Quantitative proteomics for drug toxicity

Yuan Gao, Ricky D. Holland and Li-Rong Yu

Advance Access publication date 7 April 2009

Abstract
The emerging field of toxicoproteomics has been greatly advanced by quantitative proteomic technologies and their increasing applications in toxicology. The discipline is focused on the proteomic study of toxicity caused by toxic substances, including but not limited to drugs, toxins, environmental stressors, chemicals and any other materials that induce significant pathological responses. Drug safety is a major point of concern during the development phase and clinical application. Identification of toxicity biomarkers, potential drug targets and characterization of toxicity mechanisms represent major research areas for quantitative toxicoproteomics during drug development and evaluation. Further development and application of quantitative proteomic approaches would significantly facilitate the realization of personalized medicine.

Keywords: proteomics; mass spectrometry; toxicity biomarker; drug target; toxicity mechanism; toxicoproteomics

INTRODUCTION
In the past decade, many novel proteomic technologies were developed and applied to various biological systems for the understanding of cellular activities, disease development and physiological responses to therapeutic interventions and environmental perturbations [1–3]. One of such emerging fields is the application of proteomic technologies to toxicological research, which has led to the development of a new area, termed toxicoproteomics. Toxicoproteomics, albeit in its early stage, has drawn great attention in toxicology and has a great application potential [4]. The discipline is focused on the proteomic study of toxicity caused by toxic substances, including but not limited to drugs, toxins, environmental stressors, chemicals and any other materials that may cause significant pathological responses (e.g. engineered nanomaterials [5]). The area needs to combine knowledge of toxicology, pathology and other disciplines to define, at the proteome level, the toxicity pathways, modes of action and key molecular events.

Toxicoproteomics delineates pathological responses to a specific toxicant at the protein molecule level. It requires systematic analysis of protein function changes associated with toxicant exposures, including qualitative and quantitative measurements of protein expression, protein–protein/toxicant interactions, protein modifications, protein structures, subcellular localization and so forth. The research must provide tools for better evaluation and understanding of toxicity, and generate knowledge regarding organism responses to toxic substances for the promotion of public health.
The discipline has several major goals: to develop proteomic technologies for efficient analysis of proteins, to discover and validate protein biomarkers for early detection of toxicity and disease, to elucidate molecular mechanisms of toxicity, and to generate knowledge for better risk assessment of toxic substances.

While many diseases are developed via human’s exposure to environmental toxicants either acutely or chronically, medications provide tools to correct the diseased state to a normal condition. However, safety issue is a major concern for regulatory agencies, pharmaceutical industry and patients since medications may induce unacceptable organ toxicities. Like other ‘omics’ technologies, proteomics provides novel approaches for toxicity evaluation and monitoring in drug discovery and development, in both preclinical and clinical stages. A great potential of toxicoproteomics to accelerate drug development is anticipated, and application of proteomic technologies to drug toxicological research and post-market safety assessment is expanding. Although the research subjects of toxicoproteomics are quite broad, this review is mainly focused on its application to drugs. We will briefly delineate the current quantitative proteomic technologies, biomarker discovery and validation, drug target identification and proteomic analyses of drug toxicity mechanisms.

QUANTITATIVE PROTEOMICS TECHNOLOGIES

Proteomics technologies have been developed and improved rapidly for relative or absolute quantitation of protein expression in biological systems. Considerable efforts have been made for both mass spectrometry (MS)-based and protein microarray-based measurements [1, 6]. Here the discussion is focused on the former. For MS-based methods, both matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) techniques are currently utilized to generate peptide/protein ions for mass measurements [1]. However, ESI is more commonly applied to the analysis of complex peptide/protein mixtures since it is more straightforward to be coupled online with liquid chromatography (LC) separation. Although efforts have been made to quantify disease-related molecular signatures by solely measuring the masses of peptides or proteins, for example, surface-enhanced laser desorption/ionization (SELDI) profiling [7], such an approach is less informative since the identity of molecules is not characterized. More informative approaches require the proteins to be identified using tandem MS (MS/MS). To achieve this, a peptide or protein is isolated and subjected to fragmentation using specific techniques, such as collision-induced dissociation (CID), electron capture dissociation (ECD), electron transfer dissociation (ETD), sustained off-resonance irradiation (SORI) fragmentation, and infrared multiphoton dissociation (IRMPD) [8, 9]. Peptides/proteins are then identified via a search of their MS/MS spectra against a protein database. Mass measurement accuracy and fragmentation efficiency are critical for confident protein identification. Direct analysis of intact proteins (i.e. top-down approach) requires mass spectrometers to be extremely high in resolution, for example, Fourier transform ion cyclotron resonance MS (FTICR-MS) with a high magnetic field, in order to achieve high mass accuracy. However, analysis of peptides (proteolyzed from proteins) is more flexible and allows the use of a wide range of mass spectrometers (i.e. bottom-up approach).

Strategies have been developed to quantify protein abundances and associated changes in a relative or absolute fashion (Figure 1). The absolute quantification (AQUA) method developed in Dr Gygi’s lab [10] involves the synthesis of stable-isotope labeled (heavy) peptides and spiking these peptides in known quantities into the proteome sample to be analyzed, where the ‘heavy’ peptides serve as internal standards for peptide quantitation. More recently, stable isotope labeled artificial proteins of concatenated signature peptides (i.e. QconCAT) were synthesized to allow the standards to be spiked into samples before trypsin digestion [11]. To be in line with sample preparation and pre-fractionation to a maximum degree for more robust quantitation, in vitro synthesized isotope-labeled full-length protein standards called Protein Standard Absolute Quantification (PSAQ) were introduced into proteome samples for absolute quantitation [12].

Relative protein quantitation can be achieved by incorporating stable isotope labeling into the proteome samples to be analyzed. The labeling can be performed before or after protein extraction (pre-extraction or post-extraction). The pre-extraction method mainly refers to the metabolic labeling during cell culturing using 14N/15N-enriched media
Figure 1: Strategies for relative and absolute protein quantitation using in-solution-based mass spectrometry technology. For relative protein quantitation without any labeling, each proteome sample is prepared and enzymatically digested separately, and the resulting peptides are analyzed using LC–MS–MS/MS. Pre-fractionation can be conducted at the protein and peptide levels to increase the proteome coverage. Relative quantitation can also be achieved by incorporating stable isotope labeling for the proteome samples to be compared at different stages of sample preparation (left side). To quantify the proteins absolutely, isotope-coded peptide or protein standards (AQUA, QconCAT, or PSAQ) are spiked into the peptide or protein samples with known quantities (right side). The abundance of the peptide to be analyzed is then compared with that of the isotope-coded standard. AQUA, absolute quantification; ICAT, isotope-coded affinity tag; iTRAQ, isobaric tag for relative and absolute quantitation; PSAQ, protein standard absolute quantification; QconCAT, absolute quantification using concatenated signature peptides; TMT, tandem mass tag.

[13] or stable isotope-coded amino acids [14–16]. Nitrogen and carbon elements are preferred since light and heavy versions of the labeled peptides elute off the reversed-phase LC column with little or no retention time difference. A representative example of applying metabolic labeling strategy for quantitative proteome analysis is the method called stable isotope labeling by amino acids in cell culture (SILAC) [16]. In this technique, mainly $^{13}$C-labeled lysine and/or arginine is spiked into culturing media to label cellular proteins during cell proliferation. The advantage of this approach is that the compared samples can be pooled early to avoid variations in the parallel processing of multiple samples. Although the method was originally designed for cultured cells or microorganisms, more recently SILAC was extended to label the whole animals (e.g. mice), in which complete $^{13}$C$_6$-lysine labeling was achieved in all organs tested [17]. The SILAC-mouse approach provides a versatile tool allowing protein levels to be measured from whole organs all the way down to sub-cellular organelles and individual proteins. Stable isotopic labeling can also be conducted after protein is extracted, when metabolic labeling is not applicable or intended to use. In this context, isotope-coded affinity tags (ICAT) approach is a method of choice to specifically label Cys residues and enrich Cys-containing peptides (after digestion) for pairwise comparison of relative protein abundance between two proteome samples [18]. The advantage of this technology is that sample complexity is decreased, and thus the potential to identify low abundance proteins is increased. Enzyme-catalyzed $^{16}$O/$^{18}$O exchange is another labeling strategy, in which two oxygen atoms of the C-terminal carboxylate group of each peptide can be replaced by $^{18}$O atoms in $^{18}$O-water via trypsin-catalyzed oxygen exchange mechanism [19, 20]. To improve the dynamic range of quantitation, enzyme-catalyzed $^{16}$O/$^{18}$O labeling can be used in combination with ICAT [21]. Some isotope-coded amine-reactive reagents, for example, the isobaric multiplexing tagging reagents for relative and absolute protein quantitation (iTRAQ) [22] and Tandem Mass Tags (TMT) [23] were developed to label peptides for the purpose of relative protein quantitation of multiple samples.

Label-free approach has been increasingly employed for quantitative proteomic analysis of a large number of proteome samples without applying any stable isotope labeling [24]. Proteome samples are digested using proteolytic enzymes, and the resulting peptides are analyzed using 1D or 2D LC–MS/MS. After peptide ion peaks are aligned and intensities are normalized, the abundance of each identified peptide/protein is compared across analyzed samples. However, since labeling is avoided, sample preparation and instrument analysis should be rigorously controlled. For each quantitative approach described previously, high-resolution mass spectrometers (e.g. FTICR-MS and Orbitrap-MS) are expected to achieve better quantitation and higher confidence of protein identification. These types of instruments are especially valuable for the identification and quantification of intact protein biomarkers when isotope-labeled standards and efficient fragmentation techniques for intact proteins are incorporated.
TOXICITY BIOMARKER DISCOVERY

Drug toxicity remains a major point of concern during development and upon clinical use. Incidences of toxicity in liver, kidney, heart, brain and other organs have been reported with chemotherapeutic agents. In non-targeted anticancer therapies, for example, cardiotoxicity has arisen with manifestations of heart failure, ischemia, hypotension, hypertension, edema, bradycardia, etc. [25] for a wide range of therapeutic agents [26]. Cardiotoxicity is exemplified by the most commonly used anthracycline, doxorubicin, with toxicity reported to be 14–49% in the patients treated for lymphoma [27]. It has been expected that the new generation of novel targeted anticancer therapy would provide better therapeutic efficacy and lower organ toxicity. However, recent observations suggest that this type of ‘magic bullets’, such as trastuzumab, imatinib and sunitinib, often induce adverse effects on heart in a small population of patients [28–30].

Early detection of organ toxicity provides great benefits to patients to prevent further development of adverse events and complications of disease management. Toxicity detection should be at the earliest and minimally symptomatic stage, which requires the detection tools to be very sensitive and reliable. In the case of cardiotoxicity, the most practical monitoring tool is serial assessment of left ventricular function with echocardiography or radionuclide ventriculography [28]. However, it has been found that the use of echocardiography is limited by poor reproducibility and high variability [31]. Protein biomarkers could be invaluable tools in this context. Quantitative proteomic technologies have been used for many years to identify protein biomarkers of diseases [32]. Serum and plasma represent the most common samples for biomarker discovery since these samples can be easily collected. However, it has been a big challenge to identify low abundance biomarkers for early disease or toxicity detection since serum/plasma samples have a wide abundance range across different proteins (10–11 orders of magnitude) [33]. Nonetheless, this does not impede the biomarker discovery process from the blood source since potentially biomarkers could be present in blood for all types of organs. Urine represents another type of sample for biomarker search, especially for kidney toxicity and injury. Cerebrospinal fluid is a good source for biomarker identification for neurotoxicity. Large efforts have also been made to identify biomarkers from tissues and interstitial fluids since these types of samples may contain biomarkers with higher abundance compared to the blood. Another challenge is the throughput of proteomic platforms for the analysis of hundreds or even thousands of samples that are generated from time- and dose-dependent toxicological studies. A higher throughput platform to handle a large volume of samples would largely advance the field.

BIOMARKER VALIDATION

Biomarker validation is a critical process to translate biomarker candidates to nonclinical or clinical use. While thousands of potential protein biomarkers of diseases have been identified, the number of biomarkers that are eventually validated and qualified by the U.S. Food and Drug Administration is very limited. Validation of biomarkers needs both clinical samples and measurement tools. The approaches and/or platforms that are used to measure biomarkers must be validated. This process is termed, herein, analytical validation. Analytical validation of selected biomarkers, including sample handling, instrument analysis and data processing, should provide robust, specific, accurate and precise analytical measurements along with adequate dynamic range. While antibody-based technologies have been used for quantitative measurement of biomarkers, emerging proteomic technologies, for example, protein chips and MS-based multiple reaction monitoring (MRM) [34], are being developed toward realization of these platforms as valid tools for protein biomarker validation and disease diagnosis as well as toxicity detection.

When candidate biomarkers are identified from clinical samples or from animal models, these markers must be validated in clinical settings and eventually be qualified for clinical use. The process of clinical validation is quite challenging. Typically, clinical validation of biomarkers requires a large cohort of samples as discussed previously [33], and should be performed prospectively in a blind test. Samples are independent from those used at the biomarker discovery stage. Inter- and intra-individual variations of biomarker candidates in normal conditions need to be determined (baseline variation). This baseline evaluation needs to consider as many sources of variation as possible, including age, nutrition, sex, etc. The specificity and sensitivity of biomarkers
must be determined in an acceptable size of tested population. Establishment of sophisticated biostatistics models for toxicity biomarker validation by incorporating molecular mechanisms of toxicity is beneficial for achieving high predictive power of biomarkers.

**DRUG TARGET IDENTIFICATION**

Molecular target-orientated therapy (i.e. targeted therapy) has advantages over traditional chemotherapy; however, it is challenging since many target-directed drugs often lack molecule-binding specificity. Evaluation of drug targets not only provides insights into the primary mechanism-of-action of a drug but also the understanding of the side effects or toxicity as a result of ‘off-target’ interactions, which will further provide the rationale for optimization of drug design to minimize toxicity. Evaluation of drug targets not only provides insights into the primary mechanism-of-action of a drug but also the understanding of the side effects or toxicity as a result of ‘off-target’ interactions, which will further provide the rationale for optimization of drug design to minimize toxicity.

Imatinib is a selective small molecule that inhibits multiple kinases, including Bcr-Abl, c-Kit and PDGFR [35]. It has been approved to treat chronic myelogenous leukaemia (CML) via inhibition of Bcr-Abl activity, and treat gastrointestinal stromal tumors (GIST) through inhibition of c-Kit kinase. Cardiac toxicity has been raised for this drug through its target Abl kinase [29]. Fernandez et al. [36] reengineered imatinib to delete its Abl-inhibitory activity by modifying its chemical structure and preserved inhibitory property for c-Kit kinase. A marked reduction of cardiotoxicity was observed while the efficacy to treat GIST was remained. The above study demonstrates that it is feasible to re-design a drug to minimize its toxicity and retain its efficacy based on better understanding of drug targets.

Technologies have been developed to identify and characterize drug targets at the proteome level. Basically, the approaches can be summarized into three categories: affinity chromatography, activity-based probing, and protein and small molecule microarrays (Figure 2). Sleno et al. [37] published a detailed review on these techniques. For example, Wissing et al. [38] recently developed a strategy to affinity capture targets (kinases) for the kinase inhibitors of interest. They immobilized four structurally different protein kinase inhibitors on the solid-phase beads which were used to pack four columns. Each cell lysate from three different cancer cell lines was then passed through four consecutive columns to isolate affinity-bond kinases. More than 140 protein kinases were identified using LC–MS/MS analysis of isolated protein contents. In addition, 209 unique phosphorylation sites were identified from the phosphopeptides isolated by immobilized metal affinity chromatography (IMAC) enrichment [38]. A study by Bantscheff et al. demonstrated that quantitative measurement of target kinases was achievable using iTRAQ. They used immobilized nonselective kinase inhibitors (kinobeads) to capture hundreds of endogenously expressed protein kinases and purine-binding proteins that interact with investigated kinase inhibitors. Quantitative profiling of the drug-treated K562 cells by imatinib, dasatinib and bosutinib confirmed the known targets including the Abl and Src family kinases and identified the receptor tyrosine kinase DDR1 and the oxidoreductase NQO2 as novel targets of imatinib [39]. This case study demonstrates that proteomic technologies are valuable tools for the confirmation of known drug targets and identification of novel targets. The novel targets discovered might be associated with organ toxicities, which would provide a great opportunity for further improvement.
of the original drug to prevent its interaction with the target(s) inducing toxicity.

PROTEOMICS FOR THE UNDERSTANDING OF DRUG TOXICITY MECHANISMS

Elucidation of the mechanisms involved in drug-induced toxicity would provide essential information for improving drug development and more accurate prediction of drug safety. Two key features must be clarified to delineate the toxicity mechanisms for a specific drug: first, the target responsible for toxicity and second, the signaling pathway mediating the toxicity [40] (Figure 2). While the unique capabilities of proteomic technologies have been discussed in the previous section, their potentials in quantitative measurement of pathways for drug toxicity are also obvious, in providing information on the cellular events downstream of the target(s) mediating toxicity or when the toxicities are induced by other factors (e.g. drug metabolites) not directly related to its targets. Mechanisms of drug toxicity are complicated and may need to be elucidated on a case-by-case basis since organ toxicities (e.g. liver, kidney, heart, etc.) were observed for a broad range of drugs used to treat different diseases. Discussion on different organ toxicities and toxicities induced by various disease treatments is beyond the scope of this review, so only a few will be provided as examples. Doxorubicin (DOX, a non-selective anticancer drug) induces cardiomyopathy and its mechanisms were proposed to be related to impaired myocardial adrenergic regulation, intracellular calcium overload, releases of cardiotoxic cytokines, free radical oxygen, dysfunction of mitochondria, apoptosis, and so forth [41]. Undoubtedly, quantitative proteomics can play pivotal roles in deciphering toxicity mechanisms systematically at the molecular level.

For targeted cancer therapy using anti-kinase inhibitors, issues related to cardiotoxicity have been raised as well in recent years, thus additional efforts should be made since only limited data are available to understand the molecular mechanisms responsible for this type of toxicity. Inhibition of ErbB2 signaling was proposed to be a central mechanism of trastuzumab (a monoclonal antibody targeting ErbB2)-induced cardiotoxicity; however, the pathophysiological mechanism is likely more complicated [28]. The cardiotoxicity of imatinib was recently found to be associated with activation of the endoplasmic reticulum (ER) stress response, which resulted in JNK-induced recruitment of proapoptotic proteins such as Bax to the mitochondria and mitochondrial collapse [29]. Since this type of medication directly inhibits the kinases that play key roles in tumorigenesis, phosphoproteomic approaches would be valuable tools for unraveling the signaling pathways mediating toxicity.

Proteomic approaches have been applied to drug toxicity research for the understanding of toxicity mechanisms; however, such toxicoproteomic studies have not been conducted as broadly as other proteomic branch areas, e.g. cancer proteomics. Main areas of mechanistic studies using proteomics include quantitative protein profiling of drug-induced organ toxicity, toxicity pathway mapping, assessment of drug/metabolite-induced protein modifications, and measurement of oxidative responses [42–45]. For example, an overdose of acetaminophen, N-acetyl-p-aminophenol (APAP), causes acute hepatotoxicity in humans and experimental animals. To understand the mechanism underlying APAP-induced liver injury, Fountoulakis et al. [42] employed 2D-PAGE-based proteomic approach to quantitate mouse liver proteins affected by APAP-induced hepatotoxicity. The expression levels of about 35 of the identified proteins were altered due to treatment with APAP. Many of the proteins showing changed expression levels are either known targets for covalent modification by N-acetyl-p-benzoquinoneimine (NAPQI) or involved in the regulation of mechanisms that are believed to drive APAP-induced hepatotoxicity [42]. Ruepp et al. specifically performed quantitative proteomic analysis of mitochondrial subfractions from the liver following exposure to APAP. The results revealed the decrease of chaperone proteins (Hsp10 and Hsp60), ATP synthase subunits and beta-oxidation pathway proteins [43]. Furthermore, Andringa et al. found 3-hydroxy-3-methylglutaryl coenzyme A synthase 2 (HMG-CoA synthase) had significantly declined levels of reduced thiols and decreased activity after APAP treatment as being revealed from the mitochondrial proteomic analysis. They also observed that catalase, a key enzyme in hydrogen peroxide metabolism, showed modification in protein thiol content [44]. The above studies suggest that comprehensive proteome analyses of protein expression changes, protein modifications, protein–protein/drug interactions, and specific subcellular
contents would provide novel mechanistic insights into drug-induced toxicity.

Similarly, proteome-wide investigation of protein oxidation is another strategy for mechanistic study of toxicity. MDMA (3,4-methylenedioxymethamphetamine, ecstasy) is a synthetic derivative of amphetamine and is frequently abused. Acute exposure of MDMA alone or with other substances of abuse can damage many organs including the liver and the brain. To understand the underlying mechanism of organ damage, Moon et al. designed a proteomic experiment to specifically identify oxidatively modified mitochondrial proteins as a result of MDMA exposure in rat liver [45]. The results revealed that MDMA exposure markedly increased levels of oxidatively modified proteins involved in energy supply, fat metabolism, antioxidant defense, and chaperone activities. Among these, the activities of mitochondrial aldehyde dehydrogenase, 3-ketoacyl-CoA thiolases, and ATP synthase were significantly inhibited [45]. The data demonstrate for the first time that MDMA causes the oxidative inactivation of key mitochondrial enzymes which most likely contributes to mitochondrial dysfunction and subsequent liver damage in MDMA-exposed animals. Each of the above studies emphasized a certain aspect of proteomic analyses while a complete understanding of toxicity mechanisms needs more comprehensive analysis of the toxicity systems using a variety of proteomic technologies and other systems biology approaches.

**CONCLUSIONS**

The emerging field of toxicoproteomics is greatly driven by technical advances in proteome analysis and the broad needs in toxicological studies. Many quantitative proteomic technologies are currently available for systematic analysis of drug toxicity at the molecular level. Label free or stable-isotope-based global proteome quantitation enables not only the discovery of toxicity biomarkers but also the understanding of toxicity mechanisms at the pathway level. Stable isotope labeled peptides or proteins are novel tools for both relative and absolute protein quantitation. Analytical and clinical validation of biomarkers is essential to make biomarkers useful in clinical application. This will also facilitate the development of proteomics-based diagnostic assays. Specific techniques for drug target discovery play important roles in identifying toxicity targets for further improvement of drug development and understanding toxicity mechanisms. The full bloom of toxicoproteomics in drug development and toxicity assessment requires high throughput analysis of proteome samples, further development or optimization of proteomic tools for robust quantification of biomarkers, and broad application of quantitative proteomic technologies to the understanding of toxicity mechanisms.

**Key Points**

- The aim of toxicoproteomics is to discover and validate toxicity biomarkers, elucidate toxicity mechanisms and generate knowledge for better risk assessment using proteomic technologies.
- Quantitative proteomic technologies are valuable tools for the discovery of toxicity biomarkers and mechanisms.
- Biomarker validation requires analytical validation for assay platforms and qualification of biomarkers in clinical studies.
- Identification of drug targets, protein modifications, and molecular pathways is essential for the understanding of drug toxicity mechanisms at the protein level.

**Acknowledgments**

We are grateful to Dr Donna Mendrick for critical reading and constructive comments on this manuscript. Y.G. is grateful to NCTR/FDA for postdoctoral support through the Oak Ridge Institute for Science and Education (ORISE). The views presented in this article do not necessarily reflect those of the U.S. Food and Drug Administration.

**FUNDING**

National Center for Toxicological Research, U.S. Food and Drug Administration (NCTR/FDA) (to L-R.Y.).

**References**


42. Foufoulaakis M, Berndt P, Boelsterli UA, et al. Two-dimensional database of mouse liver proteins: changes in hepatic protein levels following treatment with...

