Proteomics studies reveal important information on small molecule therapeutics: a case study on plasma proteins

Lello Zolla

Department of Environmental Science, University of Tuscia, Piazzale Università, Viterbo, Italy

The most abundant proteins in serum, such as albumin and IgG, act as molecular sponges that bind and transport low molecular weight proteins/peptides and drugs. In the near future, pharmacoproteomics, the use of proteomic technologies in the field of drug discovery and development, and interactomics, the branch of proteomics which is concerned with identifying interactions between proteins, will allow researchers to (i) know the specific protein changes that occur in biological compartments in response to drug administration; (ii) design small novel therapeutic molecules that can have extended half-lives if carried by plasma protein in the blood stream. Advances in these fields will open new avenues of tailor-made molecular therapy, reducing present limitations on treatment arising from toxicity and inefficiency.In this short review we report and discuss the most recent developments arising from the use of proteomic tools in blood plasma protein research, looking at the identification of proteins found in plasma as well as their interactions with small molecules such as drugs, peptides, organic chemicals and metals. We believe this research demonstrates that proteomic technologies, and in particular pharmacoproteomics, interactomics and post-translational modification analysis, could be instrumental in the design of new tailor-made drugs leading to substantial improvements in molecular therapy.

Plasma proteome

Blood plasma is the most complex human-derived proteome. Because of this complexity, and the enormous range of concentration encountered across the population of protein components, spanning in excess of ten orders of magnitude, whole blood plasma is the most difficult specimen to analyze, and this creates serious challenges for proteomics. Much progress has already been made in this field and new directions have been put forward and discussed as part of the HUPO Plasma Proteome Project (PPP), to focus efforts on the remaining challenges [1]. The PPP initiative has three stated long-term goals (i) to make a comprehensive analysis of the protein constituents of plasma; (ii) to determine the extent and source of variation in an individual's plasma over time; and (iii) to determine the extent of variation in plasma between individuals within and across populations [2]. Blood plasma is known to contain proteins derived from blood cells and other body tissues that may have ended up there through cell death or damage (causing proteins to be released from normal cells), or they may come from aberrant protein secretions from tumor cells. In a recent investigation [3], the examination of the plasma protein component categories revealed that many of the proteins detected in plasma are normally associated with cells (i.e. they are not known plasma proteins). These 'cellular leakage proteins' were categorized according to their original location and function. Intracellular proteins accounted for up to 42% of the proteins identified, while membrane-associated proteins, including those proteins that are membrane-based but not known to be released in plasma (i.e. receptors, coreceptors and adhesion molecules) [3] accounted for another 13%. Another 5% of the proteins were found to be of cellular origin, and are either secreted or occupy an extracellular location, and 3% were identified as specific cytokines or cytokine-related proteins. All these proteins are generally considered passenger proteins (some more transient than others) that utilize plasma for transportation, localization and mediation of cellular responses. Overall, this group is the least

E-mail address: zolla@unitus.it.

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2008.09.013

characterized but possibly the most interesting one in terms of potential to yield biomarkers, with protein concentrations believed to range from low μg/ml to pg/ml levels, and possibly extending to levels below the detection limits of traditional ELISA assays (~1 pg/ml).

The 'classic' plasma proteins, those whose activity is specifically localized in the plasma, such as human serum albumin (HSA), complement components and apolipoproteins, make up only 4% of the total protein found in plasma. Approximately 34% of the proteins identified, however, had no known function [3]. Hence, with such a diverse population of proteins derived from a range of sources we might expect that the analysis of the extracellular proteome of proteins circulating in the plasma and the cell-based proteome are necessary and complementary for an exhaustive plasma investigation. One strategy applied in several recent examples involves using a secondary tissue or fluid of interest to first identify potential candidates for biomarkers and then screening the complementary plasma sample for their presence. Such an investigation is highly desirable because disease markers present in plasma may include proteins with significant potential for early disease diagnosis, containing information that directly reflects pathophysiological states and represents an invaluable source of diagnostic information for a variety of different diseases [4]. Thus, a broad inventory of plasma proteins (both qualitative and quantitative) could be used for the identification of putative protein markers for any diagnosable disease as well as for the development of new therapeutic products [4].

Quantitatively speaking, the core plasma protein is albumin, representing about 50% of the total plasma protein content (in the order of 30–50 g/l). Immunoglobulins (Igs) represent 20–25% of the total protein mass [3]. Low-abundance plasma proteins from tissue leakage and cytokines are present in the range of picograms to nanograms per milliliter.

The most abundant proteins in serum, such as albumin, IgG and transferrin, are known to act as the carriers for hormones, lipoproteins and many other proteins, lipids and metals. They may, in fact, be described as molecular sponges, which bind and transport low molecular weight (LMW) protein/peptide species preventing their rapid clearance by the renal system, thereby extending their half-life in the blood stream. Although it has been estimated that 32% human plasma proteins do not display associations with the major proteins, about 48% have been found in association with IgG and 20% of proteins have also been found with HSA codepleted groups, suggesting the possibility of weak overlapping interactions between some proteins and both HSA and IgG (see Fig. 1). Moreover, because proteins in the circulatory system are exposed to a variety of proteases or chemical oxidations [5,6], plasma is rich in peptides, the so-called 'peptidome' (see Box 1) from which about 5000 peptides have been revealed [7], prevalently associated with the carrier proteins. If Igs are retained in the

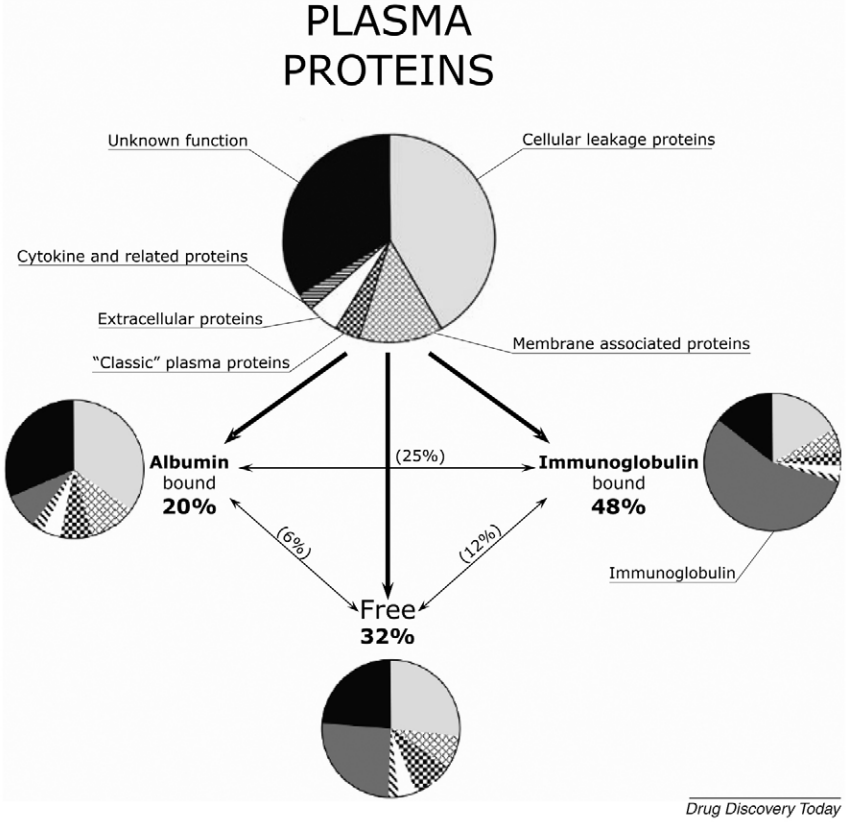

FIGURE 1

Category distributions of plasma proteins and their interaction with albumin or IgG. 'Free' proteins refer to proteins not found in either the albumin or the IgG codepleted fractions. The percentages in parentheses, listed among albumin, IgG and free, reflect the proteins overlapping between two groups.

BOX 1

Peptidome

The array of peptides normally present within the circulatory proteome is termed the 'peptidome', and could be a rich source of cancer-specific diagnostic information because it provides a record of the cellular and extracellular enzymatic events that take place at the level of the cancer-tissue microenvironment. This new information archive seems to show that most peptides *in vivo* are bound to high-abundance proteins such as albumin. Measuring panels of peptidome markers might be more sensitive and specific than conventional biomarker approaches.

Biomarker

A biomarker is a biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or a disease. A biomarker may also be used to see how well the body responds to a treatment for a disease or condition, to measure the progress of disease or the effects of treatment.

Pharmacoproteomics

The use of proteomic technologies in the field of drug discovery and development.

Toxicoproteomics

Toxicoproteomics is the use of global protein expression technologies to better understand environmental and genetic factors, both in the episodes of acute exposure to toxicants and in the long-term development of disease. Primary aims in toxicoproteomics are the discovery of key-modified proteins, the determination of affected pathways, and the development of biomarkers for eventual prediction of toxicity.

Phosphoproteomic profiling

The use of proteomic technologies and tools, such as phosphorylation-site-specific antibodies, to assess and monitor the phosphorylation state(s) of various proteins involved in signaling pathways in cells.

Isotope-coded affinity tagging

A method that allows for the quantitative measurement of proteins in two separate populations by mass spectrometry via the differential tagging of proteins in each population at cysteine residues with heavy or light isotope reagents.

Protein microarray

Protein microarrays are composed of a series of immobilized spots, containing a homogeneous or heterogeneous 'bait' molecule. A spot on the array can represent an antibody, a cell or phage lysate, a recombinant protein or peptide or a nucleic acid. The array is queried with either a probe (a labeled antibody or ligand), or an unknown biological sample (e.g. a cell lysate or serum sample) containing analytes of interest. By directly or indirectly tagging the query molecules with a signal-generating moiety, a pattern of positive and negative spots is generated.

Combinatorial ligand libraries

A combinatorial ligand library is composed of millions of diverse hexapeptide baits, able to capture aspecifically all peptides/proteins in any given proteome. In this way they concentrate the 'low-abundance' proteome, while drastically cutting the concentration of the most abundant compounds. It is based on the concept of a 'one-bead, one-peptide' approach.

Free-flow electrophoresis

Free-flow electrophoresis (FFE) is electrophoresis carried out in an aqueous medium without using any solid matrix, such as acrylamide. It is useful for the separation of a wide variety of charged analytes like low-molecular weight organic compounds, peptides, proteins, protein complexes, membranes, organelles and whole cells in aqueous media under native and denaturing conditions. The analyte is injected into a thin, laminar separation buffer film (which defines the electrophoretic mode such as zone electrophoresis, IEF or isotachophoresis) and is deflected by an electric field perpendicular to the flow direction.

Differential scanning calorimetry

Differential scanning calorimetry measures the difference in the amount of heat required to increase the temperature of a changed biological sample with respect to an unchanged one (reference) as a function of temperature. The difference in thermograms of blood plasma between normal and diseased individuals is not related to the concentrations of the most abundant plasma proteins but, rather, seems to arise from binding interactions involving as yet unknown biomarkers with the more abundant plasma proteins, particularly albumin. Such a behavior is consistent with the 'interactome' hypothesis.

Interactomics

Interactomics is a fusion science of biology, informatics and engineering which provides a global view of protein family interaction networks. It involves the study of both the interactions and the consequences of those interactions between and among proteins, and other molecules within a cell and can be used to compare networks of interaction between and within species to see how the traits of such networks are varied and conserved. The interactome network includes certain calculated parameters that weigh the reliability of a given interaction (i.e. the 'edges' of the interactome network) between two proteins, and also qualify the functional environment around any given protein (i.e. the 'nodes' of the interactome network).

sample during peptide-based proteomic studies, they can act as a generous source of random peptide sequences contributing to a large library of 'background' species that, because of their abundance (~20 mg/ml sample concentration) and variety, greatly complicate the process of peptide detection and identification.

Plasma protein interactions appear to exhibit little dependence on protein size. In fact, small plasma proteins such as the basic proline-rich peptide p-e (Ib-9, MW = 6020 Da, IgG codepleted protein) show more interaction specificity than large proteins such as cardiac titin isoforms (MW = 2,991,181 Da) that are typically found in more than one protein interaction group [3]. This kind of information, concerning the interactions between human plasma proteins, should be useful for further studies in human blood systems/network biology, such as evaluating possible protein contaminants in therapeutic protein products prepared from human plasma, and for the design of analytical approaches to deplete high abundance plasma proteins effectively. Additionally, it could be used to help design proteins with a reduced rate of clearance from the circulatory system, particularly important when small proteins are used as therapeutic agents [8]. Normally, small uncomplexed proteins and peptides (i.e. less than 40 kDa) are rapidly cleared from the circulation through enzymatic degradation, uptake by the reticuloendothelial system or by glomerular filtration, which discriminates on the basis of molecular size and charge [8]. It is believed that the circulation half-life of this LMW fraction is directly related to its binding affinity to large high abundant carrier proteins.

Some of the strategies devised to retard the clearance of therapeutic proteins include the covalent attachment of polyethylene glycol or dextran chains or by protein–protein crosslinking. Genetic modification has also been used to create chimeras of the therapeutic protein of interest with long-lived plasma proteins like albumin or IgG [9].

Over 10,000 different proteins have been estimated to be commonly present in the plasma, most of which are at very low relative abundances [10]. The dynamic range of currently available proteomic techniques such as mass spectrometry (MS) or 2D gel electrophoresis (2-DE) spans three to four orders of magnitude and as such does not approach the ten orders of magnitude represented by the plasma proteins. Thus, a significant challenge for proteomic analysis of plasma is how to reveal the low-abundance proteins. The strategies that have been most frequently used to overcome the dynamic range problem are to fractionate the plasma proteome into smaller subsets, and/or to deplete one or more of the major proteins. Immunoaffinity is the most efficient way to deplete proteins and so these methods are most widely used [11–13]. Immobilized antibodies packed in columns or cartridges capture several the most abundant proteins lowering the protein content down to 1% of the initial amount. Although nontarget proteins are removed in this process, it has been shown to be reproducible [14]. Multiple orthogonal separation steps have also been used, such as strong cation exchange chromatography followed by reverse phase-liquid chromatography (RP-LC) [8] or in-solution isoelectric focusing (IEF) [15]. Recently, a novel separation approach has been proposed, using 2D free-flow electrophoresis (FFE), separating proteins and peptides in solution according to their p*I* before LC–MS/MS [16]. A commercial combinatorial ligand library (ProteoMiner) is now available, enabling researchers to pick out the low-abundance proteins [17]. Another innovation is the MicroFlow MF10, a device to prefractionate complex low volume, low-abundance samples that can also enrich for very specific species of proteins based on charge and/or size either in native or in denaturing format [18].

All these strategies concomitantly remove proteins/peptides associated with the highly abundant proteins targeted for depletion. A recent study [19] focused on the binding properties of six of the most abundant serum proteins – to investigate the small peptides/proteins that may be bound to them. Each of the proteins targeted was found to be capable of binding several different peptide/proteins (210 proteins from a total of 378 unique peptides), many of which are clinically useful biomarkers. These results suggest that, on the one hand, albumin or IgG depletion before protein identification may actually eliminate many valuable biomarkers but, on the other hand, the identification of biomarkers, through the selective isolation of protein bound to the more abundant proteins in serum, could be a novel proteomic strategy. Because the general and specific sample losses increase with each separation stage, the best solution to this dilemma is to maximize the efficiency of each separation stage and minimize several dimensions required for the characterization of complex mixtures.

Biomarkers

As stated above, proteomic approaches may be utilized for disease classification as well as for the development of novel biomarkers related to prognosis, diagnosis and choice of therapeutic regimen. It is generally accepted that these biomarkers will not originate from classic plasma secretions but, most probably, from leakage, secretion or shedding of proteins from the specific affected tissue, cell type or cellular pathway. Nevertheless, blood is still the logical choice of biospecimen and has become the most frequently used biomarker discovery matrix to date, because the driving force of biomarker discovery is the development of blood-based assays for early detection and prediction of therapeutic response [20]. It follows that if protein biomarkers have been identified in biopsied tissues, researchers will then be able to examine the blood to determine whether these biomarkers are in circulation. The great advantage being that blood tests are much less invasive than biopsies. The downside is that once such proteins are released into around 6 liters of plasma, their concentration is extremely low. Hence, the need to extend the dynamic range of protein detection and this is where proteomics comes into its own. Although blood is a very convenient and noninvasive fluid to monitor for biomarkers, it poses many challenges from the perspective of protein detection.

Probably the greatest potential application for proteomics lies in investigating pathways that are easily targeted by small molecules or therapeutic antibodies. Among the challenges facing clinical proteomics is the ability to link protein expression profiles to specific disease phenotypes and the identification of relevant biomarkers to develop diagnostic tools [21]. In this sense, proteomics is an expansion of reductionist biology, where single proteins are analyzed in a high throughput fashion to arrive at an understanding of the entire system. The intention is that, once the human blood proteome has been fully described, several biomarkers will emerge for each particular disease state that can be used to provide a 'fingerprint' of that disease and here, there are obvious implications for blood donor testing [22]. There are already systematic searches underway to look for plasma proteins that are biological indicators or biomarkers for cancer (see Refs. [23,24]). If a suite of biomarkers were to be available for early detection, stratification into distinct subtypes and monitoring of progression or response to therapy we could expect significant improvements in clinical outcomes for cancer patients. The intention is to develop panels of biomarkers that will allow early detection of cancer and prediction of the probable response to therapy, possibly detectable in a single proteomics experiment. Despite the recent progress in proteomic technologies based on MS, however, the discovery of novel clinical assessment tools has been slow. This is partly because of the inherent difficulties in working with blood. It is hoped that a better understanding of the limitations of blood for comparative protein profiling and an appreciation of the advantages of cancer tissue or cancer cell secretomes will greatly enhance the progress [25].

A comparative proteome study of the body fluids of schizophrenia patients has been carried out to look for biomarkers or associated proteins in an effort to understand the etiology of schizophrenia. It was found that protein expression of the TTR tetramer and apolipoprotein E (ApoE) was downregulated by up to 1.68 and 3.62 times, respectively, in schizophrenia patients compared to normal controls [26]. A study that revealed potential diagnostic cancer biomarkers used blood from SJL (selected

by James Lambert) mice, in which the growth of RcsX lymphoma cells induces an inflammatory response by stimulating $V^{\beta}16$+T cells [27]. After the depletion of albumin and immunoglobulin, 1079 mouse NR proteins were identified in SJL mouse plasma in a single experiment. Most were found to be upregulated (e.g. acute phase reactants) but some proteins were found to be unique to the tumor-bearing mouse plasma (i.e. haptoglobin, proteosome subunits, fetuin-B, 14-3-3 zeta and MAGE-B4 antigen). While it remains to be seen whether a similar unique profile occurs in human lymphomas, recent studies have demonstrated that in sepsis and cancer, the inter-alpha inhibitor proteins (IaIps), a family of structurally related serine protease inhibitors, are present in relatively high concentrations in human plasma, suggesting their suitability as potential biomarkers [28].

Pharmacoproteomics

It would seem logical to assume that a better understanding of the biological mechanisms of drug toxicity and the development of therapeutic resistance would lead to the development of improved therapeutics with significant utility in clinical research. In this context, pharmacoproteomics, the use of proteomic technologies in the field of drug discovery and development (see Fig. 2), is an emerging science. Protein drug-activity biomarkers can be defined as specific protein changes in total protein content in biological compartments that occur in response to drug administration [29]. In contrast to genomic approaches, the application of pharmacoproteomics, by virtue of being stimulus-induced, may be more amenable to systematic experimental manipulation that could ultimately result not only in validation of a biomarker, but also in characterizing its behavior under specific conditions [30]. In association with chemical compound library screening strategies, pharmacoproteomics is expected to accelerate the discovery of new lead compounds for future drugs. It should also play an important part in preclinical studies by providing information on the mode of action and the eventual deleterious side effects of drugs. Hopefully, ongoing and future proteomic studies will open up new avenues of tailor-made molecular therapy, reducing present limitations associated with treatment toxicity and efficiency.

MS-based proteomics is ideal in this respect, because it offers a targeted way to identify hundreds of protein or peptide biomarkers simultaneously, with the obvious potential for discovering drug-activity markers because there is often no *a priori* knowledge of the particular proteins that are likely to change [31] in response to drug administration [32,33]. While sample enrichment narrows the dynamic range over which molecules of interest must be detected, the improvements in separation technology and MS have extended the ability to find and identify proteins over an ever wider range of relative protein concentrations. Ion trap instrumentation has been superseded by the greatly improved linear ion trap mass spectrometers. At the same time, there has been a steady decrease in the diameters of LC-capillary columns that has resulted in increased sensitivity and ability to quantify proteins. This has improved analysis by increasing, by orders of magnitude, the detectable dynamic range of a protein that would not have been picked up previously in samples of similar size. As such, at this time, LC–FTICR-MS (Fourier transform ion cyclotron

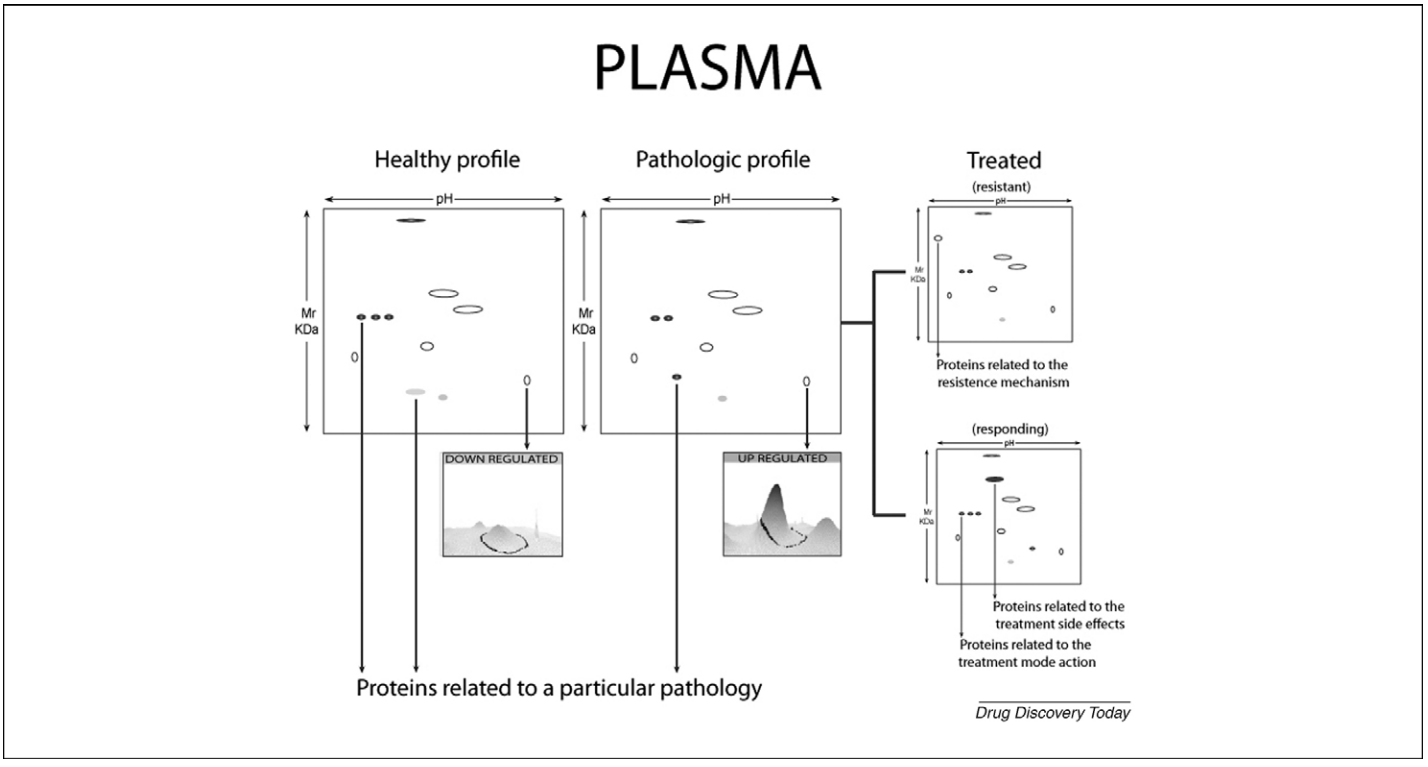

FIGURE 2

Theoretical example of differential screening of plasma proteins from healthy, ill and treated (responding and resistant) patients by 2D gel electrophoresis. The appearance or disappearance of protein spots as well as quantitative variations of protein expression can be observed.

resonance MS) technology provides the most sensitive and powerful measurement platform and though some significant potential exists for further increasing its dynamic range, this applies to those cases limited by sample complexity, rather than detection limits. Recently, however, differential scanning calorimetry has provided a new window into the plasma proteome. Thermograms of plasma from diseased individuals not only were found to differ dramatically from those of normal (control)_individuals, but also differed between sufferers of different illnesses, at least for the three diseases examined (arthritis, Lyme disease and lupus). Each disease appears to have a distinctive and characteristic thermogram. These radically different thermograms seen for different diseased states seem to arise from binding interactions involving as yet unknown biomarkers with the more abundant plasma proteins, particularly albumin. Such behavior is consistent with the 'interactome' hypothesis, suggesting a novel use for calorimetry as a diagnostic tool [34].

For a careful approach to clinical design and data analysis it is important to maximize the chance of discovering meaningful drug-activity markers. The first step is to establish a link between drug administration and the resulting biologically quantified change [35]. This strong foundation should set the stage for any data-driven hypothesis generation and testing, and further establish the utility of such a multidisciplinary tool in expanding the frontiers of clinical research. Too often there is no information available on the tissue under investigation, in these circumstances protein-profile changes in response to drug treatment can serve as evidence of cell pharmacodynamic activity, and such data can be used to establish a minimum dose for early clinical studies. More broadly, because targeted proteomics can also be hypothesis generating, detecting a drug-induced change can also lead to new lines of research and inquiry in the field of drug development or pathophysiology. A clear understanding of proteomic variability under controlled conditions is crucial to enable future clinical studies to be appropriately designed and to improve the contextual interpretation of data obtained from pharmacoproteomic studies. It is also important to establish a suitable time period between drug administration and analysis primarily to allow sufficient time for a tissue-targeted drug to reach a pharmacokinetic steady state before exploring protein profile changes [36]. This is to ensure minimal interference from extraneous factors on endogenous variability, thus maximizing the probability of identifying even subtle drug-induced changes. The variability between individual subjects could potentially affect observations in unpaired comparison studies, such as diseased versus healthy volunteers, and should be taken into account [37]. In other words, we cannot assume that pharmacodynamic responses to identical chemical or biochemical stimuli will be similar among healthy individuals and that these responses will differ significantly and consistently in patients with pathological disorders, because a variety of factors will influence the profile of each individual [37]. This aspect requires further investigation to ascertain how useful biomarkers will ultimately prove to be and what consequence this will have for drug development.

Pharmacoproteomics is not limited to the extracellular proteome analysis of blood plasma, because cell-based proteome analysis also has an important role in this field. In this context, post-translational modifications (PTM) of proteins have a significant impact. Some 300 potential modifications are known, and they include phosphorylation, glycosylation, acetylation, myristoylation, palmitoylation, methylation, sulfation, prenylation and ubiquitylation (http://www.abrf.org/index.cfm/dm.home). During the past two decades, sensitive MS methods have been refined to determine the type and site of protein modifications that can seldom be predicted from a genomic sequence [38]. Some of the protein modifications are regulatory and reversible, most notably phosphorylation which controls protein functions such as localization, complex formation, stability and activity through different mechanisms [39]. As a result of extensive research, we now know that about 30% of all human proteins can be modified by phosphorylation, while only 0.05% of the phosphorylated residues are tyrosine [40]. Protein phosphorylation is one hallmark of the protein–protein interactions (PPIs) underlying signaling networks and, in many cases, it is aberrant protein kinase activity that drives disease-associated derangements in signaling pathways. The aim would be to design drugs that effectively disrupt this aberrant protein-phosphorylation-based enzymatic activity and epigenetic phenomena. Pharmacoproteomics, or the tailoring of therapy based on proteomic knowledge, is expected to take a central role in this process [41].

A recent study describes a cell-based drug discovery platform based on phosphospecific flow cytometry, or phosphoflow, with which researchers were able to screen for inhibitors of multiple endogenous kinase signaling pathways in heterogeneous primary cell populations at the single-cell level [42]. Protein microarrays that examine protein–protein recognition events (i.e. phosphorylation) in a global, high-throughput manner have been used to profile cellular signal pathways in a way not possible with gene arrays [43,44]. The activity levels of the proteins in a particular pathway can thereby be assessed in 'real time', to tailor treatment to each patient's cellular 'circuitry' [45–47]. The advantage of protein microarrays lies in their ability to provide a 'map' of known cellular signaling proteins that generally reflect the state of information flow through protein networks in individual specimens. The identification of crucial nodes or interactions within the network (see later) is a potential starting point for drug development and the design of individual therapeutic regimens [48–50]. Accurate annotation of peptide modifications through unrestrictive database searches can reveal post-translational modifications, as well as sequence polymorphisms [51].

Monitoring different phosphoprotein levels will also help to identify treatment-acquired resistance to chemotherapy. The effect of imatinib on the tyrosine phosphoproteome in bcr-abl positive leukemia cell lines has been examined by proteomic approaches. The investigation revealed 64 sites of tyrosine phosphorylation corresponding to 32 different proteins [52]. Affinity column chromatography using an immobilized pyrido(2,3-d)pyrimidine derivative has been successfully used to select protein kinases that bound to the matrix and were then identified using MS, demonstrating the fact that pyrido(2,3-d)pyrimidine is a kinase inhibitor with low specificity [53]. A similar methodological approach was applied to study the p38 kinase inhibitor SB203580, demonstrating that

cyclin G-associated kinase (GAK) and CK1 were almost as potently inhibited as p38α [53,54].

Protein expression subsequent to treatment with the HDACI trichostatin-A has been studied in human pancreas ductal carcinoma cell lines using 2-DE and MALDI-TOF MS [55,56]. Trichostatin-A appears to upregulate proteins which promote cell death and downregulate proteins that favor cell growth. Similarly RC307 modulates proteins that are involved in proliferation, cell cycle regulation, apoptosis and gene expression in colon carcinoma cells [57]. The pharmacoproteomic approach was found to be particularly useful for the identification of molecular alterations implicated in type 2 diabetes, and for further characterization of existing or new drugs [58].

Research has shown that several physiological factors lead to changes in both the plasma proteome and the peptidome, including stress, sleep, sport training, eating meals and pregnancy. Exposure to toxic agents, including drugs, can also be detected using proteomics leading to a distinct field of study known as toxicoproteomics [59]. Recently, the proteins in the plasma of workers exposed to benzene were analyzed [60]. Two

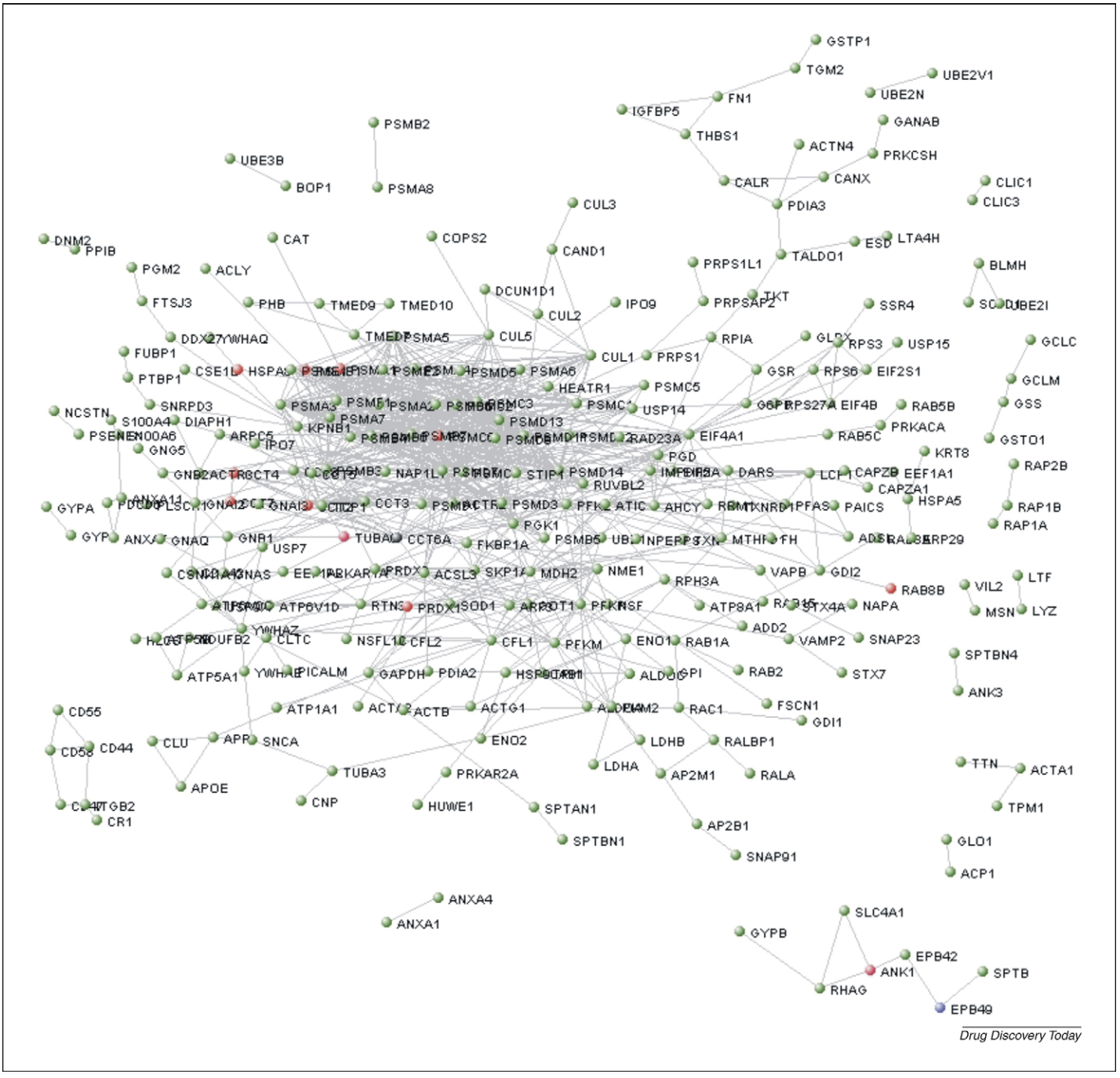

FIGURE 3

Human erythrocyte protein–protein interaction network (interactome). Nodes of the network are the known RBC proteins and links connecting the nodes correspond to the known interactions [70].

significantly upregulated protein profiles were found in workers exposed to the toxic chemical: the T cell receptor b chain and matrix metalloproteinase-13. Thus, the plasmatic T cell receptor b chain may be a useful indicator for the early detection of exposure to benzene. Comparative proteome analysis of human plasma allowed Qian *et al*. [61] to identify 32 proteins that were significantly increased after lipopolysaccharide administration. A study of the plasma of workers occupationally exposed to polycyclic hydrocarbons revealed several proteins as markers of such an exposure [62]. Using the same technique, truncated forms of α1-antitrypsin were discovered in serum samples of patients presenting with severe acute respiratory syndrome [63]. In another study involving patients with the same syndrome, samples revealed a total of 38 differential spots, most of them corresponding to acute phase proteins [64]. The authors also identified proteins such as peroxiredoxin II, not previously detected in plasma 2-DE.

The investigation of plasma protein interactions with metals remains a relatively new and challenging arena for toxicoproteomics. Current studies include the identification of the protein carrier and how to determine toxic metal concentrations, which requires sensitive analytical techniques and powerful equipment. Thankfully, the search for suitable quantitative techniques in the field of drug design and proteomics may be almost over with the development of inductively coupled plasma MS (ICP-MS) [65].

Another relevant field for pharmacoproteomic applications is that of elucidating the biological mechanisms of drug resistance. In drug-resistant cells alternative protein forms appear that prevent the drug binding to active sites and/or executing signaling effects independent of the regulated native protein forms. The methodology of proteomics seems highly applicable to the search for drug-resistant protein forms (drug-resistance proteomics). In this context the plasma proteome in aspirin (acetylsalicylic acid [ASA])-sensitive and ASA-resistant coronary ischemic patients has been analyzed. The expression of one isotype of the fibrinogen γ chain and three isotypes of haptoglobin was increased in ASA-resistant patients as well as that of three vitamin D binding protein (DBP) isotypes. It has been suggested that DBP regulates the inhibitory effect of ASA on platelets, by reducing the inhibitory effect of ASA on thromboxane A2 production [66]. In the field of oncology, proteomics is becoming widely used for the identification of tumor-specific protein markers, and pharmacoproteomics has found a place in the evaluation of chemotherapy, particularly for the characterization of drug-resistance mechanisms [58].

Interactomics

Identifying important interactions between blood proteins is emerging as another focus of blood-based proteomic research. Many proteins circulate in blood, not as single entities but as multicomponent complexes, and as such comprise the blood 'interactome' [4]. Protein-protein interactions (PPI) information can be extracted from the currently available databases of interactions (*in silico* approaches).

Little is currently known about the interactions of any given protein in the blood or whether these interactions are biologically relevant. Examining protein complexes will, however, often yield information about protein function. The identification of such interactions may bring to light important information for designing new small molecule therapeutics. This conceptual framework contrasts with the single protein (or single pathway) approach that has previously dominated biology.

The combination of proteomic and *in silico* approaches allows one to not only identify disease and/or drug-related changes in the proteome but also predict the changes in the protein interactome network that are associated with disease-related change of function. It can also predict modifications associated with PTM changes occurring as a result of a specific disease or stage of disease, as well as drug or gene therapy treatment. These linked approaches represent the future of clinical identification of a specific disease, its current stage or severity, and the effects and side effects of various treatments.

Clearly, an interactome of plasma proteins is far from reality at present, because of its complexity, while the simplicity of the human erythrocyte cell structure has instead made it an optimal cell for proteomic study. In fact, while nucleated cells contain 20,000–30,000 proteins [31,67,68], red blood cells (RBCs), which lack nuclei and other organelles, contain far fewer. A comprehensive list of RBC proteins known to date contains 751 entries obtained using proteomics technology [69]. The resulting PPI network is depicted in Fig. 3 [70]. The network presented here was derived from protein interaction data obtained from the Unified Human Interactome (UniHI) [71]. The correlation reported by UniHI is derived from gene expression experiments and represents a measure of confidence for the interaction [72]. As the knowledge of PPIs continues to increase owing to advances in proteomics research we expect that, in addition to a deeper understanding of the erythrocyte protein complement and how it relates to function, it will also be possible to understand how changes in the RBC proteome and interactome affect the development of erythrocyte disorders [69].

Conclusion

The promise of proteomic-based profiling, as opposed to gene transcript profiling alone, is that the resulting prognostic signatures are derived from drug targets (proteins) not genes, so the pathway analysis provides a direction for therapeutic mitigation. The only dark cloud on the horizon may be that extending the use of pharmacoproteomics (beyond molecular network analysis for patient-tailored therapy to include risk stratification or 'predispostional testing') is likely to raise a larger societal issue, particularly in a climate of cost containment. The powerful combination of pharmacoproteomics and interactomics may win through despite costs, furnishing new targeted therapies for disease prevention as well as for post-therapy monitoring.

In short, these techniques not only provide a window on the presence or severity of a specific disease, and drug and gene therapy treatment, but can also be used to identify disease- and/or drug-related changes in the proteome. They can also predict those changes in the protein interactome network that are associated with disease-related changes of function [73]. Without doubt, these linked approaches represent the future for the cost effective and noninvasive identification of specific clinical diseases, their prognosis and effective treatment.

Acknowledgements

This work was supported in part by The Italian National Blood Centre, Istituto Superiore di Sanità, Rome, Italy (directed by Dott. Giuliano Grazzini). The author would like to thank Dr G.M. D'Amici for his help and Dr J. Scarpa for English revision.

References

1 Cottingham, K. (2006) HUPO plasma proteome project: challenges and future directions. *J. Proteome Res.* 5, 1298
2 Omenn, G.S. (2004) The human proteome organization plasma proteome project pilot phase: reference specimens, technology platform comparisons, and standardized data submissions and analyses. *Proteomics* 4, 1235–1240
3 Shen, Y. *et al.* (2005) Characterization of the human blood plasma proteome. *Proteomics* 5, 4034–4045
4 Jacobs, J.M. (2005) Utilizing human blood plasma for proteomic biomarker discovery. *J. Proteome Res.* 4, 1073–1085
5 Barelli, S. *et al.* (2008) Oxidation of proteins: basic principles and perspectives for blood proteomics. *Proteomics Clin. Appl.* 2, 142–157
6 D'Amici, G.M. *et al.* (2007) Proteomic analysis of RBC membrane protein degradation during blood storage. *J. Proteome Res.* 6, 3242–3255
7 Richter, R. *et al.* (1999) Composition of the peptide fraction in human blood plasma: database of circulating human peptides. *J. Chromatogr. B Biomed. Sci. Appl.* 726, 25–35
8 Torchilin, V.P. and Lukyanov, A.N. (2003) Peptide and protein drug delivery to and into tumors: challenges and solutions. *Drug Discov. Today* 8, 259–266
9 Colcher, D. *et al.* (1999) Effects of genetic engineering on the pharmacokinetics of antibodies. *Q. J. Nucl. Med.* 43, 132–139
10 Adkins, J.N. *et al.* (2002) Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. *Mol. Cell. Proteomics* 1, 947–955
11 Pieper, R. *et al.* (2003) Multi-component immunoaffinity subtraction chromatography: an innovative step towards a comprehensive survey of the human plasma proteome. *Proteomics* 3, 422–432
12 Zolotarjova, N. *et al.* (2005) Differences among techniques for high-abundant protein depletion. *Proteomics* 5, 3304–3313
13 Huang, L. *et al.* (2005) Immunoaffinity separation of plasma proteins by IgY microbeads: meeting the needs of proteomic sample preparation and analysis. *Proteomics* 5, 3314–3328
14 Liu, T. *et al.* (2006) Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. *Mol. Cell. Proteomics* 5, 2167–2174
15 Tang, H.-Y. *et al.* (2005) A novel four-dimensional strategy combining protein and peptide separation methods enables detection of low-abundance proteins in human plasma and serum proteomes. *Proteomics* 5, 3329–3342
16 Nissum, M. *et al.* (2007) Two-dimensional separation of human plasma proteins using iterative free-flow electrophoresis. *Proteomics* 7, 4218–4227
17 Boschetti, E. and Righetti, P.G. (2008) The ProteoMiner in the proteomic arena: a non-depleting tool for discovering low-abundance species. *J. Proteome* 71, 255–264
18 Wasinger, V. *et al.* (2008) Prefractionation, enrichment, desalting and depleting of low volume and low abundance proteins and peptides using the MF10. *Methods Mol. Biol.* 424, 257–275
19 Zhou, M. *et al.* (2004) An investigation into the human serum interactome. *Electrophoresis* 25, 1289–1298
20 Rifai, N. *et al.* (2006) Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat. Biotechnol.* 24, 971–983
21 Marko-Varga, G. and Fehniger, T. (2004) Proteomics and disease – the challenges for technology and discovery. *J. Proteome Res.* 3, 167–178
22 Petricoin, E. *et al.* (2004) Clinical proteomics: revolutionizing disease detection and patient tailoring therapy. *J. Proteome Res.* 3, 209–217
23 Hanash, S.M. *et al.* (2008) Mining the plasma proteome for cancer biomarkers. *Nature* 452, 571–579
24 Omenn, G.S. (2006) Strategies for plasma proteomic profiling of cancers. *Proteomics* 6, 5662–5673
25 Simpson, R.J. *et al.* (2008) Proteomics-driven cancer biomarker discovery: looking to the future. *Curr. Opin. Chem. Biol.* 12, 72–77
26 Wan, C. *et al.* (2006) Dysregulation of retinoid transporters expression in body fluids of schizophrenia patients. *J. Proteome Res.* 5, 3213–3216
27 Vadiraja, B.B. (2005) Comparative plasma proteome analysis of lymphoma-bearing SJL mice. *J. Proteome Res.* 4, 1814–1825
28 Josic, D. *et al.* (2006) Proteomic characterization of inter-alpha inhibitor proteins from human plasma. *Proteomics* 6, 2874–2885
29 Sjoholt, G. *et al.* (2006) Proteomic strategies for Individualizing therapy of acute myeloid leukemia (AML). *Curr. Pharm. Biotechnol.* 7, 159–170
30 Rosenfeld, R.G. (2007) Pharmacogenomics and pharmacoproteomics in the evaluation and management of short stature. *Eur. J. Endocrinol.* 157, S27–S31
31 Tyers, M. and Mann, M. (2003) From genomics to proteomics. *Nature* 422, 193–197
32 Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* 422, 198–207
33 Pusch, W. *et al.* (2003) Mass spectrometry-based clinical proteomics. *Pharmacogenomics* 4, 463–476
34 Garbett, N.C. (2008) Calorimetry outside the box: a new window into the plasma proteome. *Biophys. J.* 94, 1377–1383
35 Hood, L.M. and Perlmutter, R.M. (2004) The impact of systems approaches on biological problems in drug discovery. *Nat. Biotechnol.* 22, 1215–1217
36 Patil, S.T. *et al.* (2007) Identifying pharmacodynamic protein markers of centrally active drugs in humans: a pilot study in a novel clinical model. *J. Proteome Res.* 6, 955–966
37 Dunning, B.E. *et al.* (2005) Alpha cell function in health and disease: influence of glucagon-like peptide-1. *Diabetologia* 48, 1700–1713
38 Jiang, X. *et al.* (2008) Technologies and methods for sample pretreatment in efficient proteome and peptidome analysis. *Proteomics* 8, 686–705
39 White Forest, M. (2008) Quantitative phosphoproteomic analysis of signaling network dynamics. *Curr. Opin. Biotechnol.* 19, 404–409
40 Hunter, T. (2000) Signaling – 2000 and beyond. *Cell* 100, 113–127
41 Wulfkuhle, J.D. *et al.* (2006) Technology insight: pharmacoproteomics for cancer – promises of patient-tailored medicine using protein microarrays. *Nat. Clin. Pract. Oncol.* 3, 256–268
42 Krutzik, P.O. *et al.* (2008) High-content single-cell drug screening with phosphospecific flow cytometry. *Nat. Chem. Biol.* 4, 132–142
43 Haab, B.B. (2005) Antibody arrays in cancer research. *Mol. Cell. Proteomics* 4, 377–383
44 Zhu, H. and Snyder, M. (2003) Protein chip technology. *Curr. Opin. Chem. Biol.* 7, 55–56
45 Paweletz, C.P. *et al.* (2001) Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene* 20, 1981–1989
46 Angenendt, P. (2005) Progress in protein and antibody microarray technology. *Drug Discov. Today* 10, 503–511
47 LaBaer, J. and Ramachandran, N. (2005) Protein microarrays as tools for functional proteomics. *Curr. Opin. Chem. Biol.* 9, 14–19
48 Eckel-Passow, J.E. *et al.* (2005) Experimental design and analysis of antibody microarrays: applying methods from cDNA arrays. *Cancer Res.* 65, 2985–2989
49 Petricoin, E.F. *et al.* (2002) Clinical proteomics: translating bench side promise into bedside reality. *Nat. Rev. Drug Discov.* 1, 683–695
50 Leuking, A. *et al.* (2005) Protein biochips: a new and versatile platform technology for molecular medicine. *Drug Discov. Today* 10, 789–794
51 Tanner, S. *et al.* (2008) Accurate annotation of peptide modifications through unrestrictive database search. *J. Proteome Res.* 7, 170–181
52 Salomon, A.R. *et al.* (2003) Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* 100, 443–448
53 Wissing, J. *et al.* (2004) Chemical proteomic analysis reveals alternative modes of action for pyrido2 3-dpyrimidine kinase inhibitors. *Mol. Cell. Proteomics* 3, 1181–1193
54 Godl, K. *et al.* (2003) An efficient proteomics method to identify the cellular targets of protein kinase inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* 100, 15434–15439
55 Cecconi, D. *et al.* (2003) Proteomic analysis of pancreatic ductal carcinoma cells treated with 5-aza-2′-deoxycytidine. *Electrophoresis* 24, 1871–1878
56 Cecconi, D. *et al.* (2007) Proteomic analysis of pancreatic endocrine tumor cell lines treated with the histone deacetylase inhibitor trichostatin A. *Proteomics* 7, 1644–1653
57 Milli, A. *et al.* (2008) A proteomic approach for evaluating the cell response to a novel histone deacetylase inhibitor in colon cancer cells. *BBA-Protein Proteom* 1784, 1702–1710
58 Chapal, N. (2004) Pharmacoproteomic approach to the study of drug mode of action, toxicity, and resistance: applications in diabetes and cancer. *Fund. Clin. Pharmacol.* 18, 413–422

59 Petricoin, E.F. *et al.* (2004) Toxicoproteomics: serum proteomic pattern diagnostics for early detection of drug induced cardiac toxicities and cardioprotection. *Toxicol. Pathol.* 32 (Suppl. 1), 122–130
60 Joo, W.A. *et al.* (2004) Proteomic analysis of plasma proteins of workers exposed to benzene. *Mutat. Res.* 558, 35–44
61 Qian, W.J. *et al.* (2005) Comparative proteome analyses of human plasma following in vivo lipopolysaccharide administration using multidimensional separations coupled with tandem mass spectrometry. *Proteomics* 5, 572–584
62 Kap-Soon, N. *et al.* (2004) Protein biomarkers in the plasma of workers occupationally exposed to polycyclic aromatic hydrocarbons. *Proteomics* 4, 3505–3513
63 Ren, Y. *et al.* (2004) The use of proteomics in the discovery of serum biomarkers from patients with severe acute respiratory syndrome. *Proteomics* 4, 3477–3484
64 Chen, J.H. *et al.* (2004) Plasma proteome of severe acute respiratory syndrome analyzed by two-dimensional gel electrophoresis and mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* 101, 17039–17044
65 Huanga, J. *et al.* (2006) The application of inductively coupled plasma mass spectrometry in pharmaceutical and biomedical analysis. *J. Pharmaceut. Biomed. Anal.* 40, 227–234
66 Lopez-Farre, A.J. *et al.* (2007) Relationship between vitamin d binding protein and aspirin resistance in coronary ischemic patients: a proteomic study. *J. Proteome Res.* 6, 2481–2487
67 Steen, H. and Mann, M. (2004) The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* 5, 699–711
68 Yates, J.R., III *et al.* (2005) Proteomics of organelles and large cellular structures. *Nat. Rev. Mol. Cell Biol.* 6, 702–714
69 Goodman, S.R. *et al.* (2007) The human red blood cell proteome and interactome. *Exp. Biol. Med.* 232, 1391–1408
70 Kurdia, A. *et al.* (2007) Centrality measures for the human red blood cell interactome. *Engineering in Medicine and Biology Workshop*, pp. 98–101, IEEE Dallas
71 Chaurasia, G. *et al.* (2007) UniHI: an entry gate to the human protein interactome. *Nucleic Acids Res.* 35 (Database issue), D590–D594
72 Balasubramanian, R. *et al.* (2004) A graph theoretic approach to testing associations between disparate sources of functional genomics data. *Bioinformatics* 20, 3353–3362
73 Liumbruno, G. *et al.* (2008) Transfusion medicine in the era of proteomics. *J. Proteomics* 71, 34–45