Computational and Structural Biotechnology Journal 11 (2014) 73–90

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Review

Proteomics for systems toxicology

Bjoern Titz *[,1], Ashraf Elamin [1], Florian Martin, Thomas Schneider, Sophie Dijon, Nikolai V. Ivanov, Julia Hoeng, Manuel C. Peitsch

Philip Morris International R&D, Philip Morris Products S.A., Quai Jeanrenaud 5, 2000 Neuchâtel, Switzerland

ARTICLE INFO

Available online 27 August 2014

Keywords:
Systems toxicology
Quantitative proteomics
Computational analysis

ABSTRACT

Current toxicology studies frequently lack measurements at molecular resolution to enable a more mechanism-based and predictive toxicological assessment. Recently, a systems toxicology assessment framework has been proposed, which combines conventional toxicological assessment strategies with system-wide measurement methods and computational analysis approaches from the field of systems biology. Proteomic measurements are an integral component of this integrative strategy because protein alterations closely mirror biological effects, such as biological stress responses or global tissue alterations. Here, we provide an overview of the technical foundations and highlight select applications of proteomics for systems toxicology studies. With a focus on mass spectrometry-based proteomics, we summarize the experimental methods for quantitative proteomics and describe the computational approaches used to derive biological/mechanistic insights from these datasets. To illustrate how proteomics has been successfully employed to address mechanistic questions in toxicology, we summarized several case studies. Overall, we provide the technical and conceptual foundation for the integration of proteomic measurements in a more comprehensive systems toxicology assessment framework. We conclude that, owing to the critical importance of protein-level measurements and recent technological advances, proteomics will be an integral part of integrative systems toxicology approaches in the future.

Contents

* Corresponding author.
E-mail address: bjorn.titz@pmi.com (B. Titz).
[1] Contributed equally.

http://dx.doi.org/10.1016/j.csbj.2014.08.004

1. Introduction

Conventional toxicological assessment of chemical substances relies heavily on in vitro assays and animal studies to test and identify exposure doses at which relevant apical endpoints are adversely affected. These apical endpoints measure major effects on animal physiology including gross developmental defects or reduction of body weight. Based on these results, recommendations for human exposure limits are derived. Although this conventional toxicological approach has clearly proven its value, more recent discussions on the future requirements for toxicological assessment have highlighted some of its shortcomings and emphasized the need to further evolve toxicology assessment with new tools and approaches (e.g., through the Tox21 and EPA ToxCast™ initiatives) [1,2]. The challenges faced by the current toxicological assessment approach include the recent explosive growth of required tests (e.g., for approximately 300 new chemicals per year in the U.S. alone), the need for new endpoints such as endocrine modulation, and the need to evaluate the effect of chemical mixtures [1]. Most important, however, is the urgent need for deeper insights into toxicological mechanisms as the basis for improved toxicity predictions for different human exposure scenarios. An important challenge in this endeavor is the selection of the right assay systems to conduct predictive studies. While we are witnessing the development of in vitro systems of increasing relevance and complexity, they can still not fully replace animal studies. This is a second reason to focus our attention on mechanistic understanding of toxicity as this opens two routes for developing more predictive assessment tools. First, mechanistic understanding allows for the identification of key events which can be replicated as discrete assays in vitro. Second, mechanistic understanding allows identifying which portion of animal biology translates to human biology and is thus adequate for toxicology testing. Related to this is the notion that the quantitative analysis of a discrete number of toxicological pathways that are causally linked to the apical endpoints could improve predictions (Pathways of Toxicity, POT) [3]. These concepts were recently summarized in a systems toxicology framework [4] where the systems biology approach with its large-scale measurements and computational modeling approaches is combined with the requirements of toxicological studies. Specifically, this integrative approach relies on extensive measurements of exposure effects at the molecular level (e.g., proteins and RNAs), at different levels of biological complexity (e.g., cells, tissues, animals), and across species (e.g., human, rat, mouse). These measurements are subsequently integrated and analyzed computationally to understand the causal chain of molecular events that leads from toxin exposure to an adverse outcome and to facilitate reliable predictive modeling of these effects.

Importantly, to capture the full complexity of toxicological responses, systems toxicology relies heavily on the integration of different data modalities to measure changes at different biological levels—ranging from changes in mRNAs (transcriptomics) to changes in proteins and protein states (proteomics) to changes in phenotypes (phenomics). Owing to the availability of well-established measurement methods, transcriptomics is often the first choice for systems-level investigations. However, protein changes can be considered to be closer to the relevant functional impact of a studied stimulus. Although mRNA and protein expression are tightly linked through translation, their correlation is limited, and mRNA transcript levels only explain about 50% of the variation of protein levels [5]. This is because of the additional levels of protein regulation including their rate of translation and degradation. Moreover, the regulation of protein activity does not stop at its expression level but is often further controlled through posttranslational modification such as phosphorylation; examples for the relevance of post-transcriptional regulation for toxicological responses include: the tight regulation of p53 and hypoxia-inducible factor (HIF) protein-levels and their rapid post-transcriptional stabilization, e.g., upon DNA damage and hypoxic conditions [6,7]; the regulation of several cellular stress responses (e.g., oxidative stress) at the level of protein translation [8]; and the extensive regulation of cellular stress response programs through protein phosphorylation cascades [9–11].

This review is intended as a practical, high-level overview on the analysis of proteomic data with a special emphasis on systems toxicology applications. It provides a general overview of possible analysis approaches and lessons that can be learned. We start with a background on the experimental aspect of proteomics and introduce common computational analyses approaches. We then present several examples of the application of proteomics for systems toxicology, including lung proteomics results from a subchronic 90-day inhalation toxicity study with mainstream smoke from the reference research cigarette 3R4F. Finally, we provide an outlook and discuss future challenges.

1.1. Experimental and computational approaches for the quantitative analysis of proteomic alterations

1.1.1. Experimental approaches for quantitative proteomics

1.1.1.1. Gel-based liquid chromatography mass spectrometry (LC MS/MS) approaches. Two-dimensional polyacrylamide gel electrophoresis (2DGE) is used to assess perturbations on the proteome based on changes in protein expression (Fig. 1A). The 2DGE workflow relies on the separation of proteins based on their pH (charge) as well as their size and has the capability to separate and visualize up to 2000 proteins in one gel. The first dimension, which is known as isoelectric focusing (IEF) separates the proteins by their isoelectric point (pI), i.e. the pH at which they exhibit a neutral charge. The second dimension further separates the proteins by their mass. State-of-the-art image acquisition and analysis software such as SamSpots (TotalLab) allow the simultaneous comparison of control and treated samples to identify the differentially regulated proteins by their relative intensity in a label-free approach. A variant of 2DGE is difference gel electrophoresis (DIGE) which is based on labeling of proteins with fluorescent cyanine dyes (Cy2, Cy3 and Cy5) of different samples resulting from e.g. different treatments. The characteristics of these dyes allow for the analysis of up to three pools of protein samples simultaneously on a single 2D gel to detect differential variances in proteins between samples [12]. The most challenging aspect of this approach has been the development of algorithms that can address gel distortion (warping). Investigators now account for gel warping by running several gels per sample and analyzing gels by principal component analysis to determine which should be excluded from further analysis [12].

Although 2DGE is a powerful tool to identify many proteins using well-established protocols and detection of posttranslational modifications (PTMs) in proteins, the approach has its limitations. The major limitation is that not all proteins can be separated by IEF, such as membrane, basic, small (<10 kDa) and large (>100 kDa) proteins. Hence, they cannot be detected by 2DGE and require a separate approach based on membrane protein purification protocols and one-dimensional gel electrophoresis. The second limitation is that less abundant proteins are often masked by the abundant proteins in the mixture [13,14].

1.1.1.2. Gel-free liquid chromatography mass spectrometry (LC MS/MS) approaches. Protein fractionation is crucial to simplify mixtures before analysis by mass spectrometry (MS). Liquid chromatography (LC) is the most commonly used method for protein fractionations in this context (Fig. 1A). The LC approach takes advantage of differences in the physiochemical properties of proteins and peptides, i.e., size, charge, and hydrophobicity. 2D-LC can be used to fractionate protein mixtures on two columns with different physiochemical properties and thereby maximize the separation of proteins and peptides in complex mixtures [15].

Mass spectrometry is widely considered to be the central technology platform for toxicoproteomics. MS has brought many advantages to the advancement of toxicoproteomics including unsurpassed sensitivity, improved speed and the ability to produce high throughput datasets. Owing to the high accuracy of MS, peptides in the femtomolar (10^{-15})