time-point sampling after $^2\text{H}_2\text{O}$ administration is also possible [47]. For the kinetic calculations, we assume that protein levels do not change during the $^2\text{H}_2\text{O}$ -metabolic labeling study period, and that there is a steady-state flux of all proteins. We have validated this assumption through quantification of plasma protein abundance using synthetic stable isotope-labeled peptides. Thus, at a steady state, the rate constant represents both the fractional synthesis rate (FSR) and the fractional catabolic rate (FCR). In this case, the fractional synthesis rate (FSR) of a protein could be

calculated based on the slope of the labeling of the tryptic peptide and precursor amino acid enrichment using the formula [15]:

$$\text{FSR} = \text{slope of } E_{\text{peptide}} / E_{\text{precursor}} \quad (1)$$

where the slope of E_{peptide} is the rate of the increase in ^2H -labeling of peptide during $^2\text{H}_2\text{O}$ administration and $E_{\text{precursor}}$ is the sum of the enrichment of the amino acids constituting the peptide sequence at the steady state. With this design, collection of multiple samples at early hours of the study enables the estimation of turnover rates of proteins with a short half-life, while extending the experiment for several days or weeks allows the estimation of the kinetics for proteins with slower turnover rates. The FSR also can be calculated based on a single time-point sampling after $^2\text{H}_2\text{O}$ administration. However, for accurate estimate of a protein synthesis rate, it is critical to select an appropriate sampling time after $^2\text{H}_2\text{O}$ exposure. Since distinct proteins have a wide range of half-lives, this approach may be satisfactory only for selected sets of individual proteins. In addition, sufficient biological and technical replicates are required to achieve good statistics based on one time-point sampling. Although this approach does not require the correction for the baseline enrichment, the net ^2H labeling can be calculated via subtraction of the total baseline enrichment before heavy water administration: $E_{\text{peptide}}(t) - E_{\text{baseline}}$. Thus, this approach is very simple and straightforward if sampling points are accurately selected based on the half-lives of the analyzed proteins.

The FSR calculation using equation (1) necessitates the analysis of amino acids labeling in specific tissues in order to determine true precursor labeling. As mentioned above, invasive tissue analysis limits the application of this technique mainly to animal studies and complicates its translation to clinical research. In order to circumvent the problems related to the measurement of intracellular amino acid labeling, we developed an algorithm for estimation of the precursor enrichment based on accessible body fluids [37]. The rationale is similar to those used for heavy water-based lipid turnover studies and based on the fact that the ^2H -labeling of body water represents the precursor enrichment. Thus, the precursor amino acid enrichment in equation (1) could be replaced with the total body water enrichment. However, since a product (analyzed peptide) incorporates multiple copies of ^2H , the denominator in equation (1) should take into account the asymptotic number of deuterium (N) incorporated into a peptide:

$$\text{FSR} = \text{slope of product labeling} / (E_{\text{water}} * N) \quad (2)$$

where E_{water} is the steady-state enrichment of total body water and N is the asymptotic number of deuterium atoms incorporated into a peptide, which is calculated using a mathematical algorithm. Since the asymptotic labeling of a peptide is a function of total body water and the number of exchangeable hydrogen atoms [$E_{\text{peptide}} = f(E_{\text{water}}, N)$], when two of the three parameters are known, the third one can be calculated. Thus, N can be calculated using a simple algorithm based on experimental measurement of a peptide's labeling (E_{peptide}) and body water

enrichment (E_{water}). For this purpose, the software models an isotopomer distribution of a peptide based on plasma $^2\text{H}_2\text{O}$ labeling and the different numbers of incorporated ^2H atoms and compares that with the experimentally measured plateau labeling of a peptide. The theoretical isotopic distribution is calculated based on the elemental composition of a peptide sequence and the number of incorporated ^2H atoms. Each calculated isotope distribution is then correlated against the measured isotopic distribution, and the best fit of N is determined based on the minimum of the sum of squares error between the theoretical isotopic distribution simulated by the program and the experimentally measured isotope distribution. Plasma $^2\text{H}_2\text{O}$ labeling is measured using an acetone exchange method, and the isotope distribution of a peptide(s) is determined using high-resolution full scan spectra. Thus, estimation of the FSR in a short-term experiment requires measurements of peptide labeling by LC-MS/MS, water labeling by GC-MS [48], and calculation of the asymptotic number of deuterium atoms incorporated into the peptide (i.e., the N) using a mathematical algorithm [37].

We demonstrated the utility of this approach by quantifying the effect of the nutritional status on the synthesis of albumin and other acute-phase response proteins in rats [15]. With this approach, protein turnover could be determined in a few hours with the total body water (TBW) enrichment of $\sim 2.5\%$. For the plateau labeling of analyzed plasma proteins, we used the data from our 10-day $^2\text{H}_2\text{O}$ experiment. Since the half-life of rat albumin is ~ 1.8 day, the number of incorporated deuterium atoms from 10-day labeling experiment (i.e., 5 half-lives of albumin) is close to the maximum possible ^2H incorporation. This short-term 7-h $^2\text{H}_2\text{O}$ labeling protocol allows measurement of the kinetics of proteins with a wide range of rate constants ($\sim 1\%/h$ for albumin and $\sim 16\%/h$ for ApoB100). Calculated half-lives of different plasma proteins observed using this approach agree with their known biological functions. For example, rapid FSRs were observed for the acute-phase response proteins haptoglobin and fibrinogen. Hemoglobin, albumin, and ApoAI which are involved in oxygen delivery, fatty acid transport, and reverse cholesterol transport, respectively, have the longest half-lives from all the studied plasma proteins. The observed half-lives are also in agreement with the N -end rule, which states that the half-life of a protein is determined by the nature of its N -terminal amino acid residue [49]. ApoB, ApoE, and haptoglobin with destabilizing amino-terminal Phe, Gln, and Asn, respectively, have shorter half-lives, while hemoglobin, albumin, ApoAI, and ceruloplasmin with Ala (albumin and ApoA I) and Gly (ceruloplasmin) have longer half-lives.

A short-term (e.g., 7-h) $^2\text{H}_2\text{O}$ -labeling experiment in rats also allows assessing the effect of nutritional status on the synthesis of plasma proteins, including albumin. Using this approach, it was determined that fasting has a divergent effect on protein synthesis in accordance with the biological function of the protein. In agreement with previous studies using amino acid tracers, it was found that fasting increases the synthesis rate of ApoB100 while reducing the synthesis rates of albumin and fibrinogen. Stimulated synthesis of ApoB100, the principal protein of very-low-density lipoprotein (VLDL), suggests increased secretion of VLDL, a well-known phenomenon in fasting. However, the synthesis rate of albumin, the most abundant plasma protein, was reduced \sim twofold in the fasting state as compared to the fed state. Presumably, this was related to the regulation of albumin synthesis by amino acid substrate availability.

2.2. The long-term heavy water protocol for protein synthesis

Although the short-term experimental design enables one to assess the turnover rates of plasma proteins in several hours, it requires the knowledge of the precursor enrichment. Alternatively, a long-term labeling protocol allows one to measure protein turnover based on modeling of the time-course labeling of analyzed peptides without knowledge of precursor enrichment; note that this is often based on the assumption of a single compartment [15, 22]. The drawback of this design is that it requires the collection of multiple samples for the curve fitting. The FSR in a long-term experiment is calculated by fitting the time-course total labeling of a peptide ($E_{\text{peptide}}(t)$) to an exponential rise curve equation:

$$E_{\text{peptide}}(t) = E_0 * (1 - e^{-kt}) \quad (3)$$

where E_0 is the calculated asymptotical total labeling and k is the rate constant. An accurate calculation of the rate constant requires at least five appropriately timed data points, and greatly depends on the accuracy of the last time-point measurement. Ideally, it is preferred that the last time point corresponds to asymptotical labeling; however, the presence of sufficient early time points will also accurately predict the theoretical E_0 . The half-life of a protein is determined based on the turnover rate constant: $t_{1/2} = \ln 2/k$.

Total labeling of a peptide will be calculated using the formula:

$$\text{MPE} = \text{MPE}M_1x_1 + \text{MPE}M_2x_2 + \dots + \text{MPE}M_ix_i \quad (4)$$

where MPE M_i is the molar percent enrichment of an isotopomer and calculated as

$$M_i = (M_i / \sum(M_0, \dots, M_i)) * 100\% \quad (5)$$

Similar to other tracer experiments, there is a time delay between $^2\text{H}_2\text{O}$ administration and the effective onset of a protein labeling. Such delays most likely reflect a lag between ribosomal protein synthesis and export. Secretory proteins are synthesized on polysomes bound to rough endoplasmic reticulum (ER) and are transported to the lumen of the ER. Before secretion, proteins are transported from the ER to the Golgi apparatus and there is a temporal delay in the transfer from the ER. This delay is especially important in calculation of FSR for relatively fast turnover proteins, such as ApoB100 [50]. It takes ~30 min for newly synthesized ApoB100 to be packaged and released into the circulation; thus, there is a time lag between protein synthesis and appearance in the plasma. To take the delay into account, the expression of $E_{\text{peptide}}(t)$ must be modified for an accurate calculation of the rate constant:

$$E_{\text{peptide}}(t) = E_0 * (1 - e^{-k(t-\tau)}) \quad (6)$$

where τ is the delay time.

In both short-term and long-term heavy water metabolic labeling experiments, the production rates (PR) for a protein is calculated as the product of FSR and the respective pool size of a given protein:

$$\text{PR} \left(\text{g} \times \text{kg}^{-1} \times \text{h}^{-1} \right) = \text{pool size} \times \text{FSR} \quad (7)$$

where the pool size is an absolute content of a protein. In the case of plasma proteins, the pool size is the product of a protein concentration and plasma volume, estimated as 45 ml/kg body weight. Plasma concentration of a protein can be measured using a standard enzyme-linked immunosorbent assay (ELISA) techniques or the isotope dilution method by mass spectrometry [51].

Although the low dose of $^2\text{H}_2\text{O}$ (~0.5% TBW enrichment) is well tolerated in humans, the transient dizziness has been observed in some subjects with the higher bolus aiming to bring TBW enrichment 1.5–2% [52]. To reach this high level of $^2\text{H}_2\text{O}$, according to the original study designs, human subjects ingested 4–5 smaller doses of $^2\text{H}_2\text{O}$ over 4–5 h. Recently, instead of a primed bolus, the gradual increase of $^2\text{H}_2\text{O}$ of TBW enrichment was proposed. According to this protocol, $^2\text{H}_2\text{O}$ enrichment of TBW exponentially increases and reaches the plateau value [25, 53]. The gradual increase of $^2\text{H}_2\text{O}$ in body fluids prevents any side effects related to ^2H -isotope effect. This nonsteady-state labeling of TBW increases the study duration and somewhat complicates the calculation. We applied this approach, i.e., slow increase of $^2\text{H}_2\text{O}$ enrichment of TBW, to study mitochondrial proteome dynamics in a rat model of heart failure [54]. We also constructed a new algorithm to calculate the time-dependent changes in heavy mass isotopomers of newly synthesized peptides. To account for the relatively slower increase in body water labeling, we fit the measured body water enrichment into an exponential curve that yields the body water turnover curve. Then, the modeled continuous body water curve was used for estimation of kinetically relevant body water enrichment required for accurate calculation of synthesis rates. We demonstrated that the calculated turnover rate constants for mitochondrial proteins using this nonsteady-state labeling protocol are very similar to those based on the steady-state bolus labeling of TBW [55]. Thus, this data analysis approach allows accurate quantification of the rate constants to analyze a protein turnover when $^2\text{H}_2\text{O}$ is administered without a priming bolus. This is of particular importance for human studies when it is preferable to increase the TBW enrichment gradually in order to eliminate concerns related to occasional transient dizziness observed with a high bolus dose of heavy water [52]. This also simplifies the study design, since small amounts of heavy water can be consumed outside of the clinical research unit without interference with the daily lifestyle of study subjects.

3. High-resolution mass spectrometry for heavy water-based proteome dynamics studies

Like other stable isotope-based turnover studies, heavy water metabolic labeling requires sensitive and reproducible measurements of isotope labeling of proteins. This necessitates an

accurate quantification of isotopomer distribution of protein-bound amino acids or tryptic peptides unique to a specific protein. Accurate and precise estimates of the isotopic ratio are critical when one aims to quantify subtle changes in protein synthesis due to diseases or an intervention.

Classical studies of protein turnover studies with heavy water utilized GC-MS to measure ^2H -incorporation into protein-bound amino acids after the hydrolysis of protein(s). Because of the low cost of GC-MS instruments, they have traditionally been more accessible than LC-MS instruments. In addition, until recent developments in high-resolution ion detections, many LC-MS instruments had lower accuracy in isotope ratio measurements compared to simple GC-MS instruments. A gas chromatography inlet enables separation of individual amino acids and quadrupole mass analyzer allows accurate measurement of isotope enrichment with $\pm 0.3\%$. In the case of ^2H -labeled compounds, heavy isotopomers enriched with ^2H are slightly shifted and eluted in front of the monoisotopic signal (M_0). This chromatographic fractionation was used for the accurate quantification of low ^2H enrichment in amino acids and other molecules. With this approach, as low as 0.01% ^2H could be accurately measured using a simple quadrupole GC-MS instrument [56]. The majority of early studies with GC-MS were focused on total body or tissue-specific mixed protein turnover without giving knowledge about individual proteins. Later on, this approach was extended to the analysis of purified individual proteins. This requires labor-intensive purification of individual proteins and permits only analysis of one protein at a time. In addition to being time consuming, these protocols suffer from potential contamination associated with protein isolation. The development of isotope-ratio mass spectrometry (IRMS) systems adds more than 100-fold increase in sensitivity for measuring of ^2H enrichment compared with GC-MS [57]. However, similar to GC-MS, IRMS instruments are limited to analysis of protein-bound amino acids.

Recently, a proteomics-based approach was applied to assess the protein turnover in a mixture of proteins [15, 25, 47]. In contrast to static proteomics, the dynamic proteomics method requires accurate quantification of the isotope distribution of peptides that requires high-resolution mass analysis. Studies by Anderson's group evaluated the utility of different type of electron spray ionization (ESI) and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry for the isotope distribution analysis [58, 59]. A Finnigan TSQ 700 or Micromas Quattro II, Thermo-Finnigan linear trap quadrupole (LTQ) linear ion-trap and Applied Biosystems Q-STAR XL hybrid quadrupole-TOF, and Bruker BiFlex III MALDI-TOF were tested [59]. Tandem spectra on the ion-trap instrument were collected in either a zoom scan or profile mode while the quadrupole instrument was operated in the selected ion monitoring (SIM) mode. It has been determined that the signal intensity is the key parameter for accurate characterization of isotope distribution. For instance, the quantification of M_1 with precision better than 5% requires intensities of the base peak $\geq 20,000$ counts in a MALDI-TOF instrument. Based on our experience, similar precision on LTQ linear ion-trap instrument can be achieved with an ion intensity of 10^4 relative to the background signal. It has been noted that MALDI-TOF slightly overestimates M_1 . When the ESI trap and quadrupole instruments were tested for the accuracy and precision of isotope distribution, the ion-trap MS performed better than the SIM quadrupole MS. Interestingly, the quadrupole instrument in SIM mode

had greater precision than MALDI-TOF MS and the accuracy of the quadrupole measurement was improved when it was operated in a profile scan mode.

When applied to protein turnover studies, the high resolution of MALDI-TOF MS allows accurate quantification of ^2H enrichment of tryptic peptides [22]. The traditional proteomics methods coupled with MALDI-TOF MS-assisted isotope distribution analysis greatly advanced protein turnover studies in a mixture of proteins. However, the absence of a chromatographic inlet enables the analysis of only the most abundant proteins and therein compromises a broader application to low abundant proteins. Also, regardless of the peptide abundance, the presence of interfering signals compromises the utility of MALDI-TOF MS for turnover studies in a complex mixture of proteins. To avoid this issue, we also evaluated the utility of the linear ion-trap LTQ instrument for measurement of the fractional synthesis rates of plasma proteins [15]. The high sensitivity of LTQ MS in zoom scan mode allows accurate measurement of the kinetics of proteins and the assessment of changes in plasma proteins synthesis rates related to animal nutritional status. One of the limitations of this instrument is that only a limited number of peptides can be targeted in each duty cycle which is limited by the scan speed. To increase the number of analyzed proteins in a single run, the chromatogram can be divided into several time segments. In this case, several peptides are analyzed in each time segment. Still, this approach allows only the quantification of isotope distribution using 10–15 peptides from 3–6 proteins using 2-h high-performance liquid chromatography (HPLC) gradient and well-designed MS method. Although several mass spectrometer platforms with liquid chromatographic inlets allow accurate quantification of the isotope distribution, only the high-resolution mass spectrometers permit measuring protein turnover on a truly proteome-wide scale.

It has been shown that quadrupole time-of-flight (Q-TOF) MS instruments have a good reproducibility and can accurately measure isotope ratios [60], and it was utilized to study the lipoprotein turnover in mice. However, Q-TOF instruments have relatively lower resolution (~30,000) that limits the isotope ratio accuracy of the isotope ratio analysis. By contrast, the hybrid Fourier transform ion cyclotron (FT-ICR) and Orbitrap mass spectrometers are characterized by unsurpassed resolution (>100,000), high mass accuracy, and sensitivity [61, 62]. The high mass accuracy of these instruments improves identification and characterization of peptides, while high resolution provides additional information for the characterization of the molecular formula based on natural enrichment. Importantly, the high resolution of these instruments, coupled with the increased scan rates, allows accurate isotope distribution analysis that enables measurement of metabolic labeling of all analyzed peptides. Recently, we demonstrated that isotopic ratios between the monoisotopic and heavy isotopic peaks are consistently lower than predicted values and the magnitude of the spectral error in the FT-ICR MS is proportional to the scan duration of the ion clouds (i.e., resolution setting) [63]. It has been shown that the logarithm of the measured isotopic ratio linearly decreases with the acquisition time, and this phenomenon has previously been used to improve the accuracy of the isotopic distribution analyses [64]. However, even at the lowest resolution setting (e.g., 7,500) a significant error (~5%) was observed with FT-ICR MS analysis. Mass accuracy and isotopic ratios may be affected by the static Coulomb repulsion of ions, so fewer ion numbers

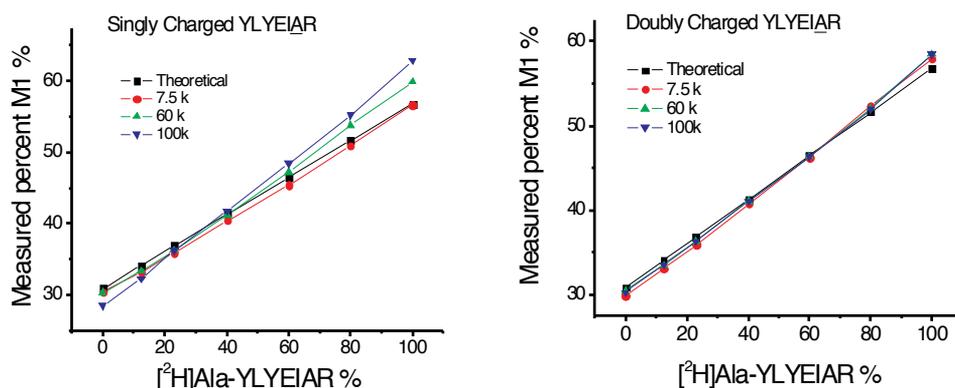


Figure 4. Calibration curves of $[2\text{-}^2\text{H}]$ alanyl-YLYEIQ enrichment (0–97%) measured at different resolutions in Orbitrap Elite instrument. More than 5% error was observed at higher resolution (100 K) for a singly charged peptide ion. Measured enrichment of doubly charged peptide ion is similar to simulated theoretical values.

could reduce the error. However, accurate quantifications of isotopomers require a sufficient number of ions. We found that ion intensities could be accurately measured with ion counts ranging from ~10,000 to 100,000. In this range, the isotopic ratios are approximately the same, while higher ion counts leads to greater error in isotope ratio measurements. To obtain accurate isotopic ratio measurements of peptides, multiple scans with different durations were performed, and the data were extrapolated to the initial moment of the ion rotation. This approach minimizes the absolute isotopic ratio error to within ~1–0.5%. In addition, we found that monitoring the parent ions in the SIM mode (mass interval is 10 Da), and the collision-induced dissociation (CID) fragments in the single reaction monitoring (SRM) mode, improves the specificity of the assay and allows selective identification of peptide and its fragments for isotopomer analysis. Using SIM and SRM experiments in the same acquisition allows reliable simultaneous quantification of the isotopic distribution of both the parent peptide and its fragment ions [37]. An accurate measurement of two consecutive peptide fragments allows one to calculate the labeling of protein-bound amino acids, including alanine, glutamine, and glutamate [37].

Next, we tested the utilities of hybrid Orbitrap Velos and Orbitrap Elite instruments for the ^2H -based metabolic labeling studies [54, 55]. To evaluate the utility of the newer generation Orbitrap Elite instrument for isotope distribution analysis, a calibration curve was constructed by adding an increasing amount of $[2\text{-}^2\text{H}]$ alanyl-YLYEIQ to a constant amount of unlabeled YLYEIQ. Interestingly, similar but lower magnitude error in isotope ratios was observed in both Orbitrap instruments. Consistent with previous studies [65], the Orbitrap also yields higher error at higher resolution setting. The Orbitrap Elite displayed the highest accurate isotope ratio measurements. We found consistent underestimation of the isotope ratio measurement when lower ^2H enrichment was measured, while overestimation was observed at the higher ^2H enrichment (Fig. 4). Interestingly, the error was less when doubly charged ions of the same peptide were analyzed.

The accuracy and precision of molar percent enrichment (MPE) determinations, calculated as the fraction of the total intensity, depends on the number of isotopomers that are used in the calculation. This is due to a lower abundance of heavy isotopomers which introduce more error in MPE calculation. To circumvent this problem, an alternative approach, i.e., M1/M0 ratio, was proposed to assess ^2H -induced changes in an isotope distribution [59]. Although this approach is useful for the modeling of the labeling data in a long-term experiment, it does not allow one to assess the total labeling of an analyzed peptide and asymptotical number of ^2H , the critical step in calculation of the FSR in a short-term experiment.

4. Data analysis in global proteome dynamics studies with heavy water

The high-resolution mass spectrometers allow one to analyze isotopic distributions of virtually all peptides, thus enabling measurement of global proteome dynamics. The bottleneck in these experiments is the data processing. Therefore, high-throughput and robust bioinformatics tools are required to extract the relative isotopomer information from time-course data for the calculation of protein turnover rate constants based on large volume and complex data sets generated by high-resolution mass spectrometers.

Several software solutions have been proposed for the tracer-specific protein turnover studies. SILACtor has been successfully used for protein turnover SILAC experiments in cell culture [66]. SILACtor is useful for *in vitro* proteome dynamic experiments when the heavy precursor is 100% enriched and the protein product labeling gradually increases from 0 to 100%. However, it is not applicable to *in vivo* experiments when only partial labeling is feasible. The Topograph software developed by Macross and colleagues is another software that analyzes data from pre-labeled amino acid experiments, and it is applicable to both *in vitro* and *in vivo* experiments [67].

The heavy water-metabolic labeling approach poses further specific challenges to data analysis software [68]. In contrast to protein turnover studies with pre-labeled amino acids that lead to substantial average mass shifts in newly synthesized proteins, the labeling with heavy water mainly affects the relative isotopomer distribution without a measurable mass shift (maximum $\sim 0.2\text{--}0.4$ Da in an average mass of tryptic peptides). Thus, the partial labeling of proteins with the overlapping isotope profiles of labeled and unlabeled species complicates routine data analysis with $^2\text{H}_2\text{O}$ -labeling approach.

Therefore, the successful implementation of the heavy water labeling experiment, in addition to improvements in mass spectrometry, sample preparation, and fractionation, depends on the efficiency of robust software for data processing. It is also preferable that the software could handle the data generated by different high-resolution instruments. Recently, several high-resolution mass spectrometer platforms have been used to study protein turnover using stable isotopes, including $^2\text{H}_2\text{O}$. Q-TOF mass spectrometer (Agilent) was applied to assess the proteome dynamics in plasma and different tissues [25]. For the data analysis, the authors used MassHunter software package (B0.4) from Agilent (Santa Clara, CA) specially designed for the isotopic distribution analysis of peptides processed in Agilent 6520 Q-TOF mass spec-

trometers. As Agilent's proprietary software, the MassHunter software package is not freely available to the public. Although this software facilitates the analysis of data generated in Agilent Q-TOF MS, for accurate isotopomer profiling, each sample is analyzed twice: during the first injection, MS/MS spectra are collected for protein identification and a second injection was performed for high-resolution full scan acquisition which doubles the instrument time per sample and limits high-throughput analysis. In addition, unlike high-resolution FT mass spectrometers, Q-TOF instruments have relatively lower resolution (~30,000 compared to 120,000 in Orbitrap Elite) that limits the accuracy of isotope ratio-based turnover measurement in this instrument.

Although currently available FT LTQ-ICR and LTQ-Orbitrap hybrid instruments allow both MS/MS scans and full scan analysis in a single acquisition with unsurpassed high resolution, in contrast to an Agilent Q-TOF instrument, they are not supported with software that could automatically extract the data from high-resolution full scan spectra. Thus, specialized software for automated high-resolution data analysis is critically needed. To advance $^2\text{H}_2\text{O}$ -metabolic labeling for *in vivo* studies of protein turnover, the new software must be robust, user friendly, accurate, and capable of producing statistically rigorous results. Recently, Ping and colleagues described a software IsotoQuan/RateQuant-ProTurn [47] for calculation of peptide isotope distribution and protein turnover rates from heavy water labeling experiments. The software has been useful for processing data sets from $^2\text{H}_2\text{O}$ -metabolic experiments. It uses manual validation for peak integration, and fixed exponential decay functions for protein turnover rates calculation.

In the original version, the software used a mass accuracy of 100 ppm and resolution of 15,000, which increases the likelihood of contamination of mass isotopomers by co-eluting signals. To avoid the complexity caused by co-elution, the mass spectrometers were operated at lower resolution (15,000) and mass accuracy (100 ppm) [69] which simply masks the interfering signals due to low resolution. A later version of the software included more stringent filtering parameters: a mass window of 75 ppm is recommended for 30,000 or 60,000 resolution (<http://www.heartproteome.org/proturn/index.html>). However, this software is not freely available to the public, and the raw data from outside investigators could be processed only with the assistance of a web administrator. So far, to the best of our knowledge, no study from outside investigators has been reported using this software.

To aid our heavy water-based proteome dynamics studies, we recently developed an alternative software [55] which is available at the University of Texas Medical Branch (UTMB) website, <https://ispace.utmb.edu/users/rgsadygo/Proteomics/HeavyWater/Version.1.0>. Although this semiautomated software still requires a skilled operator for the data analysis, to the best of our knowledge, this is the only freely available software for quantification of proteome dynamics using heavy water-based metabolic labeling approach. With this software, a routine data analysis workflow for the heavy water labeled samples starts with the peptide/protein identifications from tandem mass spectra using protein sequence databases. Thus, the software reads all peptide IDs from the MASCOT mzIdentML files and confirms each ID based on the stored MS/MS scans at every time point. The initial step is to overlay the chromatographic profiles for each LC-MS run from all time points. Then, the software generates extracted ion

chromatograms for each isotopomer for positively identified peptides from the high-resolution full MS scans within the elution time window of the corresponding MS/MS scan. In addition to peptide selection based on an exact mass and retention time, the software also filters unlabeled peptides at the baseline ($t=0$). For this purpose, theoretical masses are calculated as additional confirmation of a peptide's identity. Only peptides satisfying the modifiable filtering criteria based on exact mass (<10 ppm), peptide score (>35), signal intensity ($>10^4$), and peptides present in at least five time points of ^2H labeling are selected for quantification. The latter criterion is required to obtain sufficient data points for the kinetic modeling of the data. Although this conservative selection of peptides reduces the number of analyzed proteins, it substantially improves the accuracy of the results. The chromatographic profile of a peptide is determined by estimating the signal-to-noise ratio. The software removes peptide IDs that have chromatographic overlaps with other signals and a spectra of low quality (low signal-to-noise ratios or low MASCOT scores) based on the correlation of individual peaks across the elution profile. Also, peptides that cannot be assigned to a unique protein are excluded. All outliers are removed using appropriate statistical methods.

Next, the mass isotopic distributions for all selected peptides are quantified as a function of time. Peaks intensities are extracted from the averaged full scan by searching for an intensity that is maximum within the theoretical mass window (± 10 ppm). We then use separate software to compute FSR, and the values for the same proteins are averaged. Examination of large data sets reveals that even with using these stringent criteria, contaminating signals may result in inaccurate rate constant calculations. This could be related to contributions from minor overlapping unresolved peaks that may not be easily filtered during isotope distribution analysis by the software. Therefore, a second-line filtering step of "contaminated peptides" involves the elimination of outlier peptides based on the coefficient of variation in the protein turnover rate constant relative to the average of the other peptides. Thus, the extracted data from only those peptides that could be modeled with the regression coefficient cutoff of 0.95 for nonlinear curve fitting and coefficient variation less than 30% relative to average of other peptides are selected for final quantification of the rate constants. These stringent selection criteria combined with precise isotope distribution analysis results in accurate quantification of protein synthesis rates.

5. Biological application of heavy water-based proteome turnover studies

Recent technological advancements in bioanalytical instrumentation and their application to systems biology are starting to significantly advance our understanding of integrative physiology. These achievements would not be possible without progresses made in genomics, proteomics, and metabolomics, that is, "omics" technologies that enable comprehensive screening of the genome, proteome, and metabolome [70, 71], respectively. The immense information collected using these "omics" sciences helps to understand the diseases mechanisms and facilitates early diagnosis of the disease, along with implementation and evaluation of personalized therapy [72]. Utilization of genomics in particular enabled the discovery of several genetic diseases. Although multiple protein biomarkers have been identified using

quantitative proteomics, compared to genomics, proteomics is still lagging behind as a clinical test method. This is partly related to the complexity of the human proteome. In addition, profiling of proteins may not be sufficient to understand physiological changes in a living organism, because they have inherent limitations associated with the low sensitivity of static measurements which are the end result of the changes in dynamic flux. In general, stress-induced changes in a biological system first affect the flux of a protein(s) that may lead to more drastic changes in their pool sizes. Only an uncompensated response to stress would result in the nonsteady-state changes in the synthesis or degradation of proteins leading to alterations of their pool sizes. Importantly, the magnitude of changes in flux measured with a small amount of tracer often exceeds the changes in large pool sizes. This is why the kinetic measurements are usually more sensitive than static measurements. In addition, if the stress equally increases or decreases both synthesis and degradation, then the pool size may not change at all. Isotope-based technologies allow investigators to measure changes in flux, and recently, "fluxomics" joined the "omics" sciences. Stable isotope-assisted dynamic metabolomics helped discover previously unknown metabolic pathways [73]. While fluxomics measures large numbers of metabolite turnover, a stable isotope-assisted protein turnover investigates the dynamic genome expression through the temporal changes in a protein flux. Thus, the traditional static proteomics, coupled with a metabolic labeling approach and high-resolution mass spectrometry, is expected to provide a means for simultaneous measurements of proteome dynamics. From the tracer selection point of view, a heavy water-based metabolic labeling approach is of particular interest. For example, H is the ubiquitous element of all biological molecules, and as a universal tracer, $^2\text{H}_2\text{O}$ labels DNA, RNA, proteins, and metabolites and provides the wealth of information in integrated comprehensive "omics." Because of our focus on proteomics, we will mainly highlight the biological application of $^2\text{H}_2\text{O}$ -based proteome dynamics.

Since proteins are indispensable to life activity and involved in multiple structural functions, enzymatic, activities, signal transduction, growth, and repair functions, only minor alterations in a protein homeostasis can lead to genetic and acquired diseases. Mass spectrometry-based protein turnover studies enables the analysis of perturbations in the protein metabolism in different diseases. Recently, $^2\text{H}_2\text{O}$ -based metabolic labeling approach was applied to study proteome dynamics in whole blood, blood cell fractions, plasma, whole tissue samples, and cell organelles. Here we will focus mainly on *in vivo* animal and human studies with $^2\text{H}_2\text{O}$.

5.1. Animal studies

Early studies with the $^2\text{H}_2\text{O}$ -based labeling approach were focused on plasma albumin and mixed tissue proteins synthesis. Using a rat model, it was found that plasma protein synthesis is very sensitive to nutrient availability and ~50% of plasma albumin that was synthesized over a 24-h period was produced within ~5 h after the meal [74]. Furthermore, this study demonstrated that the heavy water approach also permits the analysis of plasma albumin synthesis during metabolic "steady-state" and "nonsteady-state" conditions corresponding to fasted and fed states. Consistent with these results, using a proteomic approach, we demonstrated albumin synthesis in rats was significantly reduced after 22-h fasting [15].

The effects of dietary factors on tissue protein synthesis were investigated in acute fasting (20 h) vs. chronic food restriction (7 days), and feeding (a single meal) conditions in rats. Both acute and chronic fasting significantly reduced mixed tissue protein synthesis in the liver and gastrocnemius muscle, while it did not affect protein synthesis in the left ventricle of the heart [32], indicating that cardiac protein synthesis is preserved in conditions of nutritional perturbations. The follow-up studies demonstrated that diet-induced obesity in mice did not affect the skeletal muscle protein synthesis; however, it did impair the response of muscle protein synthesis to nutrient supply [34].

Understanding the mitochondrial proteome is a new emerging area in proteomics analysis which is largely aimed at targeting over one thousand proteins that are critical in adenosine triphosphate (ATP) synthesis and cell signaling [75]. Mitochondrial dysfunction plays a key role in aging and different diseases associated with oxidative stress and impaired energy metabolism [54, 76]. Therefore, recent attention toward mitochondrial biogenesis [77, 78] and proteome dynamics [55, 69] became an intense area of research in mitochondrial biology. The wide range of concentration of mitochondrial proteins poses a great challenge for comprehensive analysis of the mitochondrial proteome. Nevertheless, several fractionation and enrichment methods have been used to map mitochondrial proteomes. Different labeling approaches have been applied to measure the turnover rates of mitochondrial proteins. For example, [$^2\text{H}_3$]-leucine was used to assess the *in vivo* turnover rates of mitochondrial proteome in the mouse liver and heart [67]. We utilized the $^2\text{H}_2\text{O}$ -based metabolic labeling technique to assess protein kinetics in cardiac, brain, and liver mitochondria. Adult rats were given $^2\text{H}_2\text{O}$ in the drinking water for up to 60 days. Plasma $^2\text{H}_2\text{O}$ and myocardial and hepatic tissue ^2H enrichment of amino acids were stable throughout the experimental protocol [55]. Analysis of mitochondrial protein synthesis in rat liver revealed that the half-lives of proteins range from 2 to 6 days. In the heart, the two spatially distinct subpopulations of cardiac mitochondria, subsarcolemmal (SSM, found along the perimeter of the cell) and interfibrillar (IFM, located between the myofibrils) mitochondria, were analyzed. It is well known that SSM and IFM populations have distinct biochemical functions with IFM having a greater respiratory capacity and resistance to Ca^{2+} -induced stress [79, 80]. Multiple tryptic peptides were identified from each protein in both SSM and IFM, and showed time-dependent increases in heavy mass isotopomers that was consistent within a given protein. In contrast to the liver, cardiac mitochondrial protein synthesis was relatively slow (average half-life of 30 days, or 2.4% newly made per day). Thus, the rate of synthesis of cardiac mitochondrial proteins is approximately sevenfold longer than that of the liver. Analysis of protein synthesis based on protein location within the mitochondrion revealed a shorter half-life for outer membrane proteins than inner matrix proteins in both SSM and IFM. Subunits of mitochondrial electron transport chain (ETC) complexes and proteins with other related functions displayed similar half-lives, suggesting that the differences in mitochondrial proteins turnover could be explained by their sub-complex association. Although the synthesis rates for individual proteins were correlated between IFM and SSM ($R^2=0.84$, $p<0.0001$), values in IFM were 15% less than SSM ($p<0.001$) [55]. The differences in distinct mitochondria populations may have a particular relevance to mitochondrial dysfunction in different diseases, since previous studies found differential effects of aging, diabetes, and heart failure in SSM and IFM. It has been shown that IFM are

more susceptible than SSM to disease-associated damage [81]. In particular, rats with advanced pressure overload-induced heart failure have greater dysfunction in IFM than SSM, suggesting severe impairment in protein synthesis and/or stability in IFM than in SSM.

Previously, it has been shown that the turnover rate of the total mixed mitochondrial brain proteins are slower than those of cardiac proteins [82]. When we compared the turnover rates of individual proteins in the rat brain and heart mitochondria, we found that in the brain, the turnover rate of superoxide dismutase is indeed slower than in the heart (Fig. 5). By contrast, ATP synthase F1 β has a much faster turnover rate in the brain than the heart, suggesting that the kinetics of individual proteins in each organ is determined by their functions. Consistent with previous studies [82], we found that similar to the heart, the turnover rates of all analyzed mitochondrial brain proteins had much slower turnover rates compared to those in the liver.

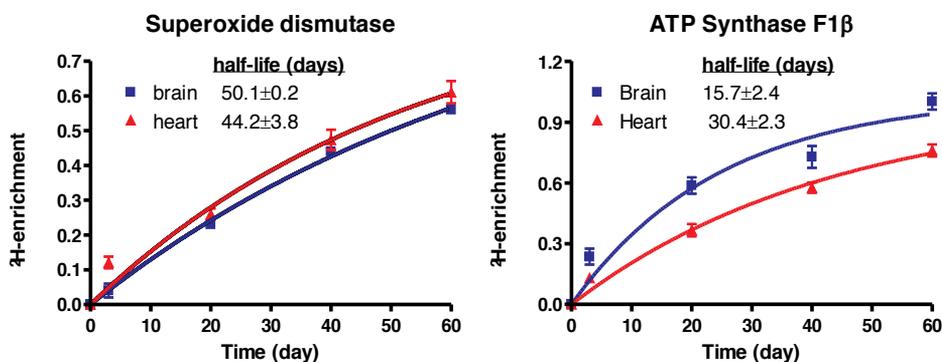


Figure 5. Comparison of half-lives of brain and heart mitochondrial proteins.

To test the effect of heart failure on the stability of cardiac mitochondrial proteins, we utilized our $^2\text{H}_2\text{O}$ approach to measure mitochondrial proteome dynamics in a well-established rat model of heart failure induced by chronic transverse aortic constriction (TAC) [54]. Decreased mitochondrial ATP generating capacity in myocardium is a hallmark of heart failure; however, the underlying mechanisms contributing to mitochondrial dysfunction in heart failure are not yet fully understood. Rats with TAC develop moderate heart failure after 22 weeks, which results in left ventricular remodeling, dysfunction, and reduced oxidative capacity in mitochondria. Heart failure caused a decrease of mitochondrial proteins and respiratory capacity in IFM, but not in SSM. We used a heavy water method to determine whether the decreased synthesis of mitochondrial proteins contribute to the respiratory dysfunction in heart failure. Although the synthesis rates of proteins in IFM tend to be higher than those in SSM, it only started to reach modest significance ($p=0.08$) in this study. Surprisingly, in spite of the changes in the mitochondrial protein content, the average rate of protein synthesis (based on the kinetics of 49 proteins with different functions) was similar in sham-treated and heart failure groups. This was due to bidirectional changes in the synthesis rate of different mitochondrial proteins. In particular, heart failure increased the turnover rate of several proteins involved in

fatty acid oxidation, electron transport chain, and ATP synthesis, while it decreased the turnover of other proteins, including pyruvate dehydrogenase subunit in IFM, but not in SSM. The study of proteome dynamics suggested that reduced respiratory capacity in IFM might be related to increased degradation of several IFM proteins involved in fatty acid oxidation and ETC. Interestingly, proteins with destabilizing *N*-terminal amino acids of mature proteins exhibited shorter half-lives compared to those with stabilizing *N*-terminal amino acids.

Thus, the kinetic measurements of mitochondrial proteins may help understand the mechanisms responsible for mitochondrial alterations in the failing heart. Taken together, utilization of the $^2\text{H}_2\text{O}$ method for mitochondrial proteome studies demonstrated that this method is robust and can distinguish subtle differences in synthetic rates between subcellular populations of mitochondria. In addition, measuring the kinetics of individual proteins enables one to uncover changes in the mitochondrial proteome due to heart disease that cannot be obtained by simply measuring their static expression at any given time point.

In a follow-up study, Lam and coworkers applied the heavy water method to determine protein kinetic signatures of β -adrenergic-induced cardiac remodeling in a mouse model [47]. Several kinetic markers of calcium signaling, energy metabolism, proteostasis, and mitochondrial dynamics were identified. Although large set of data was generated, the biological relevance of these results requires further evaluation based on protein properties and pathways that they are involved.

Hellerstein and coworkers used $^2\text{H}_2\text{O}$ labeling-based dynamic proteomics combined with the stable isotope labeling in mammals (SILAM) quantitative proteomics to explain the effect of long-term calorie restriction on longevity [83]. Through assessment of both catabolic rate and absolute synthesis of hepatic proteins, the authors demonstrated that calorie restriction reduces the turnover of most (~80%) hepatic proteins, including mitochondrial proteins. Thus, long-term calorie restriction increases the stability of proteins and reduces global protein synthetic burden that is associated with decreased mitochondrial biogenesis and mitophagy. The pathway analysis revealed that proteins with related functions display coordinated changes. *In silico* analysis identified peroxisome proliferator-activated receptor gamma coactivator 1- α as a potential regulator of altered network dynamics.

The $^2\text{H}_2\text{O}$ -labeling methods were also applied to identify kinetic biomarkers of neuronal dysfunction in mouse models of neurodegeneration [84]. After a bolus administration of $^2\text{H}_2\text{O}$, appearance of ^2H -labeled neuronal proteins with transport and cargo functions in cerebrospinal fluid was quantified. Compared to controls, the appearance of proteins in mice with neurodegeneration was delayed, which was normalized after microtubule-modulating pharmacotherapy, suggesting that the transport kinetics may provide a test method for monitoring disease progression and therapy for neurodegenerative diseases.

We applied $^2\text{H}_2\text{O}$ -based metabolic labeling approach to assess the high-density lipoprotein (HDL) proteome dynamics in a diet-induced mouse model of nonalcoholic fatty liver disease (NAFLD) [85]. HDL displays multiple functions that include reverse cholesterol transport (RCT), preventing inflammation, oxidation, platelet activation, and maintaining endothelial function. In metabolic diseases associated with insulin resistance, HDL may lose these protective functions and become dysfunctional. The reasons for these changes are not fully

understood and may be attributed to alterations of the HDL particle composition and modifications of HDL proteins. In addition to ApoAI and ApoAII (which account ~65% and ~15% of HDL protein mass, respectively), recently more than 50 less abundant HDL proteins have been identified. These HDL proteins involved in lipid metabolism, acute-phase response, innate immunity, protease inhibition, and regulation of endothelial cell apoptosis that determines HDL's anti-inflammatory, anti-atherogenic, and cell survival properties. Thus, alterations in the HDL proteome composition may be a key factor involved in HDL dysfunction.

It is well known that a Western diet (WD, high-fat diet containing cholesterol) for 12 weeks leads to insulin resistance, NAFLD (hepatic steatosis, oxidative stress, and inflammation), and atherosclerosis (aortic root lesion) in low-density lipoprotein receptor (LDLR^{-/-}) mice. Proteomics analysis of ApoB-depleted plasma revealed that a WD also altered the levels of multiple proteins known to be associated with HDL. The kinetics of 60 previously identified HDL proteins involved in lipid metabolism, thrombosis, protease inhibition, complement regulation, and acute-phase response were quantified. The analyzed HDL proteins exhibited a wide range of half-lives varying from a few hours to days. For instance, in a standard chow diet-fed LDLR^{-/-} mice, ApoE, ApoAI, and PON1 have half-lives 5, 15, and 64 h, respectively. A WD has differential effects on the turnover rates of proteins with different functions. We found that a WD results in decreased levels and increased catabolism of PON1 which is responsible for the antioxidant function of HDL. Interestingly, a WD also resulted in increased levels and turnover of phospholipid transfer protein (PLTP), which is responsible for promoting HDL remodeling through phospholipid transfer from ApoB-containing particles to HDL. Mice deficient in PLTP are protected from atherosclerosis, while HDL from mice over expressing PLTP is dysfunctional in promoting cholesterol efflux, and these mice developed higher atherosclerotic lesion compared to control mice. Thus, ²H₂O labeling allows to measure HDL proteome flux that is relevant to HDL functionality.

Since the RCT function of HDL represents the dynamic flux of cholesterol from peripheral tissues, including macrophage transfer to liver for clearance, we next applied our ²H₂O-metabolic labeling approach to assess HDL flux as an *in vivo* index of RCT [31]. Because ²H from ²H₂O incorporates into both lipids and proteins, ²H₂O allows studying the kinetics of both HDL-cholesterol (HDLc) and apoAI, the principal protein of HDL. Mice were given ²H₂O in the drinking water and serial blood samples were collected at different time points. Fractional catabolic rates (FCR) for HDLc and apoAI were assessed based on their ²H₂O-metabolic labeling. In addition, the synthetic heavy peptide of apoAI (VAPL(⁶C₁₃)GAEL(⁶C₁₃)QESAR) and [²H₆]cholesterol were used for absolute quantification of pool sizes and production rates (PR) of apoAI and HDLc, respectively. ApoE^{-/-} mice, which are prone to atherosclerosis, displayed an increased FCR ($p < 0.01$) and a reduced PR of both HDLc and apoAI ($p < 0.05$) compared to controls. In human apoAI transgenic mice (resistant to atherosclerosis), PRs of HDLc and human apoAI were strikingly higher than in wild-type mice. We also validated our HDL turnover method as an index of RCT. For this purpose, HDL turnover and macrophage-specific RCT were assessed in the same animals. Myriocin, an inhibitor of sphingolipid synthesis, was used as a modifier of HDL metabolism. Myriocin significantly increased HDL flux and macrophage-to-feces RCT, indicating compatibility of these methods. We conclude

that $^2\text{H}_2\text{O}$ labeling can be used to measure HDLc and apoAI flux *in vivo*, and to assess the role of genetic and pharmacological interventions on HDL turnover.

$^2\text{H}_2\text{O}$ labeling-based HDL turnover method also was applied to assess the effect of different isoforms of apoAI and gender on *in vivo* HDL function in wild-type human transgenic apoAI mice and mice with 4WF isoform of human apoAI, in which 4 tryptophan residues are substituted with phenylalanine [86]. The *in vitro* cholesterol efflux assay demonstrated that the 4WF isoform of apoAI was resistant to myeloperoxidase-induced loss of function while human apoAI transgenic HDL lost all ABCA1-dependent cholesterol acceptor activity. This was associated with a small, nonsignificant increase in HDL turnover *in vivo*. Male mice displayed significantly higher plasma apoAI levels than females for both isoforms of human apoAI, ascribed to increased production rate of HDL. Safety, simplicity, and low cost of the $^2\text{H}_2\text{O}$ suggest that this approach can be used for human use to study the effects of HDL-targeted therapies on both HDL proteome and HDLc dynamics.

5.2. Human studies

Although $^2\text{H}_2\text{O}$ has been used for more than 60 years in animal studies to measure a proteins' renewal rate, only in 2004 it was introduced to study protein synthesis rates in humans [87]. This first human study validated the basic underlying assumptions of $^2\text{H}_2\text{O}$ use in humans, i.e., equilibrium with total body water and amino acids is rapid and body water enrichment can be maintained constant for a long period of time. With $\sim 0.4\%$ TBW enrichment, the FSR of albumin based on albumin-bound alanine enrichment was determined to be $\sim 4\%$ /day in renal patients.

A recent study evaluated the long-term safety and hemodynamic effects of higher levels of heavy water ingestion in healthy young human subjects [53]. Subjects consumed 70% enriched $^2\text{H}_2\text{O}$ in 4 boluses of 0.51 ml/kg body weight daily during the first week of labeling. During the second week, the subjects consumed 4 boluses of 0.56 ml/kg. This protocol resulted in gradual increase of body water enrichment up to $\sim 2\%$ during the 14 days of heavy water exposure. The subjects' vital signs were monitored during $^2\text{H}_2\text{O}$ administration, and these subjects were followed up to an 8-month period. Total body water enrichment during exposure and subsequent physiological clearance from body fluids were determined during the following 2 weeks. No signs of discomfort and physiological effect were reported in these healthy young adults. After depletion of 14 of the most abundant proteins by multiple affinity columns, the tryptic digest from remaining proteins was fractionated using two-dimensional liquid chromatography separations and analyzed by the LTQ Orbitrap instrument. The turnover rates of hundreds of proteins were then determined. There was no correlation between protein turnover rates and protein abundance. Although many proteins involved in cardiovascular disease were also quantified, this proof of the concept study did not evaluate any link between protein turnover rates and disease. It was concluded that $^2\text{H}_2\text{O}$ is safe and effective tracer for large-scale human studies.

Several human studies utilized low-dose heavy water to assess the effect of exercise and cachexia on muscle protein synthesis. Gaiser and colleagues applied $^2\text{H}_2\text{O}$ ($\sim 0.3\%$ TBW $^2\text{H}_2\text{O}$ enrichment) with a single biopsy protocol to test the effect of short-term (24-h) exercise on mixed muscle protein synthesis [46]. With this approach, the effect of acute resistance exercise

on integrative myofibrillar protein synthesis in healthy young subjects was determined. Subjects performed unilateral exercise using one leg while the other leg served as a control. Interestingly, exercise did not have any effect on the FSR of mixed muscle proteins. The high-intensity resistance exercise increased myofibrillar protein synthesis in the exercising leg ($0.94 \pm 0.16\%/h$) compared to the control leg ($0.75 \pm 0.08\%/h$, $p < 0.05$), demonstrating that short-term low-level 2H_2O exposure allows one to detect subtle changes in human muscle protein synthesis.

Recently, Wilkinson and colleagues expanded on these studies and investigated the effect of long-term (8-day) exercise on mixed muscle protein synthesis with heavy water for monitoring day-to-day changes in muscle subfractions (myofibrillar, collagen, sarcoplasmic) synthesis [88]. Similar to the study by Gaiser and colleagues, the authors employed a one-legged resistance exercise that allows use of the second leg as an internal control. The longer period of exercise and heavy water administration with multiple muscle biopsies at different time points in this study allowed them assess the changes in muscle protein synthesis in response to the temporal and cumulative successive bouts of exercise. By using the highly sensitive IRMS instrument, this study validated the utility of low dose (0.16–0.24% enrichment of TBW) heavy water for quantification of diurnal changes in muscle protein synthesis and for the assessment of short-term changes in protein turnover. It was demonstrated that protein synthesis in myofibrillar and collagen fractions was increased due to both short-term and long-term exercise; however, sarcoplasmic protein synthesis remained unchanged.

Scalzo and colleagues applied heavy water-based dynamic proteomics to assess integrated and individual muscle protein synthesis response and mitochondrial biogenesis for endurance exercise in males and females after 3 weeks of sprint interval training [89]. This study utilized 3 weeks of 2H_2O -labeling protocol to achieve 1–2% TBW enrichment. It was demonstrated that due to exercise, muscle protein synthesis increased and the magnitude of change was higher in males compared with females. The increase in integrative muscle protein synthesis was associated with increased mitochondrial biogenesis assessed based on the synthesis rates of individual mitochondrial proteins and mitochondrial biogenesis signaling. It is important to note that it is unfeasible to use pre-labeled amino acid tracers for this kind of long-term studies of muscle protein synthesis, because this would require inpatient tracer infusion for several days.

Recently, a few studies utilized the heavy water method to assess the protein turnover in different diseases. A single oral dose of heavy water was applied to assess muscle protein synthesis in patients undergoing surgery for upper gastrointestinal cancer [90]. It was demonstrated that the mixed muscle protein synthesis was not decreased, rather, it was marginally increased as compared to healthy controls ($p = 0.03$), suggesting that an increase in muscle breakdown may account for muscle wasting in cancer patients.

Studies from Hellerstein's group tested the utility of the heavy water method as a diagnostic tool in patients with psoriasis diseases [91]. The epidermal kinetics was determined in patients with psoriasis using twice-daily doses of 2H_2O for 16–38 days. Keratin turnover was significantly accelerated in psoriatic lesions, suggesting that keratin synthesis could be used as a kinetic biomarker of psoriasis and other skin diseases.

These studies demonstrated that the heavy water method has a great potential for human studies.

6. Challenges and future directions

Since $^2\text{H}_2\text{O}$ can be administered to humans, the dynamic proteomics approach could be widely used for clinical studies. Proteomics centers and infrastructure, which are equipped with state-of-the-art instrumentations and bioinformatics, exist in many areas in the USA and around the world. Static quantitative proteomics is already making highlights in clinical research and patient care. It is expected that in the near future, $^2\text{H}_2\text{O}$ will complement the traditional proteomics and expand to different areas of clinical research. The most obvious application of the heavy water method would be its utilization for the assessment of dynamics of circulatory proteins. Because of the high sensitivity of existing mass spectrometers, dynamic proteome analysis using small-tissue biopsy samples is also feasible. Thus, there is a great potential of using “dynamic markers” of health and disease. However, despite the wide-range potential for use in clinical settings, the heavy water method is still lagging behind as a diagnostic tool in patient care. This is partly related to several unmet methodological, instrumental, and bioinformatics challenges associated with studies of heavy water-based proteome dynamics. Unresolved issues related to the patient-oriented test design, user-friendly software development, and challenges centered around the data interpretation currently impede the routine clinical application of this technology.

In particular, a simple study design with a minimal number of short-term samples is very critical. This also requires creation of a reference database with human protein half-lives for implementation of a simple test for the proteins of interest based on their expected half-life ranges. In terms of methodological issues, still there are no published study on the effect of posttranslational modification and damage-induced aggregation of proteins on protein turnover and stability.

Although the mass spectrometry-based hardware tools are developing very fast, the cost of existing instruments is not easily affordable for many clinical laboratories which drives the cost of any proteomics test. Therefore, the cost reduction in this direction would facilitate the dynamic proteomics application as a clinical test method.

Some additional challenges are related to data interpretation and software issues. To advance *in vivo* studies of proteome dynamics with heavy water, high-throughput, user-friendly, robust, vendor-independent, accurate software capable of producing statistically rigorous results is needed. As mentioned in previous sections, currently there is no freely available software for comprehensive proteome dynamics data analysis. Although our recent software allows high-throughput data analysis, there are still several unmet bioinformatics challenges related to heavy water data analysis. One of the technical issues is related to quality control in data analysis. Sample complexity is the major challenge for automated data analysis. Although off-line liquid chromatography and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (both 1D and 2D) separations simplify peptide mixtures, co-elution of peptides persists, even after the long-gradient chromatographic separation. This problem is

more severe with heavy water metabolic labeling, because in contrast to other tracers, $^2\text{H}_2\text{O}$ -metabolic labeling does not result in sizeable mass shifts of newly synthesized peptides, rather it leads to redistribution of incorporated ^2H among all heavy isotopomers. Therefore, it is critical to measure mass isotopomer distributions with the maximum number of heavy isotopomers. Although a simplified approach with M1/M0 has been proposed for the quantification of relative isotopomer abundances, an accurate evaluation of peptides ^2H enrichment requires tracing several isotopomers. Inclusion of all heavy isotopomers into calculations increases the chances of contamination by co-eluting species and chromatographic overlap signals. Although high mass resolution and accuracy substantially reduce the problem, sample complexity dramatically affects the turnover rate measurements, if not taken into account. Thus, the success in the computing of accurate turnover rates will depend on the availability of robust, easy-to-use software and bioinformatics tools for data analysis which would allow processing the co-elution profiles and extracting the mass profiles of the target species.

Our current software allows assessing the fractional catabolic and synthesis rates of a protein in a steady state. However, it is also critically important to know the absolute production rate of a protein and to determine whether protein abundance is regulated by the changes in a protein degradation or production. These types of measurements require simultaneous quantification of isotopic distribution and protein abundance. Also, currently used regression analysis for calculation of a rate constant(s) is based on a single compartmental model that relies on a steady-state assumption. However, amino acids and protein levels are in a non-steady state during growth, aging, and diseases [92]. The nonsteady-state calculations of protein turnover necessitate kinetic models, including data on both protein abundance and relative isotopomer distribution. The future bioinformatics tools based on multi-compartmental kinetic analysis and the quantification of absolute protein production rate in nonsteady-state condition would greatly advance proteome dynamics studies. In addition, there is currently a gap between dynamic proteomics and pathway analysis. Although several software are available for the functional analysis of data based on static proteomics data, currently there are no bioinformatics tools for system biology flux analysis using the proteome dynamics data.

Finally, clinical application of the heavy water method would necessitate fully automated data analysis. So far, existing software solutions are unconnected applications that require multiple format conversion for the input and analysis. Improvement in software cross talk between raw data inputs and data analysis applications would integrate data analysis pipelines with data acquisition and search engines. This would require software engineering development that could transform the existing algorithms to robust user-friendly software packages.

In addition, to the technical limitations highlighted above, the heavy water-based metabolic labeling approach is applicable to analysis of dynamics of proteins with a half-life of greater than ~ 2 h. This is because it takes approximately 1 h to reach the steady-state enrichment in the amino acid pool, thus it cannot be used for rapidly secreted fast turnover peptides. On the other hand, it is ideally suited to assess proteins that have a more constant rate of secretion and relative stable plasma concentrations, and a half-life of >2 h. It is also not appropriate in short-term experiments (less than 1 week) to measure proteins in plasma that are slowly

synthesized constituents of cells, such as troponin or creatine kinase, released in response to tissue injury or necrosis.

Thus, routine and widespread utilization of $^2\text{H}_2\text{O}$ as a diagnostic tool in patient care requires future advancement in several areas. As we discussed above, robust study designs complemented with facile sample preparation, multiplexed analysis, and user-friendly software package allowing high-throughput data processing and interpretation are required. As a universal tracer, heavy water could be used to measure other metabolic fluxes along with proteome dynamics. Thus, as a comprehensive diagnostic tool, the heavy water method could revolutionize personalized medicine, provided there are certain future technological advancements in this field.

Acknowledgements

We thank Dr. Vernon E. Anderson for his insight and efforts in developing early stages of our proteome dynamic studies. This work was funded by National Institutes of Health Grants 5R21RR025346 (SP, TK), R21 HL-114407 (TK), NLBIHHSN268201000037C (RGS), 1R01GM112044-01A1 (RGS, TK), American Heart Association 131RG14700011, and 15GRNNT25500004 (TK), and was supported in part by the National Institutes of Health, Clinical Translational Science Collaborative of Cleveland UL1TR000439 (TK).

The Thermo Orbitrap Elite mass spectrometer used in this study was purchased using NIH shared instrument grant 1S10RR031537-01 (BW).

Author details

T. Kasumov^{1,2*}, B. Willard³, L. Li³, R.G. Sadygov⁴ and S. Previs⁵

*Address all correspondence to: kasumov@neomed.edu or tkasumov@hotmail.com

1 Department of Gastroenterology and Hepatology, Cleveland Clinic, Cleveland, OH, USA

2 Department of Pharmaceutical Sciences, Northeast Ohio Medical University, Rootstown, OH, USA

3 Department of Research Core Services, Cleveland Clinic, Cleveland, OH, USA

4 Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, TX, USA

5 Pharmacokinetics, Pharmacodynamics and Drug Metabolism, Merck NJ, USA

References

- [1] Ratner S, Rittenberg D, Keston AS, Schoenheimer R: The Journal of Biological Chemistry, Volume 134, June 1940: Studies in protein metabolism. XIV. The chemical interaction of dietary glycine and body proteins in rats. By S. Ratner, D. Rittenberg, Albert S. Keston, and Rudolf Schoenheimer. *Nutrition reviews* 1987, 45(10):310–312.
- [2] Rittenberg D, Shemin D: The metabolism of proteins and amino acids. *Annual review of biochemistry* 1946, 15:247–272.
- [3] Schoenheimer R, Rittenberg D, Foster GL, Keston AS, Ratner S: The Application of the Nitrogen Isotope N15 for the Study of Protein Metabolism. *Science* 1938, 88(2295): 599–600.
- [4] Schoenheimer R, Ratner S, Rittenberg D: Studies in protein metabolism: X. The metabolic activity of body proteins investigated with L-(-)-leucine containing two isotopes. *The Journal of biological chemistry* 1939, 130:703–732.
- [5] Schoenheimer R, Ratner S, Rittenberg D: The Process of Continuous Deamination and Reamination of Amino Acids in the Proteins of Normal Animals. *Science* 1939, 89(2308):272–273.
- [6] Rose WC: The amino acid requirements of adult man. *Nutrition abstracts and reviews Series A: Human and experimental* 1957, 27(3):631–647.
- [7] Rittenberg D, Ponticorvo L, Borek E: Studies on the sources of the oxygen of proteins. *The journal of biological chemistry* 1961, 236:1769–1772.
- [8] Borsheim E, Aarsland A, Wolfe RR: Effect of an amino acid, protein, and carbohydrate mixture on net muscle protein balance after resistance exercise. *International journal of sport nutrition and exercise metabolism* 2004, 14(3):255–271.
- [9] Ohsumi Y: Protein turnover. *IUBMB life* 2006, 58(5–6):363–369.
- [10] James WP, Garlick PJ, Sender PM, Waterlow JC: Studies of amino acid and protein metabolism in normal man with L-[U-14C]tyrosine. *Clinical science and molecular medicine* 1976, 50(6):525–532.
- [11] Matthews DE, Motil KJ, Rohrbaugh DK, Burke JF, Young VR, Bier DM: Measurement of leucine metabolism in man from a primed, continuous infusion of L-[1-3C]leucine. *The American journal of physiology* 1980, 238(5):E473–479.
- [12] Chinkes DL, Rosenblatt J, Wolfe RR: Assessment of the mathematical issues involved in measuring the fractional synthesis rate of protein using the flooding dose technique. *Clinical science* 1993, 84(2):177–183.
- [13] Fern EB, Garlick PJ, McNurlan MA, Waterlow JC: The excretion of isotope in urea and ammonia for estimating protein turnover in man with [15N]glycine. *Clinical science* 1981, 61(2):217–228.

- [14] Ray S, Patel SK, Kumar V, Damahe J, Srivastava S: Differential expression of serum/plasma proteins in various infectious diseases: specific or nonspecific signatures. *Proteomics clinical applications* 2014, 8(1–2):53–72.
- [15] Li L, Willard B, Rachdaoui N, Kirwan JP, Sadygov RG, Stanley WC, Previs S, McCullough AJ, Kasumov T: Plasma proteome dynamics: analysis of lipoproteins and acute phase response proteins with $2\text{H}_2\text{O}$ metabolic labeling. *Molecular & cellular proteomics : MCP* 2012, 11(7):M111 014209.
- [16] Hoedt E, Zhang G, Neubert TA: Stable isotope labeling by amino acids in cell culture (SILAC) for quantitative proteomics. *Advances in experimental medicine and biology* 2014, 806:93–106.
- [17] Wu CC, MacCoss MJ, Howell KE, Matthews DE, Yates JR, 3rd: Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Analytical chemistry* 2004, 76(17):4951–4959.
- [18] Claydon AJ, Thom MD, Hurst JL, Beynon RJ: Protein turnover: measurement of proteome dynamics by whole animal metabolic labelling with stable isotope labelled amino acids. *Proteomics* 2012, 12(8):1194–1206.
- [19] Doherty MK, Hammond DE, Clague MJ, Gaskell SJ, Beynon RJ: Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC. *Journal of proteome research* 2009, 8(1):104–112.
- [20] Cayol M, Tauveron I, Rambourdin F, Prugnaud J, Gachon P, Thieblot P, Grizard J, Obled C: Whole-body protein turnover and hepatic protein synthesis are increased by vaccination in man. *Clinical science* 1995, 89(4):389–396.
- [21] Claydon AJ, Beynon R: Proteome dynamics: revisiting turnover with a global perspective. *Molecular & cellular proteomics : MCP* 2012, 11(12):1551–1565.
- [22] Rachdaoui N, Austin L, Kramer E, Previs MJ, Anderson VE, Kasumov T, Previs SF: Measuring proteome dynamics in vivo: as easy as adding water? *Molecular & cellular proteomics : MCP* 2009, 8(12):2653–2663.
- [23] Xiao GG, Garg M, Lim S, Wong D, Go VL, Lee WN: Determination of protein synthesis in vivo using labeling from deuterated water and analysis of MALDI-TOF spectrum. *Journal of applied physiology* 2008, 104(3):828–836.
- [24] Borek E, Ponticorvo L, Rittenberg D: Protein Turnover in Micro-Organisms. *Proceedings of the National Academy of Sciences of the United States of America* 1958, 44(5):369–374.
- [25] Price JC, Holmes WE, Li KW, Floreani NA, Neese RA, Turner SM, Hellerstein MK: Measurement of human plasma proteome dynamics with $(2)\text{H}(2)\text{O}$ and liquid chromatography tandem mass spectrometry. *Analytical biochemistry* 2012, 420(1):73–83.
- [26] Herath K, Bhat G, Miller PL, Wang SP, Kulick A, Andrews-Kelly G, Johnson C, Rohm RJ, Lassman ME, Previs SF *et al.*: Equilibration of $(2)\text{H}$ labeling between body

- water and free amino acids: enabling studies of proteome synthesis. *Analytical biochemistry* 2011, 415(2):197–199.
- [27] Schoenheimer R, Rittenberg D: Deuterium as an indicator in the study of intermediary metabolism. *Science* 1935, 82(2120):156–157.
- [28] Busch R, Siah IM, Gee TA, Hellerstein MK: Heavy water labeling of DNA for measurement of cell proliferation and recruitment during primary murine lymph node responses against model antigens. *Journal of immunological methods* 2008, 337(1):24–34.
- [29] Schumann WC, Gastaldelli A, Chandramouli V, Previs SF, Pettiti M, Ferrannini E, Landau BR: Determination of the enrichment of the hydrogen bound to carbon 5 of glucose on 2H₂O administration. *Analytical biochemistry* 2001, 297(2):195–197.
- [30] Brunengraber DZ, McCabe BJ, Kasumov T, Alexander JC, Chandramouli V, Previs SF: Influence of diet on the modeling of adipose tissue triglycerides during growth. *American journal of physiology, endocrinology and metabolism* 2003, 285(4):E917–925.
- [31] Kasumov T, Willard B, Li L, Li M, Conger H, Buffa JA, Previs S, McCullough A, Hazen SL, Smith JD: 2H₂O-based high-density lipoprotein turnover method for the assessment of dynamic high-density lipoprotein function in mice. *Arteriosclerosis, thrombosis, and vascular biology* 2013, 33(8):1994–2003.
- [32] Yuan CL, Sharma N, Gilge DA, Stanley WC, Li Y, Hatzoglou M, Previs SF: Preserved protein synthesis in the heart in response to acute fasting and chronic food restriction despite reductions in liver and skeletal muscle. *American journal of physiology, endocrinology and metabolism* 2008, 295(1):E216–222.
- [33] Sharma N, Okere IC, Barrows BR, Lei B, Duda MK, Yuan CL, Previs SF, Sharov VG, Azimzadeh AM, Ernsberger P *et al*: High-sugar diets increase cardiac dysfunction and mortality in hypertension compared to low-carbohydrate or high-starch diets. *Journal of hypertension* 2008, 26(7):1402–1410.
- [34] Anderson SR, Gilge DA, Steiber AL, Previs SF: Diet-induced obesity alters protein synthesis: tissue-specific effects in fasted versus fed mice. *Metabolism: clinical and experimental* 2008, 57(3):347–354.
- [35] Dufner D, Previs SF: Measuring in vivo metabolism using heavy water. *Current opinion in clinical nutrition and metabolic care* 2003, 6(5):511–517.
- [36] Busch R, Kim YK, Neese RA, Schade-Serin V, Collins M, Awada M, Gardner JL, Beyesen C, Marino ME, Misell LM *et al*: Measurement of protein turnover rates by heavy water labeling of nonessential amino acids. *Biochimica et biophysica acta* 2006, 1760(5):730–744.
- [37] Kasumov T, Ilchenko S, Li L, Rachdaoui N, Sadygov RG, Willard B, McCullough AJ, Previs S: Measuring protein synthesis using metabolic (2)H labeling, high-resolution mass spectrometry, and an algorithm. *Analytical biochemistry* 2011, 412(1):47–55.

- [38] Wallyn CS, Vidrich A, Airhart J, Khairallah EA: Analysis of the specific radioactivity of valine isolated from aminoacyl-transfer ribonucleic acid of rat liver. *The biochemical journal* 1974, 140(3):545–548.
- [39] Halliday D, Venkatesan S, Pacy P: Apolipoprotein metabolism: a stable-isotope approach. *The American journal of clinical nutrition* 1993, 57(5 Suppl):726S–730S; discussion 730S–731S.
- [40] Connell A, Calder AG, Anderson SE, Lobley GE: Hepatic protein synthesis in the sheep: effect of intake as monitored by use of stable-isotope-labelled glycine, leucine and phenylalanine. *The British journal of nutrition* 1997, 77(2):255–271.
- [41] Reeds PJ, Hachey DL, Patterson BW, Motil KJ, Klein PD: VLDL apolipoprotein B-100, a potential indicator of the isotopic labeling of the hepatic protein synthetic precursor pool in humans: studies with multiple stable isotopically labeled amino acids. *The journal of nutrition* 1992, 122(3):457–466.
- [42] Waterlow JC: Protein turnover. Cambridge, MA 02139, USA: CAB International; 2006.
- [43] Fern EB, Garlick PJ: The specific radioactivity of the precursor pool for estimates of the rate of protein synthesis. *The biochemical journal* 1973, 134(4):1127–1130.
- [44] Ahlman B, Charlton M, Fu A, Berg C, O'Brien P, Nair KS: Insulin's effect on synthesis rates of liver proteins. A swine model comparing various precursors of protein synthesis. *Diabetes* 2001, 50(5):947–954.
- [45] Belloto E, Diraison F, Basset A, Allain G, Abdallah P, Beylot M: Determination of protein replacement rates by deuterated water: validation of underlying assumptions. *American journal of physiology, endocrinology and metabolism* 2007, 292(5):E1340–1347.
- [46] Gasier HG, Riechman SE, Wiggs MP, Previs SF, Fluckey JD: A comparison of $2\text{H}_2\text{O}$ and phenylalanine flooding dose to investigate muscle protein synthesis with acute exercise in rats. *American journal of physiology, endocrinology and metabolism* 2009, 297(1):E252–259.
- [47] Lam MP, Wang D, Lau E, Liem DA, Kim AK, Ng DC, Liang X, Bleakley BJ, Liu C, Tabaraki JD *et al*: Protein kinetic signatures of the remodeling heart following isoproterenol stimulation. *The journal of clinical investigation* 2014, 124(4):1734–1744.
- [48] Shah V, Herath K, Previs SF, Hubbard BK, Roddy TP: Headspace analyses of acetone: a rapid method for measuring the 2H -labeling of water. *Analytical biochemistry* 2010, 404(2):235–237.
- [49] Varshavsky A: Discovery of cellular regulation by protein degradation. *The journal of biological chemistry* 2008, 283(50):34469–34489.

- [50] Foster DM, Barrett PH, Toffolo G, Beltz WF, Cobelli C: Estimating the fractional synthetic rate of plasma apolipoproteins and lipids from stable isotope data. *Journal of lipid research* 1993, 34(12):2193–2205.
- [51] Brun V, Masselon C, Garin J, Dupuis A: Isotope dilution strategies for absolute quantitative proteomics. *Journal of proteomics* 2009, 72(5):740–749.
- [52] Strawford A, Antelo F, Christiansen M, Hellerstein MK: Adipose tissue triglyceride turnover, de novo lipogenesis, and cell proliferation in humans measured with $2\text{H}_2\text{O}$. *American journal of physiology, endocrinology and metabolism* 2004, 286(4):E577–588.
- [53] Wang D, Liem DA, Lau E, Ng DC, Bleakley BJ, Cadeiras M, Deng MC, Lam MP, Ping P: Characterization of human plasma proteome dynamics using deuterium oxide. *Proteomics clinical applications* 2014, 8(7–8):610–619.
- [54] Shekar KC, Li L, Dabkowski ER, Xu W, Ribeiro RF, Jr., Hecker PA, Recchia FA, Sadygov RG, Willard B, Kasumov T *et al*: Cardiac mitochondrial proteome dynamics with heavy water reveals stable rate of mitochondrial protein synthesis in heart failure despite decline in mitochondrial oxidative capacity. *Journal of molecular and cellular cardiology* 2014, 75:88–97.
- [55] Kasumov T, Dabkowski ER, Shekar KC, Li L, Ribeiro RF, Jr., Walsh K, Previs SF, Sadygov RG, Willard B, Stanley WC: Assessment of cardiac proteome dynamics with heavy water: slower protein synthesis rates in interfibrillar than subsarcolemmal mitochondria. *American journal of physiology, heart and circulatory physiology* 2013, 304(9):H1201–1214.
- [56] Katanik J, McCabe BJ, Brunengraber DZ, Chandramouli V, Nishiyama FJ, Anderson VE, Previs SF: Measuring gluconeogenesis using a low dose of $2\text{H}_2\text{O}$: advantage of isotope fractionation during gas chromatography. *American journal of physiology, endocrinology and metabolism* 2003, 284(5):E1043–1048.
- [57] Hilkert AW, Douthitt CB, Schluter HJ, Brand WA: Isotope ratio monitoring gas chromatography/Mass spectrometry of D/H by high temperature conversion isotope ratio mass spectrometry. *Rapid communications in mass spectrometry : RCM* 1999, 13(13):1226–1230.
- [58] Cassano AG, Wang B, Anderson DR, Previs S, Harris ME, Anderson VE: Inaccuracies in selected ion monitoring determination of isotope ratios obviated by profile acquisition: nucleotide $18\text{O}/16\text{O}$ measurements. *Analytical biochemistry* 2007, 367(1):28–39.
- [59] Wang B, Sun G, Anderson DR, Jia M, Previs S, Anderson VE: Isotopologue distributions of peptide product ions by tandem mass spectrometry: quantitation of low levels of deuterium incorporation. *Analytical biochemistry* 2007, 367(1):40–48.
- [60] Zhou H, Li W, Wang SP, Mendoza V, Rosa R, Hubert J, Herath K, McLaughlin T, Rohm RJ, Lassman ME *et al*: Quantifying apoprotein synthesis in rodents: coupling

- LC-MS/MS analyses with the administration of labeled water. *Journal of lipid research* 2012, 53(6):1223–1231.
- [61] Marshall AG, Hendrickson CL, Jackson GS: Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass spectrometry reviews* 1998, 17(1):1–35.
- [62] Lebedev AT, Damoc E, Makarov AA, Samgina TY: Discrimination of leucine and isoleucine in peptides sequencing with Orbitrap Fusion mass spectrometer. *Analytical chemistry* 2014, 86(14):7017–7022.
- [63] Ilchenko S, Previs SF, Rachdaoui N, Willard B, McCullough AJ, Kasumov T: An improved measurement of isotopic ratios by high resolution mass spectrometry. *Journal of the American Society for Mass Spectrometry* 2013, 24(2):309–312.
- [64] Koning L, Kort C, Pinske F, Nibberring N: Segmented Fourier transform and its application to FT-ICR MC: ion abundance and mass measurements. *International journal of mass spectrometry and ion processes* 1989, 95:1–16.
- [65] Erve JC, Gu M, Wang Y, DeMaio W, Talaat RE: Spectral accuracy of molecular ions in an LTQ/Orbitrap mass spectrometer and implications for elemental composition determination. *Journal of the American Society for Mass Spectrometry* 2009, 20(11):2058–2069.
- [66] Hoopmann MR, Chavez JD, Bruce JE: SILACTor: software to enable dynamic SILAC studies. *Analytical chemistry* 2011, 83(22):8403–8410.
- [67] Hsieh EJ, Shulman NJ, Dai DF, Vincow ES, Karunadharma PP, Pallanck L, Rabinovitch PS, MacCoss MJ: Topograph, a software platform for precursor enrichment corrected global protein turnover measurements. *Molecular & cellular proteomics : MCP* 2012, 11(11):1468–1474.
- [68] Kasumov T, Willard B, Sadygov RG: Current Bioinformatics Challenges in Proteome Dynamics using Heavy Water-based Metabolic Labeling. *Journal of data mining in genomics & proteomics* 2014, 5(1):e112.
- [69] Kim TY, Wang D, Kim AK, Lau E, Lin AJ, Liem DA, Zhang J, Zong NC, Lam MP, Ping P: Metabolic labeling reveals proteome dynamics of mouse mitochondria. *Molecular & cellular proteomics : MCP* 2012, 11(12):1586–1594.
- [70] Nemutlu E, Zhang S, Juranic NO, Terzic A, Macura S, Dzeja P: 18O-assisted dynamic metabolomics for individualized diagnostics and treatment of human diseases. *Croatian medical journal* 2012, 53(6):529–534.
- [71] Filipp FV: A Gateway between Omics Data and Systems Biology. *Journal of metabolomics and systems biology* 2013, 1(1):1.
- [72] van der Greef J, Martin S, Juhasz P, Adourian A, Plasterer T, Verheij ER, McBurney RN: The art and practice of systems biology in medicine: mapping patterns of relationships. *Journal of proteome research* 2007, 6(4):1540–1559.

- [73] Zhang GF, Kombu RS, Kasumov T, Han Y, Sadhukhan S, Zhang J, Sayre LM, Ray D, Gibson KM, Anderson VA *et al*: Catabolism of 4-hydroxyacids and 4-hydroxynonenal via 4-hydroxy-4-phosphoacyl-CoAs. *The Journal of biological chemistry* 2009, 284(48):33521–33534.
- [74] Dufner DA, Bederman IR, Brunengraber DZ, Rachdaoui N, Ismail-Beigi F, Siegfried BA, Kimball SR, Previs SF: Using $2\text{H}_2\text{O}$ to study the influence of feeding on protein synthesis: effect of isotope equilibration in vivo vs. in cell culture. *American journal of physiology, endocrinology and metabolism* 2005, 288(6):E1277–1283.
- [75] Chan XC, Black CM, Lin AJ, Ping P, Lau E: Mitochondrial protein turnover: methods to measure turnover rates on a large scale. *Journal of molecular and cellular cardiology* 2015, 78:54–61.
- [76] Lesnefsky EJ, Guduz TI, Migita CT, Ikeda-Saito M, Hassan MO, Turkaly PJ, Hoppel CL: Ischemic injury to mitochondrial electron transport in the aging heart: damage to the iron-sulfur protein subunit of electron transport complex III. *Archives of biochemistry and biophysics* 2001, 385(1):117–128.
- [77] Papanicolaou KN, Ngoh GA, Dabkowski ER, O'Connell KA, Ribeiro RF, Jr., Stanley WC, Walsh K: Cardiomyocyte deletion of mitofusin-1 leads to mitochondrial fragmentation and improves tolerance to ROS-induced mitochondrial dysfunction and cell death. *American journal of physiology, heart and circulatory physiology* 2012, 302(1):H167–179.
- [78] Papanicolaou KN, Kikuchi R, Ngoh GA, Coughlan KA, Dominguez I, Stanley WC, Walsh K: Mitofusins 1 and 2 are essential for postnatal metabolic remodeling in heart. *Circulation research* 2012, 111(8):1012–1026.
- [79] Palmer JW, Tandler B, Hoppel CL: Biochemical properties of subsarcolemmal and interfibrillar mitochondria isolated from rat cardiac muscle. *The Journal of biological chemistry* 1977, 252(23):8731–8739.
- [80] Palmer JW, Tandler B, Hoppel CL: Heterogeneous response of subsarcolemmal heart mitochondria to calcium. *The American journal of physiology* 1986, 250(5 Pt 2):H741–748.
- [81] Asemu G, O'Connell KA, Cox JW, Dabkowski ER, Xu W, Ribeiro RF, Jr., Shekar KC, Hecker PA, Rastogi S, Sabbah HN *et al*: Enhanced resistance to permeability transition in interfibrillar cardiac mitochondria in dogs: effects of aging and long-term aldosterone infusion. *American journal of physiology, heart and circulatory physiology* 2013, 304(4):H514–528.
- [82] Menzies RA, Gold PH: The turnover of mitochondria in a variety of tissues of young adult and aged rats. *The journal of biological chemistry* 1971, 246(8):2425–2429.
- [83] Price JC, Khambatta CF, Li KW, Bruss MD, Shankaran M, Dalidd M, Floreani NA, Roberts LS, Turner SM, Holmes WE *et al*: The effect of long term calorie restriction on

- in vivo hepatic proteostasis: a novel combination of dynamic and quantitative proteomics. *Molecular & cellular proteomics : MCP* 2012, 11(12):1801–1814.
- [84] Fanara P, Wong PY, Husted KH, Liu S, Liu VM, Kohlstaedt LA, Riiff T, Protasio JC, Boban D, Killion S *et al*: Cerebrospinal fluid-based kinetic biomarkers of axonal transport in monitoring neurodegeneration. *The journal of clinical investigation* 2012, 122(9): 3159–3169.
- [85] Kasumov T, Li L, Willard B, Hazen S, Smith J, McCullough AJ: A western diet induced atherosclerosis in *ldlr*^{-/-} mice is associated with altered hdl proteome dynamics. *American Heart Association Meeting Dallas, TX* 2013, abstract November 16–20.
- [86] Berisha SZ, Brubaker G, Kasumov T, Hung KT, DiBello PM, Huang Y, Li L, Willard B, Pollard KA, Nagy LE *et al*: HDL from ApoA1 Transgenic Mice Expressing the 4WF Isoform Is Resistant to Oxidative Loss of Function. *Journal of lipid research* 2015.
- [87] Previs SF, Fatica R, Chandramouli V, Alexander JC, Brunengraber H, Landau BR: Quantifying rates of protein synthesis in humans by use of 2H₂O: application to patients with end-stage renal disease. *American journal of physiology, endocrinology and metabolism* 2004, 286(4):E665–672.
- [88] Wilkinson DJ, Franchi MV, Brook MS, Narici MV, Williams JP, Mitchell WK, Szewczyk NJ, Greenhaff PL, Atherton PJ, Smith K: A validation of the application of D(2)O stable isotope tracer techniques for monitoring day-to-day changes in muscle protein subfraction synthesis in humans. *American journal of physiology, endocrinology and metabolism* 2014, 306(5):E571–579.
- [89] Scalzo RL, Peltonen GL, Binns SE, Shankaran M, Giordano GR, Hartley DA, Klochak AL, Lonac MC, Paris HL, Szallar SE *et al*: Greater muscle protein synthesis and mitochondrial biogenesis in males compared with females during sprint interval training. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* 2014, 28(6):2705–2714.
- [90] MacDonald AJ, Johns N, Stephens NA, Greig CA, Ross JA, Small AC, Husi H, Fearon KC, Preston T: Habitual myofibrillar protein synthesis is normal in patients with upper GI cancer cachexia. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2014.
- [91] Emson CL, Fitzmaurice S, Lindwall G, Li KW, Hellerstein MK, Maibach HI, Liao W, Turner SM: A pilot study demonstrating a non-invasive method for the measurement of protein turnover in skin disorders: application to psoriasis. *Clinical and translational medicine* 2013, 2:12.
- [92] Tuvdendorj D, Chinkes DL, Herndon DN, Zhang XJ, Wolfe RR: A novel stable isotope tracer method to measure muscle protein fractional breakdown rate during a physiological non-steady-state condition. *American journal of physiology, endocrinology and metabolism* 2013, 304(6):E623–630.
- [93] This paper is dedicated to the memory of Dr. William C. Stanley.

Neuroproteomics – LC-MS Quantitative Approaches

Cátia Santa, Sandra I. Anjo, Vera M. Mendes and Bruno Manadas

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/61298>

Abstract

Neuroproteomics is a scientific field that aims to study all the proteins of the central nervous system, their expression, function, and interactions. The central nervous system is intricate and heterogeneous, and the study of its proteome is consequently complex, with many biological questions still requiring deep investigation. For this, mass spectrometry approaches, most often coupled with liquid chromatography (LC-MS), have been the number one choice in proteomics, and over the years it has added many important findings to the field. At this point it is important that proteomics turns to the quantitative expression of proteins instead of only identifying which proteins are present in a given sample, much because the most important alterations may be slight alterations in the quantity of a protein in a given situation. Therefore, many LC-MS quantitative approaches have been developed relying on the labeling of the proteins or even by using label-free techniques.

In this chapter, a brief description of the principles and procedures of several approaches used for relative and absolute, targeted and untargeted quantification of proteins is presented, complemented with a literature revision of their application in the neurosciences field.

Keywords: Neuroproteomics, LC-MS techniques, central nervous system, protein relative quantification, protein absolute quantification

1. Introduction

Neuroproteomics is a field that aims to study all the proteins of the central nervous system (CNS), as a whole or related to a specific condition (for example, disease, drug response, etc.). CNS is very complex, presenting a high degree of heterogeneity at several levels, such as distinct brain regions, cellular networks, and cell types [1], each one characterized by a different

proteome. Even slight perturbations of this structure can lead to CNS disorders, resulting in alterations in the proteome of all CNS constituents or of specific cellular networks.

Large-scale initiatives have been performed to sequence human and other organism's genomes [2], as well as the analysis of gene expression of the distinct regions and cells of the brain. However, although these studies have contributed with crucial information, the end-point of gene transcription is the synthesis of proteins, the effector molecules. This way, the complex and dynamic nature of the proteome has led to a paradigm shift in the neurosciences field, changing from the focus in genomic information to the analysis of the protein's expression levels, by resorting to several approaches [3, 4].

Proteomics methodologies aim to analyze a large number of proteins within a certain set of samples of an experiment [5], and the great development of this area may be attributed to the technological advances in mass spectrometry (MS), optimization in sample preparation, and computer sciences that are now able to deal with the large amount of information generated by the MS-based technologies [6, 7].

These approaches can deliver different types of data, such as identification of the protein in a sample at a given moment, expression levels of the proteins (quantitative proteomics), identification and quantification of post-translational modifications (PTM), and protein interactions (for example protein-protein interactions) [7].

Over the past years, MS-based proteomics approaches have been able to characterize proteins in complex mixtures; nonetheless, these approaches have largely been qualitative, successfully identifying a high amount of proteins from one sample but failing in quantifying the expression levels of these [5]. However, it has been pressing to turn the proteomics field to quantitative approaches, once most of the interesting biological alterations are slight differences in the amount of a protein present in a given situation [8].

The main goal of quantitative proteomics, or quantitative neuroproteomics in particular, is to measure the expression level of, theoretically, all the proteins in a given sample, preferably in a highly reproducible manner [9]. This quantitative information can be acquired in two distinct ways: absolute quantification, where the amount of the protein in the sample is calculated (for instance, in terms of concentration or copy number per cell); or relative quantification, where the amount of a given protein is expressed as a fold change for the same protein relative to another condition [5, 7]. The approaches to obtain relative quantification may be untargeted, where virtually all the proteins in the sample are quantified; or targeted where the quantification is obtained for a selected protein or a set of proteins. A brief summary of the most important methodologies is outlined in Figure 1.

The classical approach to obtain relative quantifications of a proteome was to perform a bi-dimensional electrophoresis (2DE-Isoelectric focusing followed by SDS-PAGE), where the identification of the proteins was obtained by a MS analysis and the relative quantification by measuring the staining density of matched gel spots [9]. Nonetheless, in this method, some types of proteins are underrepresented, and although hundreds to a few thousands of proteins may be detected, many proteins with lower abundance are very difficult to quantify. Also, the analysis of many samples by this method is laborious and time consuming. [9]

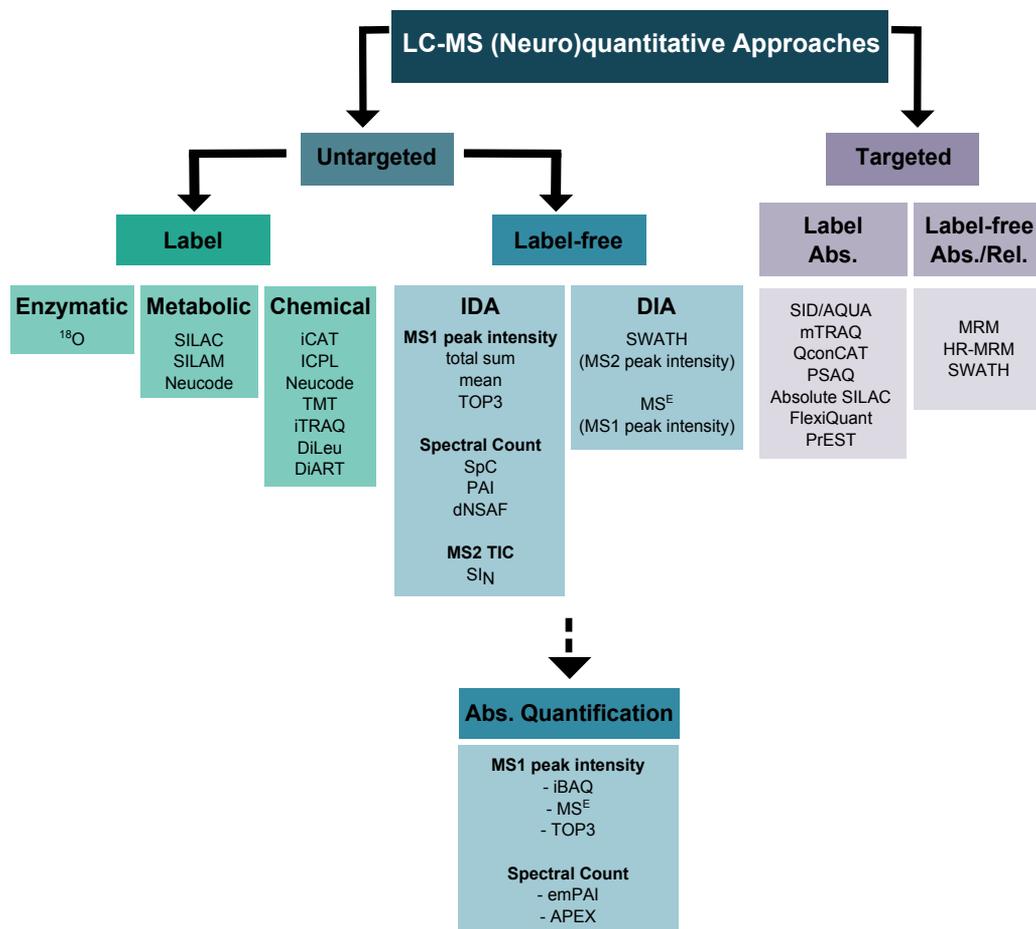


Figure 1. Diagram with brief description of the LC-MS proteomics techniques.

Therefore, over the years several methodologies were developed that support proteomic expression level quantification, and although the most popular are the so-called labeled approaches (which require the stable isotopic labeling of the samples prior to MS analysis), the label-free approaches are now gaining increasing interest mostly due to higher accuracy and sensitivity of MS instruments and improvement of the algorithms for data analysis [9].

In this chapter, a brief introduction to the different LC-MS quantitative approaches will be performed, mainly focusing on the main principle and their major achievements. Special attention will be given to the most commonly used methods in each category, and finally a revision of the literature on proteomics using those approaches will be performed, and whenever possible, examples in neuroproteomics field will be provided to elucidate the concepts.

2. Stable isotope labeling quantitative approaches

The major advantage of using MS to perform quantification instead of gel-based quantification is the possibility of slight molecular mass changes to be detectable and quantifiable by a mass spectrometer in large scale approaches and not by any other technology.

The use of stable isotopic labeling for relative protein quantification can be achieved by three different methodologies: enzymatic, as the incorporation of ^{18}O upon the protein digestion; chemical, as the incorporation of mass tags in lysines and amine-terminus of proteins or peptides; or metabolic labeling with the incorporation of heavy amino acids during protein synthesis [10].

The quantitative analysis for each approach may be performed at different levels, where some labels have mass differences that are detected (and quantified) in the precursor mass spectra (MS1), and others are based on isobaric labels that lead to peptides with the same m/z but can be distinguished (and quantified) at the fragment level (MS/MS) [9].

Each approach has its advantages and limitations, and are appropriate for different analysis depending on the biological question and on the type of sample to be used [11].

2.1. ^{18}O enzymatic labeling of peptides

The first trackable use of stable isotopes for quantification in neuronal tissue was used by Desiderio and colleagues by isotopically labeling peptide internal standards for the absolute quantification of neuropeptides [12]. To achieve this purpose, the authors used for the first time enzymatically incorporated ^{18}O (from H_2^{18}O) in the carboxylic end of the peptides [12]. Although the strategy has been used since then, it was only in 2001 that it was first reported in a study of untargeted relative quantification of the proteome of two types of adenovirus [13].

Since this first introduction, the enzymatic incorporation of ^{18}O by serine proteases has been widely used to compare the peptides produced from the protein digestion of distinct samples (usually a control sample versus a sample from the condition under study). In general, the incorporation of the heavy oxygen molecules is achieved by performing the protein digestion in H_2^{18}O using trypsin, although other enzymes such as chymotrypsin, lysine carboxylase (LysC), or GluC may also be used [13]. With this approach two oxygen atoms are introduced in the C-terminus of each generated peptide, resulting in a shift of 4Da in the mass spectra of the peptide when compared with the peptides obtained from the sample digested with regular water (Figure 2A) [14].

The advantages of ^{18}O enzymatic labeling are: the fact that virtually all the produced peptides are labeled and co-elute with the correspondent unlabeled peptide; the only reagent specifically required is H_2^{18}O ; and the procedure is easy to adapt in any proteomics lab [15, 16]. On the other hand, the procedure is labor-intensive and time-consuming; the labeling efficiency is influenced by many factors (such as pH, enzyme to be used, or the characteristics of the proteins and the peptides); and also if the ^{18}O -water to be used is less than 95% pure, some of the peptides will be labeled with only one ^{18}O , resulting in a mass spectra with both 2 Da and

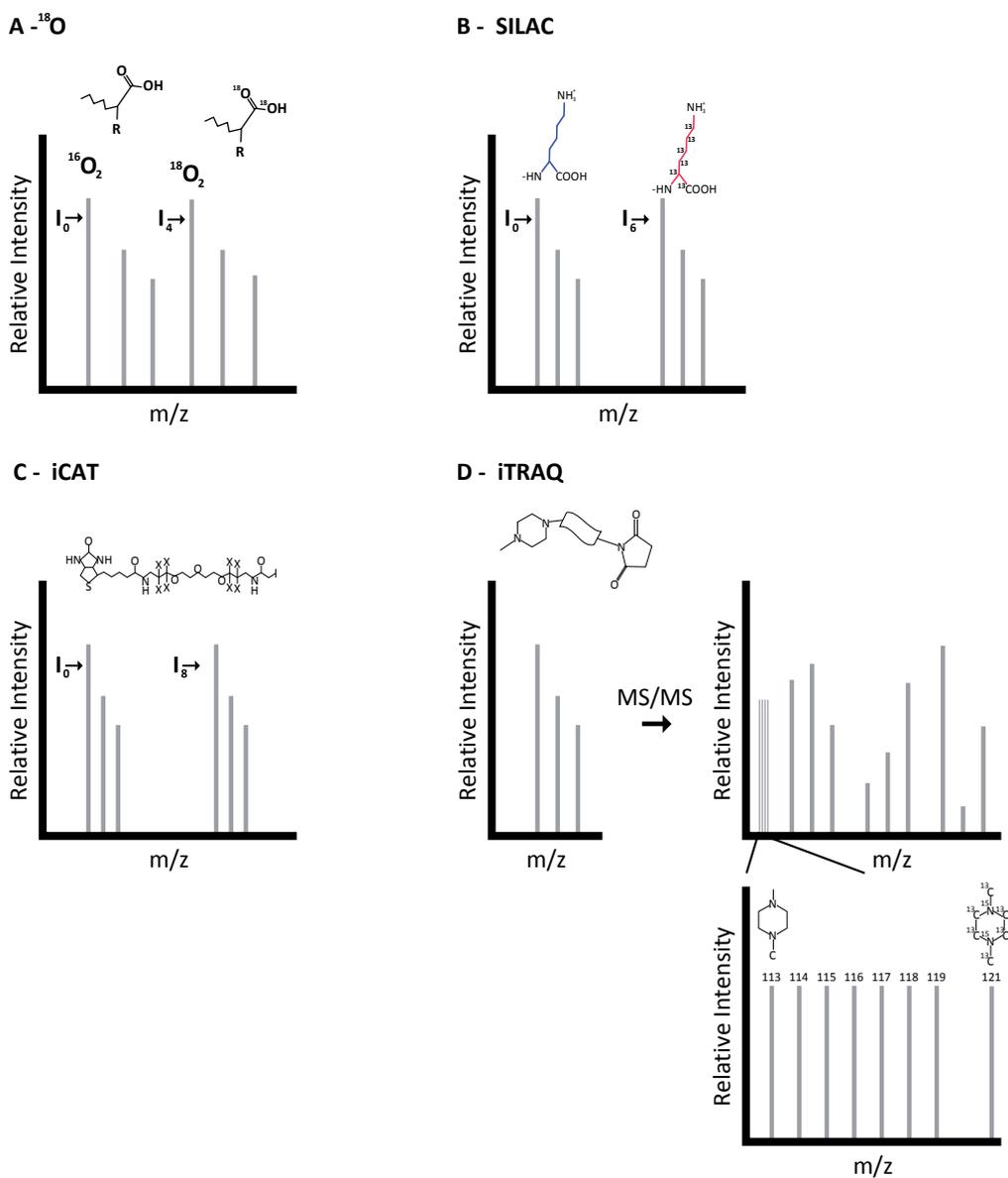


Figure 2. Representation of spectra and labelling molecules from different quantitative labeling techniques. A) Representative spectrum used for peptide identification and quantification in a ^{18}O quantitative approach with the representation of the molecules with ^{16}O and ^{18}O . B) SILAC technique with representative spectrum used for peptide identification and quantification accompanied by the representation of regular and heavy lysine. C) ICAT molecule, where the X may be hydrogen or deuterium and an example spectrum used for peptide identification and quantification. D) Representation of the iTRAQ 8-plex molecule and the spectra obtained in an iTRAQ experiment with the MS/MS spectrum that is used for identification and a zoom of the low m/z region where the reporter ions are used for quantification.

4 Da mass shifts [15]. Finally, the naturally abundant isotopes may also contribute to the peak intensities making the spectra very complex to analyze and adding the necessity for improved software for data processing [16].

In 2009, an updated ^{18}O labeling method was introduced, the acid-catalyzed labeling of the peptides, which, instead of the direct labeling of the peptides during the proteolytic digestion, it was able to separate the digestion from the labeling step, being the last performed under acidic conditions [17]. This protocol aimed to increase the distance between the unlabeled and the labeled peptide in the mass spectra (as the acidic amino acids also incorporate the heavy oxygen molecules) and also decrease the tendency of back exchange from ^{18}O to ^{16}O that was reported [17, 18]. However, the incubation in acidic conditions is prolonged, may lead to acidic hydrolysis of the peptides and deamidation of some amino acids that would increase the complexity of the spectra [18].

In order to overcome the high time consumption of the procedure, many accelerating techniques have been applied, such as heating, high pressure, or ultrasonic energy [15]. Also, other methodologies have been used, such as “inverse labeling”, which aims to decrease the influence of naturally occurring isotopes [19]. Other ^{18}O labeling approaches have been proposed throughout the years, as the incorporation of the ^{18}O molecules in cysteine (Cys) residues at the protein level by the use of ^{18}O -labeled iodoacetamide (cysteine alkylating agent) [20] or the analysis of glycoproteins after specific enrichment [21].

This ^{18}O labeling strategy has been employed to several types of samples, such as samples from the CNS, for instance, for the differential expression study of proteins in the hippocampus of rats subjected to traumatic brain injury [22] or for the quantitative profiling of CNS myelin-associated proteins in the adult mouse brain [23].

2.2. Metabolic labeling approaches

Although labeled media have been widely used in biological studies, it was only in 1999 that it was first used to evaluate protein expression by 2-DE [24] or phosphopeptides [25] in microorganisms. Nonetheless, after the introduction of SILAC (Stable Isotope Labeling by Amino Acids in Cell Culture) in 2002, metabolic labeling approach gained higher visibility [26].

Briefly, SILAC methodology consists of growing two populations of cells, one in the presence of normal (light) medium and the other in the presence of medium that contains heavy essential amino acids [27]. The labeling of the amino acids can be achieved by substituting hydrogen for deuterium, ^{12}C for ^{13}C or ^{14}N for ^{15}N [27], and this leads to an expected mass shift in the peptides coming from the heavy medium-grown cells that is visible in the mass spectra of the peptide (Figure 2B) [28]. A shift from the first report using deuterated leucine [26] to the use of labeled lysine and arginine with ^{13}C or ^{15}N has been employed, much due to the properties of the enzyme to be used (usually trypsin or LysC). In this way, virtually all peptides in the sample will be labeled [28], and also eliminates the problem of some deuterated peptides eluting at different retention times than the unlabeled analogue [29]. In order not to introduce quantitative errors in a SILAC experiment, all the proteins must be labeled; therefore, the cells must be kept in culture with medium supplemented with dialyzed serum (to avoid unlabeled

amino acids) for at least five passages in order to have at least 97% labeling [26, 28], although a study of the labeling efficiency is advisable whenever a new cell line is used [28].

The major difference between this approach and others is that the labeling of the proteins is performed metabolically, and also the mixing of the samples to be compared is performed in the first steps of sample preparation leading to less variability in the results (Figure 3) [30]. Other advantages of the use of SILAC is its ease of use and implementation and also the possibility of multiplexing (up to 5 samples per experiment) [9, 30].

While it was proposed initially that dialyzed serum should be used to avoid the presence of non-labeled amino acids, this fact posed as a challenge for some cell culture types. In contrast, many studies have already been performed with regular serum, proving that this extra caution may not be necessary [30].

Over the years SILAC has been adapted to different cell types and with many different applications, such as the analysis of protein-protein interactions [31], identification and quantification of PTMs (for example by using methyl SILAC with labeled methionine [32]) and protein modification dynamics [33], measurement of proteome translation or turnover (by applying pulsed SILAC) [34, 35], or secretome protein quantification [36, 37].

Thus, SILAC has been applied to try to answer many neurobiological questions since it was introduced; and in the last years, many studies have been published using this technique in many different areas, such as the psychiatric field with studies of alcohol abuse [38] and schizophrenia [39]; in neurodegenerative diseases by studying the functions of Parkin [40]; or apoptosis in a neuroblastoma cell line [41].

The first rationale about SILAC was that it could only be applied to immortalized cell lines and never to cultured primary cells. However, there are now many published studies that use this technique in primary cells [9, 42], namely in primary neuronal cell lines, as in a study of neuronal phosphotyrosine proteome in response to stimulation by a neurotrophic factor [43]; in a quantitative analysis of synaptic proteins from cultured cortical neurons from a mouse model of mental retardation [44]; in the analysis of microtubule dynamics in rat hippocampal neurons [45]; or even by enabling the analysis of primary cultured astrocytes proteome and secretome [46]. Also, a strategy to diminish the number of passages necessary for the complete labeling in cultured primary neurons (60% after 6 days and 90% after 10 days) was proposed by multiplexing SILAC and using labeled amino acids for all the samples so that the protein labeling incorporation rate may be the same in both samples (because both samples will have the same heavy/light incorporation ratio) [47, 48].

In neuroproteomic studies, although neuronal-derived immortalized and primary cell lines may be considered good simplified models, the use of mammal models (such as rodents) are considered to be more complete. The general principle of SILAC was to add heavy amino acids to cells in culture, making this approach incompatible with animal models. One of the first attempts to overcome this challenge was by using the SILAC approach in cultured Neuro2A cells and then mix them with mouse brain samples to work as internal standards [49]. The first mammal to have the entire proteome labeled *in vivo* was a rat being fed with protein-free diet supplemented with algal cells enriched with ¹⁵N [50, 51]. In 2008, the first mouse model to be

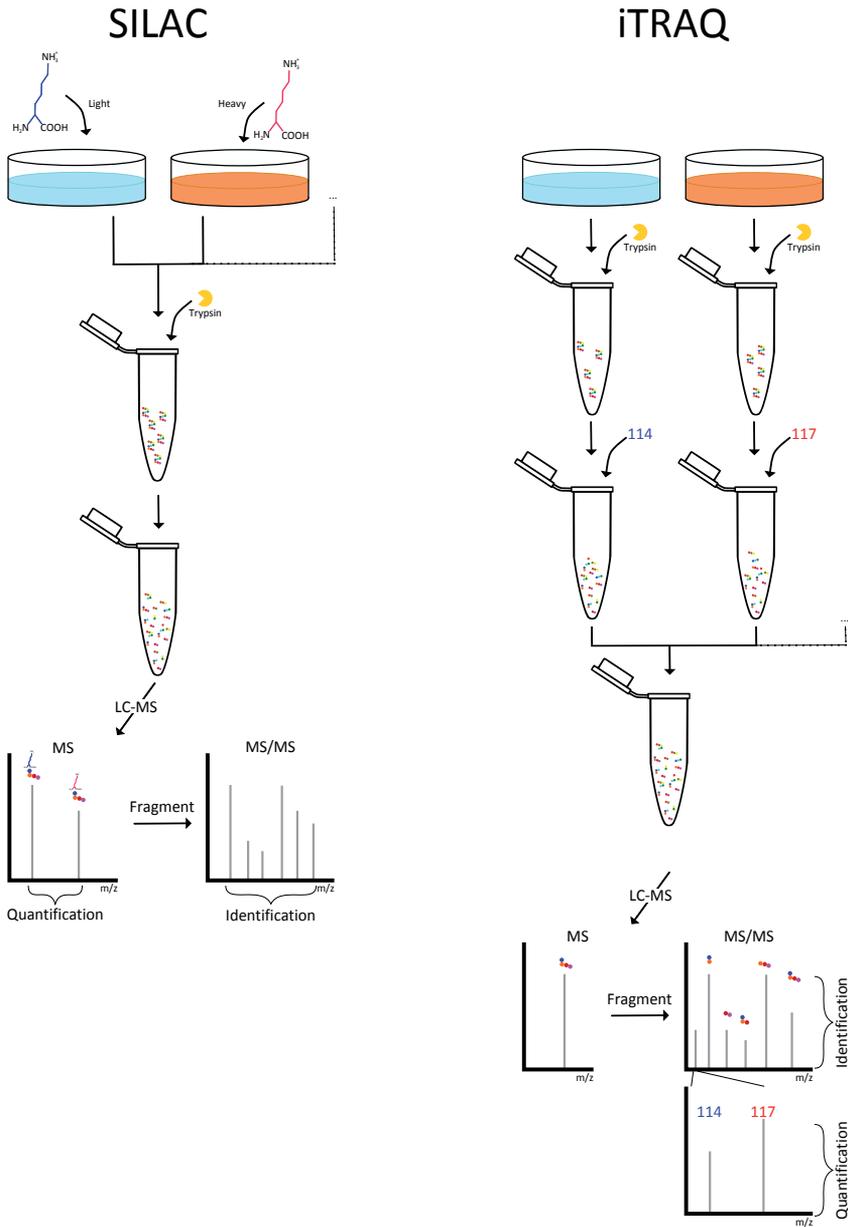


Figure 3. Comparison of the quantitative procedures of SILAC and iTRAQ.

labeled *in vivo* by using a heavy amino acid ^{13}C -Lysine [52] was introduced. This new strategy was named SILAM or Stable Isotope Labeling (by Amino Acids) in Mammals, and it has been applied in several topics of the neuroscience field as the quantification of the synaptosomal proteome of the rat cerebellum during development [53] or the proteome relative changes in barrel cortex synapses upon sensory deprivation in mice [54].

In what concerns neuroproteomics, the labeling of brain tissue *in vivo* is a great advantage, although in order to be able to completely label all proteins in the brain of rodents it is necessary to feed the animals with a special “heavy” diet at least for two generations, making this approach time-consuming and expensive [52, 55, 56]. Therefore, one of the most promising possibilities of SILAM is to use tissue from control SILAM-labeled animals as internal standards to compare between unlabeled conditions [57, 58].

Also, because of this drawback, the super-SILAC approach was introduced, where multiple cell lines are labeled with SILAC and are afterwards used as internal standards to compare with unlabeled tissue [59, 60]. This technique was firstly introduced with cancer cell lines in 2010, but it has recently been applied to the study of mitochondria from mice brain by using a super-SILAC mix of mouse brain mitochondria [61].

It was recently observed that the energy required to break down a nucleus into its component nucleons (nuclear binding energy) is different for each isotope of every element leading to a so-called “mass defect” (a mass difference of 6 mDa in the same molecule when a ^{12}C is exchanged by a ^{13}C atom and a ^{15}N for a ^{14}N) led to the hypothesis that a calculated incorporation of isotopes into proteomes would generate a MS1-centric quantification technology combining SILAC with the multiplexing capacity of isobaric tagging (see below) [62]. This new approach is named neutron encoding (NeuCode) SILAC, where peptide identifications are generated using the MS1 scans collected at 30,000 resolving power, where the same peptide with multiple labels will appear as a single peak in the spectra, whereas to obtain the quantitative information a higher resolution (480,000) MS1 scan is used, where the isotopologues can be resolved and the quantitative information extracted as for normal SILAC (with a mass shift of 36 mDa instead of 4 or 8 Da) [62]. This approach has the advantage of decreasing redundant acquisition of fragment spectra for the same precursor ion (as in classical SILAC), and because the quantitative information is acquired at the MS1 level, it is not dependent on peptides selected for MS/MS and is not subjected to dynamic range compression caused by co-isolation of precursor ions (as in isobaric labeling, see below) [62, 63].

In these first reports, the authors claim that the NeuCode approach may be used for 12-plexing by using 3-plex SILAC, each one combined with 4 isotopologues, resulting in four distinct peaks in a high-resolution spectra [62, 63], although it has already been used for 6- and 18-plex in yeast cells proteome [64]. This approach has already been used in other applications, such as C-terminal product ion annotation, based on the fact that all the y-ion in the fragment spectra will appear as doublets [65, 66], or in top-down proteomics (analysis of the intact proteins instead of peptides resulting from protein digestion) [67]. The major disadvantage of this technique is that it requires MS equipments capable of high-resolution powers ($\geq 480,000$); nonetheless, this approach is expected to be easily adapted for neuroproteomics research.

2.3. Chemical labeling approaches: Isotope techniques

The first technique using isotope labeling probes was called isotope-coded affinity tag (ICAT) and was introduced in 1999 [68]. In this approach, a specific reagent (“tag”) is added to the cysteines of proteins, once this tag has a thiol-specific reactive group, a linker with 8 deuteriums in the heavy form, and a biotin affinity tag [68]. The procedure is simple and based on some

basic steps: first the protein extracts must be isolated and the cysteines reduced, then the proteins are labeled with the heavy or light ICAT molecule and joined for protein digestion; the labeled peptides are enriched with an avidin affinity chromatography and analyzed by LC-MS, where for each precursor a pair of ions will be visible with a mass shift in MS1 mass spectra (Figure 2c) [68].

This first ICAT molecule was designed with 8 deuteriums leading most of the times to a difference in retention times of the homologue peptides, where the labeled peptide does not co-elute with its unlabeled pair, making the spectra analysis very difficult. Also, this mass difference of 8 Da may be confused with other biological modifications (such as a peptide containing 2 cysteines and an oxidation of methionine, both leading to a 16 Da mass shift). [69] On the other hand, the ICAT tag itself was quite large contributing with a mass addition sometimes bigger than advisable and leading to many fragments in the MS/MS spectra, complicating the identification of the peptides' sequence [69].

Due to these limitations of the initial approach, new strategies were introduced based on the same principles, but with a cleavable site introduced to the tag [69, 70] or also the possibility of labeling the sample in a solid-phase format [70]. This new cleavable ICAT (cICAT) has an acid-cleavable linker group connecting the biotin with the thiol-reactive isotope tag and uses 9^{13}C instead of the 8 deuterium, this way, after labeling and chromatographic enrichment, the biotin moiety is cleaved giving rise to a smaller modified peptide [69].

This ICAT strategy has already been applied for different approaches as the creation of aldehyde-reactive tags (hydrazide-functionalized) isotope-coded affinity tag (HICAT) for the identification and quantification of lipid-conjugated proteins [71].

This isotope-labeling technology has been applied in several neuroscience projects such as the study of the influence of aging in the proteome of CSF (cerebrospinal fluid) [72], the study of differential mitochondrial proteins analysis in the pathophysiology of Parkinson's [73] or Alzheimer's diseases [74], and also to aid the study of the expression of synaptosomal protein in cerebral ischemia [75], migraine mouse models [76], or in the study of addiction [77].

The greatest limitation of this approach is the fact that only peptides containing cysteines are labeled and enriched, making these the only candidates for protein identification and quantification, leading most of the times to poor sequence coverages. For this reason, a similar strategy, ICPL (isotope-coded protein labeling) was developed, which, instead of labeling sulfhydryl groups labels all free amine groups [78]. This strategy is very similar to ICAT, with the exception that it has specificity for primary amine groups (lysine side chains and N-termini), and has no biotin moiety so the option to enrich labeled peptides does not exist. On the other hand, it is expected that at least 70% of all peptides will have labeled lysines [78, 79], or virtually all the peptides if the labeling is performed after digestion (post-digest ICPL) [79, 80].

This post-digest ICPL can be combined with other fractionation methods such as IEF prior to LC-MS [81] or even with enrichment of peptides with specific PTM's as phosphorylation or glycosylation [82].

The original ICPL molecule could be multiplexed for three samples where the molecule had 0, 3, or 7 deuterium (d0, d3, and d7 molecules, respectively) [78], but is commercialized in a 4-plex version allowing the labeling with 0, 4, 6, and 10 Da mass shifts and may be labeled with deuterium or ^{13}C [9]. Although this approach is not widely used it has the capacity to be applied successfully to any protein samples, and it has already been used to study the proteome of postmortem prefrontal cortex from control and schizophrenic patients [81, 83] and in biopsy tissue samples from patients with glioblastoma [84].

As for SILAC, very recently, the NeuCode strategy described above has been applied to chemical labeling with the development of an amine-reactive mass tag that takes advantage of the differential neutron-binding energy between ^{13}C and ^{15}N isotopes that enables up to 12-plex MS1-based protein quantification [63]. Another NeuCode approach proposed is to use carbamylation of amine groups via urea isotopologues for protein/peptide labeling, and therefore relative quantification [85].

2.4. Chemical labeling approaches: Isobaric techniques

All the methods described above use isotopic labeling of the proteins or respective peptides, this way the calculation of the relative amounts is achieved by the analysis of the intensity of the precursor ion peaks at the MS1 spectra. In 2003, a revolutionary variation of these techniques was introduced where the mass tag that was added to the peptides is isobaric, making all the precursor ions from the samples in study appear as a single peak in MS1, but upon fragmentation it leads to the formation of reporter ions separated by 1 Da coming specifically from each of the samples [10]. The first approach applying this principle was called Tandem Mass Tag (TMT) and in this first report synthesized peptides with the tag were used [86]. A year later another approach was described, the isobaric tags for relative and absolute quantification (iTRAQ). This concept was applied for the first time to label global proteomes (yeast in this case) and even with the advantage of allowing the simultaneous analysis of 4 samples (iTRAQ 4-plex) [87].

The molecule used to tag the proteins or the respective peptides after digestion for both approaches, iTRAQ and TMT, has three main components and the principles are the same, although structurally different between the two methods (Figure 2D). The molecules are constituted by an amine-reactive group, which links the reagent to lysines and N-termini of the proteins or peptides; by a reporter group, which has differential labeling with isotopes (^{13}C , ^{15}N or ^{18}O) and is, upon fragmentation, the monitored ion for quantification in the MS/MS spectra; and also a balancer group, which aims to keep the overall mass of the reagent equal among all labels and is also differentially labeled with isotopes. [10]

A few years after this first introduction of TMT and iTRAQ, the neuroproteomics field had the highest multiplexing usage of these approaches, in this case by studying proteomic changes in CSF of patients with Alzheimer's disease undergoing intravenous immunoglobulin treatment with iTRAQ 8-plex [88] and by comparing CSF proteome in postmortem versus antemortem drawing of the samples using a 6-plex TMT approach [89]. In 2012, upon the substitution of a ^{13}C for a ^{15}N in two of the 6-plex tags it was noticed that the new tags were 6.32 mDa lighter, this way an 8-plex approach was developed even without changing the

structure of the molecule but only by changing the isotopologue used [90], and it was hypothesized that a 10-plex or even 18-plex approach was possible [90]. In fact, TMT is commercially available in four different kits (TMTzero, TMTduplex, TMTsixplex, and TMTtenplex), whereas iTRAQ is commercially available in two versions (iTRAQ 4-plex and iTRAQ 8-plex).

The TMT tags give origin to reporter ions in the 126–131 Da region of the MS/MS spectra and the molecules used in all the available kits have the same structure. On the other hand, the iTRAQ 4-plex and 8-plex molecules, which generate ions in the 114–117 Da and 113–121 Da (except the 120 because of phenylalanine immonium ion contamination), respectively, have different structures [10, 91].

Isobaric labeling of the proteins and quantification at the MS/MS level (outlined in Figure 3) has the advantage of each precursor ion appearing as a unique peak leading to an increase in sensitivity both at the MS and MS/MS level with no increase in mass spectra complexity [87, 92]. On the other hand, it is now known that the reporter ions of isobaric tags are prone to ratio compression, meaning that together with the target precursor ion some contaminating near-isobaric ions can be co-isolated and fragmented, contributing to reporter ion intensity and biasing of the quantitative information [92, 93]. This fact leads to a ratio compression around the unit, because when reporter ion intensity has interference from reporters coming from peptides derived from proteins with unchanged expression, the ratio between the two samples tend to be 1 [93]. To overcome this drawback, an MS3 strategy has been developed [93], as well as its combination with synchronous precursor selection (SPS) [93], and although with these strategies the accuracy and precision is enhanced it comes with the cost of a reduction in the number of proteins quantified [11].

Some comparative studies have been performed between the different isobaric methodologies, and when comparing 4-plex with 8-plex iTRAQ, the latter led to more consistent ratios without compromising peptide identifications [91]. On the other hand, in another report, when comparing TMT 6-plex with the two versions of iTRAQ, the 4-plex iTRAQ performed better in terms of peptide identifications and similarly in terms of precision of peptide-spectrum matches [94]. These discrepant results may be due to the use of different equipments and softwares for data analysis. [10]

The amine reactive tags were the first ones to be developed and are also more commonly used. However, new molecules have been developed to label other protein residues or PTMs. Both methods have been adapted for these applications, TMT has been adapted for iodoacetyl cysteine-reactive tandem mass tags (iodo-TMT) to identify and quantify S-nitrosylated peptides [95], carbonyl-reactive TMT (glyco-TMT), which may be used with two different chemistries, either aminoxy-TMT or hydrazide-TMT, and enable quantification both at the MS1 (coded with isotopes) and the MS/MS (coded with isotopic reporters) [96]. iTRAQ has also been adapted for the detection and quantification of carbonylation of proteins by means of functionalizing the iTRAQ molecule with hydrazine (iTRAQH) [97], for phosphoproteome identification and quantification (phospho-iTRAQ) [98], and also for identifying new N-termini generated by proteases in a strategy combining iTRAQ with terminal amine isotopic labeling of substrates (TAILS) [99, 100].

TMT and iTRAQ technologies have been extensively used by the scientific community to answer several biological questions and applied to almost all types of samples. In neuroproteomics, these approaches have been extensively used to characterize the differential proteomes of neuronal disorders, drug responses or brain regions. Many studies have been performed using these techniques in areas such as neurodegenerative disorders, as the study of the putamen proteome of an MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) monkey model of Parkinson's disease [101] or in the serum of Parkinson's disease patients [102], or even in the analysis of synaptosomes from cortical brain tissues from Alzheimer's disease patients [103]; in neuropharmacoproteomics, as in the examples of a study of protein quantitative alterations induced by antidepressants in the hippocampus of mice [104]; also in addition as in the evaluation of the effects of administration of plasminogen activator after ischemic injury in mice [105] or the alterations upon chronic exposure to cocaine [106]; in neuropsychiatric and other CNS disorders, such as schizophrenia, with the study of protein expression in the thalamus and CSF of patients [107] and a study of neurofibromin knockdown PC12 cell line as a model of neurofibromatosis [108]; or even in studies of neuronal function such as memory formation in hippocampus [109].

Once, these commercially available isobaric tags were expensive and laborious to produce, in 2010 two new isobaric approaches were proposed, *N,N*-Dimethyl Leucines (DiLeu) [110] and deuterium isobaric aminereactive tag (DiART) [111], which should serve as cost-effective alternatives to iTRAQ and TMT [10].

DiLeu was developed inspired by the chemical isotopic labeling by formaldehyde dimethylation of lysines [112], which is an inexpensive approach, and the aim is to combine it with isobaric labeling and quantitation at the MS/MS level [110]. This way a 4-plex set of dimethylated leucines for amine groups labeling was developed, and has a structure similar with the other isobaric approaches, with an amine-reactive group, a balance group, and a reporter group (115–118 Da) [110]. DiLeu has a labeling efficiency similar to iTRAQ and generates reporter ions with higher intensity; nonetheless, this approach requires an extra step of activation of the reagents prior to the labeling reaction because it uses a different chemistry [10, 110], and is also prone to the co-isolation of precursor ions (as iTRAQ and TMT). Recently, DiLeu was used to test if the implementation of ion mobility MS would mitigate this phenomena [113].

The DiLeu strategy has already been applied to study the neuropeptidome of a crustacean species [114], and for relative quantification of amine-containing metabolites [115]. A 12-plex DiLeu strategy has been introduced that takes advantage of changing isotopologues in the reporter groups, similarly to NeuCode or TMT 10-plex [116].

DiART was designed as a less expensive 6-plex isobaric labeling reagent to label amine groups of proteins and peptides and is, once more, based in a very similar structure as iTRAQ and TMT using an amine-reactive group, a balancer group, and a reporter group in the mass range of 114–119 Da. [111, 117] In a study comparing DiART and iTRAQ, the authors found that DiART leads to more intense reporter ions and consequently less ratio compression, however with the DiART approach, the common fragmentation method is not advisable due to easy

reporter ion fragmentation [118]. DiART has also proven to be compatible and valuable for PTM analysis as quantitative phosphoproteomic studies [119].

Although isotopic and isobaric techniques are based in different methods of quantification and have strengths and drawbacks, both have proven to be valuable for quantitative proteomics [11] and the combination of several methods has been applied to increase throughput of the analysis. This combination is called hyperplexing, because it enables the simultaneous analysis of a higher number of samples, such as with the combination of metabolic 3-plex labeling with isobaric 6-plex TMT that enables the analysis of 18 samples [120], also it is expected that by combining different strategies an even higher throughput and more reproducible results will be achieved [10]

3. Label-free approaches

As an alternative to the labeled methods, several label-free approaches (Figure 4) have emerged, some of them with comparable accuracy to the labeled methods and all of them with similar or higher proteome coverage and dynamic range [121, 122]. These methods gained popularity mainly due to their low cost, their simple sample preparation, the unlimited number of samples that can be compared, and their multiple applications [121]. These attributes turn label-free methods into a powerful technique for clinical applications and large screenings. However, as samples are analyzed separately, these types of methods are highly dependent on run-to-run reproducibility, therefore sample preparation and analyzes should be well implemented and standardized. Furthermore, the methods rely also on the software capacity for both data extraction and capacity to accommodate errors [123, 124].

In general, label-free approaches can be divided into two distinct groups according to the method used for data extraction. On one hand, the quantification can be inferred by counting the number of peptides or spectra assigned to a given protein, and therefore are generically called spectral counting methods. On the other hand, when liquid chromatography is coupled with mass spectrometry, quantitative values can be measured through the extraction of the area of the precursor ions' chromatographic peaks - area under the curve (AUC) or MS1 signal intensity methods. [121-123]

Traditionally, label-free methods were associated with the commonly used shotgun approaches, where mass spectrometry instruments operate in a data-dependent acquisition mode (DDA, also called information-dependent acquisition or IDA) (Figure 4A). Therefore, these methods have also the advantage of being used in data previously acquired for protein identification [125, 126].

In this type of experiments, the instruments are set to scan the precursor ions followed by the selection of a limited set to be fragmented, usually the most intense ones. The fragmentation spectra (MS/MS spectra) obtained will then be used for peptide identification. Independently of the method used to extract quantitative information, the mass spectrometers working on IDA mode must be fine-tuned in order to acquire enough data to perform both the identifica-

tion and the quantitative analysis [127]. This is particularly important for MS1 quantification methods, where enough points per chromatographic peak to perform an accurate extraction should be acquired, without misplacing the acquisition of good fragmentation spectra that allows peptide's identification. Although this balance is not so crucial for the spectral count methods, it is also important to have a good balance between survey and fragmentation scan in order to be able to achieve a higher proteome coverage. Therefore, the development of mass spectrometers with faster scans combined with higher resolution power has been fundamental for the increase in the use of label-free approaches [122, 125].

Label-free methods still rely on peptide identification, the IDA experiments tend to be biased to the most abundant proteins and are highly affected by sample complexity/dynamic range. Therefore, the use of data-independent acquisition (DIA) methods, where fragmentation spectra is acquired for the entire sample without any pre-selection of precursor ions, soon started to be used for label-free quantitative approaches as an alternative to the limitations of IDA experiments [122, 126].

Finally, although label-free approaches are mainly a method for relative quantification (Figure 4B), several groups have also taken efforts to evaluate the relationships between label-free measurements and absolute quantification (Figure 4C) of proteins in complex samples. And in fact, several adaptations came out as good correlations between label-free measurements with protein concentration, allowing the use of label-free methods for the determination of the absolute abundance of a protein [122, 128].

3.1. Spectral counting-based label-free methods

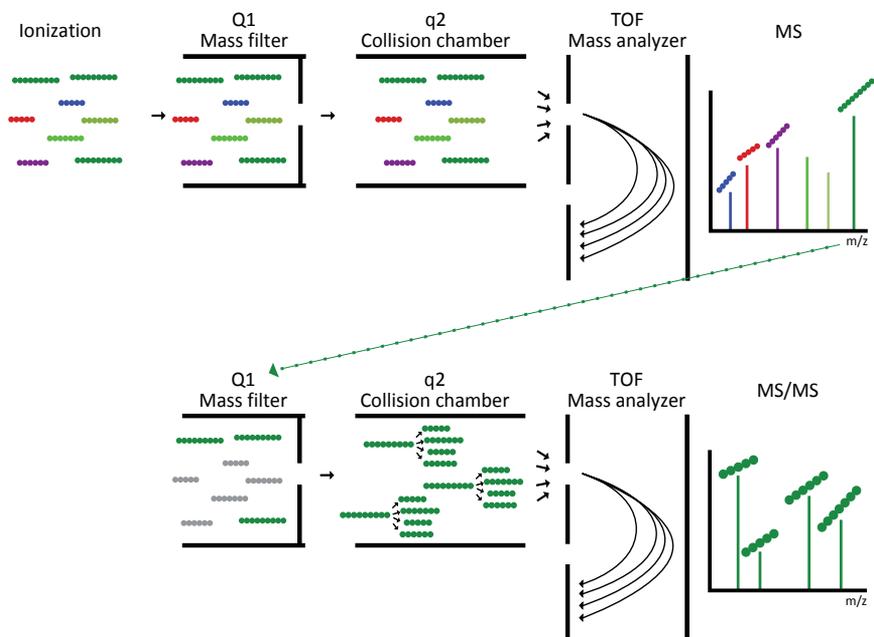
Spectral counting methods consist of simply counting of the number of peptides and/or fragmentation spectra of a particular protein, and comparing the value between conditions. Within this group of label-free methods, it is possible to distinguish some different types: 1) those that are based on unique peptide counting; 2) those based on MS/MS counting (SpC); and finally, 3) an adaptation of spectral counting, spectral TIC counting (MS2 TIC) [132].

3.1.1. Peptide counting and Spectral Counting (SpC)

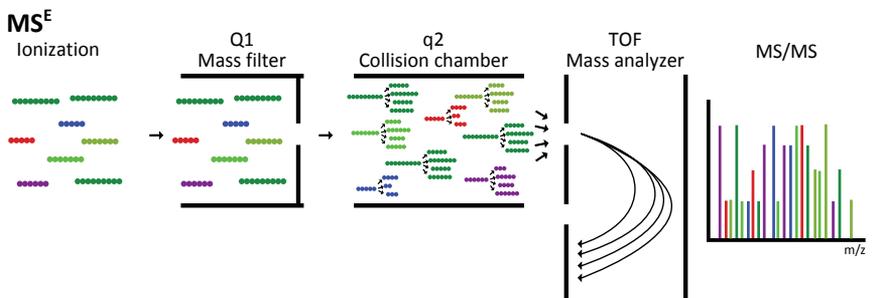
The correlation between the number of peptides acquired in an IDA experiment with the protein abundance was firstly reported in 2001 by Washburn and colleagues [133]. In this work, the authors used the codon adaptation index (CAI) as a measurement of the protein abundances, and correlated CAI ranges with the number of proteins identified and the number of peptides identified per protein. CAI relies on the evidence that mRNAs of highly expressed proteins preferably use some codons (those of which the tRNAs are present in the greatest amounts) rather than others specifying the same amino acid [134], and at that time it was already proved to correlate well with protein levels [135]. With this assessment, Washburn and colleagues were able to note that the most abundant proteins were identified with multiple peptides, while for the low abundant proteins the identification was achieved based on one or two peptides. Although no special focus at the quantitative level was performed, this evidence would be the basal principle of the spectral counting approaches [133].

A-Acquisition modes

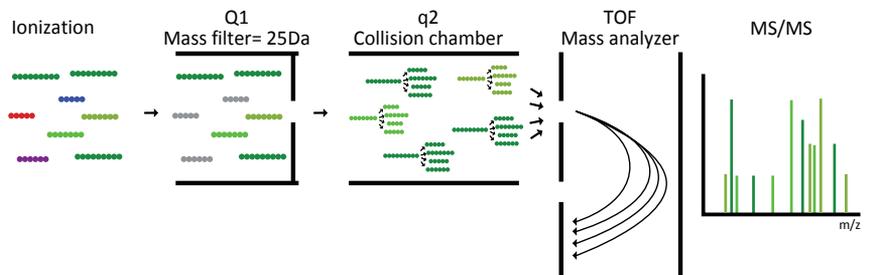
IDA



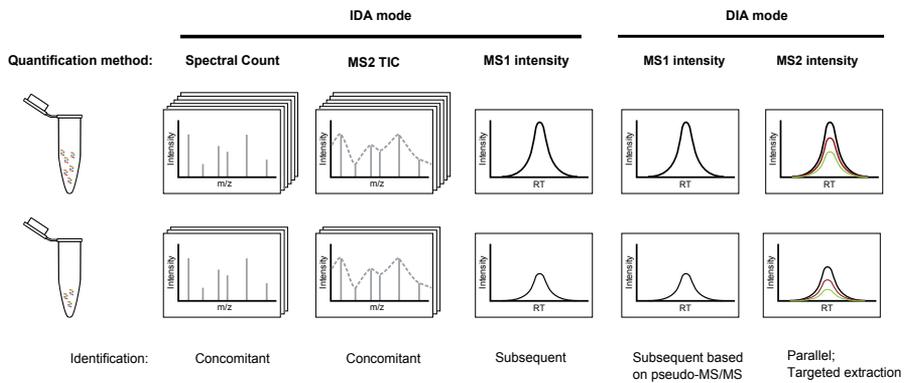
DIA



SWATH-MS



B-Relative Quantification



C-Absolute Quantification

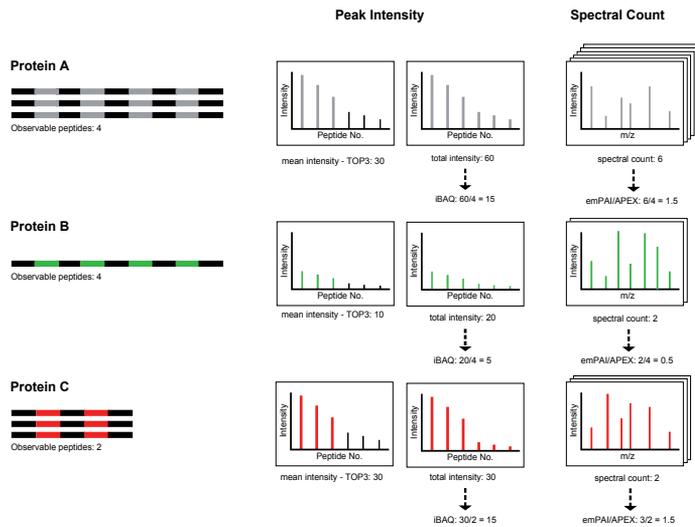


Figure 4. Overview of the label-free MS-based quantitative methods, instrumental principles and data analysis. (A) Comparison of the MS instrumental principles of the acquisition modes most commonly used in label-free approaches: 1) Information Dependent Acquisition (IDA) where fragmentation spectra are only acquired for a group of selected precursor ions based on their intensities; versus 2) Data Independent Acquisition (DIA) where fragmentation spectra is acquired for all the precursor ions independent of its intensity. Fragmentation spectra can be acquired for the entire mass range simultaneously (MS^E) or by covering the mass range in sequential smaller windows of defined size (SWATH-MS). (B) Schematic representation of the different label-free approaches for relative quantification. In the spectral-counting (SpC) approach, peptide/protein abundances can be estimated based on the number of identified MS/MS spectrum. In the MS2 TIC approach, peptide/protein abundances can be estimated based on the mean of the TIC (sum of all the fragments in a given MS/MS spectra) off all the identified MS/MS spectrum. In the precursor ion-intensity-based approach (for both IDA and MS^E method), the changes of peptide/protein abundances are determined by measuring and comparing the chromatographic peak areas of the corresponding peptides. The changing peptides

are subsequently identified based on the respective MS/MS spectra (IDA) or a recomputed pseudo-MS/MS spectra (DIA). In the SWATH-MS approach, changes in confident peptides/proteins are determined based on the fragment ion intensities (MS2 intensity), designed as peak groups of each previously identified peptide. In this example, results would indicate a higher peptide abundance in State A. (C) Representative examples of label-free methods for absolute quantitative proteomics. In the case of the strategies based on MS1 intensities, the average of the three most intense ions (TOP3) and the iBAQ index are used to generate reliable absolute quantitative data. In the strategy based on spectral count, both emPAI and APEX strategies used the number of identified peptides normalized for the expected number of peptides (to reduce the impact of protein size) as an indicator of the protein abundance. As an example, proteins A and C, present at the same abundance, have different spectral counts but they present the same normalized spectral count. Adapted from [129–131].

At the end of the same year, the first quantitative report based on the spectral counting principle was published by Pang et al. [136]. In this work, the authors introduced the concept of peptide “hit” (now known as peptide hits technology or PHT [137]) as a measure to estimate the relative changes in protein abundance. In this method, each hit corresponds to one identified peptide and the protein abundance is calculated by summing all the hits. The method assumes the principle that the coverage of the protein increases in proportion to the protein abundance, which is reflected in the number of peptide hits of a given protein. In the same report, the authors applied this quantitative method to the identification of biomarkers for inflammation in urine samples of healthy vs. disease conditions, and performed a comparison between the proposed approach and the usual quantitative 2D-gel approach. Similar quantitative results were obtained between the methods studied, with a significant increase in the number of the proteins analyzed in the gel-free approaches combined with a significant reduction in the required amount of sample and sample processing [136].

In 2003, Gao et al. [138] applied for the first time a statistical method (Student’s t-test), already widely used for gene array experiments, in peptide hits quantitative data in order to quickly assess with statistical significance the abundance changes^o between treatments/conditions. The use of such method into quantitative proteomics was evaluated in a widely used biological system by performing a comparison with the results obtained in previous reports, revealing a high degree of concordance. Therefore, the use of such statistical evaluation can quickly highlight the proteins that are in fact altered from the entire data set of proteins analyzed in larger screenings, turning the data analysis into a more automated and reliable method [138].

After the initial report using peptide hits as a quantitative measurement of protein levels [136] and following the same principle stated in that work, some adaptations to that quantitative method started to appear in order to take into account the protein characteristics that could influence the results. Matthias Mann’s group was a pioneer in the development of such adaptations, with the first adaptation appearing in 2002 by Rappsilber and collaborators [139]. In this work, the authors characterized the human spliceosome by an exhaustive identification of the constituents of that multiprotein complex, and by obtaining the relative abundance of the different classes of proteins involved. In order to do so, the authors presented a new method to quantify protein levels, the protein abundance index (PAI), which consists of the number of MS/MS spectra identified divided by the number of theoretically observable peptides, i.e., the theoretical peptides that will feat in the mass range of MS [139]. By considering the theoretical number of peptides that can be formed from a given protein, the authors compensated the impact of the protein size, since larger proteins can give rise to more peptides within

the MS mass range. However, once the authors considered all MS/MS spectra that originated positive identifications from peptides acquired with different charge states to modified peptides, the measured values also reflect the response of a given protein to the measurement procedure and not only its abundance.

Soon, label-free approaches being performed in comparative screenings and some alternative methods based on the principle stated above started to emerge. At the same time, two independent studies focused on the proteome changes observed in the development stages of the human malaria parasite *Plasmodium falciparum* were published presenting two alternative methods to evaluate these proteomics changes. While Florens and collaborators [140] compared the protein sequence coverage between the development stages to estimate protein relative abundance, Lasonder and colleagues [141] used the total number of unique peptides identified and introduced the use of the extracted ion chromatograms (XIC) of individual peptides as a method to confirm the absence or presence of a particular protein. With the use of the MS-XIC evaluation, the authors overcame one of the limitations of IDA experiments where it is possible that a peptide is not selected for fragmentation in a particular sample due to changes in sample complexity [141].

Another disadvantage of spectral counting is that the length of the protein influences the number of theoretical peptides that can be produced from tryptic digestions [142, 143]. Therefore, in order to overcome this limitation, several modifications were proposed to take into account the protein size [121]. The most widely used is the normalized spectral abundance factor (NSAF), proposed by Zybailov in 2006 [144], which consists of the normalization of the SpC of a given protein by the protein length (L). These values are further normalized by the sum of the SpC/L for all the proteins analyzed, thus taking into account the experimental variation. Furthermore, this method presents a high dynamic range (~4 orders of magnitude) and is able to measure smaller variations (lower than 50% variation) [144]. This method was revised by the same group, presenting an improved NSAF approach that is able to deal with peptides shared between proteins and the distributed normalized spectral abundance factor (dNSAF) [145].

The use of shared peptides for quantification has been a critical issue since the abundance of a peptide that is shared across proteins depends on the contributions of the multiple proteins to which it belongs [127, 146]. Therefore, it is incorrect to overestimate the protein abundance by counting the shared peptides multiple times, typically these peptides are simply ignored in protein-level quantification analysis [147]. However, this may significantly decrease the number of proteins for which it is possible to estimate its abundance (as much as 50%) [146]. Thus other approaches have been used to include these peptides. Some approaches try to assign the shared peptides for a particular protein (the most abundant of the group) by taking into account parameters such as the number of unique peptides to calculate the relative abundance of each protein [148, 149]. dNSAF is perhaps the most known example of such type of adaptation [145]. Finally, some authors also proposed to analyze the proteins that have shared peptides as a protein group and not individually. However, these proteins can present different regulatory mechanisms, therefore their combination fails to estimate the real variation [150].

3.1.2. Spectral TIC (MS2 TIC)

In 2008, Asara and colleagues [142] presented a new method for relative protein quantification that could be considered an extension of the spectral counting technique. In this approach, the average of the TIC for all of the MS/MS spectra that identified a protein was used as a quantitative measure. Each spectral count gets a unique abundance value, which consist of the sum of all the fragments in a given MS/MS spectra, instead of being just counted as one event. In this study, the authors proved that this “spectral TIC” method was effective and expanded the dynamic range of quantitative ratios allowing for larger protein abundance [142]. This would allow to overcome one of the limitations of the spectral counting, its intrinsic tendency to easily reach the saturation for the most abundant peptides, not being able to quantify properly large protein ratio differences, and limiting the dynamic range of the method [122]. In this approach, the authors counted all the MS/MS spectra that resulted in positive identification and the average was used, instead of the sum of the TIC, in order to overcome the sampling bias caused by different protein molecular weights (larger proteins generate more tryptic peptides than smaller proteins). The proposed method was tested by evaluating its capacity to reach the theoretical ratio of a known digestion mixture, and comparing it with other quantitative methods already well established. With this comparison, the authors showed that the spectral TIC has a similar accuracy to the AUC methods and is able to correctly calculate large variations [142] and detect relative changes in low abundance proteins [151].

This method had some improvements; it was combined with data from the SpC method in order to obtain a better characterization of the samples [152], also Griffin and collaborators proposed a new normalized label-free method that combines the three MS abundance features, namely the peptide and spectral counting with the TIC intensity [153]. This method, termed normalized spectral index (SI_N) combines the reproducibility already presented by spectral counting methods with an increase in the accuracy of the determination of protein abundance observed in TIC intensity methods. Furthermore, by correcting it for protein length, it also reduced the samples bias to large proteins [153].

3.2. MS1 signaling intensity or Area Under the Curve (AUC)

Bondarenko and Chelius [154, 155], in 2002, were the pioneers of the use of MS1 signal intensity as a measurement of protein levels. Bondarenko, in his technical work, tested the hypothesis that peak area of the peptides should reflect its concentration and therefore those peak areas should correlate with protein concentration. To test that, different amounts of a pure protein were analyzed, alone or spiked in a complex mixture, and the extracted peptides' areas were compared with peptides' concentration, revealing a high degree of correlation even in samples with high complexity. Furthermore, the authors also proposed the use of a correction factor designed as experiment-dependent correction factor that aimed to reduce the impact of some experimental parameters, such as differences in sample preparation, that could lead to some bias of the results. The use of such correction factor, which is determined from the mean tendency of the non-variable proteins, proves to improve the accuracy of the quantification [155]. Therefore, the use of normalization methods became a key feature in label-free quantification, and several alternatives have been proposed. Those alternatives can be divided into

two groups based on their basic principles. On one hand, some normalization methods are based on the principle that a large portion of the proteome does not change, therefore the mean tendency between experiments can be used to accommodate some experimental deviations. On the other hand, the normalization for housekeeping proteins, a protein or set of proteins known to be constant, or for an internal standard added to the samples before sample processing can be used since both will reflect the effect of sample processing [155, 156].

The MS1 intensity label-free methods are highly accurate since they require the use of high-resolution mass spectrometer in order to be able to distinguish the co-eluting species [121, 152]. Protein quantification based on AUC requires the comparative measurement of precursor ions intensity at a particular retention time, therefore this type of quantitative methods is also dependent on the power of data extraction algorithms, and several different methods are already available [121, 125, 130]. Independently of the software used, the data analysis of MS1 intensity peaks generically comprises a set of defined steps: feature detection, alignment of retention times, peak picking, noise reduction, and normalization of MS intensities [121, 130]. The detected and normalized peaks are then compared between the samples and their MS/MS spectra are used for protein identification [121]. The estimation of the protein abundance can be obtained mainly by three different strategies: by summing all the peptides considered in the analysis; by performing the mean of all the peptides; or considering only the 3 most intense peptides (usually using the mean value), the so-called TOP3 method [157, 158].

Since quantification is done at the MS1 level, the estimation of protein abundance is not dependent on the acquisition of a particular MS/MS spectra in all the experimental conditions. In fact, a given peptide can be identified in a single sample and quantified across all the remaining samples [141]. Thus, these methods are not so prone to the variability associated with variation of sample to sample complexity. This characteristic, associated with the unlimited number of samples to be compared, enables MS1 intensity methods as suitable methods for clinical biomarker discovery, which normally requires high sample throughput [125].

Due to the large number of modified methods and the generalized use of the terms spectral counting and peak intensity to include all the modifications, it is not always clear which particular method was used in a given experiment. Therefore, to simplify the categorization, the reports are commonly grouped in these two generic categories, taking into account only their basic principles.

The use of spectral counting and/or MS1 intensity methods in neuroproteomics is vast and usually alternates between the use of one type or another [9, 159, 160]. However, some reports combine the two methods to improve the results obtained, such as in the case of the interactomics study of the AMPA receptor performed in collaboration with our group [161] where both spectral counting and MS1 intensity were used to identify the truly positive interactors from the negative control. Within the areas where label-free methods are being used, it is possible to identify some general studies focused on understanding proteomics changes in brain regions, such as the evaluation of frontal cortex changes caused by frontotemporal lobar degeneration (FTLD) [162]; cell- and tumor-specific alterations, such as the comparison of astrocytes and astrocytoma proving evidences for the existence of important membrane

biomarkers capable to define the cell lineage of the tumor [163]; and characterization of protein architecture of secretory vesicles, key components that mediate intercellular signaling [164]. Although, regarding label-free approaches, the neuropsychiatric field is dominated by the use of MS^E, in the study of neurodegenerative diseases (such as Parkinson's, Alzheimer's, and Huntington's diseases) there is an evident tendency of spectral counting/peak intensity methods [9] in both the analysis of cell and animal models (both obtained by the use of chemical and genetic alterations) [165–167] and also CSF [168] and postmortem tissues (mainly in the case of Alzheimer's disease) [169, 170]. Those proteomics screenings led to the identification of several deregulated proteins, contributing to an increase in the understanding of the pathways that are altered in those disorders.

There is an inherent tendency for a larger number of spectral counting reports [162–165, 167, 169, 170] when compared with the peak intensity reports [166, 168], as observed from the examples stated above. This underestimation of peak intensity reports from IDA experiments is associated with the preference from the alternative peak intensity methods based on DIA acquisition (such as MS^E). Although in reduced numbers, there are also some reports on the use of MS2 TIC methods, more specifically SI_N, in the neuroproteomics field. As an example of its applicability, there are two studies involving brain tumors. In one study, the authors performed a characterization of the differentiation states of glioblastoma stem cells (cells responsible for tumor formation and growth [171]), in the other study the authors were focused in the analysis of the secretome of glioma cells in order to identify the proteins that could be involved in tumor cells migration [172].

Further quantitative neuroproteomics studies were already summarized in several reviews [9, 159, 160].

3.3. Data-independent acquisition methods

As stated above, in order to overcome the limitation of the use of IDA modes, alternatives with DIA are starting to be used. The major advantage of this acquisition mode relies on its ability to record fragmentation spectra from the entire set of precursors of a given sample, without any selection that can bias the acquired data. However, in these experiments, the data analysis is very challenging, since the link between the precursor and its fragments is lost [173], these methods are highly dependent on the development of algorithms capable of extracting valuable information from the data acquired [174].

These methods operate in a cyclic mode, throughout the entire liquid chromatography (LC) time range, by alternating between survey and fragment ion spectra. Generically, these methods can be divided into two distinct groups, those that acquire the fragmentation spectra of the entire mass range simultaneously, and those that scan the *m/z* range in sequential isolation windows of different widths. The use of sequential isolation windows is a way to reduce some of this complexity, by decreasing the number of concurrent ions being fragmented at a given moment [173, 174].

Usually in DIA experiments, the quantitative information is still obtained from the precursor ion signal, while the fragmentation spectra are mainly used for peptide identification by both

the use of common tools developed for DDA, or by searching pseudo MS/MS spectra reconstituted based on co-elution profiles of precursors and their potential fragments [174].

Several DIA acquisition methods were developed based on the use of different mass spectrometers and/or different dissociation methods (see Table 1) [173, 174], however, within this chapter only the most used method, LC-MS^E, and the SWATH-MS method will be presented.

Method	Instrument	m/z selection window (width)	Dissociation method	Ref
Shotgun CID	Q-ToF	Full m/z range	CID in-source	[175]
Original DIA	Ion Trap	10 m/z	CID in-collision cell	[176]
MS ^E	QqTOF	Full m/z range	CID in-collision cell	[177]
p ² CID	Q-ToF	Full m/z range	CID in-source & collision cell	[178]
PAcIFIC (Precursor Acquisition Independent From Ion Count)	Ion Trap	2.5 m/z	CID in-collision cell	[179]
All Ions Fragmentation (AIF)	Orbi-Trap	Full m/z range	CID in HDC collision cell	[180]
XDIA or DIA-ETD-CAD	IonTrap-ETD-CAD	20 m/z	ETD in cell	[181]
SWATH	QqTOF	25 m/z	CID in-collision cell	[173]
Fourier transform-all reaction monitoring (FT-ARM)	LTQ-FT or LTQ-Orbitrap	12 m/z or 100 m/z	CID in-collision cell	[182]

Table 1. List of DIA methods (adapted from [173, 174]).

3.3.1. Liquid Chromatography-Mass Spectrometry Elevated energy (LC-MS^E)

LC-MS^E was the first label-free method from DIA used in proteomics quantitative screening. This method is based in the neutral loss acquisition mode and was first reported in large datasets by Wrona and collaborators in 2005 [183] as a “all-in-one” analysis for metabolite identification. This method was further transposed to proteomics studies, mainly supported by QqTOF instruments [177, 184]. MS^E consist of the acquisition of samples in two alternate modes, first samples are acquired in a low energy mode to collect precursor ions masses (MS precursor scan) and then in a high-energy mode to induce the fragmentation of the entire samples and acquisition of all the product ions (MS/MS scan) [184]. Over the years the coupling with the continuous development of MS and LC systems (more specifically, the use of UPLC-MS^E), more reproducible and accurate quantification has been achieved. However, as an inherent issue of DIA experiments, a large amount of data acquired remains unused, therefore a considerable effort has been done in order to obtain algorithms capable to extract more information from the acquired data than that already available [174].

3.3.2. Sequential Window Acquisition of all Theoretical Fragment-Ion spectra (SWATH-MS)

In 2012, Gillet and collaborators [173] presented the SWATH-MS method, although at that time other DIA methods were already widely used. It was a method that was particularly innovator due to its proposed data extraction methodology. Here, the authors proposed a targeted data extraction by combining parallel analysis of samples with an optimized IDA method for peptide identification followed by a DIA acquisition to be used to extract quantitative information. From the IDA method, a list (called "library") containing all the information regarding a given identified peptide (such as RT, precursor m/z , and MS/MS spectra) was obtained and it was further used to extract the XICs of the specific fragment ions (called peak groups) from all the high confidence peptides identified. Thus, instead of using the precursor intensity as performed by the other methods, in SWATH-MS the use of MS2 signaling intensity-based method was introduced, which is similar to the quantification already performed in MRM and PRM experiments, for the untargeted analysis of large fractions of the proteome. Furthermore, the authors also showed that with SWATH-MS, it was possible to achieve similar reproducibility and accuracy as for the targeted methods for protein quantification [173].

For the acquisition of the fragmentation spectra of virtually all the precursor ions present in a sample, the mass spectrometer, a high-resolution Triple-TOF instrument, operates in the sequential isolation window acquisition principle introduced by previous DIA studies [174, 176]. By fractionating the sample in SWATH acquisition windows, this method leads to a reduction of the concurrently fragmented precursors and consequent reduction of the acquired MS2 spectra complexity.

As data extraction is performed by targeting the peptides already identified, the loss of precursor-fragments linkage is overcome, and a large percentage of data is effectively used. Furthermore, this targeted data extraction also allows that additional criteria, such as the transition intensity ratio, m/z error, and similarity to the identified MS/MS spectra, can be used in combination with the usual chromatographic criteria to evaluate the confidence of the peak group formed. Therefore, protein quantification is obtained from a more reliable extracted data [185].

The SWATH-MS method seems to be able to overcome the majority of the limitations of label-free methods, it is unbiased, presents a broad range of precursor ion fragmentation (covering almost the entire mass range usually analyzed), and it relies on targeted data extraction [173], thus making this method a promising strategy to be applied in large screenings, such as the discovery of biomarkers [9, 129, 186–188]. Although, being a very recent methodology, the great expectation regarding its application into the biomedical field is reflected in the several improvements already achieved into the different domains associated with this method. There are already improvements in the DIA acquisition mode with the introduction of the variable windows mode where windows with different widths are adjusted to the number of precursor ions per m/z range, thus the number of concurrent ions are reduced in the most populated regions. Moreover, several different groups, have been working on the improvement of sample preparation and library creation to increase the number of proteins quantified per sample, as well as to obtain more reproduc-

cible data [189–193]. Finally, different algorithms were also developed to address SWATH data, both in the targeted mode [194] and untargeted mode, which is mainly focused on performing protein identification directly from the SWATH data [195].

The introduction of the concept of a protein library that can be used to interrogate multiple samples has opened the door to the idea of having cell-, tissue-, and species-specific libraries containing exhaustive lists of identified proteins capable of covering the entire proteome. Those libraries can then be used in both research and clinical fields to extract larger quantitative information from the analyzed samples. Within this scope, Aebersold and co-workers have already published the first repository with 10,000 human proteins that claims to successfully detect and quantify 50.9% of all human proteins [196]. Furthermore, as the SWATH file of a given sample corresponds to the MS/MS spectra signature of that sample, that file can be interrogated any time it is required without the need to re-analyze the sample. Therefore, with these SWATH files it is possible to create a repository of samples that can be used in longitudinal studies [129].

As stated above, the use of DIA in proteomics is recent and is not a common option, therefore, this overview in the neuroproteomics field will be done for MS^E, which is the most used method, and also for SWATH-MS due to the exponential increase in the interest and development associated with this approach.

MS^E is perhaps the most used large-screen, label-free method, particularly in the neuroproteomics field, and the only DIA method that has gained enough visibility so far [9, 197]. Although MS^E was also used in different neuroproteomics areas, such as in the studies of frontotemporal lobar degeneration [198] and the profiling of phosphorylation events in different rat tissues, including the brain [199], its use was particularly potentiated by the Sabine Bahn group for the study of neuropsychiatric diseases, such as schizophrenia, major depression, and bipolar disease [160]. In general, their published works were mainly focused on differential analysis of human samples, both serum [200, 201] and postmortem tissue [202–204], from patients versus healthy controls, or including different disease groups or groups with different levels of antipsychotic medication. Those works aimed to identify differentially altered proteins that could distinguish between the disease groups, but could also contribute to a better understanding of the diseases. In fact, the authors were able to identify several different proteins that are altered between schizophrenia patients versus controls, including proteins altered in first-onset paranoid patients [201], and observed also some proteomics alterations that were dependent on the dose of antipsychotic medication [204]. Finally, to distinguish the effect of the medication from the disease alterations, Sabine Bahn's group also studied the modifications caused by some of the antipsychotic drugs in rat frontal cortex, being able to identify proteins altered by the medication, some of them altered in both types of medication used [205]. More recently, MS^E was also used to perform proteomic profiles of the first episode of major depressive disorder patients and sex-specific alterations of adults diagnosed with Asperger syndrome [206].

Being a very recent method, SWATH-MS reports are mainly associated with technical improvements, and in demonstrating its capacity to obtain large proteomics profiles with its potential use to clinical studies and biomarkers discovery, such as the study of plasma PTMs

as phosphoproteins [187] and glycoproteins [207], large screening of twins [208], and human library creation [196], and also its applicability in biopsies specimens [188]. In the neuroproteomics area, it is already possible to find some reports, such as the work published by our group [209], where we presented a pipeline for reproducible quantitative screenings using a membrane-enriched sample from rat cortex, indicating that our approach is suitable for evaluation of membrane proteins, key players in the majority of neuronal dysfunctions. There are also two other works from Fox's group regarding mitochondrial alterations: one of them corresponding to an exhaustive characterization of mitochondrial proteome from embryonic and postnatal rat brain revealing a rearrangement of proteins from glycolysis and mitochondrial trafficking/dynamics, which may suggest a development change to accommodate the required energy demands in different developmental stages [210]. Another study focused on mitochondrial functional alterations associated with deregulation of PTEN-induced kinase 1 (PINK1), a Parkinson's disease-associated protein [211].

3.4. Absolute quantification based on label-free approaches

Although the majority of the screenings are based on relative quantification, some authors started to focus on the possibility to also extend these methods to absolute quantifications [128], since the calculation of the protein abundances in a sample is essential to increase the understanding towards the biological systems and its variations [177, 212]. Overcoming the elevated cost and demanding sample preparation of an isotopic dilution-based method to perform absolute quantification, the use of label-free techniques reveals a reliable alternative (although less accurate than the referred methods). The available methods can be divided into two generic classes based on the quantification algorithms used: 1) those based on tandem MS data, e.g., protein sequence coverage or spectral counting including emPAI [213] and APEX [214]; and 2) those based on the measurement of precursor ion intensity such as MS^E [177], T3PQ [157], and iBAQ [215].

In general, all these techniques were described as having good correlation with protein amounts in both simple mixtures of proteins with known amounts (alone or spiked in complex samples) and for unknown proteins in complex samples. When complex samples were used, the accuracy of the results were confirmed by comparing the values achieved within several different techniques, such as other mass spectrometry-based quantitative methods (including isotopic labeled methods), transcriptomics analysis, and ELISA [157, 177, 212–215]. In all cases, a proper estimation of the protein abundance was achieved with or without standards.

The more cost-efficient and easier option is to exclude the standard proteins and calculate protein abundances from the fraction of each protein in the total protein pool assuming that most of the proteins that contribute to the total protein pool are identified and quantified. As examples of this quantification without standards, it is possible to find the determination of the copy number (using total protein approach (TPA) [216]) and the definition of the stoichiometry of protein complexes [217]. However, the quantification accuracy can be increased by using a standard curve from a mixture of proteins with known amounts that have different sizes and concentration [212].

Additionally, the majority of these methods are already implemented in several tools available for proteomics analysis, therefore, it is possible to combine both relative and absolute quantifications in a simple way, as in the case of the emPAI that is implemented in the MASCOT server, one of the most used servers in proteomics [214].

A brief presentation of the most used methods for label-free absolute quantification focused on the major differences between them, and some reports using these methods into the neuroproteomics field are presented below.

3.4.1. Spectral counting-based methods

Exponentially Modified Protein Abundance Index (emPAI): Mathias Mann's group presented what would be the first method for absolute quantification by proving that a transformation of the PAI values (described above) could be in fact associated with the absolute amount of a given protein [213]. In this study, the authors showed that the PAI values have a linear relation with the logarithm of protein concentration, therefore the absolute quantification of a given protein can be obtained by the exponentially modified PAI (emPAI), which is equal to the following equation $10^{\text{PAI}-1}$.

Absolute Protein Expression (APEX): In theory, the APEX method is similar to the previously proposed emPAI method since it is based on the number of peptides identified normalized for the theoretical number. However, instead of considering the redundant peptides, it relies only in the unique peptides. And furthermore, which is also its major strength, this method uses machine learning to calculate the number of theoretical peptides than can be identified in the particular experiment. To achieve the probable number of peptides, the theoretical number of peptides is normalized for a correction factor specific for the experimental settings [212, 214].

3.4.2. Intensity based methods

One of the disadvantages of spectral counting based methods already observed for relative quantification is the fact that in these methods the saturation is easily reached, therefore, failing in accurately quantifying proteins present at higher levels. On the other hand, as those methods rely on the MS/MS spectra identification, they are also biased to the most intense proteins, therefore, spectral counting-based methods are only accurate within a reduced dynamic range. Furthermore, they also present a large variability between replicates. Similar to what is observed for relative quantification with the use of MS1 intensity-based methods these limitations are overcome.

Peak intensity-based absolute quantification method (iBAQ): In this method, the amount of a given protein is calculated by the sum of the peak intensities of all peptides matching to it, divided by the number of theoretically observable peptides [215].

LC-MS^E: Silva and colleagues reported in 2006 [177] for the first time the relationship between MS signal response and protein concentration. In this work, the authors discovered that the average of the three most intense peptides is highly correlated to the effective amount of a protein in a sample. In this study the authors spiked the samples with a known amount of a

mixture of proteins (internal standards). The internal standards were then used to calculate a universal signal response factor (which was shown to be the same for all the tested proteins) that correlates the intensity calculated with the amount of the proteins, and is then used to obtain the quantification for the unknown proteins.

Three most intense peptides peak area (T3PQ): The T3PQ method is an adaptation of the method previously used in LC-MS^E [177] for IDA methods [157]. The principle of this approach relies on the evidence that for each protein (independent on its size) identified by a set of peptides, the average of the three most efficiently ionized peptides (those with the highest MS signals) directly correlated with the amount of the corresponding protein. This method proved to be more accurate and reproducible than the methods already used (in particular, when compared with the spectral counting methods) [157].

Absolute quantification methods have been used mainly in studies focusing on the understanding of complexes stoichiometry, and not in large screenings that are the most frequent assays in the neuroproteomics field. Therefore, there are only few reports using label-free-based absolute quantification, particularly in neuroproteomics, and those are mainly associated with the iBAQ method. One of the more interesting reports where iBAQ was used is perhaps the work regarding the characterization of the isolated synaptic boutons that culminate with the establishment of the amounts of the proteins that compose those vesicles [218]. iBAQ was also used to obtain a comprehensive characterization of the protein abundance in several organs, such as the brain [158], and in some experiments that focus on the determination of the amount of enriched proteins in tissue-specific (hair bundles) [219] and condition-specific (BACE1 knockouts) [220] proteomes.

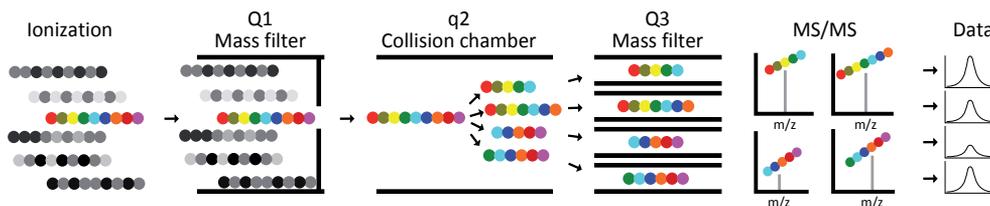
4. Multiple reaction monitoring

Multiple reaction monitoring (MRM) is a highly selective scan mode in MS that has been extensively used for the last 30 years for absolute quantification of small molecules [221]. Similarly, the knowledge acquired in small molecules targeted quantification has been transposed for targeted quantification of peptides and proteins where several reviews can be found in the literature [222–224]. Shotgun proteomics MS-based studies retrieve the identification of thousands of proteins in a single analysis, plus the relative quantification by label-free [225, 226] or isotopic labeled strategies [227, 228]. However, in these global profiling methods, low-abundance peptides may be difficult to be detected, generating “missing data” and low precision problems that can impair statistical analyses [229, 230]. Consequently, the untargeted approach has been widely used, for instance, in clinical studies of biomarker discovery to find new candidates and, the MRM targeted MS-based approach has been used in the verification/validation phase, overcoming many of the difficulties associated with antibody-based protein quantification [231, 232].

Developing and validating MRM-MS assays is a laborious process, but once constructed, it can be used for accurate and precise quantification of one or several proteins on a large scale and across laboratories [233]. The high selectivity of MRM scan mode is achieved using, most

predominantly, triple-quadrupole mass spectrometers. Quadrupoles are known as “mass filters” where in a first stage (Q1), the mass/charge ratio (m/z) of the intact peptide (precursor ion) is selected, fragmented in the collision cell (q_2), and in a second stage (Q3) a specific fragment of the precursor is selected, generating the selected reaction monitoring experiment (SRM) with one transition (precursor/ fragment), or if several fragments are being monitored, an MRM experiment with several transitions (Figure 5A) [234].

A-True MRM



B-High Resolution MRM

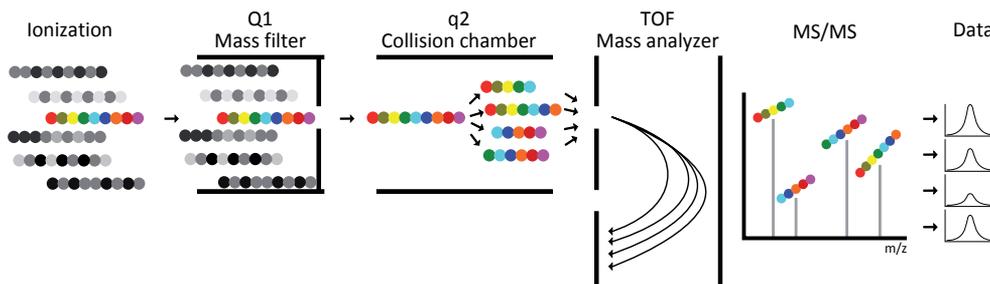


Figure 5. Schematic representation of the (A) MRM scan mode performed in a triple-quadrupole instrument and the (B) high-resolution multiple reaction monitoring (HR-MRM) scan mode performed in a QqTOF system. In classical MRM scan mode, the first quadrupole (Q1) selects the m/z of the precursor that will be fragmented in the collision cell (q_2) and one of the resulted fragments is then selected by the third quadrupole (Q3) towards the detector. The two stages of mass filters (Q1 and Q3) represent a transition and more than one can be monitored in a single run. HR-MRM works similarly at the first stage (Q1) but after fragmentation, all the fragments are scanned by the TOF mass analyzer instead of selecting only one each time that the precursor is fragmented. This will generate a high-resolution mass fragmentation spectrum where extracted ion chromatograms for each fragment can be obtained by the use of specific softwares.

The peptide sequences to be monitored must be carefully selected as they have to be unique for a given protein, where peptides with less than 8 residues and those susceptible to undergo modifications during sample processing (methionine oxidation, cysteine alkylation) must be avoided. Additionally, for protein isoforms or PTM's quantification, specific peptides should be selected for accurate measurements. [229, 235, 236]. The combination of LC separation followed by the MRM acquisition (2 m/z filters) results in high precision, sensitivity, and high selective measurements for the selected peptides and, consequently, for the protein [229]. The best candidate peptide(s) to be monitored for the quantification can be selected based on

prediction tools (in silico) or on experimental evidences [236]. Selection based on empirical data involves previous LC-MS/MS experiments from the biological sample to obtain preliminary information on the peptide characteristics such as ionization and fragmentation. After the selection of the peptides (precursors) and their specific fragments, MRM transitions are evaluated by re-analyzing the sample to help in the selection of the most selective and sensitive for each peptide of interest [237]. In order to avoid long optimization and multiple rounds of analyses, there are online repositories such as PeptideAtlas, the Global Proteome Machine Database, and Pride, which contain peptide sequences and empirical MS spectra to support MRM designing without the need of preliminary sample processing and analysis [229]. Even for proteins not found in the database, there are several in silico bioinformatic tools that select high-responding peptides from candidate proteins, such as ESP predictor [238], PeptideSieve [239], PepFly [240], MIDAS [241], among others. TIQAM is another interesting software tool that selects the proteotypic peptides based on the in silico prediction and integrates that information with the PeptideAtlas repository or other sources to generate the list of transitions based on the validated fragmentation spectrum [242].

The number of proteins monitored by an MRM experiment is usually low and the duty cycle (the time for the instrument to cycle through separation and detection of each transition) will depend on the number of peptides per protein and the number of transitions per peptide. To overcome the limited number of proteins monitored in an MRM experiment, a timed acquisition mode, termed scheduled MRM (sMRM) analysis was developed where transitions are acquired only during a defined elution time window [235]. Consequently, thousands of transitions can be monitored, allowing the quantification of hundreds of proteins in a single run. Colangelo and collaborators developed a pipeline for large scale (>1000 transitions/run), label-free LC-MRM assays for the quantification of 112 rat brain synaptic proteins [243]. The workflow began with data-dependent acquisition using 5600 Triple TOF to identify the sequences of the peptides present in the biological sample of interest. The peptide library information was then converted into thousands of MRM transitions that were easily transposed to the 5500 QTRAP (demonstrating the consistency of the fragmentation patterns between the instruments) to be acquired using the sMRM methods. To address the very short dwell times due to the high number of transitions, they presented an improvement in the sMRM methods' sensitivity and robustness using an intelligence-based MRM acquisition (termed extended or xMRM). Firstly, variable acquisition windows throughout the run can be used and secondly, a "triggered xMRM", where the secondary MRM transition for each peptide was only monitored if the primary MRM exceeded a given threshold. The xMRM enabled the reduction of the number of transitions to be monitored at a given time leading to an increase of 63–68% in the dwell times for peptides and, consequently, an increase in the sensitivity for the limiting peptide concentrations.

Although MRM is considered to be a very high selective scan mode, the possibility to have non-desired peptides with isobaric or very similar m/z values can increase with sample complexity [244]. The consequence of a non-selective method is the overestimation in concentration determination of the targeted peptide. Therefore, the use of HR-MRM can increase the method selectivity and consequently improve the accuracy of the quantification. The scan

mode works as described in the first stage (Q1) and in fragmentation (Q2) for triple quadrupoles with the difference in the last stage, where rather than focusing on a single ion fragment in Q3, fragment ions of all masses are scanned by a TOF analyzer generating high-resolution MS/MS spectra (Figure 5B). Thus, fragment ions can be extracted from the high-resolution MS/MS spectra of the targeted peptides to generate extracted ion chromatograms (XICs) of high resolution [245]. Tong and collaborators performed a targeted HR-MRM analysis for the quantification of 47 tear proteins using the Triple-TOF mass spectrometer (QqTOF) with good reproducibility ($CV < 5\%$) [245]. In addition to the improvement in selectivity, the multiple steps for the selection of the best transitions are not required as in triple-quadrupole instruments.

4.1. Absolute quantification of proteins by MRM

Beyond protein identification and relative quantification by MS, absolute quantification of proteins in biological samples has also been performed using synthetic unlabeled and/or labeled peptides [12]. Absolute protein quantification has been generally performed based on the principle of the stable-isotope dilution (SID) where stable isotope-labeled synthetic analogues are spiked into the samples to extrapolate protein amounts present in a sample. Gerber and collaborators termed this approach as the AQUA methodology, where the best candidate peptides for the quantification of a given protein are synthesized with at least one residue replaced by stable isotopes, resulting in a very similar endogenous peptide (called AQUA peptides) but with a sufficient m/z difference so that they can be distinguished by MRM [246]. Protein quantification is performed by spiking the sample with a known amount of the AQUA peptide and the peak areas ratio of the unlabeled/labeled peptides are used to determine the expression levels of the protein of interest. On the other hand, Barnidge and collaborators performed a study to compare protein quantification using two different methods, one based on the AQUA approach and the other on an external calibration curve created from successive dilutions of the synthetic unlabeled peptide [247]. Quantification based on the external calibration curve resulted in better precision and accuracy values than quantification based on the sample spiking of the analogous labeled synthetic peptide. In this study, the external calibration curve was able to accurately determine the peptide concentration however, for more complex samples, the matrix effect should be evaluated so that method accuracy is not compromised. The ideal approach for accurate and precise peptide quantification would be the use of external calibration curves prepared in the representative matrix by spiking the unlabeled synthetic peptide at various concentrations and a constant amount of the analogous stable isotope synthetic (SIS) peptide as internal standard. However, proteins or peptides of interest are usually present in the representative matrix that impairs the accuracy and precision of the quantification method if the calibration curves are performed by spiking the synthetic unlabeled peptide into the matrix. For that reason, Campbell and collaborators proposed an alternative approach called the reverse curve method where varying amounts of the labeled peptide are spiked in the representative matrix to create the calibration curve [248]. In this work, seven apolipoproteins were quantified in human plasma using the three approaches: a) spiking the sample with a known amount of the analogous synthetic-labeled peptide (AQUA approach); b) spiking the representative matrix with different amounts of the unlabeled synthetic peptide and a constant amount of the labeled peptide to create the “classical”

calibration curve; and c) spiking the representative matrix with varying amounts of the labeled synthetic peptide to create the curve and a constant amount of the unlabeled peptide to work as internal standard (reverse calibration curve). For both cases using external curves, some corrections are required due to the endogenous peptide already present in the sample. The correction for the classical calibration curve is performed by subtracting the y-interception of the curve to the determined concentration of the endogenous peptide in the sample. For the reverse curve the correction factor corresponds to an increase of each curve point by an amount proportional to the ratio of the amounts of endogenous unlabeled peptide to the spiked synthetic unlabeled peptide [29]. The AQUA approach revealed to be inaccurate for endogenous peptide quantification below and above the concentration of the IS spiked into the sample. As expected, this result demonstrates that accurate peptide quantification in samples can only be achieved if the spiked IS amount into the sample is close to the concentration of the endogenous unlabeled peptide. The reverse curve has the advantage of allowing the determination of the limits of detection and quantification (LOD and LOQ) once the representative matrix does not contain the synthetic-labeled peptide. In addition, the quantification using the reverse curve revealed to be the most accurate and precise between the three methods, thus this approach can be used with confidence to quantify endogenous peptides/proteins already present in the surrogate matrix.

Absolute quantification by MRM applied to neuroproteomics was first described by Desiderio and collaborators to quantify picomole amounts of the endogenous methionine-enkephalin (ME) in the human pituitary by comparing the response of the endogenous ME to one of the deuterated ME internal standard (d5-ME) [249]. More recently, Kheterpal and collaborators determined the concentration of MIF-1 (neuropeptide) in different regions of mouse brain by using a calibration curve prepared by successive dilutions of the unlabeled synthetic peptide in the absence of matrix [250].

There are several studies in Alzheimer's disease that involve protein/peptide quantification by MRM [251–253]. Lame and collaborators developed a UPLC-MRM method to accurately quantify $A\beta_{1-38}$, $A\beta_{1-40}$, and $A\beta_{1-42}$ in human cerebrospinal fluid that can play a crucial role in understanding disease progression and intervention [254]. The quantification was performed using calibration curves prepared with various concentrations of the synthetic peptides spiked with constant amounts of analogous ^{15}N -labeled internal standards in an artificial CSF matrix. Also, Wildsmith and collaborators described the development of an MRM assay for the absolute quantification of 39 peptides corresponding to 30 proteins to confirm previous findings for a subset of markers for Alzheimer's disease [255].

Another approach that, in combination with MRM, allows absolute quantification based on the isotope-dilution mass spectrometry or AQUA methodology is peptide labeling with non-isobaric tags reagents, the mTRAQ reagents. Originally, mTRAQ labels appeared in two versions, the lighter version (lower in mass than the iTRAQ labels by 4 Da) and the heavy version that is identical to the iTRAQ 117 label resulting in a light version ($\Delta 0$) with a monoisotopic mass of 141 Da and a heavy version ($\Delta 4$) of 145 Da. Furthermore, a new label version ($\Delta 8$) is now available called triplex mTRAQ reagents. These have been used mostly for relative quantification but DeSouza and collaborators described a method for absolute quantification

of proteins using the duplex version of mTRAQ reagent. The procedure consisted of labeling known amounts of the synthetic peptide used for protein quantification with one of the two versions while, the opposite version was used to tag the endogenous peptides that needed to be quantified [256]. At the end, these two fractions were mixed at a known amount and the resulting mixture was analyzed by unique MRM transitions for each version of the labeled peptide resultant from the different masses of the tags. The areas that resulted from each MRM transition were then used to determine the unknown concentration of the peptide in the digested sample and, consequently, the concentration of the protein of interest.

Apart from the AQUA quantification strategy, standard peptides are usually spiked at late stages of sample processing; they are poorly compatible with sample pre-fractionation; and the digestion efficiency cannot be fully determined leading to an inaccurate quantification [257]. To address these issues, other types of standards were developed known as artificial concatamers of proteotypic peptides (called QconCAT) and are generally added into the sample just before protein digestion [258]. Concatamers are artificial protein constructs that include multiple trypsin-cleavable proteotypic peptides isotopically labeled. The isotope-labeled peptides are released during protein digestion and will be used as standards for the absolute target protein quantification. The QconCAT methodology possesses the main advantage of facilitating multiplex protein quantification where typically 10–30 target analyte proteins are encoded in each QconCAT at a level of two quantotypic peptides per protein [259]. Chen and collaborators stated that a reliable quantitative approach of clusterin in brain was needed to clarify its role in Alzheimer's disease. Consequently, they developed a stable isotope-labeled concatenated peptide (QconCAT) for the quantification of clusterin in human postmortem frontal and temporal cortex [260]. Later, they applied this approach for other protein quantifications also related to AD. At this time, a multiplexed QconCAT was designed for quantification of various isoforms of amyloid precursor protein (APP). Since common tryptic peptides between all isoforms of APP were concatenated with unique tryptic peptides for specific APP isoforms, this QconCAT-MRM method allowed the clear quantification of the total APP and each protein isoform [261].

Even with the advantage of multiplexing absolute quantification, the assessment of the digestion efficiency remains undetermined once QconCAT are usually digested at high rates, not giving the true tryptic digestion efficiencies for each protein [258]. By using the isotope-labeled equivalent of the full-length target protein, the "ideal" internal standards can be added at the very beginning of sample processing, allowing the determination of the recoveries after pre-fractionation steps and the assessment of the digestion efficiencies, which is called the "Protein Standard Absolute Quantification" (PSAQ). A comparative study between AQUA, QconCAT, and PSAQ was performed for the quantification of *Staphylococcus aureus* superantigenic toxins in water and urine samples where the PSAQ strategy revealed to be more accurate than the two other methods [262]. PSAQ also revealed to be advantageous for the absolute quantification of membrane proteins that are more prone to errors on concentration determination due to protein enrichment steps usually required and incomplete digestion [263]. In this study, accurate quantification of 7 membrane proteins was achieved using as internal standards the analogous ¹⁵N-labeled full-length proteins added at an initial stage of

sample processing. There are other quantification approaches based on the addition of full-length labeled proteins internal standards for protein quantification such as FlexiQuant [264], PrEST [265], and Absolute SILAC [266].

Although, some of the approaches presented have few publications for absolute protein quantification by MRM they can be of interest for the neuroproteomics field to confirm previous findings or to find new targets with more accurate data.

Author details

Cátia Santa^{1,2}, Sandra I. Anjo^{1,3}, Vera M. Mendes^{1,4} and Bruno Manadas^{1,4*}

*Address all correspondence to: bmanadas@gmail.com

1 CNC - Center for Neuroscience and Cell Biology, University of Coimbra, Portugal

2 Institute for Interdisciplinary Research, University of Coimbra, Portugal

3 Faculty of Sciences and Technology, University of Coimbra, Portugal

4 Biocant – Biotechnology Innovation Center, Cantanhede, Portugal

References

- [1] Kitchen RR, Rozowsky JS, Gerstein MB, Nairn AC. Decoding neuroproteomics: Integrating the genome, transcriptome and functional anatomy. *Nature Neuroscience*. 2014;17(11):1491-9.
- [2] Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science*. 2001;291(5507):1304-51.
- [3] Tannu NS, Hemby SE. Methods for proteomics in neuroscience. *Progress in Brain Research*. 2006;158:41-82.
- [4] Lundberg E, Fagerberg L, Klevebring D, Matic I, Geiger T, Cox J, et al. Defining the transcriptome and proteome in three functionally different human cell lines. *Molecular Systems Biology*. 2010;6:450.
- [5] Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. *Nature Chemical Biology*. 2005;1(5):252-62.
- [6] Fountoulakis M. Application of proteomics technologies in the investigation of the brain. *Mass Spectrometry Reviews*. 2004;23(4):231-58.

- [7] Cox J, Mann M. Quantitative, high-resolution proteomics for data-driven systems biology. *Annual Review of Biochemistry*. 2011;80:273-99.
- [8] Bantscheff M, Lemeer S, Savitski MM, Kuster B. Quantitative mass spectrometry in proteomics: Critical review update from 2007 to the present. *Analytical and Bioanalytical Chemistry*. 2012;404(4):939-65.
- [9] Craft GE, Chen A, Nairn AC. Recent advances in quantitative neuroproteomics. *Methods*. 2013;61(3):186-218.
- [10] Rauniyar N, Yates JR 3rd. Isobaric labeling-based relative quantification in shotgun proteomics. *Journal of Proteome Research*. 2014;13(12):5293-309.
- [11] Altelaar AF, Frese CK, Preisinger C, Hennrich ML, Schram AW, Timmers HT, et al. Benchmarking stable isotope labeling-based quantitative proteomics. *Journal of proteomics*. 2013;88:14-26.
- [12] Desiderio DM, Kai M. Preparation of stable isotope-incorporated peptide internal standards for field desorption mass spectrometry quantification of peptides in biologic tissue. *Biomedical Mass Spectrometry*. 1983;10(8):471-9.
- [13] Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C. Proteolytic ^{18}O labeling for comparative proteomics: Model studies with two serotypes of adenovirus. *Analytical Chemistry*. 2001;73(13):2836-42.
- [14] Fenselau C, Yao X. $^{18}\text{O}_2$ -labeling in quantitative proteomic strategies: A status report. *Journal of Proteome Research*. 2009;8(5):2140-3.
- [15] Capelo JL, Carreira RJ, Fernandes L, Lodeiro C, Santos HM, Simal-Gandara J. Latest developments in sample treatment for ^{18}O -isotopic labeling for proteomics mass spectrometry-based approaches: A critical review. *Talanta*. 2010;80(4):1476-86.
- [16] Guo Y, Miyagi M, Zeng R, Sheng Q. O18Quant: A semiautomatic strategy for quantitative analysis of high-resolution $^{16}\text{O}/^{18}\text{O}$ labeled data. *BioMed Research International*. 2014;2014:971857.
- [17] Niles R, Witkowska HE, Allen S, Hall SC, Fisher SJ, Hardt M. Acid-catalyzed oxygen-18 labeling of peptides. *Analytical Chemistry*. 2009;81(7):2804-9.
- [18] Wang S, Bobst CE, Kaltashov IA. Pitfalls in protein quantitation using acid-catalyzed O^{18} labeling: Hydrolysis-driven deamidation. *Analytical chemistry*. 2011;83(18):7227-32.
- [19] Wang YK, Ma Z, Quinn DF, Fu EW. Inverse ^{18}O labeling mass spectrometry for the rapid identification of marker/target proteins. *Analytical chemistry*. 2001;73(15):3742-50.
- [20] Wang S, Kaltashov IA. A new strategy of using O^{18} -labeled iodoacetic acid for mass spectrometry-based protein quantitation. *Journal of the American Society for Mass Spectrometry*. 2012;23(7):1293-7.

- [21] Chaerkady R, Thuluvath PJ, Kim MS, Nalli A, Vivekanandan P, Simmers J, et al. O labeling for a quantitative proteomic analysis of glycoproteins in hepatocellular carcinoma. *Clinical Proteomics*. 2008;4(3-4):137-55.
- [22] Wu P, Zhao Y, Haidacher SJ, Wang E, Parsley MO, Gao J, et al. Detection of structural and metabolic changes in traumatically injured hippocampus by quantitative differential proteomics. *Journal of Neurotrauma*. 2013;30(9):775-88.
- [23] Dagley LF, White CA, Liao Y, Shi W, Smyth GK, Orian JM, et al. Quantitative proteomic profiling reveals novel region-specific markers in the adult mouse brain. *Proteomics*. 2014;14(2-3):241-61.
- [24] Lahm HW, Langen H. Mass spectrometry: A tool for the identification of proteins separated by gels. *Electrophoresis*. 2000;21(11):2105-14.
- [25] Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. Accurate quantitation of protein expression and site-specific phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America*. 1999;96(12):6591-6.
- [26] Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics : MCP*. 2002;1(5):376-86.
- [27] Mann M. Functional and quantitative proteomics using SILAC. *Nature Reviews Molecular Cell Biology*. 2006;7(12):952-8.
- [28] Ong SE, Mann M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nature Protocols*. 2006;1(6):2650-60.
- [29] Ong SE, Kratchmarova I, Mann M. Properties of ¹³C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *Journal of Proteome Research*. 2003;2(2):173-81.
- [30] Harsha HC, Molina H, Pandey A. Quantitative proteomics using stable isotope labeling with amino acids in cell culture. *Nature Protocols*. 2008;3(3):505-16.
- [31] Blagoev B, Kratchmarova I, Ong SE, Nielsen M, Foster LJ, Mann M. A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nature Biotechnology*. 2003;21(3):315-8.
- [32] Ong SE, Mann M. Identifying and quantifying sites of protein methylation by heavy methyl SILAC. *Current protocols in protein science/editorial board, John E Coligan, et al*. 2006;Chapter 14:Unit 14 9.
- [33] Olsen JV, Blagoev B, Gnäd F, Macek B, Kumar C, Mortensen P, et al. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell*. 2006;127(3):635-48.

- [34] Boisvert FM, Ahmad Y, Gierlinski M, Charriere F, Lamont D, Scott M, et al. A quantitative spatial proteomics analysis of proteome turnover in human cells. *Molecular & Cellular Proteomics : MCP*. 2012;11(3):M111 011429.
- [35] Schwanhausser B, Gossen M, Dittmar G, Selbach M. Global analysis of cellular protein translation by pulsed SILAC. *Proteomics*. 2009;9(1):205-9.
- [36] Gronborg M, Kristiansen TZ, Iwahori A, Chang R, Reddy R, Sato N, et al. Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach. *Molecular & Cellular Proteomics : MCP*. 2006;5(1):157-71.
- [37] Hathout Y. Approaches to the study of the cell secretome. *Expert Review of Proteomics*. 2007;4(2):239-48.
- [38] Liu B, Barber DS, Stevens SM, Jr. Stable isotope labeling with amino acids in cell culture-based proteomic analysis of ethanol-induced protein expression profiles in microglia. *Methods in Molecular Biology*. 2012;829:551-65.
- [39] Gokhale A, Larimore J, Werner E, So L, Moreno-De-Luca A, Lese-Martin C, et al. Quantitative proteomic and genetic analyses of the schizophrenia susceptibility factor dysbindin identify novel roles of the biogenesis of lysosome-related organelles complex 1. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*. 2012;32(11):3697-711.
- [40] Sarraf SA, Raman M, Guarani-Pereira V, Sowa ME, Huttlin EL, Gygi SP, et al. Landscape of the PARKIN-dependent ubiquitylome in response to mitochondrial depolarization. *Nature*. 2013;496(7445):372-6.
- [41] Ren X, Bai X, Zhang X, Li Z, Tang L, Zhao X, et al. Quantitative nuclear proteomics identifies that miR-137-mediated EZH2 reduction regulates resveratrol-induced apoptosis of neuroblastoma cells. *Molecular & Cellular Proteomics : MCP*. 2015;14(2): 316-28.
- [42] Qiu H, Wang Y. Quantitative analysis of surface plasma membrane proteins of primary and metastatic melanoma cells. *Journal of Proteome Research*. 2008;7(5): 1904-15.
- [43] Spellman DS, Deinhardt K, Darie CC, Chao MV, Neubert TA. Stable isotopic labeling by amino acids in cultured primary neurons: Application to brain-derived neurotrophic factor-dependent phosphotyrosine-associated signaling. *Molecular & Cellular Proteomics : MCP*. 2008;7(6):1067-76.
- [44] Liao L, Park SK, Xu T, Vanderklish P, Yates JR, 3rd. Quantitative proteomic analysis of primary neurons reveals diverse changes in synaptic protein content in *fmr1* knockout mice. *Proceedings of the National Academy of Sciences of the United States of America*. 2008;105(40):15281-6.
- [45] Polson C, Cantone JL, Wei C, Drexler DM, Meredith JE, Jr. Stable isotope labeling by amino acids in cell culture-based liquid chromatography-mass spectrometry assay to

- measure microtubule dynamics in neuronal cell cultures. *Analytical Biochemistry*. 2014;466:65-71.
- [46] Greco TM, Seeholzer SH, Mak A, Spruce L, Ischiropoulos H. Quantitative mass spectrometry-based proteomics reveals the dynamic range of primary mouse astrocyte protein secretion. *Journal of Proteome Research*. 2010;9(5):2764-74.
- [47] Zhang G, Deinhardt K, Chao MV, Neubert TA. Study of neurotrophin-3 signaling in primary cultured neurons using multiplex stable isotope labeling with amino acids in cell culture. *Journal of Proteome Research*. 2011;10(5):2546-54.
- [48] Zhang G, Deinhardt K, Neubert TA. Stable isotope labeling by amino acids in cultured primary neurons. *Methods in molecular biology*. 2014;1188:57-64.
- [49] Ishihama Y, Sato T, Tabata T, Miyamoto N, Sagane K, Nagasu T, et al. Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. *Nature Biotechnology*. 2005;23(5):617-21.
- [50] Wu CC, MacCoss MJ, Howell KE, Matthews DE, Yates JR, 3rd. Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Analytical Chemistry*. 2004;76(17):4951-9.
- [51] McClatchy DB, Dong MQ, Wu CC, Venable JD, Yates JR, 3rd. ¹⁵N metabolic labeling of mammalian tissue with slow protein turnover. *Journal of Proteome Research*. 2007;6(5):2005-10.
- [52] Kruger M, Moser M, Ussar S, Thievensen I, Luber CA, Forner F, et al. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell*. 2008;134(2):353-64.
- [53] McClatchy DB, Liao L, Park SK, Venable JD, Yates JR. Quantification of the synaptosomal proteome of the rat cerebellum during post-natal development. *Genome Research*. 2007;17(9):1378-88.
- [54] Butko MT, Savas JN, Friedman B, Delahunty C, Ebner F, Yates JR 3rd, et al. In vivo quantitative proteomics of somatosensory cortical synapses shows which protein levels are modulated by sensory deprivation. *Proceedings of the National Academy of Sciences of the United States of America*. 2013;110(8):E726-35.
- [55] Savas JN, Toyama BH, Xu T, Yates JR 3rd, Hetzer MW. Extremely long-lived nuclear pore proteins in the rat brain. *Science*. 2012;335(6071):942.
- [56] Flintoft L. Animal models: Proteomics goes live in the mouse. *Nature Reviews Genetics*. 2008;9(9):655.
- [57] Zanivan S, Krueger M, Mann M. In vivo quantitative proteomics: The SILAC mouse. *Methods in Molecular Biology*. 2012;757:435-50.

- [58] Walther DM, Mann M. Accurate quantification of more than 4000 mouse tissue proteins reveals minimal proteome changes during aging. *Molecular & Cellular Proteomics : MCP*. 2011;10(2):M110 004523.
- [59] Geiger T, Cox J, Ostasiewicz P, Wisniewski JR, Mann M. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nature Methods*. 2010;7(5):383-5.
- [60] Deeb SJ, D'Souza RC, Cox J, Schmidt-Supprian M, Mann M. Super-SILAC allows classification of diffuse large B-cell lymphoma subtypes by their protein expression profiles. *Molecular & Cellular Proteomics : MCP*. 2012;11(5):77-89.
- [61] Stauch KL, Purnell PR, Fox HS. Quantitative proteomics of synaptic and nonsynaptic mitochondria: insights for synaptic mitochondrial vulnerability. *Journal of Proteome Research*. 2014;13(5):2620-36.
- [62] Hebert AS, Merrill AE, Bailey DJ, Still AJ, Westphall MS, Strieter ER, et al. Neutron-encoded mass signatures for multiplexed proteome quantification. *Nature Methods*. 2013;10(4):332-4.
- [63] Hebert AS, Merrill AE, Stefely JA, Bailey DJ, Wenger CD, Westphall MS, et al. Amine-reactive neutron-encoded labels for highly plexed proteomic quantitation. *Molecular & Cellular Proteomics : MCP*. 2013;12(11):3360-9.
- [64] Merrill AE, Hebert AS, MacGilvray ME, Rose CM, Bailey DJ, Bradley JC, et al. Neu-Code labels for relative protein quantification. *Molecular & Cellular Proteomics : MCP*. 2014;13(9):2503-12.
- [65] Richards AL, Vincent CE, Guthals A, Rose CM, Westphall MS, Bandeira N, et al. Neutron-encoded signatures enable product ion annotation from tandem mass spectra. *Molecular & Cellular Proteomics : MCP*. 2013;12(12):3812-23.
- [66] Rose CM, Merrill AE, Bailey DJ, Hebert AS, Westphall MS, Coon JJ. Neutron encoded labeling for peptide identification. *Analytical chemistry*. 2013;85(10):5129-37.
- [67] Rhoads TW, Rose CM, Bailey DJ, Riley NM, Molden RC, Nestler AJ, et al. Neutron-encoded mass signatures for quantitative top-down proteomics. *Analytical Chemistry*. 2014;86(5):2314-9.
- [68] Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnology*. 1999;17(10):994-9.
- [69] Hansen KC, Schmitt-Ulms G, Chalkley RJ, Hirsch J, Baldwin MA, Burlingame AL. Mass spectrometric analysis of protein mixtures at low levels using cleavable ¹³C-isotope-coded affinity tag and multidimensional chromatography. *Molecular & Cellular Proteomics : MCP*. 2003;2(5):299-314.

- [70] Zhou H, Ranish JA, Watts JD, Aebersold R. Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. *Nature Biotechnology*. 2002;20(5):512-5.
- [71] Han B, Stevens JF, Maier CS. Design, synthesis, and application of a hydrazide-functionalized isotope-coded affinity tag for the quantification of oxylipid-protein conjugates. *Analytical Chemistry*. 2007;79(9):3342-54.
- [72] Zhang J, Goodlett DR, Peskind ER, Quinn JF, Zhou Y, Wang Q, et al. Quantitative proteomic analysis of age-related changes in human cerebrospinal fluid. *Neurobiology of Aging*. 2005;26(2):207-27.
- [73] Jin J, Meredith GE, Chen L, Zhou Y, Xu J, Shie FS, et al. Quantitative proteomic analysis of mitochondrial proteins: Relevance to Lewy body formation and Parkinson's disease. *Brain Research Molecular Brain Research*. 2005;134(1):119-38.
- [74] Fu YJ, Xiong S, Lovell MA, Lynn BC. Quantitative proteomic analysis of mitochondria in aging PS-1 transgenic mice. *Cellular and molecular neurobiology*. 2009;29(5):649-64.
- [75] Costain WJ, Haqqani AS, Rasquinha I, Giguere MS, Slinn J, Zurakowski B, et al. Proteomic analysis of synaptosomal protein expression reveals that cerebral ischemia alters lysosomal Psap processing. *Proteomics*. 2010;10(18):3272-91.
- [76] Klychnikov OI, Li KW, Sidorov IA, Loos M, Spijker S, Broos LA, et al. Quantitative cortical synapse proteomics of a transgenic migraine mouse model with mutated Ca(V)2.1 calcium channels. *Proteomics*. 2010;10(13):2531-5.
- [77] Bousquet-Dubouch MP, Nguyen S, Bouyssié D, Burette-Schiltz O, French SW, Monsarrat B, et al. Chronic ethanol feeding affects proteasome-interacting proteins. *Proteomics*. 2009;9(13):3609-22.
- [78] Schmidt A, Kellermann J, Lottspeich F. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics*. 2005;5(1):4-15.
- [79] Paradela A, Marcilla M, Navajas R, Ferreira L, Ramos-Fernandez A, Fernandez M, et al. Evaluation of isotope-coded protein labeling (ICPL) in the quantitative analysis of complex proteomes. *Talanta*. 2010;80(4):1496-502.
- [80] Leroy B, Rosier C, Erculisse V, Leys N, Mergeay M, Wattiez R. Differential proteomic analysis using isotope-coded protein-labeling strategies: Comparison, improvements and application to simulated microgravity effect on *Cupriavidus metallidurans* CH34. *Proteomics*. 2010;10(12):2281-91.
- [81] Maccarrone G, Turck CW, Martins-de-Souza D. Shotgun mass spectrometry workflow combining IEF and LC-MALDI-TOF/TOF. *The Protein Journal*. 2010;29(2):99-102.

- [82] Fleron M, Greffe Y, Musmeci D, Massart AC, Hennequiere V, Mazzucchelli G, et al. Novel post-digest isotope coded protein labeling method for phospho- and glycoproteome analysis. *Journal of Proteomics*. 2010;73(10):1986-2005.
- [83] Martins-de-Souza D, Gattaz WF, Schmitt A, Rewerts C, Maccarrone G, Dias-Neto E, et al. Prefrontal cortex shotgun proteome analysis reveals altered calcium homeostasis and immune system imbalance in schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience*. 2009;259(3):151-63.
- [84] Com E, Clavreul A, Lagarrigue M, Michalak S, Menei P, Pineau C. Quantitative proteomic Isotope-Coded Protein Label (ICPL) analysis reveals alteration of several functional processes in the glioblastoma. *Journal of Proteomics*. 2012;75(13):3898-913.
- [85] Ulbrich A, Merrill AE, Hebert AS, Westphall MS, Keller MP, Attie AD, et al. Neutron-encoded protein quantification by peptide carbamylation. *Journal of the American Society for Mass Spectrometry*. 2014;25(1):6-9.
- [86] Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, et al. Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical Chemistry*. 2003;75(8):1895-904.
- [87] Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular & Cellular Proteomics : MCP*. 2004;3(12):1154-69.
- [88] Choe L, D'Ascenzo M, Relkin NR, Pappin D, Ross P, Williamson B, et al. 8-plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease. *Proteomics*. 2007;7(20):3651-60.
- [89] Dayon L, Hainard A, Licker V, Turck N, Kuhn K, Hochstrasser DF, et al. Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Analytical Chemistry*. 2008;80(8):2921-31.
- [90] McAlister GC, Huttlin EL, Haas W, Ting L, Jedrychowski MP, Rogers JC, et al. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Analytical Chemistry*. 2012;84(17):7469-78.
- [91] Pottiez G, Wiederin J, Fox HS, Ciborowski P. Comparison of 4-plex to 8-plex iTRAQ quantitative measurements of proteins in human plasma samples. *Journal of Proteome Research*. 2012;11(7):3774-81.
- [92] Mertins P, Udeshi ND, Clauser KR, Mani DR, Patel J, Ong SE, et al. iTRAQ labeling is superior to mTRAQ for quantitative global proteomics and phosphoproteomics. *Molecular & Cellular Proteomics : MCP*. 2012;11(6):M111 014423.
- [93] Ting L, Rad R, Gygi SP, Haas W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nature Methods*. 2011;8(11):937-40.

- [94] Pichler P, Kocher T, Holzmann J, Mazanek M, Taus T, Ammerer G, et al. Peptide labeling with isobaric tags yields higher identification rates using iTRAQ 4-plex compared to TMT 6-plex and iTRAQ 8-plex on LTQ Orbitrap. *Analytical Chemistry*. 2010;82(15):6549-58.
- [95] Pan KT, Chen YY, Pu TH, Chao YS, Yang CY, Bomgardner RD, et al. Mass spectrometry-based quantitative proteomics for dissecting multiplexed redox cysteine modifications in nitric oxide-protected cardiomyocyte under hypoxia. *Antioxidants & redox signaling*. 2014;20(9):1365-81.
- [96] Hahne H, Neubert P, Kuhn K, Etienne C, Bomgardner R, Rogers JC, et al. Carbonyl-reactive tandem mass tags for the proteome-wide quantification of N-linked glycans. *Analytical Chemistry*. 2012;84(8):3716-24.
- [97] Palmese A, De Rosa C, Chiappetta G, Marino G, Amoresano A. Novel method to investigate protein carbonylation by iTRAQ strategy. *Analytical and Bioanalytical Chemistry*. 2012;404(6-7):1631-5.
- [98] Glibert P, Meert P, Van Steendam K, Van Nieuwerburgh F, De Coninck D, Martens L, et al. Phospho-iTRAQ: Assessing isobaric labels for the large-scale study of phosphopeptide stoichiometry. *Journal of Proteome Research*. 2015;14(2):839-49.
- [99] Prudova A, auf dem Keller U, Butler GS, Overall CM. Multiplex N-terminome analysis of MMP-2 and MMP-9 substrate degradomes by iTRAQ-TAILS quantitative proteomics. *Molecular & Cellular Proteomics : MCP*. 2010;9(5):894-911.
- [100] Kleifeld O, Doucet A, Prudova A, auf dem Keller U, Gioia M, Kizhakkedathu JN, et al. Identifying and quantifying proteolytic events and the natural N terminome by terminal amine isotopic labeling of substrates. *Nature Protocols*. 2011;6(10):1578-611.
- [101] Lin X, Shi M, Masilamoni JG, Dator R, Movius J, Aro P, et al. Proteomic profiling in MPTP monkey model for early Parkinson disease biomarker discovery. *Biochimica et Biophysica Acta*. 2015.
- [102] Zhang X, Yin X, Yu H, Liu X, Yang F, Yao J, et al. Quantitative proteomic analysis of serum proteins in patients with Parkinson's disease using an isobaric tag for relative and absolute quantification labeling, two-dimensional liquid chromatography, and tandem mass spectrometry. *The Analyst*. 2012;137(2):490-5.
- [103] Skorobogatko YV, Deuso J, Adolf-Bryfogle J, Nowak MG, Gong Y, Lippa CF, et al. Human Alzheimer's disease synaptic O-GlcNAc site mapping and iTRAQ expression proteomics with ion trap mass spectrometry. *Amino Acids*. 2011;40(3):765-79.
- [104] Malki K, Campbell J, Davies M, Keers R, Uher R, Ward M, et al. Pharmacoproteomic investigation into antidepressant response in two mouse inbred strains. *Proteomics*. 2012;12(14):2355-65.

- [105] Merali Z, Gao MM, Bowes T, Chen J, Evans K, Kassner A. Neuroproteome changes after ischemia/reperfusion injury and tissue plasminogen activator administration in rats: A quantitative iTRAQ proteomics study. *PloS one*. 2014;9(5):e98706.
- [106] Robison AJ, Vialou V, Mazei-Robison M, Feng J, Kourrich S, Collins M, et al. Behavioral and structural responses to chronic cocaine require a feedforward loop involving DeltaFosB and calcium/calmodulin-dependent protein kinase II in the nucleus accumbens shell. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*. 2013;33(10):4295-307.
- [107] Martins-de-Souza D, Maccarrone G, Wobrock T, Zerr I, Gormanns P, Reckow S, et al. Proteome analysis of the thalamus and cerebrospinal fluid reveals glycolysis dysfunction and potential biomarkers candidates for schizophrenia. *Journal of Psychiatric Research*. 2010;44(16):1176-89.
- [108] Hirayama M, Kobayashi D, Mizuguchi S, Morikawa T, Nagayama M, Midorikawa U, et al. Integrated proteomics identified novel activation of dynein IC2-GR-COX-1 signaling in neurofibromatosis type I (NF1) disease model cells. *Molecular & Cellular Proteomics : MCP*. 2013;12(5):1377-94.
- [109] Rao-Ruiz P, Carney KE, Pandya N, van der Loo RJ, Verheijen MH, van Nierop P, et al. Time-dependent changes in the mouse hippocampal synaptic membrane proteome after contextual fear conditioning. *Hippocampus*. 2015.
- [110] Xiang F, Ye H, Chen R, Fu Q, Li L. N,N-dimethyl leucines as novel isobaric tandem mass tags for quantitative proteomics and peptidomics. *Analytical Chemistry*. 2010;82(7):2817-25.
- [111] Zhang J, Wang Y, Li S. Deuterium isobaric amine-reactive tags for quantitative proteomics. *Analytical Chemistry*. 2010;82(18):7588-95.
- [112] Hsu JL, Huang SY, Chow NH, Chen SH. Stable-isotope dimethyl labeling for quantitative proteomics. *Analytical Chemistry*. 2003;75(24):6843-52.
- [113] Sturm RM, Lietz CB, Li L. Improved isobaric tandem mass tag quantification by ion mobility mass spectrometry. *Rapid Communications in Mass Spectrometry : RCM*. 2014;28(9):1051-60.
- [114] Hui L, Xiang F, Zhang Y, Li L. Mass spectrometric elucidation of the neuropeptidome of a crustacean neuroendocrine organ. *Peptides*. 2012;36(2):230-9.
- [115] Hao L, Zhong X, Greer T, Ye H, Li L. Relative quantification of amine-containing metabolites using isobaric N,N-dimethyl leucine (DiLeu) reagents via LC-ESI-MS/MS and CE-ESI-MS/MS. *The Analyst*. 2015;140(2):467-75.
- [116] Frost DC, Greer T, Li L. High-Resolution Enabled 12-Plex DiLeu Isobaric Tags for Quantitative Proteomics. *Analytical Chemistry*. 2015;87(3):1646-54.
- [117] Ramsubramaniam N, Tao F, Li S, Marten MR. Novel and cost-effective 6-plex isobaric tagging reagent, DiART, is effective for identification and relative quantification of

- complex protein mixtures using PQD fragmentation. *Journal of Mass Spectrometry* : JMS. 2013;48(9):1032-41.
- [118] Chen Z, Wang Q, Lin L, Tang Q, Edwards JL, Li S, et al. Comparative evaluation of two isobaric labeling tags, DiART and iTRAQ. *Analytical Chemistry*. 2012;84(6): 2908-15.
- [119] Ramsubramaniam N, Tao F, Li S, Marten MR. Cost-effective isobaric tagging for quantitative phosphoproteomics using DiART reagents. *Molecular bioSystems*. 2013;9(12):2981-7.
- [120] Dephoure N, Gygi SP. Hyperplexing: A method for higher-order multiplexed quantitative proteomics provides a map of the dynamic response to rapamycin in yeast. *Science Signaling*. 2012;5(217):rs2.
- [121] Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, et al. Less label, more free: Approaches in label-free quantitative mass spectrometry. *Proteomics*. 2011;11(4):535-53.
- [122] Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. *Analytical and bioanalytical chemistry*. 2007;389(4):1017-31.
- [123] Mayne J, Starr AE, Ning Z, Chen R, Chiang CK, Figeys D. Fine tuning of proteomic technologies to improve biological findings: advancements in 2011-2013. *Analytical Chemistry*. 2014;86(1):176-95.
- [124] Mueller LN, Brusniak MY, Mani DR, Aebersold R. An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *Journal of Proteome Research*. 2008;7(1):51-61.
- [125] Zhu W, Smith JW, Huang CM. Mass spectrometry-based label-free quantitative proteomics. *Journal of Biomedicine & Biotechnology*. 2010;2010:840518.
- [126] Domon B, Aebersold R. Options and considerations when selecting a quantitative proteomics strategy. *Nature Biotechnology*. 2010;28(7):710-21.
- [127] Lundgren DH, Hwang SI, Wu L, Han DK. Role of spectral counting in quantitative proteomics. *Expert Review of Proteomics*. 2010;7(1):39-53.
- [128] Ahrne E, Molzahn L, Glatter T, Schmidt A. Critical assessment of proteome-wide label-free absolute abundance estimation strategies. *Proteomics*. 2013;13(17):2567-78.
- [129] Liu Y, Huttenhain R, Collins B, Aebersold R. Mass spectrometric protein maps for biomarker discovery and clinical research. *Expert Review of Molecular Diagnostics*. 2013;13(8):811-25.
- [130] Megger DA, Bracht T, Meyer HE, Sitek B. Label-free quantification in clinical proteomics. *Biochimica et Biophysica Acta*. 2013;1834(8):1581-90.

- [131] Kito K, Ito T. Mass spectrometry-based approaches toward absolute quantitative proteomics. *Current Genomics*. 2008;9(4):263-74.
- [132] DeSouza LV, Siu KW. Mass spectrometry-based quantification. *Clinical Biochemistry*. 2013;46(6):421-31.
- [133] Washburn MP, Wolters D, Yates JR, 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology*. 2001;19(3):242-7.
- [134] Ikemura T. Codon usage and tRNA content in unicellular and multicellular organisms. *Molecular Biology and Evolution*. 1985;2(1):13-34.
- [135] Futcher B, Latter GI, Monardo P, McLaughlin CS, Garrels JI. A sampling of the yeast proteome. *Molecular and Cellular Biology*. 1999;19(11):7357-68.
- [136] Pang JX, Ginanni N, Dongre AR, Hefta SA, Opitck GJ. Biomarker discovery in urine by proteomics. *Journal of Proteome Research*. 2002;1(2):161-9.
- [137] Gao J, Friedrichs MS, Dongre AR, Opitck GJ. Guidelines for the routine application of the peptide hits technique. *Journal of the American Society for Mass Spectrometry*. 2005;16(8):1231-8.
- [138] Gao J, Opitck GJ, Friedrichs MS, Dongre AR, Hefta SA. Changes in the protein expression of yeast as a function of carbon source. *Journal of Proteome Research*. 2003;2(6):643-9.
- [139] Rappsilber J, Ryder U, Lamond AI, Mann M. Large-scale proteomic analysis of the human spliceosome. *Genome Research*. 2002;12(8):1231-45.
- [140] Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, Haynes JD, et al. A proteomic view of the *Plasmodium falciparum* life cycle. *Nature*. 2002;419(6906):520-6.
- [141] Lasonder E, Ishihama Y, Andersen JS, Vermunt AM, Pain A, Sauerwein RW, et al. Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature*. 2002;419(6906):537-42.
- [142] Asara JM, Christofk HR, Freimark LM, Cantley LC. A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. *Proteomics*. 2008;8(5):994-9.
- [143] Florens L, Carozza MJ, Swanson SK, Fournier M, Coleman MK, Workman JL, et al. Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. *Methods*. 2006;40(4):303-11.
- [144] Zybilov B, Mosley AL, Sardi ME, Coleman MK, Florens L, Washburn MP. Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *Journal of Proteome Research*. 2006;5(9):2339-47.

- [145] Zhang Y, Wen Z, Washburn MP, Florens L. Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Analytical Chemistry*. 2010;82(6):2272-81.
- [146] Dost B, Bandeira N, Li X, Shen Z, Briggs SP, Bafna V. Accurate mass spectrometry based protein quantification via shared peptides. *Journal of Computational biology: A Journal of Computational Molecular Cell Biology*. 2012;19(4):337-48.
- [147] Dost A, Rohrer T, Fussenegger J, Vogel C, Schenk B, Wabitsch M, et al. Bone maturation in 1788 children and adolescents with diabetes mellitus type 1. *Journal of Pediatric Endocrinology & Metabolism : JPEM*. 2010;23(9):891-8.
- [148] Liu WL, Coleman RA, Grob P, King DS, Florens L, Washburn MP, et al. Structural changes in TAF4b-TFIID correlate with promoter selectivity. *Molecular Cell*. 2008;29(1):81-91.
- [149] Zybaylov B, Rutschow H, Friso G, Rudella A, Emanuelsson O, Sun Q, et al. Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS one*. 2008;3(4):e1994.
- [150] Jin S, Daly DS, Springer DL, Miller JH. The effects of shared peptides on protein quantitation in label-free proteomics by LC/MS/MS. *Journal of proteome research*. 2008;7(1):164-9.
- [151] Wu Q, Zhao Q, Liang Z, Qu Y, Zhang L, Zhang Y. NSI and NSMT: Usages of MS/MS fragment ion intensity for sensitive differential proteome detection and accurate protein fold change calculation in relative label-free proteome quantification. *The Analyst*. 2012;137(13):3146-53.
- [152] Freund DM, Prenni JE. Improved detection of quantitative differences using a combination of spectral counting and MS/MS total ion current. *Journal of Proteome Research*. 2013;12(4):1996-2004.
- [153] Griffin NM, Yu J, Long F, Oh P, Shore S, Li Y, et al. Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nature Biotechnology*. 2010;28(1):83-9.
- [154] Bondarenko PV, Chelius D, Shaler TA. Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. *Analytical Chemistry*. 2002;74(18):4741-9.
- [155] Chelius D, Bondarenko PV. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *Journal of Proteome Research*. 2002;1(4):317-23.
- [156] Wang W, Zhou H, Lin H, Roy S, Shaler TA, Hill LR, et al. Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Analytical Chemistry*. 2003;75(18):4818-26.

- [157] Grossmann J, Roschitzki B, Panse C, Fortes C, Barkow-Oesterreicher S, Rutishauser D, et al. Implementation and evaluation of relative and absolute quantification in shotgun proteomics with label-free methods. *Journal of Proteomics*. 2010;73(9):1740-6.
- [158] Wilhelm M, Schlegl J, Hahne H, Moghaddas Gholami A, Lieberenz M, Savitski MM, et al. Mass-spectrometry-based draft of the human proteome. *Nature*. 2014;509(7502):582-7.
- [159] Shevchenko G, Konzer A, Musunuri S, Bergquist J. Neuroproteomics tools in clinical practice. *Biochimica et Biophysica Acta*. 2015.
- [160] Filiou MD, Martins-de-Souza D, Guest PC, Bahn S, Turck CW. To label or not to label: Applications of quantitative proteomics in neuroscience research. *Proteomics*. 2012;12(4-5):736-47.
- [161] Santos SD, Manadas B, Duarte CB, Carvalho AL. Proteomic analysis of an interactome for long-form AMPA receptor subunits. *Journal of Proteome Research*. 2010;9(4):1670-82.
- [162] Seyfried NT, Gozal YM, Donovan LE, Herskowitz JH, Dammer EB, Xia Q, et al. Quantitative analysis of the detergent-insoluble brain proteome in frontotemporal lobar degeneration using SILAC internal standards. *Journal of Proteome Research*. 2012;11(5):2721-38.
- [163] Seyfried NT, Huysentruyt LC, Atwood JA, 3rd, Xia Q, Seyfried TN, Orlando R. Up-regulation of NG2 proteoglycan and interferon-induced transmembrane proteins 1 and 3 in mouse astrocytoma: A membrane proteomics approach. *Cancer Letters*. 2008;263(2):243-52.
- [164] Bark SJ, Wegrzyn J, Taupenot L, Ziegler M, O'Connor DT, Ma Q, et al. The protein architecture of human secretory vesicles reveals differential regulation of signaling molecule secretion by protein kinases. *PloS One*. 2012;7(8):e41134.
- [165] Zhang X, Zhou JY, Chin MH, Schepmoes AA, Petyuk VA, Weitz KK, et al. Region-specific protein abundance changes in the brain of MPTP-induced Parkinson's disease mouse model. *Journal of Proteome Research*. 2010;9(3):1496-509.
- [166] Tao Y, Fang L, Yang Y, Jiang H, Yang H, Zhang H, et al. Quantitative proteomic analysis reveals the neuroprotective effects of huperzine A for amyloid beta treated neuroblastoma N2a cells. *Proteomics*. 2013;13(8):1314-24.
- [167] Culver BP, Savas JN, Park SK, Choi JH, Zheng S, Zeitlin SO, et al. Proteomic analysis of wild-type and mutant huntingtin-associated proteins in mouse brains identifies unique interactions and involvement in protein synthesis. *The Journal of Biological Chemistry*. 2012;287(26):21599-614.

- [168] Ringman JM, Schulman H, Becker C, Jones T, Bai Y, Immermann F, et al. Proteomic changes in cerebrospinal fluid of presymptomatic and affected persons carrying familial Alzheimer disease mutations. *Archives of Neurology*. 2012;69(1):96-104.
- [169] Xia Q, Liao L, Cheng D, Duong DM, Gearing M, Lah JJ, et al. Proteomic identification of novel proteins associated with Lewy bodies. *Frontiers in Bioscience: A Journal and Virtual Library*. 2008;13:3850-6.
- [170] Andreev VP, Petyuk VA, Brewer HM, Karpievitch YV, Xie F, Clarke J, et al. Label-free quantitative LC-MS proteomics of Alzheimer's disease and normally aged human brains. *Journal of Proteome Research*. 2012;11(6):3053-67.
- [171] He J, Liu Y, Zhu TS, Xie X, Costello MA, Talsma CE, et al. Glycoproteomic analysis of glioblastoma stem cell differentiation. *Journal of Proteome Research*. 2011;10(1):330-8.
- [172] Yoon JH, Kim J, Kim KL, Kim DH, Jung SJ, Lee H, et al. Proteomic analysis of hypoxia-induced U373MG glioma secretome reveals novel hypoxia-dependent migration factors. *Proteomics*. 2014;14(12):1494-502.
- [173] Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: A new concept for consistent and accurate proteome analysis. *Molecular & Cellular Proteomics : MCP*. 2012;11(6):O111 016717.
- [174] Chapman JD, Goodlett DR, Masselon CD. Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. *Mass Spectrometry Reviews*. 2014;33(6):452-70.
- [175] Purvine S, Eppel JT, Yi EC, Goodlett DR. Shotgun collision-induced dissociation of peptides using a time of flight mass analyzer. *Proteomics*. 2003;3(6):847-50.
- [176] Venable JD, Dong MQ, Wohlschlegel J, Dillin A, Yates JR. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nature Methods*. 2004;1(1):39-45.
- [177] Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ. Absolute quantification of proteins by LCMSE: A virtue of parallel MS acquisition. *Molecular & Cellular Proteomics : MCP*. 2006;5(1):144-56.
- [178] Ramos AA, Yang H, Rosen LE, Yao X. Tandem parallel fragmentation of peptides for mass spectrometry. *Analytical Chemistry*. 2006;78(18):6391-7.
- [179] Panchaud A, Scherl A, Shaffer SA, von Haller PD, Kulasekara HD, Miller SI, et al. Precursor acquisition independent from ion count: How to dive deeper into the proteomics ocean. *Analytical Chemistry*. 2009;81(15):6481-8.

- [180] Geiger T, Cox J, Mann M. Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Molecular & Cellular Proteomics* : MCP. 2010;9(10):2252-61.
- [181] Carvalho PC, Han X, Xu T, Cociorva D, Carvalho Mda G, Barbosa VC, et al. XDIA: Improving on the label-free data-independent analysis. *Bioinformatics*. 2010;26(6):847-8.
- [182] Weisbrod CR, Eng JK, Hoopmann MR, Baker T, Bruce JE. Accurate peptide fragment mass analysis: Multiplexed peptide identification and quantification. *Journal of Proteome Research*. 2012;11(3):1621-32.
- [183] Wrona M, Mauriala T, Bateman KP, Mortishire-Smith RJ, O'Connor D. 'All-in-one' analysis for metabolite identification using liquid chromatography/hybrid quadrupole time-of-flight mass spectrometry with collision energy switching. *Rapid Communications in Mass Spectrometry* : RCM. 2005;19(18):2597-602.
- [184] Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, Li GZ, et al. Quantitative proteomic analysis by accurate mass retention time pairs. *Analytical Chemistry*. 2005;77(7):2187-200.
- [185] Collins BC, Gillet LC, Rosenberger G, Rost HL, Vichalkovski A, Gstaiger M, et al. Quantifying protein interaction dynamics by SWATH mass spectrometry: Application to the 14-3-3 system. *Nature Methods*. 2013;10(12):1246-53.
- [186] Roemmelt AT, Steuer AE, Poetsch M, Kraemer T. Liquid chromatography, in combination with a quadrupole time-of-flight instrument (LC QTOF), with sequential window acquisition of all theoretical fragment-ion spectra (SWATH) acquisition: Systematic studies on its use for screenings in clinical and forensic toxicology and comparison with information-dependent acquisition (IDA). *Analytical Chemistry*. 2014;86(23):11742-9.
- [187] Zawadzka AM, Schilling B, Held JM, Sahu AK, Cusack MP, Drake PM, et al. Variation and quantification among a target set of phosphopeptides in human plasma by multiple reaction monitoring and SWATH-MS2 data-independent acquisition. *Electrophoresis*. 2014;35(24):3487-97.
- [188] Guo T, Kouvonen P, Koh CC, Gillet LC, Wolski WE, Rost HL, et al. Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps. *Nature Medicine*. 2015; Advance online publication.
- [189] Anjo SI, Santa C, Manadas B. Short GeLC-SWATH: A fast and reliable quantitative approach for proteomic screenings. *Proteomics*. 2014;DOI: 10.1002/pmic.201400221.
- [190] Schubert OT, Gillet LC, Collins BC, Navarro P, Rosenberger G, Wolski WE, et al. Building high-quality assay libraries for targeted analysis of SWATH MS data. *Nature Protocols*. 2015;10(3):426-41.

- [191] Huang Q, Yang L, Luo J, Guo L, Wang Z, Yang X, et al. SWATH enables precise label-free quantification on proteome-scale. *Proteomics*. 2015.
- [192] Zi J, Zhang S, Zhou R, Zhou B, Xu S, Hou G, et al. Expansion of the ion library for mining SWATH-MS data through fractionation proteomics. *Analytical Chemistry*. 2014;86(15):7242-6.
- [193] Vowinckel J, Capuano F, Campbell K, Deery MJ, Lilley KS, Ralser M. The beauty of being (label)-free: sample preparation methods for SWATH-MS and next-generation targeted proteomics. *F1000Research*. 2013;2:272.
- [194] Rost HL, Rosenberger G, Navarro P, Gillet L, Miladinovic SM, Schubert OT, et al. OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nat Biotech*. 2014;32(3):219-23.
- [195] Tsou CC, Avtonomov D, Larsen B, Tucholska M, Choi H, Gingras AC, et al. DIA-Umpire: Comprehensive computational framework for data-independent acquisition proteomics. *Nature Methods*. 2015;12(3):258-64.
- [196] Rosenberger G, Koh CC, Guo T, Röst HL, Kouvonen P, Collins BC, et al. A repository of assays to quantify 10,000 human proteins by SWATH-MS. *Scientific Data*. 2014;1.
- [197] Shevchenko NM, Anastuyk SD, Menshova RV, Vishchuk OS, Isakov VI, Zadorozhny PA, et al. Further studies on structure of fucoidan from brown alga *Saccharina gurjanovae*. *Carbohydrate Polymers*. 2015;121:207-16.
- [198] Martins-de-Souza D, Guest PC, Mann DM, Roeber S, Rahmoune H, Bauder C, et al. Proteomic analysis identifies dysfunction in cellular transport, energy, and protein metabolism in different brain regions of atypical frontotemporal lobar degeneration. *Journal of Proteome Research*. 2012;11(4):2533-43.
- [199] Lundby A, Secher A, Lage K, Nordborg NB, Dmytriyev A, Lundby C, et al. Quantitative maps of protein phosphorylation sites across 14 different rat organs and tissues. *Nature Communications*. 2012;3:876.
- [200] Jaros JA, Martins-de-Souza D, Rahmoune H, Rothermundt M, Leweke FM, Guest PC, et al. Protein phosphorylation patterns in serum from schizophrenia patients and healthy controls. *Journal of Proteomics*. 2012;76 Spec No.:43-55.
- [201] Levin Y, Wang L, Schwarz E, Koethe D, Leweke FM, Bahn S. Global proteomic profiling reveals altered proteomic signature in schizophrenia serum. *Molecular Psychiatry*. 2010;15(11):1088-100.
- [202] Martins-de-Souza D, Guest PC, Harris LW, Vanattou-Saifoudine N, Webster MJ, Rahmoune H, et al. Identification of proteomic signatures associated with depression and psychotic depression in post-mortem brains from major depression patients. *Translational Psychiatry*. 2012;2:e87.
- [203] Martins-de-Souza D, Guest PC, Vanattou-Saifoudine N, Rahmoune H, Bahn S. Phosphoproteomic differences in major depressive disorder postmortem brains indicate

- effects on synaptic function. *European Archives of Psychiatry and Clinical Neuroscience*. 2012;262(8):657-66.
- [204] Chan MK, Tsang TM, Harris LW, Guest PC, Holmes E, Bahn S. Evidence for disease and antipsychotic medication effects in post-mortem brain from schizophrenia patients. *Molecular Psychiatry*. 2011;16(12):1189-202.
- [205] Ma D, Guest PC, Bahn S. Metabonomic studies of schizophrenia and psychotropic medications: Focus on alterations in CNS energy homeostasis. *Bioanalysis*. 2009;1(9):1615-26.
- [206] Steeb H, Ramsey JM, Guest PC, Stocki P, Cooper JD, Rahmoune H, et al. Serum proteomic analysis identifies sex-specific differences in lipid metabolism and inflammation profiles in adults diagnosed with Asperger syndrome. *Molecular Autism*. 2014;5(1):4.
- [207] Liu Y, Huttenhain R, Surinova S, Gillet LC, Mouritsen J, Brunner R, et al. Quantitative measurements of N-linked glycoproteins in human plasma by SWATH-MS. *Proteomics*. 2013;13(8):1247-56.
- [208] Liu Y, Buil A, Collins BC, Gillet LC, Blum LC, Cheng LY, et al. Quantitative variability of 342 plasma proteins in a human twin population. *Molecular systems biology*. 2015;11(2):786.
- [209] Anjo SI, Santa C, Manadas B. Short GeLC-SWATH: A fast and reliable quantitative approach for proteomic screenings. *Proteomics*. 2015;15(4):757-62.
- [210] Villeneuve LM, Stauch KL, Fox HS. Proteomic analysis of the mitochondria from embryonic and postnatal rat brains reveals response to developmental changes in energy demands. *Journal of Proteomics*. 2014;109:228-39.
- [211] Villeneuve LM, Purnell PR, Boska MD, Fox HS. Early expression of Parkinson's disease-related mitochondrial abnormalities in PINK1 knockout rats. *Molecular Neurobiology*. 2014.
- [212] Arike L, Valgepea K, Peil L, Nahku R, Adamberg K, Vilu R. Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. *Journal of Proteomics*. 2012;75(17):5437-48.
- [213] Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, et al. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Molecular & Cellular Proteomics : MCP*. 2005;4(9):1265-72.
- [214] Lu P, Vogel C, Wang R, Yao X, Marcotte EM. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nature Biotechnology*. 2007;25(1):117-24.

- [215] Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473(7347):337-42.
- [216] Wisniewski JR, Ostasiewicz P, Dus K, Zielinska DF, Gnad F, Mann M. Extensive quantitative remodeling of the proteome between normal colon tissue and adenocarcinoma. *Molecular Systems Biology*. 2012;8:611.
- [217] Fabre B, Lambour T, Bouyssié D, Menneteau T, Monsarrat B, Burlet-Schiltz O, et al. Comparison of label-free quantification methods for the determination of protein complexes subunits stoichiometry. *EuPA Open Proteomics*. 2014;4(0):82-6.
- [218] Wilhelm BG, Mandad S, Truckenbrodt S, Krohnert K, Schafer C, Rammner B, et al. Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins. *Science*. 2014;344(6187):1023-8.
- [219] Shin JB, Krey JF, Hassan A, Metlagel Z, Tauscher AN, Pagana JM, et al. Molecular architecture of the chick vestibular hair bundle. *Nature Neuroscience*. 2013;16(3):365-74.
- [220] Hög S, van Bebbber F, Dislich B, Kuhn PH, Haass C, Schmid B, et al. Label-free quantitative analysis of the membrane proteome of Bace1 protease knock-out zebrafish brains. *Proteomics*. 2013;13(9):1519-27.
- [221] Finlay EM, Games DE, Startin JR, Gilbert J. Screening, confirmation, and quantification of sulphonamide residues in pig kidney by tandem mass spectrometry of crude extracts. *Biomed Environ Mass Spectrom*. 1986;13(11):633-9.
- [222] Prasad B, Unadkat JD. Optimized approaches for quantification of drug transporters in tissues and cells by MRM proteomics. *AAPS J*. 2014;16(4):634-48.
- [223] Chambers AG, Percy AJ, Simon R, Borchers CH. MRM for the verification of cancer biomarker proteins: Recent applications to human plasma and serum. *Expert Rev Proteomics*. 2014;11(2):137-48.
- [224] Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: Workflows, potential, pitfalls and future directions. *Nat Methods*. 2012;9(6):555-66.
- [225] Sajic T, Liu Y, Aebersold R. Using data-independent, high resolution mass spectrometry in protein biomarker research: Perspectives and clinical applications. *Proteomics Clin Appl*. 2014.
- [226] Ngounou Wetie AG, Wormwood KL, Russell S, Ryan JP, Darie CC, Woods AG. A Pilot Proteomic Analysis of Salivary Biomarkers in Autism Spectrum Disorder. *Autism Res*. 2015.
- [227] White NM, Masui O, Desouza LV, Krakovska O, Metias S, Romaschin AD, et al. Quantitative proteomic analysis reveals potential diagnostic markers and pathways involved in pathogenesis of renal cell carcinoma. *Oncotarget*. 2014;5(2):506-18.
- [228] Dittmar G, Selbach M. SILAC for biomarker discovery. *Proteomics Clin Appl*. 2014.

- [229] Liebler DC, Zimmerman LJ. Targeted quantitation of proteins by mass spectrometry. *Biochemistry*. 2013;52(22):3797-806.
- [230] Mahoney DW, Therneau TM, Heppelmann CJ, Higgins L, Benson LM, Zenka RM, et al. Relative quantification: characterization of bias, variability and fold changes in mass spectrometry data from iTRAQ-labeled peptides. *J Proteome Res*. 2011;10(9):4325-33.
- [231] Smith RD. Mass spectrometry in biomarker applications: From untargeted discovery to targeted verification, and implications for platform convergence and clinical application. *Clin Chem*. 2012;58(3):528-30.
- [232] Boja ES, Fehniger TE, Baker MS, Marko-Varga G, Rodriguez H. Analytical validation considerations of multiplex mass-spectrometry-based proteomic platforms for measuring protein biomarkers. *J Proteome Res*. 2014;13(12):5325-32.
- [233] Boja ES, Rodriguez H. Mass spectrometry-based targeted quantitative proteomics: Achieving sensitive and reproducible detection of proteins. *Proteomics*. 2012;12(8):1093-110.
- [234] Anderson L, Hunter CL. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol Cell Proteomics*. 2006;5(4):573-88.
- [235] Lange V, Picotti P, Domon B, Aebersold R. Selected reaction monitoring for quantitative proteomics: A tutorial. *Mol Syst Biol*. 2008;4:222.
- [236] Gallien S, Duriez E, Domon B. Selected reaction monitoring applied to proteomics. *J Mass Spectrom*. 2011;46(3):298-312.
- [237] Yocum AK, Chinnaiyan AM. Current affairs in quantitative targeted proteomics: Multiple reaction monitoring-mass spectrometry. *Brief Funct Genomic Proteomic*. 2009;8(2):145-57.
- [238] Fusaro VA, Mani DR, Mesirov JP, Carr SA. Prediction of high-responding peptides for targeted protein assays by mass spectrometry. *Nat Biotechnol*. 2009;27(2):190-8.
- [239] Mallick P, Schirle M, Chen SS, Flory MR, Lee H, Martin D, et al. Computational prediction of proteotypic peptides for quantitative proteomics. *Nat Biotechnol*. 2007;25(1):125-31.
- [240] Sanders WS, Bridges SM, McCarthy FM, Nanduri B, Burgess SC. Prediction of peptides observable by mass spectrometry applied at the experimental set level. *BMC Bioinformatics*. 2007;8 Suppl 7:S23.
- [241] Mollah S, Wertz IE, Phung Q, Arnott D, Dixit VM, Lill JR. Targeted mass spectrometric strategy for global mapping of ubiquitination on proteins. *Rapid Commun Mass Spectrom*. 2007;21(20):3357-64.

- [242] Lange V, Malmstrom JA, Didion J, King NL, Johansson BP, Schafer J, et al. Targeted quantitative analysis of *Streptococcus pyogenes* virulence factors by multiple reaction monitoring. *Mol Cell Proteomics*. 2008;7(8):1489-500.
- [243] Colangelo CM, Ivosev G, Chung L, Abbott T, Shifman M, Sakaue F, et al. Development of a highly automated and multiplexed targeted proteome pipeline and assay for 112 rat brain synaptic proteins. *Proteomics*. 2014.
- [244] Makawita S, Diamandis EP. The bottleneck in the cancer biomarker pipeline and protein quantification through mass spectrometry-based approaches: Current strategies for candidate verification. *Clin Chem*. 2010;56(2):212-22.
- [245] Tong L, Zhou XY, Jylha A, Aapola U, Liu DN, Koh SK, et al. Quantitation of 47 human tear proteins using high resolution multiple reaction monitoring (HR-MRM) based-mass spectrometry. *J Proteomics*. 2015;115:36-48.
- [246] Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A*. 2003;100(12):6940-5.
- [247] Barnidge DR, Dratz EA, Martin T, Bonilla LE, Moran LB, Lindall A. Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. *Anal Chem*. 2003;75(3):445-51.
- [248] Campbell J, Rezai T, Prakash A, Krastins B, Dayon L, Ward M, et al. Evaluation of absolute peptide quantitation strategies using selected reaction monitoring. *Proteomics*. 2011;11(6):1148-52.
- [249] Kusmierz JJ, Sumrada R, Desiderio DM. Fast atom bombardment mass spectrometric quantitative analysis of methionine-enkephalin in human pituitary tissues. *Anal Chem*. 1990;62(21):2395-400.
- [250] Kheterpal I, Kastin AJ, Mollah S, Yu C, Hsuchou H, Pan W. Mass spectrometric quantification of MIF-1 in mouse brain by multiple reaction monitoring. *Peptides*. 2009;30(7):1276-81.
- [251] Chang RY, Etheridge N, Dodd P, Nouwens A. Quantitative multiple reaction monitoring analysis of synaptic proteins from human brain. *J Neurosci Methods*. 2014;227:189-210.
- [252] Choi YS, Hou S, Choe LH, Lee KH. Targeted human cerebrospinal fluid proteomics for the validation of multiple Alzheimer's disease biomarker candidates. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2013;930:129-35.
- [253] IJsselstijn L, Dekker LJ, Koudstaal PJ, Hofman A, Sillevius Smitt PA, Breteler MM, et al. Serum clusterin levels are not increased in presymptomatic Alzheimer's disease. *J Proteome Res*. 2011;10(4):2006-10.
- [254] Lame ME, Chambers EE, Blatnik M. Quantitation of amyloid beta peptides A β (1-38), A β (1-40), and A β (1-42) in human cerebrospinal fluid by ultra-perform-

- ance liquid chromatography-tandem mass spectrometry. *Anal Biochem.* 2011;419(2):133-9.
- [255] Wildsmith KR, Schauer SP, Smith AM, Arnott D, Zhu Y, Haznedar J, et al. Identification of longitudinally dynamic biomarkers in Alzheimer's disease cerebrospinal fluid by targeted proteomics. *Mol Neurodegener.* 2014;9:22.
- [256] DeSouza LV, Taylor AM, Li W, Minkoff MS, Romaschin AD, Colgan TJ, et al. Multiple reaction monitoring of mTRAQ-labeled peptides enables absolute quantification of endogenous levels of a potential cancer marker in cancerous and normal endometrial tissues. *J Proteome Res.* 2008;7(8):3525-34.
- [257] Brun V, Masselon C, Garin J, Dupuis A. Isotope dilution strategies for absolute quantitative proteomics. *J Proteomics.* 2009;72(5):740-9.
- [258] Pratt JM, Simpson DM, Doherty MK, Rivers J, Gaskell SJ, Beynon RJ. Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nat Protoc.* 2006;1(2):1029-43.
- [259] Simpson DM, Beynon RJ. QconCATs: design and expression of concatenated protein standards for multiplexed protein quantification. *Anal Bioanal Chem.* 2012;404(4):977-89.
- [260] Chen J, Wang M, Turko IV. Mass spectrometry quantification of clusterin in the human brain. *Mol Neurodegener.* 2012;7:41.
- [261] Chen J, Wang M, Turko IV. Quantification of amyloid precursor protein isoforms using quantification concatamer internal standard. *Anal Chem.* 2013;85(1):303-7.
- [262] Brun V, Dupuis A, Adrait A, Marcellin M, Thomas D, Court M, et al. Isotope-labeled protein standards: Toward absolute quantitative proteomics. *Mol Cell Proteomics.* 2007;6(12):2139-49.
- [263] Wang M, Heo GY, Omarova S, Pikuleva IA, Turko IV. Sample prefractionation for mass spectrometry quantification of low-abundance membrane proteins. *Anal Chem.* 2012;84(12):5186-91.
- [264] Singh S, Springer M, Steen J, Kirschner MW, Steen H. FLEXIQuant: a novel tool for the absolute quantification of proteins, and the simultaneous identification and quantification of potentially modified peptides. *J Proteome Res.* 2009;8(5):2201-10.
- [265] Zeiler M, Straube WL, Lundberg E, Uhlen M, Mann M. A Protein Epitope Signature Tag (PrEST) library allows SILAC-based absolute quantification and multiplexed determination of protein copy numbers in cell lines. *Mol Cell Proteomics.* 2012;11(3):O111 009613.
- [266] Hanke S, Besir H, Oesterhelt D, Mann M. Absolute SILAC for accurate quantitation of proteins in complex mixtures down to the attomole level. *J Proteome Res.* 2008;7(3):1118-30.

Symbiotic Proteomics – State of the Art in Plant–Mycorrhizal Fungi Interactions

Marco Chiapello, Silvia Perotto and Raffaella Balestrini

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/61331>

Abstract

Mycorrhizae are symbiotic associations between soil fungi belonging to diverse taxa and the roots of about 90% of all terrestrial plant species. The mutualistic nature of these symbioses is based on the nutritional exchanges between the partners. However, the benefits to the plant partner are not limited to an improved mineral nutrition because they also include a general increase in stress tolerance and health. Because of these benefits, mycorrhizae are of great interest in sustainable agriculture and forestry. In the past few years, the development of high-throughput molecular tools, in addition to the advancements in microscopy techniques, has allowed us to gain a deeper insight on the molecular mechanisms underlying the establishment and functioning of these symbioses. In this chapter, we focus on the use of proteomic tools to better understand the molecular bases of cell communication and the regulation of developmental and metabolic pathways in mycorrhizal associations.

Keywords: Proteomics, mycorrhizal associations, laser microdissection

1. Introduction

Plants cannot move away from unfavourable environments, or run away from hungry eaters, or escape from detrimental microorganisms. Fortunately, not all environments and all organisms are a threat to plants, and plants have also evolved strategies to survive adverse environmental conditions. In fact, plants have adapted to most environments, they have learned how to avoid risky relationships with detrimental microorganisms, how to be unconcerned by neutral microorganisms and how to develop intimate affairs with beneficial partners.

The latter type of interaction is referred to as a 'mutualistic symbiosis'. Symbiosis was first defined at the end of the nineteenth century by Anton De Bary as a term that simply described the regular coexistence of taxonomically different organisms [1]. A mutualistic symbiosis means more than 'regular coexistence', as it includes all relationships in which both partners can benefit from the association, and where benefits can be measured in terms of fitness and nutrient exchange.

Among mutualistic symbioses, the association of plants with nitrogen-fixing rhizobia [2] and with mycorrhizal fungi, in particular, are the result of a long co-evolution and co-operation between plants and soil microbes [1]. Different types of mycorrhizae have been found in nature: ectomycorrhiza (ECM) is predominant in forest soils and is characterized by the fact that the fungal hyphae remain outside the plant cell; endomycorrhiza comprises orchid, ericoid and arbuscular mycorrhiza (AM) and derives its name from the fact that the fungal hyphae are able to enter into the plant root cells [3].

ECM fungi have evolved from wood- and litter-decaying fungal ancestors, without any obvious reversal to saprotrophy [4]. Although the oldest ECM root fossils date back to 50 million years ago (MYA) [5], molecular analyses place the origin of ECM fungi in the Cretaceous [6] and suggest that they probably played a role in the migration of plants from the tropics to the poorer temperate regions [7]. ECM fungi mostly belong to Basidiomycota and Ascomycota, and they form symbioses with a relative small number of plant species [4]. They play an important role in forest establishment and in the successful reforestation of harsh environments, such as saline areas [8]. Moreover, ECM fungi can form fruiting bodies which have an important economic impact, such as truffles.

The AM symbiosis involves the majority of crop plants and results from the successful interaction between fungi in the Glomeromycota and the roots of about 80% of terrestrial plants [9]. This symbiosis is one of the oldest biotrophic interactions, dating back 400–450 MYA and is thought to have played a pivotal role in the water-to-land transition during plant evolution [10] (Figure 1). AM fungi have become so intimately dependent on plants that they are obligate biotrophs.

The evolutionary success of mycorrhizal symbioses likely derives from the bidirectional nutrient exchange that takes place between the two partners in most associations: fungi deliver mineral nutrients to the plants, while receiving sugars in return. It has been estimated that up to 20% of the photosynthesis-derived compounds of terrestrial plants (approximately 5 billion tons of carbon per year) are consumed by symbiotic fungi [11]. On the other hand, for example, 70% of the overall Pi acquired by arbuscular mycorrhizal rice plants is delivered *via* the symbiotic route [12]. Mycorrhizal plants benefit from their interaction with symbiotic fungi not only in terms of improved mineral nutrition, with an increased biomass production, but also in better protection against pathogens and abiotic stresses [13].

Because of the importance of mycorrhizal symbioses in plant health, several studies have focused on the biology, evolution and biodiversity of mycorrhizal associations [14]. In particular, the recent development of high-throughput molecular tools has allowed us to gain deeper knowledge on the molecular mechanisms governing the plant–fungus interaction [14],

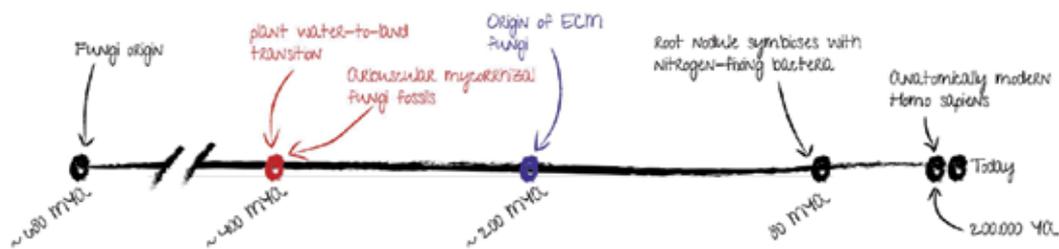


Figure 1. Schematic timeline of the root symbiosis development.

providing useful information for the application of these beneficial fungal agents to optimize plant health, nutrition and yields in sustainable agriculture and forestry.

Although examples will be given for all mycorrhizal types, this chapter mainly focuses on AM associations due to the following reasons. First, fossil and molecular records indicate for AM fungi a very long co-evolution with plants, with an unchanged morphology over 400 million years [15]. This observation opens several interesting questions such as: when has this symbiosis evolved? Has the molecular machinery that regulates this symbiosis evolved over time, or are we looking at the same situation fixed millions years ago? Understanding the biology of this obligate biotrophic interaction is a scientific challenge, but it would allow us to unravel the molecular mechanisms of the oldest known symbiosis [16]. The second reason is the high relevance of AM symbiosis for crop plants; better knowledge of these associations would have agro-environmental applications, with consequent economic and social impact.

2. Plant–symbiotic fungi interactions

The plant AM fungal interaction starts in the soil surrounding the plant roots, a region termed rhizosphere, where both plants and fungi release chemical signals in a pre-symbiotic molecular dialogue [17]. Among their root exudates, plants release in the rhizosphere signals such as strigolactones and cutin monomers, which elicit hyphal branching in AM fungi as well as apical growth of fungal hyphae towards the root surface, following the gradient of plant molecules (Figure 2A). Although the fungal receptors for these plant molecules remain unknown, it has been proved that they are perceived by the fungus, causing a signal cascade.

Fungal signal molecules have been identified in the past few years as being lipo-chito-oligosaccharides (LCOs) [18], the same type of signal molecules produced by rhizobia when interacting with legume plants, and chito-oligosaccharides (COs) [19]. Although the plant receptors for the fungal signal molecules have not been identified yet, large families of receptors are predicted to potentially bind these molecules.

Thanks to the exchange of these plant and fungal signal molecules, the plant and the AM fungus recognize each other and begin a more intimate phase of the interaction, with the fungus starting root colonization. The plant paves the way for fungal colonization by building up the so-called pre-penetration apparatus (PPA) [20], a transient assembly that defines the

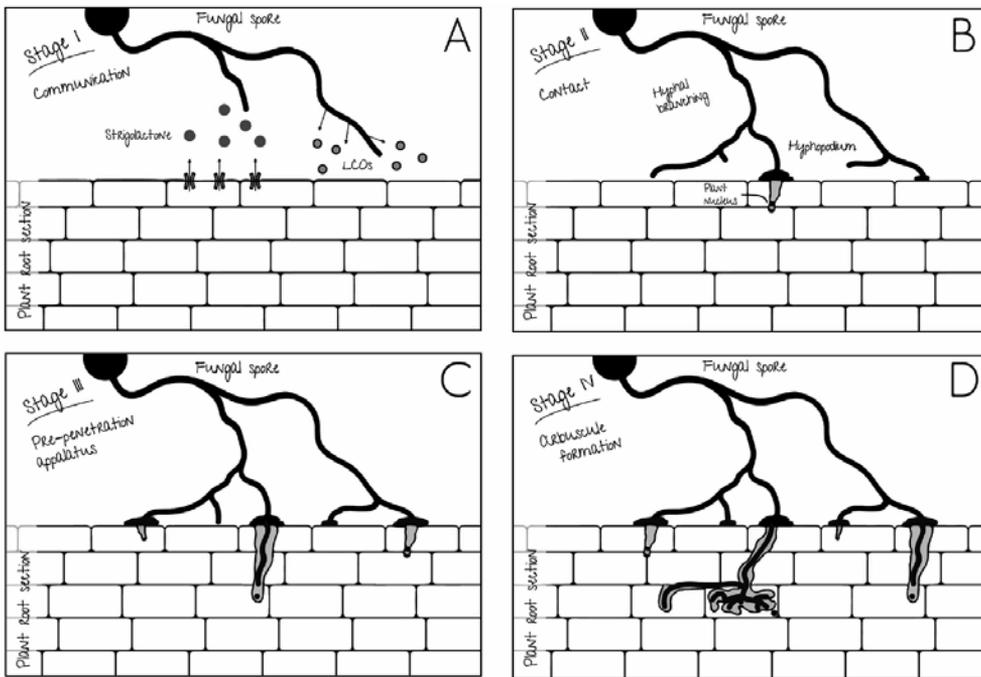


Figure 2. The boxes represent the four phases of formation of the plant–fungus association in the AM symbiosis. (A) Plant roots exude strigolactones and induce hyphal branching, while the fungus releases LCOs, perceived by the plant. (B) The AM fungus contacts the root surface and forms hyphopodia. (C) Epidermal plant cells produce a pre-penetration apparatus; the AM fungus starts to grow inside the plant and reaches the cortex. (D) The AM fungal hypha branches inside the cortex cells and form the arbuscules.

path followed by the fungal hyphae toward the inner root layers (Figure 2B–C). The AM fungus follows the path created by the PPA until the root cortex, where it starts to form a tree-like structure called 'arbuscule' (Figure 2D). The arbuscule is the core of a functional AM symbiosis, shaped as a highly branched structure where each hyphal branch is surrounded by the plant cell membrane. The contact surface between the plant and the AM fungus greatly increases around the arbuscule, thus increasing the area of nutrient exchanges. After few (ca. 4–5) days, the arbuscule collapses [21] and is replaced by a new one in the same or in another cortical cell. During AM fungal colonization, 'early stage' indicates the phase occurring prior to and during the initial contact between the two symbionts. This stage ends with the formation of the arbuscules that mark the transition to the 'late stage' of the symbiosis. However, fungal colonization of plant roots occurs at many access points not normally synchronized and the mycorrhizal symbiosis is highly dynamic, meaning that when new access points are created, arbuscules are forming and collapsing. Therefore, early and late stages can be really distinguished only after the first contacts between plant and fungus. The intracellular accommodation of unbranched hyphae (during the early stage) and of arbuscules (at a later stage) is a coordinated developmental process between the plant and the fungal cells: it involves an intricate and largely unexplored signal exchange, intense secretory activity related to the biogenesis of the perifungal membrane and an overall reorganization of the cell architecture.

Several plant genes are known to be required for the establishment and functioning of the AM symbiosis. Some of them encode proteins that are components of the so-called 'SYM pathway' and are essential for early signalling and root colonization [10]. Other genes are likely involved in nutrient exchange during arbuscule functioning, such as the *Medicago truncatula* gene coding for a phosphate transporter (PT4) specifically induced in arbuscule-containing cells [22]. However, despite molecular and cellular evidence of the expression of these genes in arbuscule-containing cells, the corresponding proteins have not been identified through proteomic approaches until very recently [23], most likely because of their accumulation in a small sub-population of root cells, those harbouring the arbuscules, and because of technical difficulties with membrane protein extraction.

Whereas the main signal molecules involved in the AM fungus–host plant dialogue have been identified, little information is so far available on the recognition events and on the long-maintenance factors involved in the ECM symbiosis (Garcia et al. 2015), although auxin and ethylene have been identified as some of the signals exchanged between the two partners in ECM [24,25]. By contrast, nothing is known concerning this aspect in the ericoid and the orchid mycorrhiza. The colonization steps have a very different morphology in ECM and AM symbioses. During the symbiotic phase, ECM fungi form a fungal sheath (the mantle) that develops outside the root. From the inner layers of the mantle, some hyphae penetrate between the epidermal and the outer cortical cells to form an intercellular hyphal network (the Hartig net) inside the root tissues [1]. Mycorrhiza-induced small secreted proteins (MiSSPs) are fungal proteins known to be involved during the formation and maintenance of the symbiosis between ECM fungi and their host plant [26,27].

The identification of the key molecular players in mycorrhizal symbioses is mandatory to understand the complex interactions between the symbiotic partners and the ways to improve and fully exploit their symbiotic potential in sustainable crop and forest management. Genome sequences of several mycorrhizal fungal species are now available and provide a great opportunity to increase our knowledge on the mycorrhizal lifestyle, on the metabolic capabilities of these symbioses and on the molecular dialogue between the two symbiotic partners [28].

3. The symbiotic proteomics of mycorrhizal interactions

Proteomics is the large-scale study of proteins from a specific proteome in order to understand cellular processes, and includes assessment of protein abundance, protein modifications, along with identification of interacting partners and networks. As the aim of proteomics is the identification of proteins, the molecular components actually taking part in cellular processes, rather than their genetic information, proteomics could be the main technique to unravel the key players of mycorrhizal symbioses. However, when the number of proteomics studies is compared with those using genomics, transcriptomics or microscopy, the gap is very significant (Table 1). One of the reasons is that the methods for protein identification are based, nowadays, mainly on mass spectrometry, a more complex and expensive technology than

high-throughput DNA or RNA sequencing. Protein identification is made by matching the peptides masses to corresponding masses calculated by the software on proteins or translated gene sequences available in databases. If sequences are not found in databases, protein identification fails. Over the years, many loopholes have been found, among them the easiest was to use sequences from other species. Concerning the identification of plant proteins, the best-studied and well-sequenced plant is *Arabidopsis thaliana*. Unfortunately, this plant is not able to form any type of mycorrhizal symbiosis. In the past few years, DNA sequencing has become cheaper and almost a routine technique, allowing the genome sequence of many organisms to become available.

Another aspect that has hindered the use of proteomics in the study of mycorrhizal interactions is the fact that the mycorrhizal symbiosis involves a small percentage of plant root cells, that may contain fungal structures at different developmental stages and with different putative roles [29]. For example, arbuscules are limited to the root cortical cells in the AM symbiosis, a tissue where not all plant cells are colonized. In addition, the majority of key proteins are likely to be membrane proteins. Taken together, this means that protein extraction from AM roots will lead to a very small percentage of proteins expressed in symbiosis. This 'dilution effect' has made it very hard to identify the key proteins directly involved in plant–fungus interactions.

In summary, the lack of sequence databases of reference organisms and the difficulties in protein extraction have characterized the first decade of proteomics applications to mycorrhizal symbiosis, and explain the limited results obtained.

4. Proteomics in action

The first proteomic investigation of the plant–mycorrhizal fungus symbiosis was published by Dumas et al. (1990) and used mono-dimensional polyacrylamide gel electrophoresis (PAGE) to separate soluble proteins from non-mycorrhizal roots and from roots infected by different AM fungi [30]. After this pioneering study, and because of improvements in sample extraction, sample purification and in the technological performance of the equipment, many studies have aimed to identify the key players involved in mycorrhizal interactions (Figure 3). Many strategies have been set up, depending on the target mycorrhizal type, on the symbiotic stage of interest and subcellular localization [30].

4.1. Proteomics on the early stages of the mycorrhizal symbiosis

The studies on the early stages of the symbiosis coincide with the earlier studies in symbiotic proteomics. Burgess and collaborators set up a complex experiment to identify proteins either induced, enhanced or inhibited during the early stages of ECM development [31]. They compared, over a time-course, the profiles of proteins expressed in roots inoculated with three different isolates of *Pisolithus tinctorius* showing different degrees of root colonization: isolate H2144 exhibited a very high infectivity, isolate 441 showed moderate infectivity, while isolate H506 was not able to induce ECM [31]. They used two-dimensional electrophoresis (2DE)-

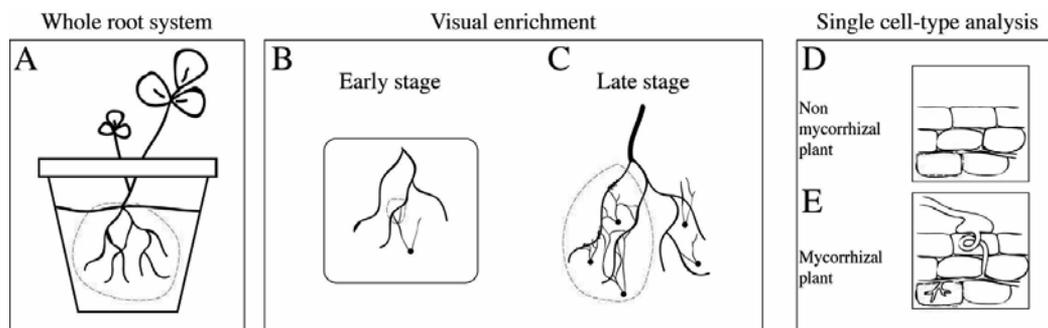


Figure 3. The figure represents different approaches that can be used to collect biological material for AM roots proteomics. (A) Collection of the whole root is the easiest and quickest approach, but it has a drawback due to the fact that the amount of proteins involved in the symbiosis is a very small percentage of the all extracted proteins. The visual enrichment approach allows the collection only of the roots in contact with the fungus. (B) The root organ culture system allows to study the early stages of colonization by following the root and hypha growth until they contact each other and then collect only the root pieces reached by the hyphae. (C) By microscope inspection, roots with the higher percentage of fungal hyphae in contact can be selected. (D–E) Using a laser microdissection (LMD) approach, it is possible to select non-colonized cells (D) and colonized cells (E) from root sections, and to collect them separately, thus avoiding the dilution effect caused by the heterogeneous situation of a mycorrhizal root.

PAGE, a gel electrophoresis technique introduced in the middle 1970s by O'Farrell and Klose and able to separate with high-resolution proteins by two orthogonal properties: iso-electrical point and molecular weight [32]. With this approach, Burgess et al. (1995) found that the morphological changes observed in the inoculated plant roots were linked with massive changes in protein composition, and claimed that these changes commenced at the time of contact between the two partners [31]. It was later discovered [33,34] that these morphological changes during the establishment of ECM were not caused by the contact between plant and fungus, because molecular signals released in the rhizosphere were sufficient to trigger them. The work by Burgess et al. (1995) was nevertheless important because it revealed some plant and fungal symbiosis-related polypeptides and demonstrated that their upregulation was tightly correlated with fungal infectivity.

Five years later, another paper on the early stage of the ECM symbiosis was published by Laurent et al. (1999). They used the same methodological approach, 2DE-PAGE separation, but they focused their attention on the cell wall polypeptides, in order to identify cell surface proteins involved in ECM symbiosis development. It was a huge sampling effort not only because they analyzed the early stage of the symbiosis but also because they had to enrich samples for cell wall polypeptides (CWP). One of the main results was the observation of the enhanced synthesis of several immunologically related 31- and 32-kDa fungal polypeptides, called symbiosis-regulated acidic polypeptides (SRAPs) [35]. As gene expression studies were also carried out, these proteomic data also highlighted the fact that expression of SRAP-32 was regulated at transcriptional level, suggesting that the synthesis of new hyphal proteins is an important process during symbiosis formation [35].

In AM symbiosis, the phase between the first contact and the formation of the first arbuscule, a period ranging between few hours and 1–2 weeks after inoculation, is normally consid-

ered as an early stage of the interaction [36]. Focus on this particular phase is important to understand the cross-talk between the partners and how the proteomes of the two organisms change during the colonization events. A study by the group of Dumas-Gaudot [37] focused on the early stages of the AM symbiosis in three different genotypes of the model plant *M. truncatula*: the wild-type (J5), a mycorrhiza-defective (TRV25, *dmi3*) and an autoregulation-defective (TR122, *sun1*) genotype. The study was aimed at investigating changes in the root proteome elicited in response to appressorium formation by *Glomus intraradices*. For this purpose, the authors compared by 2DE-PAGE the root proteome from non-inoculated roots and from roots synchronized for appressorium formation by *G. intraradices*. The authors showed that proteins that responded to appressorium formation were differentially expressed in different genotypes. This paper was important because it also reported, for the first time, the identification of plant root proteins involved in mycorrhizal symbioses by matrix-assisted laser desorption ionization (MALDI) time-of-flight (TOF) mass spectrometry. This technique revealed appressorium-responsive proteins that were previously unknown on the basis of transcriptome analyses, demonstrating that proteomics and transcriptomics are complementary approaches [37].

The early stage of the mycorrhizal symbiosis represents a challenging, but also a very attractive, stage for proteomics. The sampling time and sampling method are crucial for the experiment's outcome, and synchronization of the root colonization events would enrich root samples in the proteins of interest. Lopez-Meyer and Harrison (2006) first proposed a system to synchronize AM fungal spore germination and root penetration events [38]. Despite this technical improvement, the amount of proteins involved in the plant response to the AM fungus and in the PPA formation, as compared to the amount of proteins in the whole root, is expected to be very low and at the limit of detectability. Therefore, new strategies for sample preparation and experimental designs are required to reveal the key components of this fundamental phase of the mycorrhizal symbiosis.

4.2. Proteomics on late stage

Late stage of the AM symbiosis commonly indicates the phase in which the fungus has already build up the arbuscules. In nature, the symbiosis is highly dynamic and very complex because, while arbuscules are forming, new penetration events occur.

Bestel-Corre et al. [39] reported the first mass spectrometry (MS)-based identification of mycorrhiza-related proteins. These authors studied the response of *M. truncatula* inoculated either with the AM fungus *Glomus mosseae* (current name *Funneliformis mosseae*) or with the nitrogen-fixing bacterium *Sinorhizobium meliloti*. Proteins were separated by 2DE-PAGE and image analyses, with precise quantification of spots volume performed to identify differentially expressed protein spots. Those spots were excised from the gels and analyzed by mass spectrometry. Notably, only plant or bacterial proteins were identified, may be due to difficulties in extracting fungal proteins. The authors identified several proteins related to defence responses, root physiology and respiratory pathway. However, none seemed to be a key protein in the mycorrhizal symbiosis.

The proteomic analysis of the late stage of the AM symbiosis using the whole root system as starting material never allowed very good results. To overcome this problem, many studies on the late stage of the symbiosis focused their attention on a specific sub-cellular compartment. With this approach, the same group published two other papers few years later, using sub-cellular fractionation methods, reporting more remarkable results described in the next paragraphs.

Mycorrhizal systems different from ECM and AM have been seldom investigated with proteomic approaches. However, recent results have been published for orchid mycorrhiza by Valadares et al. (2014). For orchids, the association with symbiotic fungi is required for seed germination and seedling development, when the plant relies on the fungus also for carbon supply (a strategy termed mycoheterotrophy). Recently, 2D-LC-MS/MS (two-dimensional liquid chromatography MS/MS) coupled to isobaric tagging for relative and absolute quantification has been used to identify proteins with differential accumulation in the orchid species *Oncidium sphacelatum* at different stages of plant development after seed inoculation with a *Ceratobasidium* sp. fungal isolate. Eighty-eight proteins, including proteins putatively involved in energy metabolism, cell rescue and defence, molecular signalling and secondary metabolism, have been identified and quantified. These results suggest profound metabolic changes during the development of mycorrhizal orchids, likely related to a switch from the fully mycoheterotrophic to the photosynthetic stages [40].

4.3. Proteomics on sub-cellular compartment

Although 30% of naturally occurring proteins are predicted to be embedded in biological membranes [41], comprehensive membrane proteomics is technically difficult due to the hydrophobicity, heterogeneity and lower abundance of membrane proteins. In mycorrhizal symbioses, membrane proteins are very important because they likely include the receptors that control the fungal–plant dialogue as well as the transporters that mediate nutrient exchange. For these reasons, many authors have focused on membrane proteins using different enrichment protocols.

The first study on sub-cellular fractionation of membrane proteins was conducted by Benabdellah et al. (1998). They isolated the microsomal protein fraction from colonized tomato roots, where they found several differentially expressed proteins [42]. Few years later, the same group identified for the first time a protein related to mycorrhizal symbiosis by Edman N-terminal sequencing, after plasma membrane enrichment and 2DE-PAGE protein separation [43].

The years between 2000 and 2010 saw the rapid increase of mass spectrometry as the main proteomic technique for protein identification and quantification, replacing other techniques previously used. The difficulties and the low number of proteins identified with the Edman N-terminal sequencing were overcome with the advent of mass spectrometry, and new intriguing possibilities were opened.

Valot et al. (2005) also used a sub-cellular proteomic approach to monitor membrane-associated protein modifications during the AM symbiosis [44]. 2DE-PAGE of root microsomes

revealed some mycorrhiza-responsive proteins including 15 induced, 3 up-regulated, and also 18 down-regulated proteins. Among those 36 regulated proteins, 25 were identified using the MALDI-TOF. Except for an acid phosphatase and a lectin, none of them was previously reported as being regulated during the AM symbiosis. This sub-cellular proteomic approach allowed for the first time the identification of fungal proteins expressed *in planta*. In their final conclusion, the authors pinpointed the next challenge: the identification of membrane proteins located in and around the arbuscule, the mycorrhizal symbiosis-specific fungal structure.

Arbuscules are ephemeral structures that form continuously and collapse at the end of a short life-span. The identification of proteins temporarily present in a sub-set of cell types remains, at the present time, a technical challenge for quantitative proteomics. Moreover, membranes associated with the arbuscules are significantly less, relative to the overall root membranes, as also suggested by the low amount of fungal RNA found in extensively colonized AM roots, maximally reaching 12% of the total RNA extracted [45]. Despite these considerations, the same group attempted to enrich samples for plasma membranes using a discontinuous sucrose gradient method [46]. In this chapter, two complementary proteomics methodologies for protein fractionation and identification were applied for the first time to the plant-AM fungus symbiotic association: an automated 2D liquid chromatography-tandem mass chromatography (LC-MS/MS) using a strong cation exchange and reverse phase chromatography, and SDS-PAGE combined with a systematic LC-MS/MS analysis. The enrichment for plasma membrane proteins helped to reduce the sample complexity, and both proteomic approaches involved a pre-fractionation step before MS analysis, another step that reduced further sample complexity. Only proteins consistently retrieved with the two methodologies were taken into account, resulting in the identification of 78 proteins. Of those proteins, 56% were predicted to contain one or more transmembrane domains, while 30% were already known to be localized on the plasma membrane. Very stringent criteria were applied to detect only proteins that were exclusively found in the plasma membrane of mycorrhizal plants. Only two proteins passed this severe threshold: a plasma membrane proton-efflux P-type ATPase (Mtha1) and a blue copper-binding protein (MtBcp1). Considering the highly stringent criteria, Valot et al. (2006) concluded that these two proteins were biologically relevant and deserved further investigations. The importance of these proteins was in fact revealed in subsequent studies. Even though they did not identify the specific function of MtBcp1, Pumplin and Harrison (2009) suggested the presence of at least two distinct domains in the peri-arbuscular membrane (PAM): an 'arbuscule branch domain' that contains the symbiosis-specific phosphate transporter, MtPT4, and an 'arbuscule trunk domain' that contains MtBcp1 [47]. Concerning the other protein specifically induced in arbuscule-containing cells, Wang et al. (2014) showed that H⁺-ATPases are required for enhanced proton pumping activity in membrane vesicles. Functional impairment of this gene led to impairment in the host plant nutrient uptake through the mycorrhizal symbiosis, whereas its overexpression increased both phosphate uptake and plasma membrane potential, suggesting that this H⁺-ATPase plays a key role in energizing the peri-arbuscular membrane, thereby facilitating nutrient exchange in arbusculated plant cells [48].

In the past few years, proteomics has seen great technical advances, especially in the mass spectrometry equipment and bioinformatics resources, with the development of new separa-

tion techniques, new multi-dimensional procedures, new searching algorithms, new mass spectrometers and the availability of more databases. Owing to these new technologies, Abdallah et al. (2014) were able to identify 1,226 root membrane proteins and to report for the first time the proteomic identification of several symbiosis marker genes: MtPt4, a mycorrhiza-specific phosphate transporter [22], the AM-inducible ammonium transporter GmAMT4.1 in soybean [49], STR half-ABC transporters [50] and vesicle-associated membrane proteins VAMP721d/e [51].

In the experiments of Abdallah et al. (2014), proteins were quantified by label-free counting. Protein quantification in label-free experiments is generally based on two types of measurements: peptide peak intensity and spectral count. These parameters are measured for individual LC-MS/MS or LC/LC-MS/MS runs and changes in protein abundance are calculated *via* a direct comparison between different analyses [52]. The spectral counting strategy used by Abdallah et al. (2014) suggests that accommodation of AM fungi within root cortical cells implies both a dynamic reorganization of the root membrane proteome and the *de novo* synthesis of AM-related proteins [23]. This study, beside the identification of proteins corresponding to key genes already identified in mycorrhizal symbiosis, also reported new proteins, many of which support the importance of membrane trafficking during mycorrhiza colonization.

In summary, sub-cellular and peptide fractionations led to the identification of many key proteins involved in AM symbiosis. Despite the recent contributions of proteomics to the study of the plant–fungus mycorrhizal interactions were substantial, the role of many of them as actors in the symbiosis is still to be fully understood.

4.4. Proteomics to identify fungal proteins

Most proteomic studies on the AM interaction have focused on plant proteins. Identification of fungal proteins is more challenging due to the impossibility to grow AM fungi in axenic cultures, to the lower amount of fungal biomass, as compared to plant material, and to the more scanty sequence information. However, pioneering studies have been carried out in France by Dumas-Gaudot and Recorbet [53,54].

Dumas-Gaudot et al. (2004) used the root organ culture method [55] to enrich for proteins expressed in the extra-radical mycelium of the AM species *G. intraradices*. They successfully produced, for the first time, a 2DE reference map for the extra-radical proteome of an AM fungus. After the selection of the most intense protein spots, they tried to identify them by mass spectrometry. Unfortunately, only very few proteins from filamentous fungi were known and present in public databases, and the only available genome sequence, at that time, was from *Neurospora crassa*, phylogenetically very far from Glomeromycota. In spite of that, identification was possible for 8 proteins out of the 14 analyzed, and homologies were found for 4 of them.

Few years later, the same group attempted again the identification of fungal proteins expressed in the AM association [54]. They maintained the same experimental system, but they used the GeLC-MS/MS method and could identify 92 different fungal proteins. GeLC-MS/MS approach

combines a mono-dimensional gel (1D-PAGE) and a nano-scale capillary liquid chromatography-MS/MS. Briefly, after the 1D-PAGE separation, the mono-dimensional gel is cut in several pieces; in-gel digestion results in different protein fractions that are separated and analyzed by LC-MS/MS. Using the MetaCyc database, a collection of more than a thousand metabolic pathways [57], these authors grouped those proteins in 11 pathways that span energy, metabolism and cell rescue processes. These data, together with previous identifications of putative homologues of cell-cycle gene in *G. mosseae* and *G. intraradices* [56,57], suggest that signalling pathways known in model species may also operate in AM fungi. Although the GeLC-MS/MS strategy opened the possibility to large-scale proteomics of mycorrhizal fungi, no further data have been published with this technique.

Secreted fungal proteins play key roles in host plant colonization and symbiosis development in ECM interactions. Vincent et al. (2012) have identified the extracellular proteins secreted in the growth medium by the free-living mycelium of the ECM fungus *L. bicolor* using 2-DE, IPG-IEF shotgun (IPG strip was cut into fractions and tryptic peptides were eluted from the each fraction) and SDS-PAGE shotgun, with the aim to validate predicted secreted proteins and identify putative novel effectors of the symbiosis. Among the 224 proteins identified, there were carbohydrate-active enzymes (CAZymes), probably involved in cell wall remodelling during hyphal growth, as well as secreted proteases. Additionally, the involvement of some of these proteins in the establishment of the mycorrhizal symbiosis was supported by transcriptomic analyses of ECM roots [58].

5. Mycorrhizal fungi and heavy metals

In plants, stress tolerance to soil pollution can be increased by their interaction with mycorrhizal fungi [59]. Six out of ten of the most polluted soils in the world are contaminated by heavy metals [60], and mycorrhizal symbioses have been found to reduce metal toxicity to the host in soils with potentially toxic amounts of soluble and insoluble metals. Phytoremediation, the plant-mediated reclamation of polluted soils, is receiving increasing attention as a natural method to restore the biological features of the soil. Mycorrhizal fungi can have an important part in this process. Many studies have been carried out on the benefits of mycorrhizal plant-fungus interaction in heavy-metals-polluted soils, but only few of them have used a proteomic approach to identify the key proteins potentially involved in mycorrhiza-mediated stress tolerance. Researches on this topic have analyzed different plant organs, like leaves or root, but also focused on the symbiotic fungus.

Bona et al. (2010) studied the leaf proteome of the arsenic hyperaccumulator fern *Pteris vittata* inoculated with two fungi (*Glomus mosseae* and *Gigaspora margarita*), with and without arsenic treatment [59]. The symbiosis with both fungi decreased arsenic concentration compared with non-mycorrhizal plants, indicating the protective effect of mycorrhizal fungi. Interestingly, the plant protein expression profile was different when the plant was inoculated with *G. mosseae* or *G. margarita*. Although they studied a different biological system, Canga-huala-Inocente and colleagues (2011) identified instead a core of 25 proteins, supporting the

existence of conserved plant responses to *Glomus irregulare* and *Glomus mosseae*, at least in a woody perennial species such as grapevine [61].

A study on root proteomics has been conducted by Aloui et al. [62]. They reported, using a 2DE approach followed by MS/MS, the protective effect conferred by *G. intraradices* to the model legume *M. truncatula* in the presence of Cd. They identified 36 mycorrhiza-related proteins, but only 6 displayed changes in abundance upon Cd exposure. These proteins – a cyclophilin, a guanine nucleotide-binding protein, an ubiquitin carboxyl-terminal hydrolase, a thiazole biosynthetic enzyme, an annexin, a glutathione S-transferase (GST)-like protein and a S-adenosylmethionine (SAM) synthase – seem to have a function in oxidative stress alleviation [62]. The authors also suggested that antioxidant enzymes and non-enzymatic antioxidants could be probably involved both in arbuscule senescence [63] and in plant protection against oxidative damage caused by Cd.

In addition to plant proteomics, mycorrhizal fungi have been also investigated for changes in their protein profiles when exposed to heavy metals. Adaptive metal tolerance has been reported for mycorrhizal fungi isolated from polluted soils [64], although the underlying cellular and molecular mechanisms have been seldom identified [57].

Chiapello et al. (2015) used gel-based and gel-free techniques as a complementary approach to study the proteome of *Oidiodendron maius* Zn, an ericoid mycorrhizal fungus isolated from a polluted soil [65] and showing adaptive tolerance to zinc and cadmium [66]. *O. maius* Zn can establish endomycorrhizal symbiosis with the roots of ericaceous plants also in heavily contaminated soils [67]. The aim of the study was to understand the response of this metal-tolerant fungus to Cd and Zn ions and to reveal common and/or specific cellular and molecular mechanisms to counteract heavy metal stress caused by these metals. The authors concluded that Cd and Zn induce common as well as specific responses. Among the common induced proteins, agmatinase, an enzyme involved in polyamines biosynthesis, represents a novel finding in relation to heavy metal responses in fungi.

6. Conclusion and future perspective

Proteomics has allowed us to identify proteins expressed and regulated during the development and functioning of mycorrhizal symbioses, therefore contributing to a better understanding of the events occurring at the cellular level.

Protein identification is strongly dependent on gene and protein sequences available in databases, and the constant increase in the number of sequenced genomes in the past decade, together with improvement of mass spectrometry technology, has helped scientists to obtain more reliable data. In the past few years, a specialized fungal genomics portal, called MycoCosm (<http://genome.jgi.doe.gov/fungi>), has been created by the US Department of Energy (DOE) Joint Genome Institute (JGI), offering an access point to the data from all the sequencing genome project managed by the DOE JGI [68,69]. Starting from the three first genome projects on *Laccaria bicolor*, *Tuber melanosporum* and *Rhizophagus intraradices* [70,71,72], several more

genomes from symbiotic fungi, including ECM, AM, orchid and ericoid fungi, have been recently sequenced, with the aim to determine the diversity of the molecular processes involved in the interaction [28]. Despite the more powerful techniques and wider reference datasets for protein identification, current limitations exist in the application of proteomics to the study of plant–microbe interactions. In particular, new extraction methods, microsomal studies, sub-cellular enrichment, gel-free separation methods, pre-fractioning separations and new mass spectrometry are still far from being fully explored [73].

In order to identify proteins from a very small subset of target cells, Gaude et al. (2012) combined laser capture microdissection (LCM) and LC-MS/MS [74]. Laser microdissection permits the rapid isolation, from sections of a heterogeneous tissue, of a selected cell population in a manner compatible with the extraction of DNA, RNA or proteins [29]. Using LCM, arbuscule-containing cortical cells and cortical cells from non-mycorrhizal *M. truncatula* roots were isolated. Proteomic analyses on these cells revealed a number of proteins involved in lipid metabolism, most likely related to the synthesis of the PAM. This targeted analysis on a specific subset of colonized cells, those harbouring the arbuscules, curiously did not identify known PAM marker proteins, thus suggesting that either sample preparation or instrument capability were not sensitive enough. Although this first use of LCM in the proteomic investigation of the AM mycorrhizal symbiosis did not identify known marker proteins, it highlighted the PAM as an important carbon sink. The LCM technique coupled with MS/MS techniques could be a powerful combination to investigate the protein profiles of specific cells at specific time-points. Moreover, LCM can help to overcome the problem of asynchronous fungal development and arbuscule maturation in mycorrhizal roots. To be able to combine LCM samples of synchronous arbusculated cells with sub-cellular enrichment, peptide pre-fractionation and analysis with powerful MS instruments such as Orbitrap Velos (Thermo Fischer company) may reveal an unexpected specificity during the development of this symbiotic structure.

Novel MS/MS techniques developed in the past few years in other research fields could also be applied to investigate plant–fungus symbiotic interactions. For example, selected reaction monitoring (SRM) is a targeted MS technique used to complement untargeted shotgun methods. SRM is used to measure across multiple samples – in a consistent, reproducible and quantitatively precise manner – a set of candidate proteins involved in a particular cellular process [75]. Based on known data from the literature or previous experiments, a set of target peptides that optimally represent the protein are selected and after their validation they are used for protein quantification. Unfortunately, the sensitivity of SRM is limited and it cannot cover the entire proteome of an organism. Nevertheless, this technique is really promising for the fine protein quantification in different cell types or conditions. Taylor et al. (2014) used SRM in plant science to confirm protein abundance in *Arabidopsis* mutant lines, even when discrimination between very similar proteins was needed [76]. However, the application of this technique for the identification of OsPT11, homologue to MtPT4, from wild-type and mutant lines did not work, probably due to the method's sensitivity (Chiapello, 2013, unpublished data). Another promising technique to further investigate the proteome of arbuscule-containing cells is the single-cell imaging mass spectrometry (IMS), a powerful technique used to map the distribution of endogenous biomolecules with subcellular resolution [77].

	Year	Plant	Fungus	Separation technique	Identification technique	Reference
ECTO	1995	<i>Eucalyptus grandis</i>	<i>Pisolithus tinctorius</i>	2DE-PAGE	---	Burgess et al., 1995
	1999	<i>Eucalyptus globulus</i> subsp. <i>bicostata</i>	<i>Pisolithus tinctorius</i>	2DE-PAGE	---	Laurent et al., 1999
	2007	---	<i>Boletus edulis</i>	2DE-PAGE	ESI-Q-TOF MS	Liang et al., 2007
	2012	---	<i>Laccaria bicolor</i>	IPG-IEF / 1-DE and 2-DE -PAGE	LCQ IT	Vincent et al., 2012
AM	2000	<i>Lycopersicon esculentum</i>	<i>Glomus mosseae</i>	2DE-PAGE	N-terminal sequencing	Benabdellah et al., 2000
	2001	<i>Medicago truncatula</i>	<i>Glomus mosseae</i>	2DE-PAGE	MALDI-TOF	Bestel-Corre et al., 2001
	2004	---	<i>Glomus intraradices</i>	2DE-PAGE	Q-ToF2	Dumas-Gaudot et al., 2004
	2005	<i>Medicago truncatula</i>	<i>Glomus intraradices</i>	2DE-PAGE	MALDI-TOF	Valot et al., 2005
	2006	<i>Medicago truncatula</i>	<i>Glomus intraradices</i>	2DE-PAGE	MALDI-TOF	Amiour et al., 2006
	2006	<i>Medicago truncatula</i>	<i>Glomus intraradices</i>	1-DE and 2-DE -PAGE / 2D-LC	LCQ Deca XP+	Valot et al., 2006
	2009	---	<i>Glomus intraradices</i>	1-DE-PAGE	ESI-Q-TOF MS	Recorbet et al., 2009
	2009	<i>Medicago truncatula</i>	<i>Glomus intraradices</i>	2DE-PAGE	LCQ Deca XP+	Aloui et al., 2009
	2010	<i>Pteris vittata</i>	<i>Glomus mosseae</i> / <i>Gigaspora margarita</i>	2DE-PAGE	QSTAR XL hybrid quadrupole-TOF	Bona et al., 2010
	2011	<i>Vitis vinifera</i>	<i>Glomus mosseae</i> / <i>Glomus intraradices</i>	2DE-PAGE	MALDI-TOF	Cangahuala-Inocente et al., 2011
2012	<i>Populus alba</i>	<i>Glomus intraradices</i>	2DE-PAGE	ESI-Q-TOF MS	Lingua et al., 2012	
2014	<i>Medicago truncatula</i>	<i>Rhizophagus irregularis</i>	1-DE-PAGE	LTQ XL ion trap	Abdallah et al., 2014	
Others	2015	---	<i>Oidiodendron maius</i>	2DE-PAGE / 2D-LC	MALDI-TOF / QSTAR MS/MS	Chiapello et al., 2015
	2014	<i>Oncidium sphacelatum</i>	<i>Ceratobasidium</i> sp. Isolate	2D-LC	ESI-Q-TOF MS	Valadares et al., 2014

Table 1. List of papers in which proteomics has been applied to study mycorrhizal symbiosis. For simplicity, AM fungi have been indicated with the names used in the original articles, despite the relatively recent taxonomic revision (Redecker D1, Schüssler A, Stockinger H, Stürmer SL, Morton JB, Walker C. 2013. An evidence-based consensus for the classification of arbuscular mycorrhizal fungi (Glomeromycota). *Mycorrhiza* 23:515-31).

The ability to analyze a single-cell proteome is exciting, but also extremely challenging. The first difficulty is the sensitivity, both correlated with the sample itself and with the mass

spectrometry detection capability. Every single cell can contain proteins in a range of few to million copies per cell. However, mass spectrometers are now really powerful, and even with an attomole detection limit, only the most abundant proteins are detectable [78]. The estimated number and concentration of proteins in a single mammalian cell is 33 attomole/cell for the most abundant and 830 yotomole/cell for the less abundant [77]. The second challenge is the inherent limitation associated with the imaging modality itself. Even if further development is needed to obtain the combined resolution and sensitivity required, IMS stands up as a very promising technique to analyze specific cell types or conditions. By employing IMS, Ye et al. (2013) detected a large array of organic acids, amino acids, sugars, lipids, flavonoids in roots and root nodules of *M. truncatula* during nitrogen fixation [79]. They demonstrated that IMS can obtain unique information on the identity and spatial distribution of plant metabolites, although high-resolution MALDI-MS is required to fully resolve the metabolic differences in nodule chemistry.

In conclusion, similarly to other 'omics' approaches, proteomics has also made rapid progress in the recent year, thus making this approach a very useful one to complement information on gene expression in mycorrhizal tissues. At this speed of technological developments, methods that allow us to easily assign proteins up-regulated during symbiosis to specific cell types and sub-cellular compartments may not be too far ahead. These proteomic techniques will be powerful tools to unravel the molecular component involved in plant–mycorrhizal fungal interactions.

Author details

Marco Chiapello¹, Silvia Perotto^{1,2} and Raffaella Balestrini^{2*}

*Address all correspondence to: raffaella.balestrini@ipspp.cnr.it

¹ Department of Life Sciences and Systems Biology, University of Torino, Torino, Italy

² Institute for Sustainable Plant Protection – CNR, UOS Torino, Torino, Italy

References

- [1] Smith SE, Read DJ. Mineral nutrition, toxic element accumulation and water relations of arbuscular mycorrhizal plants. *Mycorrhizal Symbiosis*. 3rd edn. Academic Press, London, 2008. pp.145–148.
- [2] Markmann K, Parniske M. Evolution of root endosymbiosis with bacteria: how novel are nodules? *Trends Plant Sci* 2009;14:77–86.

- [3] Girlanda M, Perotto S, Bonfante P. Mycorrhizal fungi: their habitats and nutritional strategies. In: Christian P, Kubicek I, Druzhinina S. (ed.) *Environmental and Microbial Relationships (The Mycota)*. 2nd ed. Springer; 2015. pp. 229–256.
- [4] Tedersoo L, May TW, Smith ME. Ectomycorrhizal lifestyle in fungi: global diversity, distribution, and evolution of phylogenetic lineages. *Mycorrhiza* 2009;20:217–63.
- [5] LePage BA. The evolution, biogeography, and palaeoecology of the Pinaceae based on fossil and extant representative. *Acta Horticult* 2003;615:29–52.
- [6] Bruns TD, Shefferson RP. Evolutionary studies of ectomycorrhizal fungi: recent advances and future directions. *Can J Bot* 2004;82:1122–32.
- [7] E Smith S, J Read D. *Mycorrhizal Symbiosis*. 2nd ed. San Diego, California: Academic Press, 1997. p. 605.
- [8] Liang Y, Chen H, Tang M, Shen S. Proteome analysis of an ectomycorrhizal fungus *Boletus edulis* under salt shock. *Mycol Res* 2007;111:939–46.
- [9] Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson OE, et al. A higher-level phylogenetic classification of the Fungi. *Mycol Res* 2007;111:509–47.
- [10] Parniske M. Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nature* 2008;6:763–75.
- [11] Bago B, Pfeffer PE, Shachar-Hill Y. Carbon metabolism and transport in arbuscular mycorrhizas. *Plant Physiol* 2000;124:949–58.
- [12] Yang S-Y, Grønlund M, Jakobsen I, Grottemeyer MS, Rentsch D, Miyao A, et al. Non-redundant regulation of rice arbuscular mycorrhizal symbiosis by two members of the phosphate transporter1 gene family. *Plant Cell* 2012;24:4236–51.
- [13] Smith SE, Smith FA. Roles of arbuscular mycorrhizas in plant nutrition and growth: new paradigms from cellular to ecosystem scales. *Ann Rev Plant Biol* 2011;62:227–50.
- [14] Gutjahr C, Parniske M. Cell and developmental biology of arbuscular mycorrhiza symbiosis. *Annu Rev Cell Dev Biol* 2013;29:593–617.
- [15] Krings M, Taylor TN, Hass H, Kerp H, Dotzler N, Hermsen EJ. Fungal endophytes in a 400-million-yr-old land plant: infection pathways, spatial distribution, and host responses. *New Phytol* 2006;174:648–57.
- [16] Taylor TN, Remy W, Hass H, Kerp H. Fossil arbuscular mycorrhizae from the early Devonian. *Mycologia* 1995;87:560.
- [17] Bonfante P, Genre A. Arbuscular mycorrhizal dialogues: do you speak ‘plantish’ or ‘fungish’? *Trends Plant Sci* 2015;20:150-4.
- [18] Mailliet F, Poinso V, André O, Puech-Pagès V, Haouy A, Gueunier M, et al. Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. *Nature* 2011;469:58–63.

- [19] Genre A, Chabaud M, Balzergue C, Puech-Pagès V, Novero M, Rey T, et al. Short-chain chitin oligomers from arbuscular mycorrhizal fungi trigger nuclear Ca²⁺ spiking in *Medicago truncatula* roots and their production is enhanced by strigolactone. *New Phytol* 2013;198:179–89.
- [20] Genre A, Chabaud M, Timmers T, Bonfante P, Barker DG. Arbuscular mycorrhizal fungi elicit a novel intracellular apparatus in *Medicago truncatula* root epidermal cells before infection. *Plant Cell* 2005;17:3489–99.
- [21] Denison RF, Kiers ET. Life histories of symbiotic rhizobia and mycorrhizal fungi. *Curr Biol* 2011;21:R775–85.
- [22] Harrison MJM, Dewbre GR, Liu J. A phosphate transporter from *Medicago truncatula* involved in the acquisition of phosphate released by arbuscular mycorrhizal fungi. *Plant Cell* 2002;14:2413–29.
- [23] Abdallah C, Valot B, Guillier C, Mounier A, Balliau T, Zivy M, et al. The membrane proteome of *Medicago truncatula* roots displays qualitative and quantitative changes in response to arbuscular mycorrhizal symbiosis. *J Proteomics* 2014;108:354–68.
- [24] Raudaskoski M, Kothe E. Novel findings on the role of signal exchange in arbuscular and ectomycorrhizal symbioses. *Mycorrhiza* 2014;25:1–10.
- [25] Splivallo R, Fischer U, Gobel C, Feussner I, Karlovsky P. Truffles regulate plant root morphogenesis via the production of auxin and ethylene. *Plant Physiol* 2009;150:2018–29.
- [26] Plett JM, Martin F. Blurred boundaries: lifestyle lessons from ectomycorrhizal fungal genomes. *Trends Genet* 2011;27:14–22.
- [27] Plett JM, Daguette Y, Wittulsky S, Vayssières A, Deveau A, Melton SJ, et al. Effector MiSSP7 of the mutualistic fungus *Laccaria bicolor* stabilizes the *Populus* JAZ6 protein and represses jasmonic acid (JA) responsive genes. *Proc Natl Acad Sci USA* 2014;111:8299–304.
- [28] Kohler A, Kuo A, Nagy LG, Morin E, Barry KW, Buscot F, et al. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nat Genet* 2015;47:410–5.
- [29] Balestrini R, Gómez-Ariza J, Klink VP, Bonfante P. Application of Laser Microdissection to plant pathogenic and symbiotic interactions. *J Plant Interact* 2009;4:81–92.
- [30] Recorbet G, Dumas-Gaudot E. Proteomics of biotrophic plant–microbe interactions: symbioses lead the march. Hoboken, NJ, USA: John Wiley & Sons, Inc; 2008. p. 764.
- [31] Burgess T, Laurent P, Dell B, Malajczuk N, Martin F. Effect of fungal-isolate aggressivity on the biosynthesis of symbiosis-related polypeptides in differentiating euca-lypt ectomycorrhizas. *Planta* 1995;195:408–17.

- [32] O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 1975;250:4007–21.
- [33] Felten J, Martin F, Legué V. Signalling in Ectomycorrhizal Symbiosis. In: Perotto S, Baluška F. (eds.) *Signaling and Communication in Plant Symbiosis*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. pp. 123–42.
- [34] Ditengou FA, Müller A, Rosenkranz M, Felten J, Lasok H, van Doorn MM, et al. Volatile signalling by sesquiterpenes from ectomycorrhizal fungi reprogrammes root architecture. *Nature Comm* 2014;6:6279–9.
- [35] Laurent P, Voiblet C, Tagu D, de Carvalho D, Nehls U, De Bellis R, et al. A novel class of ectomycorrhiza-regulated cell wall polypeptides in *Pisolithus tinctorius*. *Mol Plant Microbe Interact* 1999;12:862–71.
- [36] Bonfante P, Genre A. Mechanisms underlying beneficial plant–fungus interactions in mycorrhizal symbiosis. *Nature Comm* 2010;1:1–11.
- [37] Amiour N, Recorbet G, Robert F, Gianinazzi S, Dumas-Gaudot E. Mutations in DMI3 and SUNN modify the appressorium-responsive root proteome in arbuscular mycorrhiza. *Mol Plant Microbe Interact* 2006;19:988–97.
- [38] Lopez-Meyer M, Harrison MJ. An experimental system to synchronize the early events of development of the arbuscular mycorrhizal symbiosis. In: Sánchez F, Quinto C, López-Lara IM, Geiger O. (eds.) *Molecular Plant–Microbe Interactions*. Saint Paul: International Society for Plant-Microbe Interactions; 2006. pp. 546–551.
- [39] Bestel-Corre G, Dumas-Gaudot E, Poinsoit V, Dieu M, Dierick J-F, van Tuinen D, et al. Proteome analysis and identification of symbiosis-related proteins from *Medicago truncatula* Gaertn. by two-dimensional electrophoresis and mass spectrometry. *Electrophoresis* 2001;23:122–37.
- [40] Valadares RBS, Perotto S, Santos EC, Lambais MR. Proteome changes in *Oncidium sphacelatum* (Orchidaceae) at different trophic stages of symbiotic germination. *Mycorrhiza* 2014;24:349–60.
- [41] Tan S, Tan HT, Chung MCM. Membrane proteins and membrane proteomics. *PROTEOMICS* 2008;8:3924–32.
- [42] Benabdellah K, Azcón-Aguilar C, Ferrol N. Soluble and membrane symbiosis-related polypeptides associated with the development of arbuscular mycorrhizas in tomato (*Lycopersicon esculentum*). *New Phytol* 1998;140:135–43.
- [43] Benabdellah K, Azcon-Aguilar C, Ferrol N. Alterations in the plasma membrane polypeptide pattern of tomato roots (*Lycopersicon esculentum*) during the development of arbuscular mycorrhiza. *J Exp Bot* 2000;51:747–54.

- [44] Valot B, Dieu M, Recorbet G, Raes M, Gianinazzi S, Dumas-Gaudot E. Identification of membrane-associated proteins regulated by the arbuscular mycorrhizal symbiosis. *Plant Mol Biol* 2005;59:565–80.
- [45] Maldonado-Mendoza IE, Dewbre GR, van Buuren ML, Versaw WK, Harrison MJM. Methods to estimate the proportion of plant and fungal RNA in an arbuscular mycorrhiza. *Mycorrhiza* 2002;12:67–74.
- [46] Valot B, Negroni L, Zivy M, Gianinazzi S, Dumas-Gaudot E. A mass spectrometric approach to identify arbuscular mycorrhiza-related proteins in root plasma membrane fractions. *PROTEOMICS* 2006;6(Suppl 1):S145–55.
- [47] Pumpilin N, Harrison MJM. Live-cell imaging reveals periarbuscular membrane domains and organelle location in *Medicago truncatula* roots during arbuscular mycorrhizal symbiosis. *Plant Physiol* 2009;151:809–19.
- [48] Wang E, Yu N, Bano SA, Liu C, Miller AJ, Cousins D, et al. A H⁺-ATPase that energizes nutrient uptake during mycorrhizal symbioses in rice and *Medicago truncatula*. *Plant Cell* 2014;26:1818–30.
- [49] Kobae Y, Hata S. Dynamics of periarbuscular membranes visualized with a fluorescent phosphate transporter in arbuscular mycorrhizal roots of rice. *Plant Cell Physiol* 2010;51:341–53.
- [50] Zhang Q, Blaylock LA, Harrison MJM. Two *Medicago truncatula* half-ABC transporters are essential for arbuscule development in arbuscular mycorrhizal symbiosis. *Plant Cell* 2010;22:1483–97.
- [51] Ivanov S, Fedorova EE, Limpens E, De Mita S, Genre A, Bonfante P, et al. *Rhizobium-legume* symbiosis shares an exocytotic pathway required for arbuscule formation. *Proc Natl Acad Sci USA* 2012;109:8316–21.
- [52] Zhu W, Smith JW, Huang C-M. Mass spectrometry-based label-free quantitative proteomics. *J Biomed Biotechnol* 2010;2010:1–6.
- [53] Dumas-Gaudot E, Valot BT, Bestel-Corre GNL, Recorbet G, St-Arnaud M, Fontaine B, et al. Proteomics as a way to identify extra-radicular fungal proteins from *Glomus intraradices*-RiT-DNA carrot root mycorrhizas. *Fems Microbiol Ecol* 2004;48:401–11.
- [54] Recorbet G, Rogniaux H, Gianinazzi-Pearson V, Dumasgaudot E. Fungal proteins in the extra-radical phase of arbuscular mycorrhiza: a shotgun proteomic picture. *New Phytol* 2009;181:248–60.
- [55] Stougaard J, Abildsten D, Marcker KA. The *Agrobacterium rhizogenes* pRi TL-DNA segment as a gene vector system for transformation of plants. *Mol Gen Genet* 1987;207:251–5.

- [56] Requena N, Mann P, Franken P. A homologue of the cell cycle check point TOR2 from *Saccharomyces cerevisiae* exists in the arbuscular mycorrhizal fungus *Glomus mosseae*. *Protoplasma* 2000;212:89–98.
- [57] Jun J, Abubaker J, Rehrer C, Pfeffer PE, Shachar-Hill Y, Lammers PJ. Expression in an arbuscular mycorrhizal fungus of genes putatively involved in metabolism, transport, the cytoskeleton and the cell cycle. In: Smith SE, Smith FA. (eds.) *Diversity and Integration in Mycorrhizas*. Dordrecht: Springer Netherlands; 2002. pp. 141–148.
- [58] Vincent D, Kohler A, Claverol S, Solier E, Joets J, Gibon J, et al. Secretome of the free-living mycelium from the ectomycorrhizal basidiomycete *Laccaria bicolor*. *J Proteome Res* 2012;11:157–71.
- [59] Bona E, Cattaneo C, Cesaro P, Marsano F, Lingua G, Cavaletto M, et al. Proteomic analysis of *Pteris vittata* fronds: two arbuscular mycorrhizal fungi differentially modulate protein expression under arsenic contamination. *PROTEOMICS* 2010;10:3811–34.
- [60] Lingua G, Bona E, Todeschini V, Cattaneo C, Marsano F, Berta G, et al. Effects of heavy metals and arbuscular mycorrhiza on the leaf proteome of a selected poplar clone: a time course analysis. *PLoS ONE* 2012;7:e38662.
- [61] Cangahuala-Inocente GC, Da Silva MF, Johnson J-M, Manga A, van TD, Henry C, et al. Arbuscular mycorrhizal symbiosis elicits proteome responses opposite of P-starvation in SO₄ grapevine rootstock upon root colonisation with two *Glomus* species. *Mycorrhiza* 2011;21:473–93.
- [62] Aloui A, Recorbet G, Gollotte A, Robert F, Valot B, Gianinazzi-Pearson V, et al. On the mechanisms of cadmium stress alleviation in *Medicago truncatula* by arbuscular mycorrhizal symbiosis: a root proteomic study. *PROTEOMICS* 2009;9:420–33.
- [63] Fester T, Hause G. Accumulation of reactive oxygen species in arbuscular mycorrhizal roots. *Mycorrhiza* 2005;15:373–9.
- [64] Colpaert JV, Muller LA, Lambaerts M, Adriaensen K, Vangronsveld J. Evolutionary adaptation to Zn toxicity in populations of Suilloid fungi. *New Phytol* 2004;162:549–59.
- [65] Turnau K. The influence of industrial dust on the mycoflora of Pino-Quercetum forest near Cracow. III. The influence of industrial dust on the mycorrhiza of *Vaccinium myrtillus*. *Sci Pap Krakow Agr Acad* 1988;226:135–45.
- [66] Vallino M, Zampieri E, Murat C, Girlanda M, Picarella S, Pitet M, et al. Specific regions in the Sod1 locus of the ericoid mycorrhizal fungus *Oidiodendron maius* from metal-enriched soils show a different sequence polymorphism. *Fems Microbiol Ecol* 2011;75:321–31.
- [67] Perotto S, Actis-Perino E, Perugini, Bonfante P. Molecular diversity of fungi from ericoid mycorrhizal roots. *Mol Ecol* 1996;5:123–31.

- [68] Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, Goodstein D, et al. The genome portal of the Department of Energy Joint Genome Institute. *Nucleic Acids Res* 2012;40:D26–32.
- [69] Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otilar R, et al. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res* 2014;42:D699–704.
- [70] Martin F, Aerts A, Ahrén D, Brun A, Danchin EGJ, Duchaussoy F, et al. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* 2008;452:88–92.
- [71] Martin F, Kohler A, Murat C, Balestrini R, Coutinho PM, Jaillon O, et al. Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* 2010;464:1033–8.
- [72] Tisserant E, Malbreil M, Kuo A, Kohler A, Symeonidi A, Balestrini R, et al. Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proc Natl Acad Sci USA* 2013;110:20117–22.
- [73] Couto MSR, Lovato PE, Wipf D, Dumas-Gaudot E. Proteomic studies of arbuscular mycorrhizal associations. *Adv Biol Chem* 2013;2013:48–58.
- [74] Gaude N, Bortfeld S, Duensing N, Lohse M, Krajinski F. Arbuscule-containing and non-colonized cortical cells of mycorrhizal roots undergo extensive and specific re-programming during arbuscular mycorrhizal development. *Plant J* 2012;69:510–28.
- [75] Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat Methods* 2012;9:555–66.
- [76] Taylor NL, Fenske R, Castleden I, Tomaz T, Nelson CJ, Millar AH. Selected reaction monitoring to determine protein abundance in *Arabidopsis* using the *Arabidopsis* proteotypic predictor. *Plant Physiol* 2014;164:525–36.
- [77] Passarelli MK, Ewing AG. Single-cell imaging mass spectrometry. *Curr Opin Chem Biol* 2013;17:854–9.
- [78] McDonnell LA, Corthals GL, Willems SM, van Remoortere A, van Zeijl RJM, Deelder AM. Peptide and protein imaging mass spectrometry in cancer research. *J Proteomics* 2010;73:1921–44.
- [79] Ye H, Gemperline E, Venkateshwaran M, Chen R, Delaux P-M, Howes-Podoll M, et al. MALDI mass spectrometry-assisted molecular imaging of metabolites during nitrogen fixation in the *Medicago truncatula*–*Sinorhizobium meliloti* symbiosis. *Plant J* 2013;75:130–45.

Targeted Proteomics in Translational and Clinical Studies

Eslam Nouri-Nigjeh, Ru Chen and Sheng Pan

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/61710>

Abstract

This chapter provides a concise overview on the methods and applications of targeted proteomics in the context of translational and clinical studies. Mass spectrometry-based targeted proteomics has emerged as a promising technique for protein and peptide quantification, presenting a great potential for clinical applications. While significant amount of discovery works have been carried out in both genomics and proteomics for an assortment of diseases, it has been challenging in further characterizing individual protein targets for their biological significance and clinical value due to the lack of effective and “universal” techniques. The development of targeted proteomics approach opened a unique avenue to bridge the discovery-based genomics and proteomics with candidate-based protein analysis, which is clinically highly relevant. Targeted proteomics analysis has been implemented on a variety of instrument platforms, and applied for a wide range of studies, from blood biomarker detection to pathway-driven mechanistic investigations, with the triple quadrupole-based selected reaction monitoring (SRM) technique being the most widely used method. With a right combination of calibration approach, internal standards, and sample preparation strategies, mass spectrometry-based targeted analysis has proven to be of inter-laboratory reproducibility and sensitivity in analyzing many clinical specimens. More recently, the advent of mass spectrometry with high frequencies and resolutions yielded the data independent acquisition (DIA) techniques, such as sequential window acquisition of all theoretical fragment ion spectra (SWATH). The unbiased nature of DIA methods would enable a wider analytical scope and a greater robustness in targeted analysis, representing a paradigm shift in targeted proteomics.

Keywords: Proteomics, Targeted proteomics, Mass spectrometry, Data independent acquisition

1. Introduction

The introduction of soft ionization techniques in mass spectrometry has ushered in a fascinating era in the analysis of large biomolecules, including metabolites and proteolytic peptides

and proteins from complex biological matrixes [1,2]. Mass spectrometry-based proteomics is nowadays utilized in a wide arena of translational and clinical applications for global profiling of biological matrixes to explore disease mechanisms and to discover novel biomarkers [3,4]. Quantitative mass spectrometry confers the highly sensitive and reproducible targeted proteomics for the multiplexed quantification of already existing targeted proteins and putative biomarkers [5,6]. These target proteins can be either a single putative protein biomarker or a set of proteins involved in a specific cell signaling or a metabolic pathway.

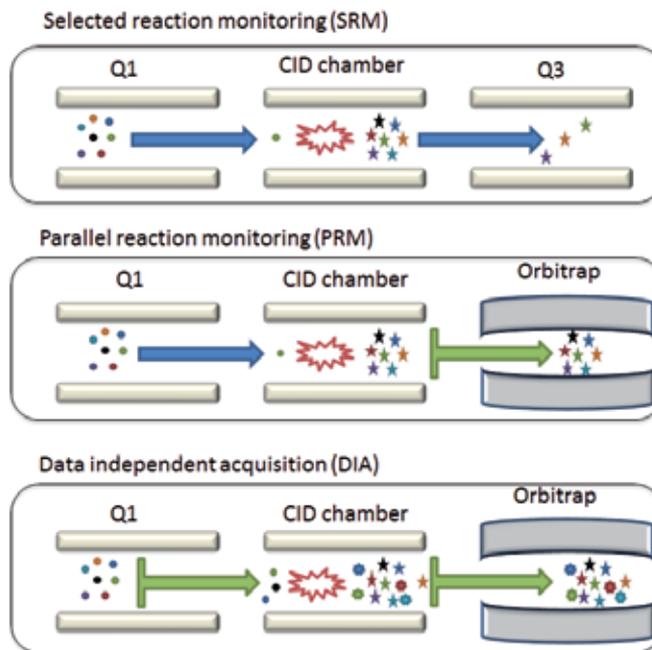
While conventional antibody-based assays, such as ELISA, offer several benefits for the protein quantification, including ease-of-use and simpler instrumentation, ELISA, however, suffers from cross-reactivities and protein/protein interactions that would alter the quantification results based on the level of carrier proteins and based on the free and conjugated levels of the target molecules [7]. Complementarily, mass spectrometry-based targeted proteomics provides a different mechanism for multiplexed protein quantification, and has indisputable advantages in the analysis of genetic changes, polymorphisms, alternative splicing, protein isoforms, and post-translational modifications [6]. In these circumstances, having an antibody with high resolution and specificity for each of these diversities, even if not impossible, would be very difficult to attain. Hence, mass spectrometry-based targeted proteomics can be of complementary importance for the antibody-based quantifications, in particular, for the instances of validating novel protein biomarkers when the corresponding antibodies are not available, or for the cases of multiplexed interrogation of hypothesis-driven key proteins [6,8].

Unlike the quantification of small metabolites, proteins are complex macromolecules, constituting large masses with multiple charges and various dynamic conformations, preventing them to be effectively separated in the gas phase within a mass analyzer or being detected with high mass accuracies. A general theme in targeted quantification, which is widely known as bottom-up proteomics, is to digest proteins by a proteinase enzyme, usually trypsin, with a high specificity to cleave the basic amino acid residues namely arginine and lysine to generate smaller tryptic peptides for facile separations and comprehensive mass spectrometric analysis [2].

A set of different targeted approaches have been already applied for the targeted quantification of proteins [9]. These approaches are based on the survey of precursor ions, survey of product ions, neutral loss, or a fragmentation pattern using a variety of instruments [9]. The earlier targeted proteomic approach, which was called selected ion monitoring (SIM), was based on the generation of an inclusion list and extraction of the exact ion masses of the targeted molecules for analysis [9]. Though this technique was simple to operate, it suffered from low selectivity as many different ions could have similar masses with low level of sensitivities. In contrast, selected reaction monitoring (SRM) built on a triple quadrupole mass spectrometer generates higher sensitivities and specificities, especially when combined with stable isotope dilution. The SRM technique was built based on the unique fragmentation pattern for each targeted molecule that can be mutually specific and provides high sensitivity.

Triple quadrupole mass spectrometer-based SRM technique has been the most widely used targeted proteomics approach to date, in which targeted analytes are selected in the first quadrupole, fragmented in the collision chamber via collision-induced dissociation (CID), and the produced transitions are further separated by the second quadrupole for detection (See

Scheme 1). In such a setting, a combination of selected transitions generated from the corresponding peptide under optimized collision conditions can provide unique identification and accuracy for targeted peptide measurement. The inclusion of multiple product ions makes the SRM analysis more specific in ion selection compared to the inclusion list-based interrogation, and minimizes the interference from a complex background of biological sample via selection of small mass intervals that leads to higher sensitivities. For the optimal use of this technique, there is an immense need to identify the suitable signature peptides that would be highly stable under prolonged digestion and storage and to be highly sensitive through the gas-phase transitions [10].



Scheme 1. Three different approaches in mass spectrometry-based targeted proteomics (a) selected reaction monitoring (SRM) where selected fragment ions from a single precursor are measured for the quantification, (b) parallel reaction monitoring (PRM) where a single precursor ion and entire fragment ions are selected, and (c) data independent acquisition (DIA) where multiple precursors ions are fragmented simultaneously and the entire fragment ions are monitored. The presented PRM and DIA technologies are based on Orbitrap mass spectrometers, such as Q-Exactive.

The concept of stable isotope dilution, which was originally developed for the quantification of metabolites, have been implemented for targeted proteomics analysis using stable isotope-labeled synthetic peptides as internal standards to facilitate mass spectrometric quantifications. Though using different instruments, with different elution, ionization, and collision conditions would impact the intensities of gas-phase transitions, the use of heavy isotope-labeled internal standards can circumvent the variations generated due to the differences in instruments and settings, and provide more robust quantification [11]. Stable isotope-labeled peptides can be used for the absolute quantification of targeted peptides and their post-

translational modifications through the synthesis of different combinations that can contain post-translational modifications [12]. An inter-laboratory study have pointed out that the SRM-based targeted proteomics using common stable isotope-labeled peptide internal standards and calibration approaches can be of high reproducibility and reliability [4].

LC-MALDI-TOF/TOF analysis is a different platform for biomarkers discovery and detection with its own unique characteristics [13,14]. In LC-MALDI-TOF/TOF setting, a peptide separation module is used to generate an array of peptides from complex mixtures in the presence of stable isotope-labeled internal standards on the sample plate; targeted proteomics is then carried out by specific interrogation of selected candidates using MALDI-TOF/TOF mass spectrometer. Such an approach involves detached MS and MS/MS acquisition, allowing repeat interrogation of a wide range of peptide targets with minimal assay development.

Since the fragmentation patterns in SRM-based targeted proteomics can be dependent on the vendor types, parallel reaction monitoring (PRM) is devised to improve the identification and quantification of the targeted peptides. In this technique, as shown in Scheme 1, all the detectable fragmentation ions from the pre-selected substrate are recorded and used for the quantification [15]. The second quadrupole is replaced with a high-resolution and high-frequency Orbitrap mass spectrometer. More recently, the advent of fast and high-resolution mass spectrometers have made a hybrid discovery and targeted proteomics possible through data independent acquisition (DIA) [16], in which, multiple precursor ions are surveyed together (See Scheme 1), rendering a new strategy for targeted mass spectrometric analysis. Using-high resolution mass spectrometer ensures efficient resolving of complex matrices, and higher frequencies enrich quantification profiles.

2. Targeted proteomics in translational and clinical investigations

Genomic and proteomic studies have already introduced a large number of putative protein biomarker candidates for an assortment of diseases [17–22]. This assortment signifies the need for a universal high-throughput targeted proteomics in order to link the putative protein biomarkers with clinical trials and to perform their verification and validation in the large-scale cohort studies [23–25]. The application of the targeted quantitative proteomics in clinical analysis covers extensive objects ranging from the quantification of proteins, multiplexed monitoring of key proteins in a pathway, targeted analysis of post-translational modifications, and examination of the expression of genetic changes.

2.1. Targeted quantification of protein level

Targeted proteomics is widely used for protein quantification. The putative protein biomarkers are designated to be quantified in a large cohort of clinical samples. This form of targeted quantification can bridge the discovery-based proteomics with the pathways analysis through high-throughput quantification of the predefined protein biomarkers. The targeted quantifications of putative protein biomarkers are based on the quantification of signature peptides after exhaustive extraction of the proteins from complex clinical matrixes. For the complex

clinical biofluids and blood samples, the reduction of complexity is of prime importance (Section 3). The development of highly multiplexed quantitative targeted assays based on the exploration of suitable signature peptides, optimal transitions, and isotope-labeled internal standards are presented in Section 4.

The use of isotope dilution allows absolute protein quantification and improves the analytical accuracy by providing internal standards and compensating the changes occurring through sample preparation and analysis. A variety of studies on targeted quantification of protein biomarkers have been reported, including quantification of C-reactive protein from plasma samples [26,27], quantification of immunoglobulin G and its glycoforms from plasma [28], and multiplexed targeted detection of protein biomarkers in plasma from pancreatic cancer patients [29]. In most of the studies, certain types of sample preparation strategies were applied to reduce the sample complexity or enrich the targeted analytes. Without using prior affinity depletion or enrichment, a study showed the feasibility of absolute quantitation of 45 endogenous proteins, including 31 putative biomarkers of cardiovascular disease, in human plasma using mass spectrometric targeted approach [30].

2.2. Targeted monitoring of key proteins in a pathway network

Targeted proteomics can be utilized for the concomitant quantification of a set of proteins that are involved in a clinical condition, or a biological process [31]. Targeted proteomics has successfully quantified 464 proteins with known or suspected roles in transcriptional regulation at RNA polymerase II transcribed promoters in *Saccharomyces cerevisiae* [32]. A list of 1,261 proteins considered to be differentially expressed in human cancer was compiled from literature and other sources [33]. Some of these cancer-related proteins were analyzed in plasma from cancer patients, and 182 proteins were detected in depleted plasma, spanning five orders of magnitude in abundance and reaching a detection sensitivity of 10 ng/mL [11].

Sentinel proteins report the activation of specific cellular processes. In a study, 570 potentially suitable sentinels for *Saccharomyces cerevisiae* from available biological data were selected for the specific proteins, phosphorylation sites, or protein degradation products that report on four general classes of biological relationships [34]. Quantitative SRM assays were developed for 157 sentinel proteins and 152 sentinel phosphopeptides that could simultaneously probe 188 distinct biological processes in *Saccharomyces cerevisiae* in response to a set of environmental perturbations.

2.3. Targeted quantification of post-translational modifications

Post-translational modifications (PTMs) are playing a significant role in the activation or inhibition of biological processes, and their changes would be indicative for a clinical condition. Among the PTMs that are investigated frequently in clinical studies are phosphorylation and glycosylation. Glycoproteins unequivocally comprise the major biomolecules involved in extracellular processes and found mostly in secretome, such as growth factors, signaling proteins for cellular communication, enzymes, and proteases for on- and off-site processing [35–37]. Glycoproteomics have been used for the discovery of biomarkers in lung and pan-

creatic cancer [38,39]. Phosphorylated secreted proteins of tumor cells have been studied as source of candidates for breast cancer biomarkers in plasma [40].

For some PTMs, such as phosphorylation, methylation, and acetylation, synthetic reference peptides can be prepared with covalent modifications to mimic naturally occurring post-translational modifications [12]. Unlike the total protein quantification, the interrogation of PTMs status relies on the measurement of the targeted peptides that have undergone the desired modification. In this type of quantification, the subproteome is typically enriched using affinity columns or other separation techniques to enhance analytical sensitivity. For example, in glycosylation analysis, enrichment of N-glycosylated peptides coupled with targeted proteomics was applied to quantify the disease-responsive proteins in the sera of prostate cancer patients [41].

2.4. Targeted quantification of genetic changes

Genetic changes may have distinct effects at protein level. It may influence the expression level of proteins, modify their sequences through single nucleotide polymorphisms, the occurrence of allelic variants, or may impact the alternative splicing events [42]. Each individual may carry thousands of nonsynonymous single nucleotide variants in the genome, corresponding to various amino acid polymorphisms in the encoded proteins [43]. In global proteomic analysis, it is challenging to identify and quantify all protein variants in complex biological samples [42]. Targeted proteomics can be used in the quantification of protein isoforms, alternative splicing, SNPs, and other genetic mutations that result in changes in protein sequence. In such studies, the selected signature peptide should be unique and representing the targeted changes and should be suitable for mass spectrometric analysis [44].

In a study, which utilized targeted proteomics to quantify single amino acid polymorphisms, the absolute concentrations of three selected single amino acid polymorphism-peptides were measured in plasma from multiple individuals using SRM with the aid of heavy isotope-labeled peptide internal standards [44]. In a different study, a strategy for the comparative analysis of single amino acid polymorphism was developed by integration of stable isotope dimethyl labeling with a variation-associated database search approach. The technique could discover as many as 282 unique variation sites and quantify them in the human liver tissues. Although the identifications were restricted to the known genomic sequence variations, the use of a concise database improved the identification of variants at the protein level [45].

3. Reducing sample complexity — Blood analysis

Blood is a highly informative clinical matrix, which has been widely used in clinical analysis. In proteomics, the major challenge associated with the plasma or serum analysis is not only the sample complexity but also the enormous dynamic range (more than 11 orders of magnitude) in protein concentration [46]. The presence of high abundance proteins, such as albumin and IgG, can significantly mask the detection of low abundance proteins. Without prior sample treatment, the reported lower detection limit for plasma analysis using targeted proteomics

was at $\mu\text{g/mL}$ level [27], which is not suitable for measuring the majority of low and medium abundant proteins.

The depletion of high abundance proteins, fractionations at either protein or peptide levels, enrichment of target proteins, peptides, or sub-proteomes are among the suitable techniques that can be used to reduce the complexity of blood samples. A useful and convenient reduction of blood complexity should be performed based on the purpose of study and target molecules that are needed to be quantified [5].

3.1. Immuno-depletion of the high-abundance proteins

Immuno-depletion of the high-abundance proteins has been widely used to reduce the blood sample complexities. By depletion of the major plasma proteins, targeted mass spectrometric analysis could reach the lower limit of detection between 1 and 10 ng/mL [5]. The number of high-abundance proteins to be depleted varies and depends on the purpose of studies. Potential loss of non-target binding proteins associated with immuno-depletion may be a concern in some cases [47,48]. It has been proven that such a simple treatment of sample is an effective way to reduce complex matrix background and to highlight the candidate analytes for targeted analysis in a high-throughput manner [29,47,49].

Candidate protein biomarkers at low ng/mL to pg/mL levels were detected in serum after removing the 12 most abundant and 77 moderately abundant proteins from serum samples using antibody affinity columns [50]. Using immuno-depletion approach, proteins with 100 ng/ml or higher concentrations are readily accessible by targeted MS in plasma without antibody enrichment [51].

3.2. Fractionation of the plasma samples at protein or peptide levels

Besides immuno-depletion of high-abundance proteins, the fractionation of the proteins by size exclusion chromatography or using 2D electrophoresis can reduce blood complexity. On the other hand, tryptic peptides from the shotgun proteomics can be separated at peptide level using orthogonal separations, such as ion exchange chromatography coupled with reversed phase LC separation (e.g. MudPIT – multidimensional protein identification technology [52]), to obtain a better resolution of the eluting peptides. Online peptide fractionation strategies were also introduced to enhance quantitative analysis [53].

3.3. Targeted enrichment of proteins, peptides, or sub-proteome

Besides the fractionation practices and immuno-depletion, target proteins or peptides or sub-proteome can be enriched from the complex matrices using affinity or chemical methods to facilitate targeted analysis.

The method of stable isotope standards and capture by anti-peptide antibodies (SISCAPA) can reach a LOD as low as 0.1 ng/mL for plasma detection [54,55]. Rabbit polyclonal antibodies raised against the selected peptide sequences were covalently immobilized on POROS supports for enrichment of target peptides along with their heavy isotope labeled counterparts,

which were spiked in as internal standards for absolute quantification [54]. The technique has proven to enrich the target peptides against the background peptides by more than 100 times, and can be used to achieve high-throughput analysis using SPE-MS/MS technique [56].

Similar to the enrichment of specific target proteins or peptides, a sub-proteome at protein or peptide level can be enriched from the complex blood samples. Among the sub-proteome that are widely enriched from blood samples are N-linked glycoproteins/glycopeptides and phosphorylated residues. For N-glycoproteome analysis, lectin affinity and hydrazide chemistry have been the most widely utilized methods for the enrichment of glycoproteins or glycopeptides. TiO₂ columns are able to selectively purify phosphorylated peptides and sialic acid-containing N-glycopeptides [57].

Lectins are glycan-binding proteins that can bind to their target glycan moiety with high specificity [58,59]. The lectin affinity of sugar-containing residues can help their affinity enrichment at protein or peptide levels [60], which may be followed with the site-directed tagging of N-glycosylation sites by ¹⁸O during the elution with N-glycosidase [61]. Hydrazide chemistry occurs when certain sugars of the glycoproteins are oxidized to form reactive carbonyl groups. These carbonyls can then be conjugated to hydrazide-activated cross-linkers. The conjugated peptides/proteins are digested by PNGase F enzyme to cleave glycans from protein N-glycosylated sites, causing a mass shift of 0.98 Da due to the conversion of asparagine to aspartic acid [62,63]. This specific mass shift can be used for targeted interrogation of N-glycopeptides to identify N-glycosylation sites [64] and to monitor the glycosylation levels associated with the corresponding N-glycosylation sites [65]. Using such an approach, studies have demonstrated a LOD in the low ng/mL range and an analytical dynamic range over 5 orders of magnitude for plasma detection [66].

TiO₂ column is able to selectively purify phosphorylated peptides and sialic acid-containing N-glycopeptides. A method that combines an optimized TiO₂ protocol and hydrophilic interaction liquid chromatography to simultaneously enrich, identify, and quantify phosphopeptides and formerly N-linked sialylated glycopeptides to monitor changes associated with cell signaling in brain tissues has been reported [57]. Head-to-head comparison of several serum fractionation schemes, including N-linked glycopeptide enrichment, cysteinyl-peptide enrichment, magnetic bead separation, size fractionation, and immuno-depletion of abundant serum proteins have been performed. The analysis showed that immuno-subtraction was the most effective way to simplify the serum proteome while maintaining reasonable sample throughput [67].

3.4. Other instrumental innovations in reducing the blood complexity

High-pressure, high-resolution separations coupled with intelligent selection and multiplexing (PRISM) is an antibody-free strategy to reduce the plasma complexity for SRM analysis [68]. The strategy capitalizes on high-resolution reversed-phase liquid chromatographic separations for analyte enrichment, intelligent selection of target fractions via online SRM monitoring of internal standards, and fraction multiplexing before nano-liquid chromatography-SRM quantification. With the aid of the depletion of the 14 most abundant proteins, a study demonstrated that this method could detect AGR2 protein in human serum with the

concentration in the range of 50–100 pg/mL [69]. It is also reported that without the upfront immuno-depletion of the high-abundance proteins, the PRISM technique can reach limit of detections at low ng/mL range [70]. In addition to sample preparation strategies, ion mobility separation has been used for analyzing plasma samples, capitalizing on the gas phase separation of the co-eluting ions [71].

4. SRM assay development

Targeted quantitative proteomics requires development of high-throughput assays [72] to effectively detect a wide range of proteins in a biological sample with high reproducibility and robustness [73]. SRM-based methods have been the gold standard for MS-based protein quantification and have been widely applied in various studies. The development of an SRM assay typically involves an appropriate sample preparation, an optimal selection of signature peptides, and a well-calibrated MS protocol [74–76]. In the analysis of blood and other biofluids, especially when targeting low-abundance proteins, an effective sample preparation is almost mandated to reduce sample complexity or/and enrich targeted analytes, as aforementioned in Section 3.

4.1. Exploration of the most suitable signature peptide

An optimal assay should include the most sensitive and the most stable unique signature peptides to represent the target proteins. Ideally, multiple signature peptides that are belonging to different domains of the protein are preferred to quantify the target protein for the reasons of reliability of the quantifications. This is because various domains may have different efficiencies for trypsin digestion. The results of quantification for each signature peptide may differ, which in this case might be indicative for the truncation of the target protein or degradation besides the different digestion rates from different domains [10].

Evaluation of candidate signature peptides from the target proteins is of importance to obtain a sensitive and reliable quantification. The uniqueness of the signature peptides can be verified by comparison with the protein databases using alignment software such as protein BLAST, and empirically verified from matrix. Moreover, the unique peptides should be evaluated for their stability, the absence of labile residues, and the risk of incomplete digestion, PTMs, having appropriate length, hydrophilicity, and other relevant parameters, such as their chromatographic and mass spectrometric characteristics [77]. Human plasma proteome project have already identified 20,433 distinct peptides, from which a highly nonredundant set of 1,929 protein sequences at a false discovery rate of 1% are inferred [78]. In addition, collections of peptide spectral libraries, such as PeptideAtlas [79] and SRMATlas [80], provide empirical data to facilitate signature peptide selection.

4.2. Optimization of collision energy and most sensitive transitions

Selection of the most robust transitions is essential for quantification of signature peptides. Usually, multiple transitions are selected for the verification of a same signature peptide. In

the case of presence of pure signature peptides, the fragmentation pattern for each of them can be performed empirically and the most stable and sensitive transitions can be selected for assay development. But in the absence of pure signature peptides, using the existing spectral libraries such as PeptideAtlas [79,81] or SRMAtlas [80] can provide useful information in evaluating the mass spectral characteristics of a peptide. These libraries are based on the computational or experimental data resulted from the collision fragmentation of a large number of synthesized peptides. However, it should be noted that the fragmentation patterns depend on the instrument types and may differ by different vendors. One possible solution can be the use of on-the-fly orthogonal-array optimization of the collision energies and transitions for any given signature peptide, especially in the absence of the pure signature peptides [82]. Another approach is using PRM, in which all the fragmented ions obtained from the same substrate are monitored together to obtain a more reliable quantification result [15], as illustrated in Scheme 1.

To assist transition selection, a novel algorithm was presented to allow the construction of SRM assays from the sequence of the targeted proteins alone. This approach relies on combinatorial optimization with machine learning techniques to predict proteotypicity, retention time, and fragmentation of peptides, enabling rapid development of a targeted SRM experiment [83]. Using the contemporary MS capabilities, instrument parameters can be optimized for each peptide for any given retention time and transition. A study has shown that the optimal collision energies for each respective charge-state can be predicted using linear equations based on the peptide precursor mass. These charge-state-dependent equations for predicting the optimal collision energies are embedded within Skyline software [84].

It is also worthy to mention that in triple Q based SRM methods, there is a reverse relation between increased dwell time to obtain higher sensitivities, and a reliable peptide profiling. Spending longer times for each analysis means less number of quantified points, and poor peptide time profiles. This issue can be partly addressed by using scheduled SRM acquisition and restricted time window for the known peptides expected to elute in the corresponding time interval [85]. Higher resolution separation with high reproducibility and longer gradient times would increase the number of target peptides to be quantified with high sensitivity and reliability within a single run.

4.3. Exploration of the most suitable internal standards

Targeted proteomics can be used for either relative or absolute quantification. In the case of absolute quantification, there is a need for appropriate calibration set up next to the isotope dilution mass spectrometry. Individual heavy isotope-labeled internal standards, which are spiked in a sample with known amounts, would serve as internal standards for the corresponding endogenous peptides for specific quantification. On the other hand, with less quantitative accuracy, a single internal standard or fluorinated internal standards can be used. In addition, stable isotope-labeled proteins, such as QconCATs, can be used as internal standards, having the advantage of circumventing the variations caused during digestion [86]. With optimal settings and a stringent quality control, SRM-based targeted proteomics can be highly reproducible within and across laboratories [4]

5. Data Independent Acquisition (DIA) for targeted analysis

Advent of high-frequency and high-resolution mass spectrometry has provided the potential for data independent acquisition (DIA) [16,87]. While conventional data dependent analysis precludes the analysis of some eluted peptides [88], in DIA acquisition, MS generates virtually all the MS/MS fragmentation spectra from all precursor ions that are falling into a predefined m/z range. Hence, each recorded MS/MS fragmentation spectrum is a multiplexed recording of the fragment ions derived from all peptides eluting in real time within the predefined m/z range of the precursor window [87]. Scheme 1 illustrates the main elements in DIA technique. Due to the unbiased fragmentation of precursor ions, DIA approach provides a high multiplexing capability, high reproducibility, and wide analytical scope. Conceptually, DIA-based mass spectrometric analysis can be viewed as an SRM assay on all peptides detected, allowing extraction of pseudo SRM data for any peptide of interest within the mass spectrometric detection limit. The design of a DIA method may be dependent on study purpose and sample type, and requires an optimal balance of multiple instrument parameters, including targeted mass range, DIA window width, duty cycle time, and automated gain control, etc.

DIA is a generic term encompassing a wide range of recently developed techniques that are built on the analysis of a non-predefined set of precursor ions. The early DIA technique, PAcIFIC, was based on the multiple LC/MS runs at limited mass ranges [89,90]. The technique suffered from prolonged analysis times. Recently, a variety of DIA techniques have been explored and implemented using different mass spectrometers, including triple TOF based sequential window acquisition of all theoretical fragment ion spectra (SWATH), Q/TOF based MS^E, and Orbitrap based multiplexing strategy (MSX) [87]. These DIA techniques differ in the instrument platforms and using isolation windows of various widths, depending on different study purposes and instrument settings [91,92].

Coupled with hydrazide-based solid phase extraction for N-glycosylation enrichment, SWATH has been applied to analyze deglycosylated N-glycopeptide in human plasma. While the sensitivity of SWATH was slightly less than SRM, the study demonstrated that SWATH could reach a detection limit of 5 ng/mL in plasma and quantify N-glycopeptides with a concentration range of 4 orders of magnitude [93]. The same approach (using N-glycopeptide enrichment) was successfully applied to analyze prostate cancer tissues and identified 1,430 N-glycosylation sites from each sample in average, including 220 proteins that showed quantitative changes associated with tumor aggressiveness [94].

A recent study has suggested that more than 10,000 human proteins (the majority of human proteins from UniProt database) could potentially be covered using SWATH-MS technique that can be of high value for clinical studies [95]. In this study, a variety of human cell types and depleted human plasma samples were analyzed with the aid of various sample preparation techniques, including affinity purification, size exclusion chromatography, strong anion exchange, and gel electrophoresis [95]. In the quantitative study of human twin population, the plasma samples from twins are used to explore the impact of longitudinal factors in blood proteomic changes. This study included the identification of some genetic changes that occurred by time [96].

For phosphoproteomics, SRM and SWATH have shown similar performance in the determination of changes of phosphopeptide levels extracted from human plasma [97]. The general theme in the DIA analysis of phosphorylation and glycosylation is the selective enrichment of the corresponding sub-proteome [15,34,94]. A DIA method, namely combination hyper-reaction monitoring (HRM), used retention time normalized (iRT) spectral libraries for spectral identification. Using a controlled sample set, the HRM outperformed shotgun proteomics both in the number of consistently identified peptides across multiple measurements and quantification of differentially abundant proteins when it profiled acetaminophen (APAP)-treated 3D human liver microtissues [98].

6. Software used in targeted proteomics

A variety of software has been developed to assist targeted proteomic data analysis. MRMer is an interactive open source and cross-platform system for data extraction and visualization of multiple reaction monitoring experiments [99]. MRMer parses and extracts information from MS files encoded in the platform-independent mzXML data format. mProphet is an automated data processing and statistical validation tool for large-scale SRM experiments [100]. Skyline can be used for analyzing a variety of targeted proteomic data, including SRM- and DIA-based data [101]. The extraction of pseudo-SRM profiles from DIA data requires a spectral library, which can be built using global profiling data for peptide and protein identification. Skyline is also capable of analyzing MSX (Multiplexed MS/MS) based DIA data [102]

DIA-Umpire is a software program that has been recently developed and performs the data extraction based on the co-elution of the substrate and its corresponding fragmentation to build a pseudo-MS/MS library, which later can become useful in identification and targeted quantification [103]. Spectronaut extends the limits of quantitative proteome profiling with DIA [98].

7. Current status and further research

Currently, SRM is the gold standard for mass spectrometry-based targeted analysis and has been widely applied in a broad range of translational and clinical studies. While SRM provides high sensitivity, one major drawback from using SRM is that a SRM assay is dependent on the geometry of instruments and the instrumental settings. Thus, it would require an extensive effort in assay development for each specific group of analytes on a particular instrument. The number of SRM assays is also limited as there is a reverse correlation between the number of transitions (selectivity) and the quality of quantification (sensitivity). On the other hand, the advent of DIA has introduced a virtually unlimited pseudo-SRM analysis that can be run once and used for the extraction of any given data within the detection limit. The technique, which, in a way, hybrids the technical characters of discovery-based proteomics and targeted analysis, is undergoing a rapid progress and represents a paradigm shift in targeted proteomics. The

non-biased nature and the highly multiplexing capacity that is enhanced by DIA will render a universal approach for targeted proteomics in translational and clinical investigations.

Acknowledgements

We are grateful to the supports from the National Institutes of Health under grants K25CA137222, R01CA180949 and R21CA164548.

Author details

Eslam Nouri-Nigjeh, Ru Chen and Sheng Pan*

*Address all correspondence to: shengp@medicine.washington.edu

Department of Medicine, University of Washington, Seattle, WA, USA

References

- [1] Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. *Science* 1989;246(4926):64–71.
- [2] Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003;422(6928):198–207.
- [3] Aldred S, Grant MM, Griffiths HR. The use of proteomics for the assessment of clinical samples in research. *Clin Biochem* 2004;37(11):943–52.
- [4] Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, et al. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring based measurements of proteins in plasma. *Nat Biotechnol* 2009;27(7):633–41.
- [5] Hüttenhain R, Malmström J, Picotti P, Aebersold R. Perspectives of targeted mass spectrometry for protein biomarker verification. *Curr Opin Chem Biol* 2009;13(5–6):518–25.
- [6] Hale JE. Advantageous uses of mass spectrometry for the quantification of proteins. *Int J Proteomics* 2013;2013:219452 DOI:10.1155/2013/219452.
- [7] Lomax DP, Roubal WT, Moore JD, Johnson LL. An enzyme-linked immunosorbent assay (ELISA) for measuring vitellogenin in English sole (*Pleuronectes vetulus*): development, validation and cross-reactivity with other pleuronectids. *Compar Biochem Physiol B Biochem Molecul Biol* 1998;121(4):425–36.

- [8] Lin D, Alborn WE, Slebos RJC, Liebler DC. Comparison of protein immunoprecipitation-multiple reaction monitoring with ELISA for assay of biomarker candidates in plasma. *J Proteome Res* 2013;12(12):5996–6003.
- [9] Domon B, Aebersold R. Mass Spectrometry and protein analysis. *Science* 2006;312(5771):212–7.
- [10] Nouri-Nigjeh E, Zhang M, Ji T, Yu H, An B, Duan X, et al. Effects of calibration approaches on the accuracy for LC/MS targeted quantification of therapeutic protein. *Analyt Chem* 2014;86(7):3575–84.
- [11] Hüttenhain R, Soste M, Selevsek N, Rost H, Sethi A, Carapito C, et al. Reproducible quantification of cancer-associated proteins in body fluids using targeted proteomics. *Sci Translat Med* 2012;4:142–94.
- [12] Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Nat Acad Sci* 2003;100(12):6940–5.
- [13] Benkali K, Marquet P, Rerolle JP, Le Meur Y, Gastinel LN. A new strategy for faster urinary biomarkers identification by Nano-LC-MALDI-TOF/TOF mass spectrometry. *BMC Genom* 2008;9(1):541–50.
- [14] Pan S, Zhang H, Rush J, Eng J, Zhang N, Patterson D, et al. High throughput proteome screening for biomarker detection. *Molecul Cellul Proteom* 2005;4(2):182–90.
- [15] Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Molecul Cellul Proteom* 2012;11(11):1475–88.
- [16] Sajic T, Liu Y, Aebersold R. Using data-independent, high resolution mass spectrometry in protein biomarker research: perspectives and clinical applications. *Proteom Clin Applic* 2014;00:1–15.
- [17] Addona TA, Shi X, Keshishian H, Mani DR, Burgess M, Gillette MA, et al. A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease. *Nat Biotechnol* 2011;29(7):635–43.
- [18] Pogue-Geile KL, Chen R, Bronner MP, Crnogorac-Jurcevic T, Moyes KW, Downen S, et al. Palladin mutation causes familial pancreatic cancer and suggests a new cancer mechanism. *PLoS Med* 2006;3(12):2216–28.
- [19] Whiteaker JR, Lin C, Kennedy J, Hou L, Trute M, Sokal I, et al. A targeted proteomics-based pipeline for verification of biomarkers in plasma. *Nat Biotechnol* 2011;29:625–34.
- [20] Gortzak-Uzan L, Ignatchenko A, Evangelou AI, Agochiya M, Brown KA, St Onge P, et al. A proteome resource of ovarian cancer ascites: integrated proteomic and bioinformatic analyses to identify putative biomarkers. *J Proteome Res* 2008;7:339–51.

- [21] Chaerkady R, Harsha HC, Nalli A, Gucek M, Vivekanandan P, Akhtar J, et al. A quantitative proteomic approach for identification of potential biomarkers in hepatocellular carcinoma. *J Proteome Res* 2008;7:4289–98.
- [22] Ludwig JA, Weinstein JN. Biomarkers in cancer staging, prognosis and treatment selection. *Nat Rev Cancer* 2005;5(11):845–56.
- [23] Anderson NL. The clinical plasma proteome: a survey of clinical assays for proteins in plasma and serum. *Clin Chem* 2010;56(2):177–85.
- [24] Domon B, Aebersold R. Options and considerations when selecting a quantitative proteomics strategy. *Nat Biotechnol* 2010;28(7):710–21.
- [25] Paulovich AG, Whiteaker JR, Hoofnagle AN, Wang P. The interface between biomarker discovery and clinical validation: the tar pit of the protein biomarker pipeline. *Proteomics Clin Applic* 2008;2:1386–402.
- [26] Kuhn E, Wu J, Karl J, Liao H, Zolg W, Guild B. Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and ¹³C-labeled peptide standards. *Proteomics* 2004;4(4):1175–86.
- [27] Williams DK, Muddiman DC. Absolute quantification of C-reactive protein in human plasma derived from patients with epithelial ovarian cancer utilizing protein cleavage isotope dilution mass spectrometry. *J Proteome Res* 2009;8:1085–90.
- [28] Hong Q, Lebrilla CB, Miyamoto S, Ruhaak LR. Absolute quantitation of immunoglobulin G and its glycoforms using multiple reaction monitoring. *Anal Chem* 2013;85:8585–93.
- [29] Pan S, Chen R, Brand RE, Hawley S, Tamura Y, Gafken PR, et al. Multiplex targeted proteomic assay for biomarker detection in plasma: a pancreatic cancer biomarker case study. *J Proteome Res* 2012;11(3):1937–48.
- [30] Kuzyk MA, Smith D, Yang J, Cross TJ, Jackson AM, Hardie DB, et al. Multiple reaction monitoring-based, multiplexed, absolute quantitation of 45 proteins in human plasma. *Molecul Cellul Proteomics* 2009;8(8):1860–77.
- [31] Chen R, Yi EC, Donohoe S, Pan S, Eng J, Cooke K, et al. Pancreatic cancer proteome: the proteins that underlie invasion, metastasis, and immunologic escape. *Gastroenterology* 2005;129(4):1187–97.
- [32] Mirzaei H, Knijnenburg TA, Kim B, Robinson M, Picotti P, Carter GW, et al. Systematic measurement of transcription factor-DNA interactions by targeted mass spectrometry identifies candidate gene regulatory proteins. *Proc Nat Acad Sci USA* 2013;110(9):3645–50.
- [33] Polanski M, Anderson NL. A list of candidate cancer biomarkers for targeted proteomics. *Biomark Insights* 2006;1:1–48.

- [34] Soste M, Hrabakova R, Wanka S, Melnik A, Boersema P, Maiolica A, et al. A sentinel protein assay for simultaneously quantifying cellular processes. *Nat Methods* 2014;11(10):1045–8.
- [35] Doerr A. Glycoproteomics. *Nat Methods* 2012;9:36.
- [36] Kim EH, Misek DE. Glycoproteomics-based identification of cancer biomarkers. *Int J Proteomics* 2011;2011:601937 DOI:10.1155/2011/601937.
- [37] Tian Y, Zhang H. Glycoproteomics and clinical applications. *Proteom Clin Applic* 2010;4(2):124–32.
- [38] Kay Li Q, Gabrielson E, Zhang H. Application of glycoproteomics for the discovery of biomarkers in lung cancer. *Proteom Clin Applic* 2012;6:244–56.
- [39] Pan S, Chen R, Aebersold R, Brentnall TA. Mass spectrometry based glycoproteomics from a proteomics perspective. *Molecul Cellul Proteom* 2011;10:1–14.
- [40] Zawadzka AM, Schilling B, Cusack MP, Sahu AK, Drake P, Fisher SJ, et al. Phospho-protein secretome of tumor cells as a source of candidates for breast cancer biomarkers in plasma. *Molecul Cellul Proteom* 2014;13(4):1034–49.
- [41] Cima I, Schiess R, Wild P, Kaelin M, Schüffler P, Lange V, et al. Cancer genetics-guided discovery of serum biomarker signatures for diagnosis and prognosis of prostate cancer. *Proc Nat Acad Sci USA* 2011;108(8):3342–7.
- [42] Horvatovich P, Franke L, Bischoff R. Proteomic studies related to genetic determinants of variability in protein concentrations. *J Proteome Res* 2014;13(1):5–14.
- [43] Sheynkman GM, Shortreed MR, Frey BL, Scalf M, Smith LM. Large-scale mass spectrometric detection of variant peptides resulting from nonsynonymous nucleotide differences. *J Proteome Res* 2014;13(1):228–40.
- [44] Su ZD, Sun L, Yu DX, Li RX, Li HX, Yu ZJ, et al. Quantitative detection of single amino acid polymorphisms by targeted proteomics. *J Molecul Cell Biol* 2011;3(5):309–15.
- [45] Song C, Wang F, Cheng K, Wei X, Bian Y, Wang K, et al. Large-scale quantification of single amino-acid variations by a variation-associated database search strategy. *J Proteome Res* 2014;13(1):241–8.
- [46] Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Molecul Cellul Proteom* 2007;6(12):2212–29.
- [47] Bellei E, Bergamini S, Monari E, Fantoni L, Cuoghi A, Ozben T, et al. High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins. *Amino Acids* 2011;40(1):145–56.

- [48] Tu C, Rudnick PA, Martinez MY, Cheek KL, Stein SE, Slebos RJC, et al. Depletion of abundant plasma proteins and limitations of plasma proteomics. *J Proteome Res* 2010;9(10):4982–91.
- [49] Amon LM, Law W, Fitzgibbon MP, Gross JA, O'Briant K, Peterson A, et al. Integrative proteomic analysis of serum and peritoneal fluids helps identify proteins that are up-regulated in serum of women with ovarian cancer. *PLoS ONE* 2010;5(6):e11137.
- [50] Lin B, White JT, Wu J, Lele S, Old LJ, Hood L, et al. Deep depletion of abundant serum proteins reveals low-abundant proteins as potential biomarkers for human ovarian cancer. *Proteom Clin Applic* 2009;3:853–61.
- [51] Anderson NL, Anderson NG, Pearson TW, Borchers CH, Paulovich AG, Patterson SD, et al. A Human proteome detection and quantitation project. *Molecul Cellul Proteom* 2009;8(5):883–6.
- [52] Magdelin S, Moresco JJ, Yamamoto T, Yates JR. Off-line multidimensional liquid chromatography and auto sampling result in sample loss in LC/LC–MS/MS. *J Proteome Res* 2014;13(8):3826–36.
- [53] Magdelin S, Elguoshy A, Yoshida Y, Hitrao Y, Xu B, Zhang Y, et al. Complementary protein and peptide OFFGEL fractionation for high-throughput proteomic analysis. *Anal Chem* 2015;87(16):8481–8.
- [54] Anderson NL, Anderson NG, Haines LR, Hardie DB, Olafson RW, Pearson TW. Mass spectrometric quantitation of peptides and proteins using stable isotope standards and capture by anti-peptide antibodies (SISCAPA). *J Proteome Res* 2004;3(2):235–44.
- [55] Neubert H, Muirhead D, Kabir M, Grace C, Cleton A, Arends R. Sequential protein and peptide immunoaffinity capture for mass spectrometry-based quantification of total human β -nerve growth factor. *Anal Chem* 2013;85(3):1719–26.
- [56] Razavi M, Frick LE, LaMarr WA, Pope ME, Miller CA, Anderson NL, et al. High-throughput SISCAPA quantitation of peptides from human plasma digests by ultra-fast, liquid chromatography-free mass spectrometry. *J Proteome Res* 2012;11(12):5642–9.
- [57] Palmisano G, Parker BL, Engholm-Keller K, Lendal SE, Kulej K, Schulz M, et al. A novel method for the simultaneous enrichment, identification, and quantification of phosphopeptides and sialylated glycopeptides applied to a temporal profile of mouse brain development. *Molecul Cellul Proteom* 2012;11:1191–202.
- [58] Lai ZW, Nice EC, Schilling O. Glyco-capture-based proteomics for secretome analysis. *Proteomics* 2013;13(3–4):512–25.
- [59] Ongay S, Boichenko A, Govorukhina N, Bischoff R. Glycopeptide enrichment and separation for protein glycosylation analysis. *J Separation Sci* 2012;35(18):2341–72.

- [60] Ahn YH, Kim KH, Shin PM, Ji ES, Kim H, Yoo JS. Identification of low-abundance cancer biomarker candidate TIMP1 from serum with lectin fractionation and peptide affinity enrichment by ultrahigh-resolution mass spectrometry. *Anal Chem* 2012;84:1425–31.
- [61] Ueda K, Takami S, Saichi N, Daigo Y, Ishikawa N, Kohno N, et al. Development of serum glycoproteomic profiling technique; simultaneous identification of glycosylation sites and site-specific quantification of glycan structure changes. *Molecul Cellul Proteom* 2010;9:1819–28.
- [62] Zhang H, Li X, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* 2003;21:660–6.
- [63] Wang L, Aryal UK, Dai Z, Mason AC, Monroe ME, Tian ZX, et al. Mapping N-linked glycosylation sites in the secretome and whole cells of *Aspergillus niger* using hydrazide chemistry and mass spectrometry. *J Proteome Res* 2011;11:143–56.
- [64] Zhang H, Liu AY, Loriaux P, Wollscheid B, Zhou Y, Watts JD, et al. Mass spectrometric detection of tissue proteins in plasma. *Molecul Cellul Proteom* 2007;6(1):64–71.
- [65] Pan S, Chen R, Tamura Y, Crispin DA, Lai LA, May DH, et al. Quantitative glycoproteomics analysis reveals changes in N-glycosylation level associated with pancreatic ductal adenocarcinoma. *J Proteome Res* 2014;13(3):1293–306.
- [66] Stahl-Zeng J, Lange V, Ossola R, Eckhardt K, Krek W, Aebersold R, et al. High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Molecul Cellul Proteom* 2007;6(10):1809–17.
- [67] Whiteaker JR, Zhang H, Eng JK, Fang R, Piening BD, Feng LC, et al. Head-to-Head comparison of serum fractionation techniques. *J Proteome Res* 2006;6(2):828–36.
- [68] Shi T, Fillmore TL, Sun X, Zhao R, Schepmoes AA, Hossain M, et al. Antibody-free, targeted mass-spectrometric approach for quantification of proteins at low picogram per milliliter levels in human plasma/serum. *Proc Nat Acad Sci* 2012;109(38):15395–400.
- [69] Shi T, Gao Y, Quek SI, Fillmore TL, Nicora CD, Su D, et al. A highly sensitive targeted mass spectrometric assay for quantification of AGR2 protein in human urine and serum. *J Proteome Res* 2014;13(2):875–82.
- [70] Shi T, Sun X, Gao Y, Fillmore TL, Schepmoes AA, Zhao R, et al. Targeted quantification of low ng/mL level proteins in human serum without immunoaffinity depletion. *J Proteome Res* 2013;12(7):3353–61.
- [71] Daly CE, Ng LL, Hakimi A, Willingale R, Jones DJL. Qualitative and quantitative characterization of plasma proteins when incorporating traveling wave ion mobility into a liquid chromatography-mass spectrometry workflow for biomarker discovery:

- use of product ion quantitation as an alternative data analysis tool for label free quantitation. *Anal Chem* 2014;86(4):1972–9.
- [72] Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol* 2006;24(8):971–83.
- [73] Surinova S, Schiess R, Hüttenhain R, Cerciello F, Wollscheid B, Aebersold R. On the development of plasma protein biomarkers. *J Proteome Res* 2010;10(1):5–16.
- [74] Pan S, Aebersold R, Chen R, Rush J, Goodlett DR, McIntosh MW, et al. Mass spectrometry based targeted protein quantification: methods and applications. *J Proteome Res* 2008;8(2):787–97.
- [75] Lange V, Picotti P, Domon B, Aebersold R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Molecul Syst Biol* 2008;4:222.
- [76] Picotti P, Aebersold R. Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat Methods* 2012;9(6):555–66.
- [77] Mallick P, Schirle M, Chen SS, Flory MR, Lee H, Martin D, et al. Computational prediction of proteotypic peptides for quantitative proteomics. *Nat Biotechnol* 2007;25(1):125–31.
- [78] Farrah T, Deutsch EW, Omenn GS, Campbell DS, Sun Z, Bletz JA, et al. A high-confidence human plasma proteome reference set with estimated concentrations in PeptideAtlas. *Molecul Cellul Proteomics* 2011;10(9):M110.
- [79] Deutsch EW, Lam H, Aebersold R. PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. *EMBO Rep* 2008;9(5):429–34.
- [80] Rost H, Malmstrom L, Aebersold R. A computational tool to detect and avoid redundancy in selected reaction monitoring. *Molecul Cellul Proteom* 2012;11(8):540–9.
- [81] Farrah T, Deutsch E, Aebersold R. Using the human plasma PeptideAtlas to study human plasma proteins. In: Simpson RJ, Greening DW (Eds.) *Method Molecul Biol* 2011;728:349–74.
- [82] Cao J, Covarrubias VM, Straubinger RM, Wang H, Duan X, Yu H, et al. A rapid, reproducible, on-the-fly orthogonal array optimization method for targeted protein quantification by LC/MS and its application for accurate and sensitive quantification of carbonyl reductases in human liver. *Anal Chem* 2010;82(7):2680–9.
- [83] Bertsch A, Jung S, Zerck A, Pfeifer N, Nahnsen S, Henneges C, et al. Optimal de novo design of MRM experiments for rapid assay development in targeted proteomics. *J Proteome Res* 2010;9(5):2696–704.
- [84] MacLean B, Tomazela DM, Abbatiello SE, Zhang S, Whiteaker JR, Paulovich AG, et al. Effect of collision energy optimization on the measurement of peptides by selected reaction monitoring (SRM) mass spectrometry. *Anal Chem* 2010;82(24):10116–24.

- [85] Fillâtre Y, Rondeau D, Jadas-Hècart A, Communal PY. Advantages of the scheduled selected reaction monitoring algorithm in liquid chromatography/electrospray ionization tandem mass spectrometry multi-residue analysis of 242 pesticides: a comparative approach with classical selected reaction monitoring mode. *Rapid Commun Mass Spectro* 2010;24(16):2453–61.
- [86] Brownridge P, Holman SW, Gaskell SJ, Grant CM, Harman VM, Hubbard SJ, et al. Global absolute quantification of a proteome: challenges in the deployment of a QconCAT strategy. *Proteomics* 2011;11(15):2957–70.
- [87] Chapman JD, Goodlett DR, Masselon CD. Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. *Mass Spectro Rev* 2014;33(6):452–70.
- [88] Michalski A, Cox J, Mann M. More than 100,000 Detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC/MS/MS. *J Proteome Res* 2011;10(4):1785–93.
- [89] Panchaud A, Jung S, Shaffer SA, Aitchison JD, Goodlett DR. PAcIFIC goes faster, quantitative and accurate. *Anal Chem* 2011;83(6):2250–7.
- [90] Panchaud A, Scherl A, Shaffer SA, von Haller PD, Kulasekara HD, Miller SI, et al. PAcIFIC: how to dive deeper into the proteomics ocean. *Anal Chem* 2009;81(15):6481–8.
- [91] Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Molecul Cellul Proteom* 2012;11(6).
- [92] Law KP, Lim YP. Recent advances in mass spectrometry: data independent analysis and hyper reaction monitoring. *Exp Rev Proteom* 2013;10(6):551–66.
- [93] Liu Y, Hüttenhain R, Surinova S, Gillet LC, Mouritsen J, Brunner R, et al. Quantitative measurements of N-linked glycoproteins in human plasma by SWATH-MS. *Proteomics* 2013;13(8):1247–56.
- [94] Liu Y, Chen J, Sethi A, Li QK, Chen L, Collins B, et al. Glycoproteomic analysis of prostate cancer tissues by SWATH mass spectrometry discovers N-acylethanolamine acid amidase and protein tyrosine kinase 7 as signatures for tumor aggressiveness. *Molecul Cellul Proteom* 2014;13(7):1753–68.
- [95] Rosenberger G, Koh CC, Guo T, Röst HL, Kouvonen P, Collins BC, et al. A repository of assays to quantify 10,000 human proteins by SWATH-MS. *Scientific Data* 2014;1:14003 DOI:10.1038/sdata.2014.31.
- [96] Liu Y, Buil A, Collins BC, Gillet LC, Blum LC, Cheng LY, et al. Quantitative variability of 342 plasma proteins in a human twin population. *Molecul Syst Biol* 2015;11(786):1_18DOI: 10.15252/msb.20145728.

- [97] Zawadzka AM, Schilling B, Held JM, Sahu AK, Cusack MP, Drake PM, et al. Variation and quantification among a target set of phosphopeptides in human plasma by multiple reaction monitoring and SWATH-MS2 data-independent acquisition. *Electrophoresis* 2014;35(24):3487–97.
- [98] Bruderer R, Bernhardt OM, Gandhi T, Miladinovic SM, Cheng LY, Messner S, et al. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen treated 3D liver microtissues. *Molecular Cell Proteomics* 2015; 14(5):1400_10 DOI: 10.1074/mcp.M114.044305.
- [99] Martin DB, Holzman T, May D, Peterson A, Eastham A, Eng J, et al. MRMer, an interactive open source and cross-platform system for data extraction and visualization of multiple reaction monitoring experiments. *Molecular Cell Proteomics* 2008;7(11): 2270–8.
- [100] Reiter L, Rinner O, Picotti P, Hüttenhain R, Beck M, Brusniak MY, et al. mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat Methods* 2011;8(5):430–5.
- [101] MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 2010;26(7):966–8.
- [102] Egertson JD, Kuehn A, Merrihew GE, Bateman NW, MacLean BX, Ting YS, et al. Multiplexed MS/MS for improved data-independent acquisition. *Nat Methods* 2013;10(8):744–6.
- [103] Tsou CC, Avtonomov D, Larsen B, Tucholska M, Choi H, Gingras AC, et al. DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nat Methods* 2015;12:258–64.



Edited by Sameh Magdeldin

Proteomics refers to the entire complement of proteins, including modification. This promising discipline has enabled us to study proteins from a massive and comprehensive point of view. The book *Recent Advances in Proteomics Research* describes in five sections some of the applications of proteomics. This fine research has been written by leading experts worldwide. This book is aimed mainly at those interested in proteins and in the field of proteins, particularly biochemists, biologists, pharmacists, advanced graduate students and postgraduate researchers.

Photo by Naked King / iStock

IntechOpen

