Mini 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

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.

© 2014 Titz et al. Published by Elsevier B.V. on behalf of the Research Network of Computational and Structural Biotechnology.

Contents

1. Introduction .............................................................. 74
   1.1. Experimental and computational approaches for the quantitative analysis of proteomic alterations ........................................... 74
       1.1.1. Experimental approaches for quantitative proteomics .............................................................. 74
       1.1.2. Computational approaches for quantitative proteomics .............................................................. 76
   1.2. How to derive biological insights from proteomic data .............................................................. 78
       1.2.1. Deriving insights protein-by-protein .............................................................. 78
       1.2.2. Deriving insights through functional modules .............................................................. 78
       1.2.3. Deriving insights through network analyses .............................................................. 80
       1.2.4. Deriving insights through data integration .............................................................. 81
   1.3. Applying proteomics for systems toxicology .............................................................. 83
       1.3.1. Proteomics in the EU framework 6 project .............................................................. 83
       1.3.2. Proteomics for liver toxicity determinations .............................................................. 83
       1.3.3. Proteomics in pulmonary toxicology: 90-day rat inhalation study to assess the effects of cigarette smoke exposure on the lung proteome .............................................................. 85
       1.3.4. Phosphoproteomics for toxicological assessment .............................................................. 86
   1.4. General discussion and future prospects .............................................................. 86
       1.4.1. The future of systems toxicology .............................................................. 86
       1.4.2. The future of proteomics in systems toxicology .............................................................. 86
2. Conclusions .............................................................. 87
References .............................................................. 87

* Corresponding author.
E-mail address: bjoern.titz@pmi.com (B. Titz).
1 Contributed equally.
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 toxicity 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 toxicity 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 modifications 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 (ToalLab) 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 (DGE) 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)
to attomolar (10^{-18}) range can be detected in tissues and biological matrices with an accuracy level of less than 10 ppm [16]. This is greatly beneficial in comparative analysis where simultaneous comparisons between control and treated samples are a key to increasing understanding of how stimuli affect the proteome and the subsequent identification of potential biomarkers [15].

The two approaches that are widely used for differential protein quantification are label-free and label-based quantitation. In the label-
free approach, proteins or peptides of each sample are separated by LC and subsequently analyzed by MS. The main advantages of this approach are: 1) comparison of multiple samples is possible (no restriction in sample number), 2) it covers a broad dynamic range of concentrations, and 3) no further sample treatment is required. This approach is, however, error-prone and requires long analysis time and large computational power to perform the data analysis. In the label-based approach, samples are modified prior to analysis. One of the most common label-based techniques is the use of isobaric tags with the iTRAQ or TMT method. The main advantages of isobaric-tag based quantification are: 1) simultaneous comparison of large numbers of samples (up to eight for iTRAQ, up to ten for TMT) 2) reduction of required MS runs (reduction of analysis time) as samples are pooled before MS analysis, and 3) low probability of introducing experimental errors during analysis due to pooling. The limitations of the technique are the limited dynamic range and the fact that the protein profiles must be similar [17].

In summary, the major advantages of the gel-free approaches are: 1) lower sample volumes can be analyzed, 2) less abundant proteins can be detected, 3) high-throughput sample analysis and data generation are possible, and 4) different classes of the proteins can be analyzed.

1.1.1.3. Targeted mass spectrometry (LC MS/MS) approaches. Because system biology requires accurate quantification of a specified set of peptides/proteins across multiple samples, targeted approaches have been developed for biomarker quantification (Fig. 1B). Selected reaction monitoring (SRM) was developed to reliably deliver precise quantitative data for defined sets of proteins, across multiple samples using the unique properties of MS. SRM measures peptides produced by the enzymatic digestion of the proteome as surrogates to their corresponding proteins in triple quadrupole MS.

An SRM-based proteomic experiment workflow begins with the selection of a list of target proteins, derived from previous experimental datasets and/or prior knowledge such as a pathway map or literature. This step is followed by: 1) selection of the proteotypic target peptides (at least two) that optimally and uniquely represent the protein target (e.g., using the SRMAtlas [18]), 2) selection of a set of suitable SRM transitions for each target peptide, 3) detection of the selected peptide transitions in a sample, 4) optimization of SRM assay parameters if some of the transitions cannot be detected, and 5) application of the assays to the detection and quantification of the proteins/peptides [19].

The major advantages of the SRM technique are: 1) multiplexing of tens to hundreds of proteins that can be monitored during the same run, 2) absolute and relative quantification is possible, 3) the method is highly reproducible, and 4) the method yields absolute molecular specificity. The limitations of this technique include: 1) only a limited number of measurable proteins can be included in the same run (the system cannot monitor thousands of proteins per run or analysis) and 2) even with its high sensitivity it cannot reach all the proteins present in an organism (limit of detection is at the attomolar level) [20].

A new MS-based targeted approach called parallel reaction monitoring (PRM) has been developed that is centered on the use of next-generation, quadrupole-equipped high-resolution and accurate mass instruments (mainly the Orbitrap MS system) (Fig. 1B). This approach is closely related to SRM, but allows for the measurement of all fragmentation products of a given peptide in parallel. The major advantages over SRM are: 1) the generated data can be easily interpreted, and the analysis can be automated, 2) higher dynamic range, and 3) quantitative information can be determined from datasets of complex samples resulting in extraction of high-quality data [21].

1.1.1.4. Posttranslational modifications. Posttranslational modifications (PTMs) represents an important mechanism for diversifying and regulating the cellular proteome. PTMs are chemical modifications that play a role in functional proteomics, by regulating activity, localization and interactions with other cellular biomolecules. The identification and characterization of protein substrates and their PTM sites are very important to the biochemical understanding of the PTM pathways and to provide deeper insights into the possible regulation of the cellular physiology induced by PTM. Examples of PTMs include phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation, lipidation and proteolysis [22].

During the past decade, MS-based proteomics has demonstrated that it is a powerful technique for the identification and mapping of PTMs that replaces the traditional biochemical techniques such as Western blots, using radioactive isotope-labeled substrates and protein microarrays. The MS-based approaches took great advantage from the advancement in MS instrumentation that allow for higher sensitivity, accuracy and resolution for the detection of less abundant proteins. For the scope of this review, only 2 PTMs will be discussed, which are the most commonly studied in disease research.

1. Analysis of phosphorylation changes
Phosphorylation represents an important posttranslational modification of proteins; in eukaryotes, approximately 30% of cellular proteins contain covalently bound phosphate. It is involved in most cellular events in which the complex interplay between protein kinases and phosphatases strictly controls biological processes such as proliferation, differentiation, and apoptosis. Phosphorylation is a key mode of signal transduction, a central mechanism in the modulation of protein function that is capable of regulating almost all aspects of cell life. Destructive or altered signaling pathways often result in abnormalities leading to various diseases including cancer [23,24], emphasizing the importance of understanding protein phosphorylation. The importance of protein phosphorylation is illustrated by the hundreds of protein kinases and phosphatases present in eukaryotic genomes [25].

2DGE was commonly used for assessing wide-scale changes in phosphorylation. However, because of the many limitations of the approach, MS approaches were developed as an alternative to 2DGE to overcome the limitations and increase the sensitivity of the detection of phospho-proteins. Today, most phospho-proteomic studies are conducted by MS strategies in combination with phospho-specific enrichment (Fig. 1C).

Because of sensitivity issues phospho-peptides need to be separated from non-phosphorylated peptides before analysis. A commonly used phospho-peptide enrichment strategy is using TiO₂, which is highly selective for phospho-peptides. It is extremely tolerant toward most buffers and salts, and thus is a robust method for the enrichment of phospho-peptides. The enriched peptides are then analyzed using MS for identification and phosphorylation site determinations [26].

2. Analysis of ubiquitylation changes
Modification of proteins by ubiquitylation is a reversible regulatory mechanism that is well conserved in eukaryotic organisms. The role of ubiquitylation is extensively studied in the ubiquitin proteasome system (UPS) as well as in cellular process such as DNA damage repair, DNA replication, cell surface receptor endocytosis, and innate immune system [27–29]. The clinical use of the proteasome inhibitor bortezomib, and the ongoing clinical trials of several other inhibitors illustrate the importance of ubiquitylation for human health [30,31].

The experimental procedure is similar to the phospho–proteomics approach (Fig. 1C). The major difference is that for the enrichment step di-Gly-lysine-specific antibodies are used [32]. Direct immunoenrichment of ubiquitylated peptides, together with high resolution LC MS/MS allows for the in-depth analysis of putative ubiquitylation sites.

1.1.2. Computational approaches for quantitative proteomics
Following the acquisition of the mass spectrometry data, the first goal of a quantitative proteomics experiment is to derive a protein expression matrix (proteins vs. samples) and identify differentially expressed proteins between selected sample groups. The path to achieve this goal can be divided into three steps: 1) peptide/protein
identification, 2) peptide/protein quantification, and 3) identification of differentially expressed proteins.

1.1.2.1. Software for processing mass spectrometry data. Several software packages support these tasks including the freely available Trans-Proteomic Pipeline [33], the CPAS system [34], the OpenMS framework [35], and MaxQuant [36] (Table 1). Each of these packages has their advantages and shortcomings, and a detailed discussion goes beyond the scope of this review. For example, MaxQuant is limited to data files from a specific MS manufacturer (raw files, Thermo Scientific), whereas the other software solutions work directly or after conversion with data from all manufacturers. An important consideration is also how well the employed quantification approach is supported by the software (for example, see Nahnsen et al. for label-free quantification software [37] and Leemer et al. for both label-free and label-based quantification tools [38]). Another important consideration is the adaptability of the selected software because processing approaches of proteomic datasets are still rapidly evolving (see examples below). While most of these software packages require the user to rely on the implemented functionality, OpenMS is different. It offers a modular approach that allows for the creation of personal processing workflows and processing modules thanks to its python scripting language interface, and can be integrated with other data processing modules within the KNIME data analysis system [39,40]. In addition, the open-source R statistical environment is very well suited for the creation of custom data processing solutions [41].

1.1.2.2. Identification of peptides and proteins. The first step for the analysis of a proteomic MS dataset is the identification of peptides and proteins. Three general approaches exist: 1) matching of measured to theoretical peptide fragmentation spectra, 2) matching to pre-existing spectral libraries, and 3) de novo peptide sequencing.

The first approach is the most commonly used. For this, a relevant protein database is selected (e.g., all predicted human proteins based on the genome sequence), the proteins are digested in silico using the cleavage specificity of the protease used during the actual sample digestion step (e.g., trypsin), and for each computationally derived peptide, a theoretical MS2 fragmentation spectrum is calculated. Taking the measured (MS1) precursor mass into account, each measured spectrum in the datasets is then compared with the theoretical spectra of the proteome, and the best match is identified. The most commonly used tools for this step include Sequest [42], Mascot [43], X!Tandem [44], and OMSSA [45]. The identified spectrum to peptide matches provided by these tools are associated with scores that reflect the match quality (e.g., a cross-correlation score [46]), which do not necessarily have an absolute meaning. Thus, it is critically important to convert these scores into probability p-values. After multiple testing correction, these probabilities are then used to control for the false discovery rate (FDR) of the identifications (often at the 1% or 5% level). For this statistical assessment, a commonly used approach is to compare the obtained identification scores for the actual analysis with results obtained for a randomized (decoy) protein database [47]. For example, this approach is taken by Percolator [48,49] combined with machine learning to best separate true from false hits based on the scores of the search algorithm. Although the estimation of false-discovery rates is generally well established for peptide identification [50], protein FDR estimates are less mature [51,52] and constantly evolving (e.g., [53,54]). Another question is how the results from different search engines can be effectively combined toward higher sensitivity, while maintaining the specificity of the identifications (e.g., [51,55]).

The second group of algorithms, spectral library matching (e.g., using the SpectralST algorithm), relies on the availability of high-quality spectrum libraries for the biological system of interest [56-58]. Here, the identified spectra are directly matched to the spectra in these libraries, which allows for a high processing speed and improved identification sensitivity, especially for lower-quality spectra [59]. The major limitation of spectral library matching is that it is limited by the spectra in the library.

The third identification approach, de novo sequencing [60], does not use any predefined spectrum library but makes direct use of the MS2 peak pattern to derive partial peptide sequences [61,62]. For example, the PEAKS software was developed around the idea of de novo sequencing [63] and has generated more spectrum matches at the same FDR-cutoff level than the classical Mascot and Sequest algorithms [64]. Eventually an integrated search approaches that combine these three different methods could be beneficial [51].

1.1.2.3. Quantification of mass spectrometry data. Following peptide/protein identification, quantification of the MS data is the next step. As seen above, we can select from several quantification approaches (either label-dependent or label-free), which pose both method-specific and general challenges for computational analysis. Here, we will only highlight some of these challenges. Data analysis of quantitative proteomic data is still rapidly evolving, which is an important fact to keep in mind when using standard processing software or deriving personal processing workflows. An important general consideration is which normalization method to use [65]. For example, Callister et al. and Kultima et al. compared several normalization methods for label-free quantification and identified intensity-dependent linear regression normalization as a generally good option [66,67]. However, the optimal normalization method is dataset specific, and a tool called Normalizer for the rapid evaluation of normalization methods has been published recently [68].

Computational considerations specific to quantification with isobaric tags (iTRAQ, TMT) include the question how to cope with the ratio compression effect and whether to use a common reference mix. The term ratio compression refers to the observation that protein expression ratios measured by isobaric approaches are generally lower than expected. This effect has been explained by the co-isolation of other labeled peptide ions with similar parent mass for the MS2 fragmentation and reporter ion quantification step. Because these co-isolated peptides tend to be not differentially regulated, they generate a common reporter ion background signal that decreases the ratios calculated for any pair of reporter ions. Approaches to cope with this phenomenon computationally include filtering out spectra with a high percentage of co-isolated peptides (e.g., above 30%) [69] or an approach that attempts to directly correct for the measured co-isolation percentage [70]. The inclusion of a common reference sample is a standard procedure for isobaric-tag quantification. The central idea is to express all measured values as ratios to the common reference sample to cancel out differences in ionization efficiencies and between sample runs. However, recently it has been demonstrated that this reliance on a single sample can increase the overall variance and that alternatively, it is beneficial to use the median of all measured reporter ions for spectrum normalization [71]. Importantly, when applying this approach to diverse sample sets (e.g., human patient samples) the comparability of these median values need to be ensured. Similarly, other quantification methods come with their own challenges, e.g., label-free approaches based on peak integration are dependent on a reliable run-to-run alignment and consistent integrations (e.g., [72,73]).

1.1.2.4. Identification of differentially expressed proteins. The results of these efforts are a protein-by-sample expression matrix, and the next analysis step often aims to identify differentially expressed proteins. Here, important considerations involve the selection of the protein-level statistics for differential abundance and how multiple hypothesis testing is taken into account. For example, Ting et al. tested a fold change approach, Student’s t-test, and empirical Bayes moderated t-test as the protein-level statistics [74]. The authors also used the common approach in RNA microarray experiments to construct linear models that captured the relevant experimental factors. They concluded that applying the empirical Bayes moderated t-test in the linear model framework resulted in a high-quality list of statistically significant differentially abundant proteins. A summary of the essential multiple hypothesis correction methods to control the FDR is given in [75]. Of these, the most commonly used method is likely the Benjamini–Hochberg approach [76].
1.1.2.5. Comparison of methods. As we have observed, many software and processing options are available for the analysis of MS data. As argued by Yates et al., it is vital to define benchmarking standards and more extensively compare the available tools [77] to allow for an evidence-based selection of the available software tools. A few comparative studies for quantitative proteomics are already available. For example, Altemar et al. compared SILAC, dimethyl and (isobaric tag) TMT labeling strategies and found that all methods reach a similar analysis depth; TMT resulted in the highest ratio of quantified-to-identified proteins and the highest measurement precision, but ratios were most affected by ratio compression [78]. Similarly, Li et al. compared label-free (spectral counting), metabolic labeling ($^{15}N/^{15}N$), and isobaric tag labeling (TMT and iTRAQ) and found the isobaric tag-based approaches to be the most precise and reproducible [79].

1.1.2.6. Computational resources for data processing. All steps of proteomics computational analysis, including protein identification, protein quantification and identification of differentially expressed proteins, require an access to high performance computational resources [80]. Software tools that match peptide masses to genome-based protein databases or spectra to spectral libraries directly can often be run in a parallelized mode to accelerate the data analysis. Classical parallelization solutions such as computing clusters are widely used and more cutting edge implementations such as cloud computing [81] or graphics processing unit (GPU) servers [82] are on the rise. The latter work demonstrated acceleration of the peptide searches within proteome database up to 60-fold compared to conventional CPU-based architecture and reflects a recent trend of using CPU-based clusters in computational systems biology.

After generation of a reliable quantitative proteomic dataset, the main challenge is to turn the data into biological knowledge. In the next section, we focus on four categories of computational approaches (protein-by-protein, functional module-based, biological network-based, and through data integration), which taken together support a comprehensive biological interpretation of the results (Fig. 2).

1.2. How to derive biological insights from proteomic data

1.2.1. Deriving insights protein-by-protein

In many cases, the first result obtained when analyzing a quantitative proteomics dataset is a list of differentially expressed proteins in the condition of interest. Initially, these proteins are often only sparsely annotated, and expansion of this annotation is a helpful first step for biological interpretation and filtering. Protein annotations can be directly derived from databases (e.g., UniProtKB) or dynamically generated for a specific biological question through text-mining approaches.

1.2.1.1. Protein databases. The UniProt Knowledgebase (UniProtKB) is the central resource for protein-centric information [83]. It consists of a high-quality, manually reviewed section (UniProtKB/ Swiss-Prot) and an automatically generated, unreviewed section (UniProtKB/TrEMBL). The available data include protein functions, catalytic activity, pathway information, and associated phenotypes and diseases. UniProt facilitates the annotation of protein lists through its own ID mapping service, batch retrieval tools, and by supporting more extensive and automated queries through BioMart [84]. For human proteins, UniProt is extended by the netXPro knowledgebase, which is still under development [85], which provides an extended view of the proteins by incorporating additional data sources such as high-throughput protein expression and protein localization experiments. Although these databases offer extensive coverage of overall protein function, the functional information for specific protein modifications is sparse. Thus, more dedicated databases are advantageous when analyzing proteomic datasets of posttranslational modifications such as phosphorylation. For example, the PhosphoSite database offers extensive annotations of phosphorylation sites for human, mouse, and rat [86], and NetPhorest allows for predictions of potential upstream kinases [87]. In addition, for toxicological assessments it can be revealing to investigate the links between the identified proteins and chemicals and chemical toxins. The STITCH database is an extensive database of protein–chemical interactions and provides convenient data access through downloadable files and an application programming interface [88]. The toxin and toxin target database (T3DB) is specifically focused on mechanisms of toxicity and targets and currently contains information for approximately 3000 toxins [89].

1.2.1.2. Text-mining approaches. The annotations derived from these sources depend on the specific scope and curation depth of these databases. To associate the identified protein list with the most up-to-date information and with specific biology/disease concepts (e.g., the disease under investigation), text-mining approaches are worth considering [90, 91]. The iHop database provides a precomputed network of protein–protein and protein–protein and protein–concept interactions, and is derived through automated text mining of the scientific literature [92]. Its main strength is to comprehensively and concisely collect up-to-date information on a given protein. However, it does not allow querying specific concepts and biological contexts. This is supported by specific text-mining tools such as EBImed [93], SciMiner [94], and PolySearch [95]. EBImed and SciMiner accept free literature queries as the input and automatically identify and associate the proteins, functions, and drugs reported in the identified literature. PolySearch more specifically handles associative queries such as “given a disease/protein/drug, find all associated diseases/proteins/drugs.” For example, PolySearch was used to support the identification and annotation of toxin-target relationships in the T3DB [89]. All three tools are especially useful when evaluating the discovered differentially regulated proteins in the context of what is already known about the process under study.

1.2.2. Deriving insights through functional modules

Although these protein-level annotations support the manual systematic interpretation of a dataset, they do not allow for a direct statistical assessment of the affected biological functions. Given a list of differentially expressed proteins, we often ask, “What are the functional categories/modules/pathways that are significantly enriched for differentially expressed proteins?” The basis for these analyses is the modular organization and regulation of biological systems [96–98]. For example, upon a certain cellular stress such as oxidative stress, we can expect to observe the coordinated up-regulation of a specific stress response protein module or the activation of a particular signaling pathway [99]. Three main components are needed to identify functional modules that are significantly enriched for affected proteins: 1) a metric to score the level of perturbation for each protein (protein-level statistic), 2) a database of relevant protein modules/sets, and 3) an algorithm to score and evaluate the statistical significance of the module enrichment (module-level statistic).

1.2.2.1. Protein-level statistics. Threshold-based approaches can, for example, be based on a multiple testing corrected t-test p-value (see above). Threshold-free approaches rely on continuous protein-level statistics such as the fold-change, a t-score, or the signal-to-noise ratio. An important consideration for the selection of the protein-level statistics is, whether up- and down-regulation of module components are considered together or as distinct effects [100].

1.2.2.2. Gene/protein set databases. Several functional module (gene/protein set) databases are available. The most prominent is the nSigDB database from the Broad Institute [101]. Here, the functional sets are grouped in different categories that range from canonical pathways (e.g., the KEGG and Reactome database [102, 103]) to gene ontologies [104]. Other functional module databases include GeneSigDB [105], which contains manually extracted signatures/modules from the literature, and PAGED [106], which combines these and other functional module databases. However, depending on the analyzed biological context,
more specialized gene set databases—such as the liver-cancer related database, Liverome [107], or even self-defined databases—can be beneficial.

1.2.2.3. Three classes of enrichment algorithms. Finally, algorithms to evaluate module enrichment are needed. These can be grouped into three categories: 1) over-representation analysis (ORA) approaches, 2) functional class scoring (FCS) approaches, and 3) pathway topology (PT) approaches [108].

ORA approaches depend on a threshold to select a list of differentially expressed proteins for the conditions of interest. These can be grouped into three categories: 1) over-representation analysis (ORA) approaches, 2) functional class scoring (FCS) approaches, and 3) pathway topology (PT) approaches [108].

ORA approaches rely on a fixed threshold to select a list of differentially expressed proteins for the conditions of interest. Subsequently, the overlap between this protein list and each functional module in the database is calculated and statistically assessed (e.g., using the Fisher exact test and multiple hypothesis correction). The advantages of ORA are simplicity, relatively quick run times, and availability (e.g., through the DAVID Bioinformatics Resources [109] or Enricher tool [110]). Because these methods rely on a fixed threshold, they disregard differences in the extent of differential regulation and do not consider weakly, but consistently regulated genes.

The second class of algorithms are FCS approaches. The most prominent of these methods is the traditional and still commonly used gene set enrichment analysis (GSEA) [111]. Here, the proteins are ranked based on a continuous protein-level metric (such as fold-change or SNR), and the enrichment of the functional modules in the database at the top or the bottom of the ranked list is statistically evaluated. Beyond the classical assessment of enrichment by the GSEA algorithm, several alternative module-level statistics have been employed (e.g., Kolmogorov–Smirnov statistic, sum, mean, or median, and the maxmean statistic) [108]. The advantages of FCS methodologies are that they do not rely on fixed thresholds and the correlation structure (between genes) can be taken into account by the employed permutation-based significance tests, depending on the null hypothesis under consideration.

The third type of approach, PT, goes beyond the FCS approach by taking the actual topology of the pathways/modules into account. For example, the signaling pathway impact analysis (SPIA) combines two types of evidence to assess the perturbation of a signaling pathway: a classical overrepresentation measure and a topology-dependent measure of the abnormal perturbation of the pathway, which takes the actual wiring of the pathway into account. A second PT algorithm, the network perturbation amplitude (NPA) approach scores the activation of a given causal biological network model [112]. Here, the employed causal network models consist of two tiers. The upper tier (backbone) is similar to a classical pathway diagram causally describing a specific biological process (e.g., NF-kB activation). The nodes in this network are causally linked to downstream gene expression nodes in the lower tier. These links describe the causal effect (positive/negative) of a given backbone node on the gene expression of a lower tier node. For the actual calculation of NPA scores, the observed gene expression changes are mapped.
onto the lower tier. The topology of the backbone network is then considered together with the direction of the links between the tiers to quantitatively summarize whether the expression changes of the downstream nodes are overall consistent with pathway activation, inactivation, or no change. For example, this approach could closely recapitulate the experimentally measured activation of NF-κB in TNFα-stimulated human bronchial epithelial cells. The increased availability of high-quality, context-specific network models will broaden the applicability of these PT approaches in the future [108].

1.2.2.4. Application to proteomic data. Although all these methods have been initially devised for gene expression data, proteomic datasets can be analyzed with the same algorithms and functional gene/protein module databases. For example, Gharib et al. used GO-term enrichment to compare the BALF proteomes of human and mice and identified shared (e.g., protease inhibitor activity) and distinct (e.g., antioxidant activity) functional processes [113]; Chin et al. used GSEA to identify proteome alteration in mouse models of Parkinson’s disease [114].

1.2.2.5. Interpretation of results. The main challenge for the interpretation of results from these methods is often the high overlap (and partial redundancy) of the protein modules. Enrichment maps have been presented as a solution to better visualize enriched gene sets/modules and their relationships (Fig. 2, right panel) [115]. Here, the enrichment results are presented as a similarity network, where the nodes represent gene sets, the links the overlap between sets, and the node color encodes the enrichment strength. The usefulness of this approach for the analysis of a proteomic dataset is highlighted by a recent study on dilated cardiomyopathy by Isserlin et al. [116]. In this study, the enrichment maps supported the interpretation of proteomic changes at an early- and mid-stage time point of the cardiomyopathy model and identified a strong up-regulation of apoptosis, proteasome and RNA processing/splicing apparatus at the mid-stage time point. In another example, Meierhofer et al. demonstrated the power of protein set analysis to gain insights into the regulation of cell and tissue homeostasis during high-fat diet feeding and medication with two anti-diabetic compounds [117], GSEA allowed for more sensitive detection of low-level but coordinated protein expression changes, and the functional modules showed a higher correlation than individual genes/proteins when comparing proteomics and transcriptomics data.

1.2.3. Deriving insights through network analyses

As discussed in the previous section, the functional categorization of genes/proteins into functional classes is an effective approach to systematically and functionally understand effects in biological systems [118]. An even more holistic viewpoint is taken by network biology approaches [119]. Here, the biological entities (e.g., transcripts, proteins) are viewed as the nodes of complex, interconnected networks. The links between these nodes can represent actual physical associations (e.g., protein–protein interactions) or functional interactions (e.g., proteins involved in the same biological process). For example, network biology approaches can highlight highly perturbed protein subnetworks that warrant further investigation [120]; they help to understand the modular organization of the cell [119], and can be applied for improved diagnostics and therapies [121,122].

1.2.3.1. Biological network models. Comprehensive and high-quality biological network models are the basis for these analyses. The available resources for network models differ in their scope, quality, and availability. The STRING database is one of the most comprehensive, freely available databases for functional protein–protein links for a broad range of species [123]. It is based on a probabilistic model that scores each link based on its experimental or predicted support from diverse sources such as physical protein interaction databases, text mining, and genomic associations. The Reactome database is a manually curated database with a narrower scope of human canonical pathways [124]. Recently, however, Reactome data have been supplemented with predicted functional protein associations from several sources including protein–protein interaction databases and co-expression data (Reactome Functional Interaction network) [125]. Several commercial curated network databases exist including KEGG, the Ingenuity® Knowledge Base and MetaCore®. At its core, the KEGG database provides metabolic pathway maps but more recently has added pathways of other biological processes (e.g., signaling pathways) [126]. The Ingenuity® Knowledge Base and MetaCore® are comprehensive resources for expert curated functional links from the literature, and are also often employed for the analysis of proteomic datasets [127–129]. These databases are well suited for generic network analyses. However, currently, their coverage of relevant mechanisms is often insufficient for tissue- and biological context-specific modeling approaches. For this, specific mechanistic network models curated by experts of the specific field of study are required. Very detailed NKB models are examples that recapitulate complex signaling and drug treatment responses [130]. For systems toxicology applications, we have developed and published a collection of mechanistic network models [131]. These models range from xenobiotic, to oxidative stress, to inflammation-related, and to cell cycle models [132–135]. The networks are described in the Biological Expression Language (BEL), which enables the development of computable network models based on cause and effect relationships [136]. Ensuring high-quality and independent validation of these network models is especially important when these models are used within a systems toxicology assessment framework. An effective approach that has been used for these networks for systems toxicology makes use of the wisdom of the crowd [137–139]. Here, within the sbv IMPROVER validation process, the derived networks are presented to the crowd on a web platform (bionet.sbvimprover.com), and classical incentives and gamification principles are used to motivate the participants to challenge and improve the presented network models. The results of this challenge are further discussed in a jamboree session with select participants, and finally the improved network models are disseminated for public use.

1.2.3.2. The Cytoscape platform. Although the approaches for possible network analyses can be overwhelming, they are facilitated by the availability of the common network analysis software platform Cytoscape [140]. At its core, Cytoscape allows for import, annotation, visualization, and basic analysis of molecular interaction networks. However, its functionality is expanded by many plugins/apps for extended data visualization, handling, and analysis capabilities. Saito et al. provided a “travel guide” to Cytoscape plugins [141]. Common analysis workflows involve the identification of functional protein networks for differentially expressed proteins or the identification of especially strongly perturbed regions (modules) in biological networks.

1.2.3.3. Functional context networks. To understand the biology altered under a specific condition, it is often helpful to visualize and analyze how the differentially expressed proteins are functionally connected and whether they form specific functional clusters. Several Cytoscape® apps support this integration of protein lists with different functional network resources. The Reactome Functional Interaction (FI) app (Reactome Fls) allows construction of a subnetwork of the extensive Reactome Fl network for a given set of genes/proteins [125]. For example, Chen et al. used Reactome Fl for data interpretation of an integrative personal omics profile from a single individual over a 14-month period [142]. The Agilent Literature Search plugin generates biological networks for protein lists based on queries of the scientific literature [143]. Outside the Cytoscape environment, the String database directly allows network generation of protein sets and provides basic analysis tools of the generated networks including clustering and functional enrichment analyses. Similar functional network analyses are also a central component of the commercial Ingenuity and MetaCore analysis tools [127–129]. For example, Chang et al. used Ingenuity Pathway
1.2.4.3. Integration approaches. Transcriptomics and proteomics data can be integrated at different biological levels ranging from individual molecules (proteins/transcripts) to functional modules to biological networks (Fig. 3, right panel). The most basic, but reasonable first data integration approach is the intersection of differentially expressed gene/protein lists (e.g., [164,165]). However, p-values of differential expression can also be integrated directly. The advantage of this approach is that consistently, but weakly regulated molecules across datasets can be identified. Traditional approaches for the integration of p-values (or z-scores) are Fisher’s [166] and Stouffer’s [167] method and Brown’s method for the integration of dependent p-values [168]. Integration at the functional level (pathway-, gene set-) level can reveal similarities that are not apparent at the level of individual biomolecules. For example, Buschow et al. conducted an integrated transcriptomics/proteomics analysis to study the activation of dendritic cells (DC). While only limited correlation at the level of individual transcripts and proteins was observed, the correlation for relevant DC activation pathways was significantly higher [165]. Kaever et al. presented a general framework for the meta analysis of p-values obtained from pathway enrichment analyses for transcriptomics and metabolomics data [169]. Moreover, data can be integrated at the biological network level. For example, Nibbe et al. used a proteomics-first approach to integrate proteomics and transcriptomics data for the identification of functional subnetworks in colorectal cancer (CRC) [170]. First, candidate subnetworks were identified by mapping differentially expressed proteins in CRC samples onto a human protein interaction network. Subsequently, independent transcriptomics datasets were employed to identify those (proteomics–defined) subnetworks that were significantly associated with the tumor/normal classification. Finally, different integration approaches can be combined. For example, Balbin et al. generated transcriptomics, proteomics, and phosphoproteomics profiles for KRAS-mutated non-small cell lung cancer cell lines [171]. These three data sources were first integrated using an approach based on Stouffer’s method. Subsequently, a network-based approach was used to identify a novel targetable subnetwork in KRAS-dependent CRC cell lines centered on LCK kinase.
D node colour
3R4F vs. fresh air

inflammatory response
unfolded protein response
translation
xenobiotic metabolism
response to oxidative stress
fatty acid oxidation

GSEA: perturbed gene/protein sets
Oxidative phosphorylation
Respiratory electron transport
Tumour field effect
Mast cell
Acute lung injury
Neutrophil granule constituents
Fatty acid metabolism
Myeloid cell development

Sod2: Superoxide dismutase 2
Ctsd: Cathepsin D
Nqo1: NAD(P)H dehydrogenase 1
Chia: Acidic mammalian chitinase

proteomics
lung tissue
exposure
recovery

up-regulated by 3R4F, high

functional protein clusters
metabolism
response to oxidative stress
xenobiotic metabolism
unfolded response
fatty acid oxidation
oxidative phosphorylation

protein
mRNA

% mRNA up-regulated (q-value < 0.05)
1.3. Applying proteomics for systems toxicology

1.3.1. Proteomics in the EU framework 6 project

Suter et al. characterized the effect of 16 test compounds using conventional toxicological parameters in the integrated EU Framework 6 Project: Predictive Toxicology (PredTox) [172]. The project, partly funded by the EU, was performed by a consortium of 15 pharmaceutical companies, 2 private companies, and 3 universities. The three major observed toxicities, liver hypertrophy, bile duct necrosis and/or cholestasis, and kidney proximal tubular damage, were analyzed in detail. The outcome of this program indicates that omics technologies can help toxicologists make better-informed decisions during exploratory toxicological studies.

The compounds included 14 proprietary drug candidates from participating companies and 2 reference toxic compounds: gentamicin and troglitazone. Following 2-week exposure in rats, conventional toxicological endpoints were collected, and transcriptomics, proteomics, and metabolomics profiles were evaluated.

Most of the proteomics data that were supportive of the mechanistic hypothesis were based on tissue proteomics (2D-DIGE). This technology, despite offering relatively low throughput, provided valuable information that was complementary to other data sources. The authors concluded that the main changes that led to a mechanistic interpretation of the findings were similarly interpreted using different statistical and pathway mapping tools, highlighting the robustness of the outcome. With this, the authors strongly supported the use of these data for confirmation of mechanistic hypotheses and discovery of putative biomarkers as very tangible outcomes of integrated omics analysis.

1.3.2. Proteomics for liver toxicity determinations

Drug-induced liver injury is a leading cause of acute liver failure, thus constituting a major reason for drug candidate failure during development or withdrawal from the market. Because of drug-related toxicity, many drug candidates that may otherwise be potentially efficacious in the treatment of diseases have been discontinued; this represents a major setback to a larger population, which may benefit from further development of these drug candidates. In addition, from the pharmaceutical industry’s perspective, the resultant regulatory actions have increased development costs to meet acceptable safety requirements. Troglitazone, a once-marketed first-generation thiazolidinedione used for the treatment of type-II diabetes mellitus, was withdrawn from the market owing to unacceptable idiosyncratic hepatotoxicity risks even though troglitazone did not cause hepatotoxicity in normal healthy rodents and monkeys in preclinical drug safety assessments and long-term studies.

To understand idiosyncratic hepatotoxicity mechanistically, Lee et al. used MS-based proteomics to characterize mitochondrial protein changes to track the involvement of specific mitochondrial proteins in troglitazone-induced hepatotoxicity in a mouse model [173].

By combining high-throughput MS-based mitoproteome-wide profiling, biochemical endpoints, and network biology, the authors demonstrated that the hepatic mitochondrial proteome followed a two-phase response to repeated troglitazone administration that culminated in liver injuries by the fourth week. This integrative approach identified the combined deterioration of key fragile nodes and a dysfunctional mitochondrial GSH transport system that lead to the eventual toxicity of troglitazone. They concluded that this approach might represent a powerful step forward in using a systems toxicology approach to advance the understanding of the risk factors of idiosyncratic toxic drugs.

Overall, as discussed by Van Summen et al., many studies within the last 5 years have successfully employed proteomic approaches to identify mechanisms and biomarkers of drug-induced hepatotoxicity [174] (see Table. 2). These studies performed proteomic analysis on different subsets of proteins such as whole tissue; cellular fractions, such as mitochondria, endoplasmic reticulum, microsomes, and serum/plasma; and also employed in vitro systems for proteomic analysis. Van Summen et al. are generally optimistic that proteomic analysis will aid in the description of toxicity mechanisms. Proteomics investigations revealed promising results upon the classification of hepatotoxic compounds and showed opportunities for the identification of protein biomarkers underlying this classification. However, the detection of idiosyncratic hepatotoxins with the currently available in vitro methods will remain challenging because these reactions are unpredictable and mostly immune mediated. For non-idiosyncratic hepatotoxins, proteomics can be used to gain insight into the mechanistic processes underlying drug-induced hepatotoxicity. Despite these promising results with a toxicoproteomics approach, the development of a panel of biomarkers will require the testing of several well-characterized model hepatotoxins. The authors state that by testing classified compounds, common patterns of toxicity can be distinguished from compound-specific mechanisms. Importantly, in their opinion, the value of proteome data can be increased by comparison with data from complementary transcriptomics and metabolomics experiments using a systems biology approach.

1.3.3. Proteomics in pulmonary toxicity: 90-day rat inhalation study to assess the effects of cigarette smoke exposure on the lung proteome

Proteomic analyses are an important component of our overall systems toxicology framework for the assessment of smoke exposure effects. Within our comprehensive assessment framework, both proteomics and transcriptomics analyses complement the more traditional toxicological parameters such as gross pathology and pulmonary histopathology as required by the OECD test guideline 413 (OECD TG 413) for a 90-day subchronic inhalation toxicity study. These systems-level measurements constitute the “OECD plus” part of the study [175] and provide the basis for deeper insights into toxicological mechanisms, which enable the identification of causal links between exposure and observed toxic effects as well as the translation between different test systems and species (see Introduction).

Here, we report on the high-level results for the proteomic component of a 90-day rat smoke inhalation study. Sprague Dawley rats were exposed to fresh air or two concentrations of a reference cigarette (3RF4) aerosol [8 μg/L (low) and 23 μg/L (high) nicotine] for 90 days (5 days per week, 6 h per day) (Fig. 3A). This exposure period was followed by a 42-day recovery period with fresh air exposure. Lung tissue was collected and analyzed by quantitative MS using a multiplexed iTRAQ approach (6 animals per group). At the level of individual differentially expressed proteins, the 90-day cigarette exposure clearly induced major alterations in the rat lung proteome compared with fresh air exposure (Fig. 3B). These alterations were significantly attenuated after the 42-day recovery period. The high 3RF4 dose showed an overall higher impact and remaining perturbations after the recovery period than the

---

Fig. 3. Impact of cigarette smoke exposure on the rat lung proteome. (A) Summary of rat exposure study. (B) Tobacco smoke exposure showed strong overall impact on the lung proteome. Heatmap shows significantly altered proteins (FDR-adjusted p-value ≤ 0.05) in at least one cigarette smoke exposure condition. Each row represents a protein, each column a sample (six biological replicates), and the log2 fold-change expression values compared with sham (fresh air) exposure is color-coded. (C) Gene set enrichment analysis (GSEA) shows a significant alteration in protein clusters and exposure conditions. The bar plot indicates the percent of the genes that show consistent, statistically significant up-regulation of the mRNA transcript upon 90-day smoke exposure (q-value < 0.05). Note that—while overall consistent—the “translation” and “unfolded protein response” clusters show less mRNA up-regulation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
<table>
<thead>
<tr>
<th>Tool</th>
<th>Comment</th>
<th>References/links</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MS raw data processing</strong></td>
<td>Flexible workflows for MS raw data processing</td>
<td>tools.proteomecenter.org/software.php [33]</td>
</tr>
<tr>
<td>Trans-Proteomic Pipeline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPAS</td>
<td>MS raw data processing</td>
<td><a href="http://www.labkey.org">www.labkey.org</a> [34]</td>
</tr>
<tr>
<td>OpenMS</td>
<td>Flexible workflows for MS raw data processing</td>
<td><a href="http://www.openms.de">www.openms.de</a> [35]</td>
</tr>
<tr>
<td>MaxQuant</td>
<td>Integrated package for quantitative proteomics analysis</td>
<td><a href="http://www.maxquant.org">www.maxquant.org</a> [36]</td>
</tr>
<tr>
<td>Sequest</td>
<td>Spectra to peptide matching</td>
<td>[42]</td>
</tr>
<tr>
<td>Mascot</td>
<td>Spectra to peptide matching</td>
<td>[43]</td>
</tr>
<tr>
<td>XTandem</td>
<td>Spectra to peptide matching</td>
<td>[44]</td>
</tr>
<tr>
<td>OMSSA</td>
<td>Spectra to peptide matching</td>
<td>[45]</td>
</tr>
<tr>
<td>Normalizer</td>
<td>Evaluation of data normalization procedures</td>
<td>quantitativeproteomics.org/normalizer [68]</td>
</tr>
<tr>
<td><strong>Protein-by-protein</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UniProt KB</td>
<td>Comprehensive protein database</td>
<td><a href="http://www.uniprot.org">www.uniprot.org</a> [83]</td>
</tr>
<tr>
<td>BioMart</td>
<td>Open source database system for unified access to biological data</td>
<td><a href="http://www.biomart.org">www.biomart.org</a> [84]</td>
</tr>
<tr>
<td>neXtProt</td>
<td>Database for human proteins</td>
<td><a href="http://www.nextprot.org">www.nextprot.org</a> [85]</td>
</tr>
<tr>
<td>PhosphoSite</td>
<td>Comprehensive phospho-protein database</td>
<td><a href="http://www.phosphosite.org">www.phosphosite.org</a> [86]</td>
</tr>
<tr>
<td>NetPhorest</td>
<td>Database of phosphorylation-specific sequence-based classifiers</td>
<td>netphorest.info [87]</td>
</tr>
<tr>
<td>STITCH</td>
<td>Database of chemical–protein interactions</td>
<td>stitch.embl.de [88]</td>
</tr>
<tr>
<td>T3DB</td>
<td>Database of toxins and toxin-target links</td>
<td><a href="http://www.t3db.org">www.t3db.org</a> [89]</td>
</tr>
<tr>
<td>iHOP</td>
<td>Database of text-mined protein–protein and protein–concept links</td>
<td><a href="http://www.ihop-net.org">www.ihop-net.org</a> [92]</td>
</tr>
<tr>
<td>EBIMed</td>
<td>Text-mining tool</td>
<td><a href="http://www.ebi.ac.uk/Rebholz-srv/ebimed/">www.ebi.ac.uk/Rebholz-srv/ebimed/</a> [93]</td>
</tr>
<tr>
<td>SciMiner</td>
<td>Text-mining tool</td>
<td>jdrf.neurology.med.umich.edu/SciMiner [94]</td>
</tr>
<tr>
<td>PolySearch</td>
<td>Text-mining tool</td>
<td>wishart.biology.ualberta.ca/polysearch/ [95]</td>
</tr>
<tr>
<td><strong>Functional modules</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAVID</td>
<td>Comprehensive functional classification resource (ORA method)</td>
<td>david.abcc.ncifcrf.gov [109]</td>
</tr>
<tr>
<td>Enricher</td>
<td>Comprehensive functional classification resource (ORA method)</td>
<td>amp.pharm.mssm.edu/Enrichr [110]</td>
</tr>
<tr>
<td>TOPPGene</td>
<td>Comprehensive functional classification resource (ORA method)</td>
<td>toppgene.cchmc.org [209]</td>
</tr>
<tr>
<td>GSEA</td>
<td>Classical FCS module enrichment method</td>
<td><a href="http://www.broadinstitute.org/gsea">www.broadinstitute.org/gsea</a> [111]</td>
</tr>
<tr>
<td>SPIA</td>
<td>Topology-based pathway enrichment method</td>
<td><a href="http://www.sysbio.se/piano">www.sysbio.se/piano</a> [210]</td>
</tr>
<tr>
<td>Piano</td>
<td>Module enrichment package for the R environment</td>
<td><a href="http://www.sysbio.se/piano">www.sysbio.se/piano</a> [100]</td>
</tr>
<tr>
<td>mSigDB</td>
<td>Comprehensive gene set database</td>
<td><a href="http://www.broadinstitute.org/gsea">www.broadinstitute.org/gsea</a> [101]</td>
</tr>
<tr>
<td>GeneSigDB</td>
<td>Database of gene sets manually curated from the literature</td>
<td><a href="http://www.genesigdb.org">www.genesigdb.org</a> [105]</td>
</tr>
<tr>
<td>PAGED</td>
<td>Integrated gene set database</td>
<td>bioinformatics.iupui.edu/PAGED [106]</td>
</tr>
<tr>
<td><strong>Network analyses</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>String DB</td>
<td>Database of confidence scored functional protein interactions</td>
<td>string.embl.de [123]</td>
</tr>
<tr>
<td>KEGG DB</td>
<td>Pathway database</td>
<td><a href="http://www.genome.jp/kegg">www.genome.jp/kegg</a> [102]</td>
</tr>
<tr>
<td>Ingenuity Pathway Analysis</td>
<td>Commercial knowledgebase and functional analysis system</td>
<td><a href="http://www.ingenuity.com">www.ingenuity.com</a></td>
</tr>
<tr>
<td>Metacore</td>
<td>Commercial knowledgebase and functional analysis system</td>
<td>thomsonreuters.com/metacore</td>
</tr>
<tr>
<td>Reactome FI</td>
<td>Extended Reactome functional interaction database (Cytoscape plugin available)</td>
<td><a href="http://www.reactome.org">www.reactome.org</a> [103]</td>
</tr>
<tr>
<td>Agilent Literature Search tool</td>
<td>CYTOSCAPE plugin for text-mining analysis</td>
<td><a href="http://www.agilent.com/labs/research/litsearch.html">www.agilent.com/labs/research/litsearch.html</a></td>
</tr>
<tr>
<td>jActiveModule</td>
<td>CYTOSCAPE plugin for the identification of network modules</td>
<td>apps.cytoscape.org/apps/jactivemodules [120]</td>
</tr>
<tr>
<td><strong>Data integration</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pride</td>
<td>Repository for MS data</td>
<td><a href="http://www.ebi.ac.uk/pride">www.ebi.ac.uk/pride</a> [157]</td>
</tr>
<tr>
<td>MOPED</td>
<td>Repository for MS data</td>
<td><a href="http://www.kolkerlab.org/projects/statistics-bioinformatics/moped">www.kolkerlab.org/projects/statistics-bioinformatics/moped</a> [158]</td>
</tr>
<tr>
<td>POINTILLIST</td>
<td>Integration of p-values</td>
<td>magnet.systemsbiology.net/software/Pointillist/ [211]</td>
</tr>
</tbody>
</table>
low 3R4F dose. GSEA can support the identification of affected biological functions and potentially allows for the more sensitive detection of low, but concerted alterations. Applying GSEA to this dataset overall recapitulated the results observed for individual differentially expressed proteins: an increase in the number of significantly affected gene sets from low to high 3R4F exposure and an overall decrease after the 42-day recovery period. Visually the recovery for the high 3R4F exposure appears more pronounced on the protein- than gene-set level. However, one functional class cluster remained significantly affected even after the 42-day recovery period. Gene sets in this cluster include neutrophil granule constituents and up-regulated genes in acute lung injury, which indicate a less reversible perturbation of these processes. Two functional clusters clearly emerge (Fig. 3D). These include the activation of inflammation- and metabolism-related processes. The detailed interpretation of affected biological processes is challenging owing to the large number of affected, overlapping gene sets that are not necessarily specific to the process under investigation.

As discussed above, methods such as enrichment maps have been developed [115] that facilitate the interpretation of complex GSEA result sets. Here, we complement GSEA with a functional network approach, which supports the identification and interpretation of perturbed functional modules (Fig. 3D). The main idea is to reduce the complexity of data interpretation by first linking the selected proteins by their functional protein interactions and then identifying and functionally interpreting the emerging functional clusters. Specifically, we make use of the STRING database, which is a comprehensive resource of confidence-scored functional protein interactions based on a range of evidence including pathway databases, text-mining, and co-expression (see above) [123]. From the functional interaction network derived for the proteins significantly up-regulated upon 90-day high 3R4F exposure, several functional clusters clearly emerge (Fig. 3D). These include the expected up-regulation of xenobiotic metabolism and oxidative stress response proteins and of proteins associated with an inflammatory response [135,132]. Another component of the stress response is the up-regulation of proteins related to the unfolded protein response (UPR). This response has been previously reported and is thought to reflect a compensatory mechanism to cope with the adverse impact of oxidative stress on protein folding in the endoplasmatic reticulum [176, 177]. Finally, several metabolism clusters are up-regulated including oxidative phosphorylation and fatty acid oxidation, which is in line with the GSEA results. This likely reflects the major metabolic alterations that are triggered in response to smoke exposure, e.g., to cope with the altered oxidative balance. For example, Agarwal has recently investigated metabolic changes in mouse lungs upon short-term cigarette smoke exposure and found up-regulation of oxidative phosphorylation [178]. Here, the authors suggested that this is part of an overall metabolic switch, which involves down-regulation of glycolysis, up-regulation of the pentose-phosphate pathway for increased NADPH generation, and a compensatory increase in the mitochondrial energy-transducing capacity. Interestingly, in this context the observed up-regulation of fatty acid oxidation could play a similar role.

Finally, we compared the differential expression response of the proteins in the identified clusters and their corresponding mRNA transcripts (Fig. 3E). Overall, these functional clusters demonstrate consistent up-regulation of the mRNA transcripts. While this is generally in line with the remark by Lefebvre et al. that in equilibrium the proteome generally reflects the transcriptome [179], clear differences between mRNA and protein expression exist. For example, we observe differences in the regulation of the functional clusters: whereas protein up-regulation of the xenobiotic cluster is well reflected on the mRNA level, no significant mRNA up-regulation is detected for the translation and unfolded protein response cluster. For example, among the proteins in the latter clusters are TPT1 (Tumor Protein, Translocation-Controlled 1) and Grp78 (Hspa5) – two proteins known to be posttranscriptionally regulated [180,181].

In summary, we have conducted a 90-day rat smoke exposure study including a 42-day recovery period. Although the quantitative proteomic analysis of lung tissue is only one component of our comprehensive assessment strategy within an overarching systems toxicology framework, it already provides an extensive view of the biological impact of cigarette smoke exposure. Globally, the impact of cigarette smoke on the protein and gene set level and the extent of recovery after subsequent 42-day fresh air exposure are apparent. Here, we especially highlight the inflammatory, xenobiotic metabolism, and oxidative stress response. Importantly, these results complement the conclusions from our recent transcriptomic analysis for a 28-day rat cigarette smoke inhalation
study [175]. Moreover, the direct comparison with transcriptomic data for the 90-day rat study revealed overall consistency between the mRNA and protein response, but also highlighted relevant differences likely due to posttranscriptional regulation. In addition, we provide further evidence for the complex compensatory metabolic switch in response to cigarette smoke exposure, which involves the up-regulation of oxidative phosphorylation and fatty acid oxidation enzymes, possibly to cope with the changing cellular energy requirements [178].

1.3.4. Phosphoproteomics for toxicological assessment

Global expression proteomics mainly captures the alterations in effector functions that cope with a specific cellular stress (e.g., up-regulation of xenobiotic enzymes) and gross alterations in the tissue composition (e.g., invasion of immune cells). Cells use a sophisticated signaling network to sense and process cellular stresses and changes in this network can be considered early indicators of a toxicological stress. Of the methods for the analysis of signaling networks, phosphoproteomics can be considered the most established (see above), but only a few studies have already used this technique to assess toxicological mechanisms.

Caruso et al. employed a systems toxicology approach to assess the impact of mercury on a B lymphocyte cell model [182]. Mercury is a potent neurotoxin, but has also been found to contribute to autoimmune diseases at low concentrations, which do not invoke neurotoxicity. To further understand this phenomenon, the authors exposed WEHI-231 cells, a murine B-cell line, for 10 min with mercury and conducted a mass-spectrometry based phospho-proteome analysis. Interestingly, the B cell receptor pathway with the Lyn kinase as the key node was identified as the most affected signaling pathway. This finding was followed up with a targeted mass-spectrometry assay and the involvement of Lyn was confirmed. From this, the authors concluded that Lyn could represent an important contributor to mercury induced autoimmune diseases.

Chen et al. used quantitative expression and phospho-proteomics to analyze the cellular response to the alkylating chemical model MNNNG (N-methyl-N′-nitro-N-nitrosoguanidine) [183]. They focused on the nuclear (phospho-) proteome and compared the response of a lab-generated cell line pair. Both cell lines had a defect in a direct detoxification enzyme for MNNNG (MGMT), but in addition one cell line was deficient in the mismatch repair system (MMR). Chen et al. and found a larger phosphorylation response in the MMR proficient cell line and identified a signaling response network that involved ATM/ATR, CDK2, Casein kinase II, and MAP kinases.

Pan et al. employed a phospho-proteomics strategy to analyze and better understand the impact of deoxynivalenol (DON) on the mouse spleen [184]. The mycotoxin DON is frequently found in human an animal food and shows immunotoxic effects that are associated with a ribotoxic stress response. Quantitative phospho-profiling revealed 90 differentially regulated phospho-proteins upon DON exposure. Both the MAP-kinase and PI3K/AKT signaling axes were affected and several additional pathways that likely contribute to immune dysregulation were identified. From this, the authors concluded that phospho-proteomics helped to further unravel the complex effect of DON on the immune system and their study will serve as a template to better understand the toxic effects of DON in the future.

1.4. General discussion and future prospects

1.4.1. The future of systems toxicology

Framed in a systems analysis context, physiological homeostasis is maintained by a hierarchy of functional domains (genetic sequence, gene transcription, transcriptional regulation, protein function and interaction, organelles, cells, and organs) that are interconnected at each level of functional organization and across levels [185]. Exposure to chemicals and xenobiotics may simply be viewed as a perturbation that alters this system. Thus, an advanced mechanistic understanding of exposure effects requires systems toxicology approaches that capture these effects on different levels of this hierarchy and eventual integrate them into quantitative (and predictive) mathematical models [4]. This perspective is already a central element of the EU framework 6 program to further aid the understanding of the mechanisms of drugs actions and drug-mediated toxicities [186]. An example is the creation of joint data repositories for the complex datasets generated by a number of EU projects, which include aging- or toxicology-related projects assembling genomic, transcriptomic, proteomic, and functional data from a variety of models.

In the context of chemical risk assessment, Wilson et al. especially emphasize the need for integrative systems-level studies (e.g., proteomics, metabolomics, transcriptomics) to generate hypotheses and test mechanisms of action, which are then used as supporting information for a particular mode of action in EPA risk assessment [187]. Overall, such integrative approaches will be instrumental in understanding the complexities of toxicokinetic and toxicodynamic steps in multiple, and possibly interacting, pathways affected by a single chemical or mixtures of chemicals in human health risk assessment.

1.4.2. The future of proteomics in systems toxicology

Mass spectrometry-based proteomics methods are evolving rapidly toward higher sensitivity, higher throughput, higher coverage, and highly accurate quantification, and thus will constitute a central component of future integrative systems toxicology approaches [188]. Specifically, these advances include new highly accurate and fast mass spectrometer instruments [189–191]; improved methods and much expanded resources for targeted proteomic measurements (SRM, PRM) [192] [193] [194]; the novel (still exploratory) SWATH technology, which combines the strengths of targeted (sensitivity, dynamic range) and untargeted measurement principles (coverage) [195]; and advances in label-free quantification approaches [196]. Considering these advances, it has recently been suggested by Aebersold et al. that—at least for the analysis of proteins—it is “time to turn the tables” [197]; MS-based measurements are now more reliable than classical antibody-based western blot methods and should be considered the gold standard method of the field.

With MS instrumentation becoming more and more mature, Van Vliet especially emphasized the need to further develop computational analysis tools for toxicoproteomic data including data integration and interpretation methods [198]. Analysis methods developed for transcriptomic data such as GSEA [111] have already been successfully used in several proteomic studies. However, when developing (or applying) analysis methods for proteomic data, it is important to keep the main differences between transcriptomic and proteomic data in mind. These include sampling differences (sampling biases, missing values) [199,200], differences in the coverage of proteomic and transcriptomic measurements [199], and the fundamentally different functional roles and modes of regulation of proteins and mRNAs.

For example, improving the integration of transcriptomic and proteomic data for toxicological risk assessment has been identified as an important topic for future computational method development [198,201]. In this review, we have presented several possible data integration approaches including some that have already been successfully applied for the integration of transcriptomic and proteomic data (see Fig. 2 and “Deriving insights through data integration” section) [170,171]. Overall, the question is still open how to best integrate these different data modalities to reliably summarize the biological impact of a potential toxicant. However, the concept of Pathways of Toxicity (PoT) [3] combined with a rigorous quantitative framework could guide a solution. Recently, we have published on a computational method that uses transcriptomics data to predict the activity state of causal biological networks that fall under the PoT category [202]. It can be imagined that such an approach can be further expanded by directly utilizing data on (phospho-) protein nodes in these networks; PoTs measured with proteomic techniques. While proteomic and transcriptomic data can already be considered as complementary for toxicological assessment (e.g., Fig. 3E),
such integrative models would yield truly synergistic results on the biological impact across biological levels. In addition, most current toxicoproteomics studies focus on the measurement of whole protein expression. However, the relevance of posttranslational modifications such as protein phosphorylation for toxicological mechanisms is well appreciated and especially the analysis of phospho-proteomes has matured (see above) [203,204]. With this, phosphoproteomics (and the measurement of other PTMs) has great potential to significantly contribute to integrative toxicological assessment strategies in the future.

When using model systems, the crucial question is how the measured molecular effects translate between species; most importantly, from animal models to human. For example, Black et al. compared the transcriptomic response of rat and human hepatocytes to 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) and found similarities, but also significant differences in the response of the two species [205]. Recently, we have co-organized the species translation challenge. For this, we generated a cross-species data set, which captures the exposure response of both human and rat epithelial cells to 52 different stimuli [206]. The molecular response was measured both by transcriptomics and by targeted phospho-proteomics. Within the sib IMPROVER framework, different computational groups were engaged to assess the predictability of exposure effects within and between the two species (Rhrissorraekh et al., submitted) [207]. Again, while overall translatability was demonstrated, the accuracy of translation was stimulus and biological process dependent. Interestingly, however, for this dataset the phospho-proteomics measurements demonstrated higher translatability than the transcriptomics results. For future toxicological applications, it will be important to further assess the translatability of transcriptomics and (phospho-) proteomics responses. Especially, it will be interesting to further evaluate, whether the reported higher conservation of the proteome vs. the transcriptome holds for relevant toxic challenges [5].

2. Conclusions

Toxicology is increasingly moving beyond the sole measurement of apical endpoints, and in the future it will be crucial to gain a better understanding of the causal chain of molecular events linking exposures with adverse outcomes (i.e., apical endpoints) toward improved predictive risk assessment [4]. Toward this overall goal, systems toxicology combines large-scale measurements (e.g., transcriptomics and proteomics) with mathematical modeling. As discussed in this review, MS-based proteomics is maturing into a robust technology for the measurement of proteome-wide exposure effects. The benefits of including proteomic data to understand exposure effects have already been demonstrated in several case studies. Although some challenges still exist to make full use of the richness of proteomic datasets [198,201,208], there is overall a great opportunity for proteomics to contribute to an improved understanding of toxicant action, the linkages to accompanying dysfunction and pathology, and the development of predictive biomarkers and signatures of toxicity. Assembling a generally accepted, robust, and integrative systems toxicology assessment framework will benefit from collaborative efforts with the active participation of industry, academia, research institutes, and regulatory bodies.

References


Li Y, Chi H, Xia L, Chu X. Accelerating the scoring module of mass spectrometry-based peptide identification using GPUs. BMC Bioinformatics 2014;15:121.


