1

Proteomic Strategies for Analyzing Body Fluids

Sung-Min Ahn and Richard J. Simpson

Summary

The rapid development of molecular and cell biology in the latter part of the last century has led us to the understanding that many diseases, including cancer, are caused by perturbations of cellular networks, which are triggered by genetic changes and/or environmental challenges. These perturbations manifest by changing cellular protein profiles, which, in turn, alter the quantitative relationship of tissue-specific proteins shed into the tissue/organ microenvironment. Such altered protein expression profiles in body fluids constitute molecular signatures or fingerprints that reflect the original perturbation of cellular networks. The exciting challenge of modern proteomics is to identify such signatures for various disease states—then the body fluids will become windows into disease and potential biospecimen sources for biomarkers of disease. ***(1)***.

Key Words: Body fluids; biomarkers; proteome; expression proteomics; targeted proteomics; sample collection; enrichment; prefractionation.

1. Introduction

1.1. Historical Perspective and Biological Context

Claude Bernard, the 19th century physiologist, introduced the concept of *milieu intérieur* or the internal environment, defining it as the circulating organic liquid that surrounds and bathes all tissue elements ***(2)***. Bernard regarded extracellular fluid as the internal environment of the body and emphasized the importance of maintaining the constancy of that environment. This notion still holds true today, especially in the era of proteomics. Indeed, the detection of protein perturbations in the internal environment is one of the major goals of the fledgling field of proteomics.

Approximately 60% of the adult human body is fluid, which is mainly distributed between two compartments: the extracellular space and the intracellular

From: *Proteomics of Human Body Fluids: Principles, Methods, and Applications*
Edited by: V. Thongboonkerd © Humana Press Inc., Totowa, NJ

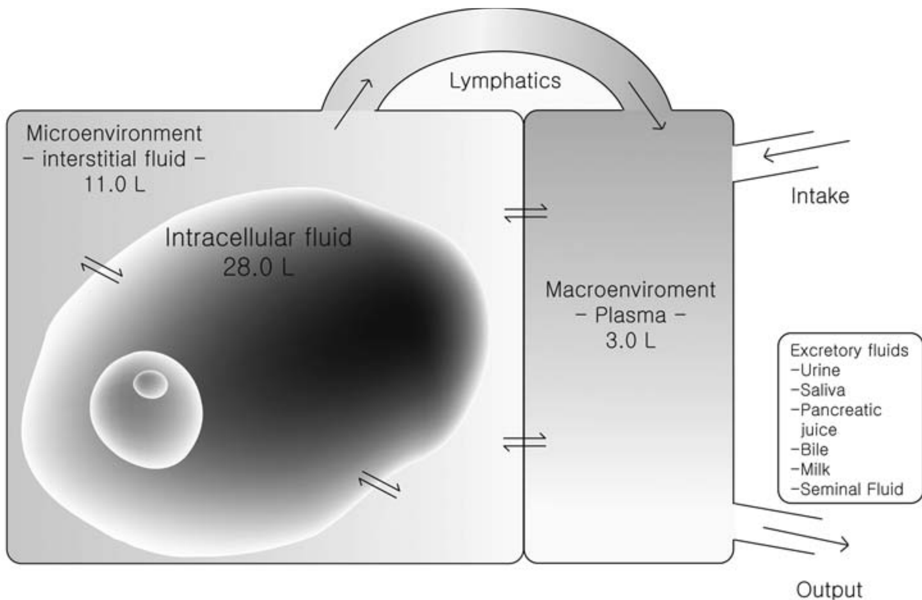

Fig. 1. Overview of body fluids and the internal environment. In an average 70-kg person, the total body water is about 60% of the body weight (approx 42.0 L). The intracellular fluid constitutes about two-thirds of the total body water (approx 28.0 L), whereas the extracellular fluid constitutes the rest (approx 14.0 L). The interstitial fluid comprises more than 75% of the extracellular fluid (approx 11.0 L), whereas plasma comprises the remaining 25% (approx 3.0 L) ***(3)***.

space ***(3)***. The extracellular fluid is broadly divided into the interstitial fluid and the blood plasma, which can be referred to as the microenvironment and the macroenvironment, respectively **(Fig. 1)**. Tissues consist of cellular elements (parenchymal and stromal cells) and extracellular elements (extracellular matrix and tissue interstitial fluid). In the literature, the term *tissue microenvironment* usually refers to both cellular and extracellular elements ***(4)***. In this chapter, however, microenvironments are limited to tissue interstitial fluid (TIF) only, which surrounds and bathes tissues. Since parenchyma, stroma, and blood all contribute to the microenvironment, their individual secreted or shed protein profiles are reflected together in the overall protein profile of the microenvironment. The microenvironment, the interstitial fluid, is in direct contact with cells, exchanging molecules with the intracellular fluid, whereas the macroenvironment, the plasma, continuously communicates with all microenvironments throughout the body, delivering nutrients and signals and receiving feedback directly or indirectly via the lymphatics.

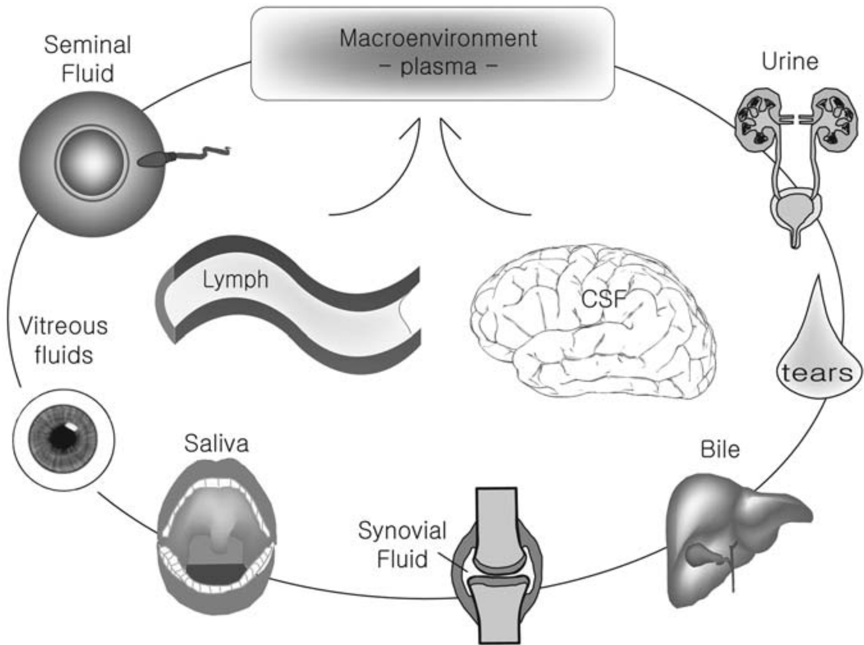

Fig. 2. Relationship between plasma and other body fluids. Tissue interstitial fluid (TIF) enters terminal lymphatics and becomes lymph. (Thus, the composition of lymph in the terminal lympatics is almost identical to that of TIF.) Hence, all lymph in the body empties into the venous system at the junctures of internal jugular veins and subclavian veins ***(94)***. CSF is secreted from the choroid plexus at a rate of about 500 mL/d, which is approximately three to four times as much as its total volume (about 150 mL). Excess CSF is absorbed by the arachnoidal villi, which have vesicular holes large enough to allow the free flow of CSF, proteins, and even particles as large as red and white blood cells into the venous system ***(95)***. Since lymph and CSF drain into the blood, theoretically all proteins in lymph and CSF are present in the blood. For this reason, lymph and CSF proteomes are considered subproteomes of the plasma. Other body fluids such as urine and tears represent plasma to varying extents while having unique characteristics of their own.

Plasma is important in any proteomic analysis of human body fluids, not only because every cell in the body leaves a record of its physiological state in the products it sheds into the blood ***(5)***, but also because it influences most other body fluids. Therefore, it is important to understand the proteomes of various body fluids in the context of plasma. Theoretically, lymph and cerebrospinal fluid (CSF) proteomes are subproteomes of plasma since these body fluids eventually drain into plasma. Other body fluids represent plasma to varying extents while having unique characteristics of their own (**Fig. 2**).

1.2. Differential Enrichment of Biomarkers in Body Fluids

One of the main challenges of proteomics is to find molecular signatures or biomarkers of disease. In plasma, high-abundance proteins such as albumin and transferrin constitute approx 99% of the total protein and the remaining 1% is assumed to include many potential biomarkers that are typically of low abundance ***(6)***. Therefore, removal of high-abundance proteins has become a common practice to enrich for low-abundance proteins in plasma. (This issue will be discussed in more details later in this chapter, as well as in other chapters.) However, before trying to remove high-abundance proteins from plasma, the concept of differential enrichment of biomarkers in various body fluids needs to be considered. **Figure 3** illustrates a simplified relationship of the concentration of secreted or membrane-shed cellular proteins in TIF, lymph, and blood. For example, if there is a cancer in the sigmoid region of the colon, cancer cells will secrete or shed cancer-specific proteins into the microenvironment. Such proteins traffic from the TIF to the lymph, being diluted during the process. Lymph fluids from various regions of the body merge and eventually drain into the circulatory system. Approximately 2.5 L of lymph drains into the systemic circulation per day, whereas about 3 L of plasma (approx 5 L of blood) is ejected from the heart every minute. Therefore, the dilution factor is at least 1.5×10^3. (Lymph fluids from different tissues have different tissue-specific proteins. This additional consideration is not included here). Given that only a 10-fold enrichment can be achieved by removing the top six most abundant proteins in plasma, the advantage of using TIF ***(7)*** or lymph ***(8)*** rather than plasma seems considerable in discovering biomarkers. For example, the study of Sedlaczek and colleagues ***(9)*** highlights the differential enrichment of CA125, an ovarian cancer marker, in different body fluids from patients with ovarian carcinoma. **Table 1** summarizes their comparative analysis of CA125 in sera, cyst fluids, and ascites. According to this study, the median value of CA125 is approx 64-fold higher in cyst fluid than in serum.

Malignant ascites is another example of differential enrichment of secreted or membrane-shed proteins. Some cancers such as colorectal and ovarian cancers can be seeded onto peritoneal cavity and cause malignant ascites via various mechanisms. According to Trape and colleagues ***(10)***, carcinoembryonic antigen (CEA) levels in malignant ascites are in the range of 33,540 ng/mL maximum, which is more than 5×10^3-fold higher than the normal plasma level of CEA ($<$5 ng/mL). Although the availability of clinical specimens often becomes the bottleneck of body fluid research owing to a paucity of clinical specimens and ethical considerations, understanding and utilizing the differential enrichment of biomarkers may open a new window of opportunity for discovering otherwise undetectable low-abundance biomarkers.

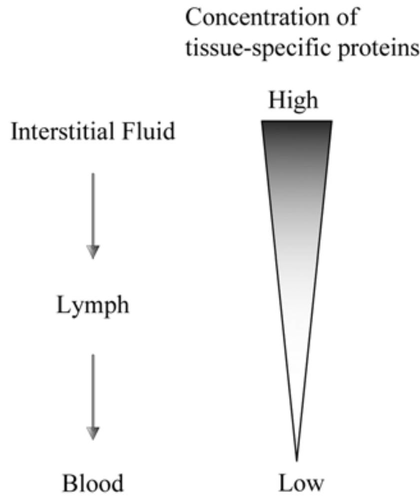

Fig. 3. Dilution of tissue-specific proteins in TIF, lymph, and blood.

Table 1
Median Levels of Ca125 in Serum, Cyst Fluid, and Ascites From Patients With Ovarian Neoplasms

	CA125 (U/mL)		
Histologic type	Serum	Ascites	Cyst fluid
Serous carcinoma	696.0	18,563.0	44,850.0
Endometroid carcinoma	661.0	14,415.5	32,150.0
Mucinous adenocarcinoma	67.0	3521.5	3930.5
Undifferentiated carcinoma	860.7	3909.5	—
Serous cystadenoma	7.1	—	42150.0
Serous cyst	4.8	—	6851.5
Mucinous adenoma	10.8	—	5691.5

From Sedlaczek et al. *(9)*, with permission.

2. Proteomic Approaches for Studying Human Body Fluids

Proteomics, a newly emerging postgenomic technology that allows one to unravel the biological complexity encoded by the genome at the protein level, is built on technologies that allow one to analyze large numbers of proteins in a single experiment. Broadly, there are two main facets of proteomics research:

1. *Expression proteomics,* which aims to catalog the proteome, i.e., the full complement of proteins expressed by the genome in any given cell, tissue, or body fluid at a given time.

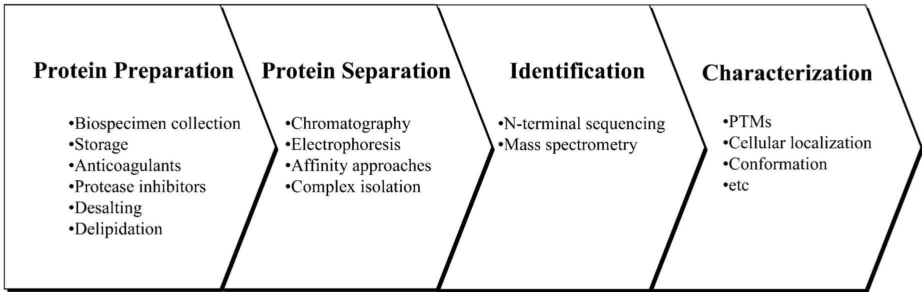

Fig. 4. Pipeline of technologies used in the field of proteomics. The overall success of qualitative and quantitative proteomics relies on the success of the individual technologies involved in a proteomic pipeline. PTMs, posttranslational modifications.

2. *Targeted proteomics,* which strives to determine the cellular functions of genes directly at the protein level (e.g., protein-protein interactions, posttranslational modifications, protein localizations within cells) ***(11)***.

Currently, the major focus of proteomics of body fluids is expression proteomics, especially the quantitative differences in protein profiles between physiological and pathological states.

Figure 4 summarizes the pipeline of technologies that comprise the field of proteomics strategies. Each step in the pipeline involves defined technologies, each of which is technically challenging and of equal importance. Needless to say, the overall success of any proteomics research depends on the success of the individual step in the proteomic pipeline. For solving specific biological questions, the combination of various options in each step provides much flexibility in experimental design. In this chapter, we address current issues and technologies involved in each step of proteomics with an emphasis on their application to body fluid research.

2.1. Biospecimen Collection and Storage

Success in proteomics very much depends on careful biospecimen preparation. In clinical chemistry, many factors are known to cause variations in biospecimen precollection, collection, and postcollection stages ***(12)***. Therefore, a standardized protocol for sample collection, processing, and storage is essential for reproducible experiments within a given laboratory and, especially, from one laboratory to another. (When we compare a large set of data from different laboratories, it is important that we are actually comparing "apples with apples!"). This was one of the major technical issues addressed in the pilot phase of the Human Plasma Proteome Project, the first systematic international effort devoted to analyzing a body fluid ***(13)***.

Among the various body fluids, blood is the most sampled and studied, yet its optimal sample preparation is still problematic. For proteomic analysis, blood can be collected as serum or plasma. When blood is removed from the body and allowed to clot, it separates into a solid clot containing blood cells and fibrin, and a liquid phase termed serum. If an anticoagulant such as heparin or EDTA is added, the liquid phase is termed plasma ***(12)***. From a clinical chemistry perspective, serum differs from plasma only in that it lacks fibrinogen. From a proteomics perspective, however, the differences between serum and plasma can be considerable. The physiological and biochemical difference between serum and plasma is demarcated by the activation of the coagulation cascade, which involves the sequential activation of proteases ***(14)***. The activated proteases during this process will in turn have proteolytic effects on other proteins. According to a recent report of the Human Plasma Proteome Organisation (HUPO), a significant number of peptides differed between serum and plasma specimens (especially intracellular, coagulation-dependent, and enzymatic activity-derived peptides) ***(15)***. The issue of coagulation can also be applied to other body fluids. Extravascular coagulations are observed in lymph ***(16)*** and synovial fluid ***(17)***, and it is likely that most of the internal body fluids have coagulation factors from blood to some extent.

Hulmes and colleagues ***(18)*** have addressed questions regarding plasma collection, stabilization, and storage procedures for proteomic analysis of clinical samples. According to their research, addition of a protease inhibitor cocktail directly to plasma collection tubes prior to phlebotomy, centrifugation within 1 h of blood draw, snap-freezing aliquots immediately in a dry ice/alcohol bath, and storing frozen aliquots in a −70°C freezer can improve sample qualities for proteomic analysis. This recommendation is supported in the report of the HUPO Plasma Proteome Project on specimen collection and handling ***(15)***.

There are a number of anticoagulants that prevent the coagulation of blood. In clinical chemistry, EDTA, heparin, and citrate are the most widely used, and the choice of anticoagulants is important since the manner in which they behave differs. Unlike EDTA and heparin, citrate is used as a concentrated solution in a ratio of 1 part to 9 parts of blood ***(19)***, which itself introduces dilution effects and variation. Heparin is a highly charged molecule, thus being able to prevent binding of molecules to charged surfaces ***(15)***. Although EDTA can interfere with assays when divalent cations are necessary, it does not have dilution effects nor does it interfere with charged molecules. Therefore, EDTA seems to be the anticoagulant of choice for proteomic analysis of body fluids when the primary aim is to catalog and quantitate proteins. However, the choice of anticoagulants may also depend on the specific aim or protein targets of experiments since anticoagulants can affect the stability of some proteins, if not all (e.g., osteocalcin) ***(20,21)***. Heparin, citrate, and EDTA have been reported to yield no obvious *m/z* (mass per charge) peaks in typical proteome analysis, yet some types of blood

collection tubes designed to reduce protein degradation contain aprotinin or other protease inhibitors that will appear as *m/z* peaks and pose a potential problem with the interpretation of mass spectra if they are not recognized as exogenous additives to a specimen ***(22)***.

With regard to the use of protease inhibitors, all data from the HUPO Plasma Proteome Project on specimen collection and handling are consistent with the benefits of blocking protease activity and, perhaps more importantly, of blocking this activity immediately, during sample acquisition ***(15)***.

Finally, the limitations of current storage methods using −70 to −80°C freezers are worth mentioning. It has been reported that some degree of degradation occurs over time in coagulation factors of stored plasma samples, presumably owing to renewal of enzymatic activity, albeit minimal, even at −80°C ***(23)***. In this context, Rouy and colleagues ***(24)*** reported that the plasma level of metalloproteinase-9 (MMP-9) decreased by 90% after 2 yr of storage at −80°C, whereas those of MMP-2 remained constant. It is surprising that two enzymes, which share many properties, behave in different manners under the same storage condition. Therefore, careful validation and interpretation are essential when we analyze a large set of body fluid samples stored in tissue banks over a period of time since at least some proteins may show different levels of stability. Topics relating to specimen collection and storage of other body fluids are dealt with in other chapters of this book.

2.2. Sample Loading: How and What to Compare

Typically, when cell or tissue lysates are subjected to proteomic analysis, equal amounts of protein are compared (e.g., 100 μg protein from each sample for 2D electrophoresis [2-DE]). In body fluid research, however, the analysis of samples based on equal protein load may cause serious problems because even the normal interval of total protein levels is very wide (e.g., it ranges from 68.0 to 86.0 mg/mL in plasma). To illustrate this potential problem, let us consider these two hypothetical patients.

Patient A

Total plasma protein	86 mg/mL
CEA	4.9 ng/mL

Patient B

Total plasma protein	68 mg/mL
CEA	4.9 ng/mL

(CEA is a tumor marker for colorectal cancer; the normal range is <5 ng/mL.)

In current medical practice, total plasma protein levels are not considered when we interpret individual protein levels (i.e., they are treated as independent variables). Therefore, CEA levels of both patients will be regarded as normal.

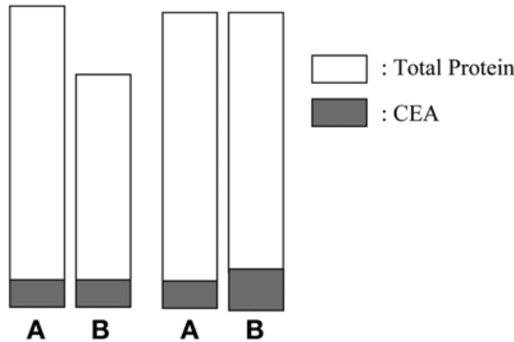

Fig. 5. Volume-based vs protein quantity-based comparison. Left panel illustrates the comparison of the same volumes of plasma from patients A and B. Total protein levels differ, yet carcinoembroyonic antigen (CEA) levels are the same. If we were to analyze the same quantities of proteins instead (i.e., 86 μg of protein for each patient), patterns will look like the right panel. In this case, total protein levels are the same (since we loaded the same quantities), yet CEA levels differ. When the same quantities of samples are compared, it is assumed that the total protein levels in patients' plasma are more or less identical. This assumption does not hold true for most body fluids.

Now, let us assume that we perform 2-DE using plasma from these two patients and that the dynamic range of detection is approx 10^{12}. (In reality, it is about 10^4.) Then, to load 86 μg of each sample to immobilized pH gradient (IPG) strips, we load 1 μL of plasma for patient A, which contains 4.9 pg of CEA, and 1.26 μL of plasma for patient B, which contains 6.2 pg of CEA. After 2-DE, the CEA spots will be selected as differentially expressed spots. (Differential gel electrophoresis [DIGE] can detect quantitative changes as low as 10% [25]). In this approach, over 20% of variation is introduced because the interpretation of CEA levels is dependent on total protein levels. In other words, CEA levels or other biomarker levels, which we try to detect, can vary according to the total protein concentrations **(Fig. 5)**.

When we deal with cell or tissue lysate, the situation is totally different. For example, if we compare radiation-treated with nontreated CaCo2 cell lines using 2-DE, we are trying to detect changes in essentially identical systems, and loading the same amount of samples (e.g., 100 μg protein from each sample) will be a reasonable way to guarantee equal comparison.

There are two ways of solving this problem. The first is by simply loading the same volume of body fluids. Although this method is perfectly compatible with the current practice of clinical laboratories, it may not be an ideal solution for expression proteomics of body fluids wherein prefractionation, such as depletion of high-abundance proteins, is commonly required. When depletion is used, for example, volumetric information is difficult to preserve and invariably lost. If we

use multiple affinity columns for depleting high-abundance proteins, volumetric information pertaining to original samples would be lost, first, by dilution during chromatographic separation and, second, by a desalting and concentrating step.

The second solution is to normalize data based on the total protein concentrations. This method provides information about relative concentration of proteins, and data can be normalized even after extensive prefractionation since quantitative information is easier to preserve.

Let us go back to the example of CEA above. In that example, more than 20% of variation was introduced just because we loaded the same quantities of protein from each sample assuming, willingly or unwillingly, that total protein levels of two samples were identical. This systematic variation can be easily corrected by calculating a normalization factor from total protein levels of each sample and applying it to the data.

If we select the sample from patient A as a baseline, the normalization factor will be 68 mg/mL divided by 86 mg/mL. Then, the CEA level of patient B will be corrected to be 4.9 pg by multiplying the compensation factor to the original data $[(68/86) \times 6.2 = 4.9]$.

If we keep track of quantitative information in each prefractionation step, it is possible to calculate proper normalization factors. For example, after depleting high-abundance proteins using affinity chromatography, we can get information about how much protein is depleted from the total proteome (e.g., 85% is depleted), which can be used for normalizing data.

2.3. Prefractionation and Fractionation

The development of proteomics technologies has enabled us to analyze a large number of proteins simultaneously. 2-DE, arguably the most widely used separation technique in proteomics ***(26,27)***, can resolve more than 5000 proteins in one gel and detect less than 1 ng of proteins per spot ***(28)***. Nevertheless, at least two technical challenges need to be overcome before proteomics can realize its full potential for protein expression profiling of body fluids. First, body fluids contain an enormous number of proteins. For example, it is claimed that plasma alone contains more than 1 million protein forms ***(6)***. Another important consideration is the problem of the dynamic range of protein abundances ***(29)***. In plasma, the dynamic range of protein abundances can extend up to 12 orders of magnitude ***(30)***, which far exceeds the current dynamic range of 2-DE (approx 10^4) ***(31)***. To circumvent these problems, good separation strategies are essential. The essence of prefractionation is the enrichment of the target population of proteins (e.g., removal of high-abundance proteins and/or the isolation of subpopulation of proteins—e.g., glycoproteins, phosphoproteins, glycosylphosphatidylinisotol (GPI)-anchored proteins, cysteine-containing proteins) whereas the essence of fractionation is the maximal separation of a complex protein

mixture into its individual components (e.g., 2-DE and multidimensional chromatography). Prefractionation and fractionation technologies are too complicated for a single review. For more detailed reviews, *see* Simpson ***(32)*** and Righetti et al. ***(33)***.

2.3.1. Reduction of Dynamic Range of Protein Abundances

Various human body fluids including plasma, CSF, ascites, and lymph are characterized by the presence of high-abundance proteins, which preclude effective analysis of low-abundance proteins (akin to searching for "needles" in a haystack). For example, 22 high-abundance proteins represent about 99% of the total proteins in plasma ***(34)***. Therefore, in any proteomic strategy for analyzing body fluids, the reduction of the dynamic range of protein abundances must be addressed in order to "drill down" to the low-abundance proteins for analysis. Two opposite approaches will be briefly introduced here. One approach reduces the dynamic range by depleting high-abundance proteins; the other achieves the goal by increasing the relative copy numbers of low-abundance proteins. These two may be called the *yin* and *yang* approaches to reducing the dynamic range of protein abundances.

2.3.1.1. YIN APPROACH OF REDUCING DYNAMIC RANGE OF PROTEIN ABUNDANCES

2.3.1.1.1. Depletion of High-Abundance Proteins. Depletion of high-abundance proteins is probably the most commonly used prefractionation technique for body fluid research. (Govorukhina and Bischoff discuss it in more detail in Chapter 2.) As just mentioned, this approach aims to reduce the dynamic range of protein abundances by removing high-abundance proteins, and it has been successfully adopted for body fluid research. For example, Pieper and colleagues ***(35)*** have shown that immunoaffinity subtraction chromatography can improve the resolution of low-abundance proteins in plasma. Although this approach is very useful in body fluid research, there are two issues that require careful consideration. The first is the limitation of this approach, which becomes evident if we take albumin, a major high-abundance protein in various body fluids, with about 50 mg/mL in plasma, as an example ***(36)***. If any depletion strategy were able to remove 99.9% of albumin from the plasma (according to a recent report, the efficiency ranges from 96.0 to 99.4% ***[37]***), the remaining (contaminating) concentration of albumin would still be approx 50 μg/mL. This concentration is 1×10^4 fold higher than CEA levels (approx 5 ng/mL) and 5×10^6 fold higher than levels of interleukin-6 (IL-6; approx 10 pg/mL). Considering that the current dynamic range of 2-DE is approx 10^4 ***(31)***, this simple comparison shows that in addition to the depletion of high-abundance proteins, technologies for the efficient separation and

enrichment of low-abundance proteins to detectable levels have to be further utilized and developed ***(38)***.

The second issue is the possibility that depletion of high-abundance proteins may diminish the chances of finding low-abundance proteins bound to and carried by high-abundance carrier proteins such as albumin ***(39)***. Although this concept is still controversial and has not proved its importance, we should be careful not to lose extra information when subtracting a portion of proteome before initial analysis.

2.3.1.2. Yang Approach of Reducing Dynamic Range of Protein Abundances

2.3.1.2.1. Reduction of Dynamic Range With a Peptide Library. This approach (enrichment of the general population of proteins to the same degree) involves constructing a large peptide library via combinatorial chemistry. Using just 20 natural amino acids and making six reaction steps, 20^6 peptide ligands can be made; owing to this enormous diversity, there is theoretically a ligand for every protein, antibody, and peptide. When a complex protein mixture such as plasma is incubated with this library under large overloading conditions, high-abundance proteins saturate their specific affinity ligands and excess is removed during the washing step, whereas low-abundance proteins continue to concentrate on their specific affinity ligands. After processing, the eluate has all the representatives of the original mixture, but with much reduced dynamic range since high-abundance proteins are significantly diluted and low-abundance proteins are concentrated ***(33)***. Although the efficiency and efficacy of this approach are not yet clear in this early stage, it is free from potential problems associated with depletion and may work as a complementary method to depletion strategy.

2.3.1.2.2. Reduction of Dynamic Range With Selective Capture Methods. This approach involves enrichment of the selective population of proteins. The peptide library mentioned in **Subheading 2.3.1.2.1.** reduces the dynamic range of protein abundances by enriching the general population of proteins to the same degree. There are, however, other methods by which we can enrich selective populations of proteins using the unique characteristics of each group. Here, we briefly introduce two examples, which are important for analyzing body fluids.

Immunoprecipitation is the most classical example of enriching a selective population of proteins. This technique is based on the immunoaffinity between antibodies and their target proteins. Immunoprecipitation has been successfully applied to the analysis of protein isoforms, phosphorylated proteins, and protein-protein complexes ***(33)***.

Glycoprotein capture is another good example. This method, specifically targeting glycoproteins, is highly relevant to body fluid research, since protein

glycosylation is prevalent in extracellular proteins, and many clinical biomarkers in body fluids are also glycoproteins ***(40)***. Currently, there are two main approaches for capturing glycoproteins. Kaji and colleagues ***(41)*** combined the classical lectin chromatography with isotope-coded tagging and mass spectrometry (MS). In this method, glycopeptides, generated by tryptic digestion of protein mixture, are captured by the lectin column. Then captured glycopeptides are isotope-tagged with ^{18}O and identified by multidimensional liquid chromatography (LC) MS. In the other approach, Zhang and colleagues ***(40)*** used hydrazide chemistry to capture glycoproteins through conjugation, which is followed by isotope labeling and identification/quantitation by tandem mass spectrometry (MS/MS).

2.3.2. Electrophoresis and Liquid Chromatography

Electrophoresis and liquid chromatography are two main streams of separation technology. In this section, only a limited number of topics will be discussed, with an emphasis on their application to body fluids. For more detailed reviews, *see* Westermeier and Grona ***(42)***, Simpson ***(43)***, and Mant and Hodges ***(44)***.

2.3.2.1. Isoelectric Focusing

Proteins are amphoteric molecules that carry a positive, negative, or zero net charge, depending on the pH of their surroundings. Therefore, when placed in a pH gradient within an electric field, proteins will migrate to the pH where they have no net charge. Isoelectric focusing (IEF) takes advantage of this phenomenon ***(45)***.

In addition to its well-known application as the first dimension analysis of 2-DE, IEF can also be used as a prefractionation technique or in combination with non-gel-based technologies such as liquid chromatography. Here we will briefly introduce three applications of IEF: prefractionation IEF for narrow-range IPGs, free-flow electrophoresis (FFE), and chromatofocusing.

2.3.2.1.1. Prefractionation-IEF for Narrow-Range Immobilized ph Gradients. The use of multiple narrow overlapping IPGs is the best remedy for increasing the resolution of 2-DE to avoid multiple proteins in a single spot for unambiguous protein identification and to facilitate the application of higher protein amounts for the detection of minor components ***(46)***.

When narrow-range IPGs are loaded with a body fluid (e.g., plasma), a massive disturbance of the focusing process ensues, stemming from two main problems. The first problem is that it is usually not possible to achieve high loads of protein, actually focused, on narrow pH gradients since most of the loaded proteins have p*I*s outside the pH range of the IPG. The second problem is the severe disturbance caused by extraneous proteins, which migrate to the

ends of the strip, where they collect in highly concentrated zones in charged states. Therefore, it is essential to prefractionate body fluids into isoelectric fractions that correspond to the pH ranges of the IPGs ***(47)***.

To achieve this goal, various liquid-phase IEF apparati such as the Rotofor™ ***(48)***, the multicompartment electrolyzer ***(49)***, and the zoom fractionator ***(50)*** have been developed. In general, these apparati, except the Rotofor, have multiple compartments separated by separation barriers with a defined pH, and the pH range of a fraction in a compartment is determined by the pH of separation barriers at both ends. After IEF, each fraction can be loaded to corresponding narrow-pH IPG strips. For a more detailed discussion of each technique, see Righetti et al. ***(33)*** and Zuo and Speicher ***(51)***. Recently, Tang and colleagues ***(52)*** reported on four-dimensional analysis, which combines the depletion of high-abundance proteins, liquid-phase IEF, and 1-DE, followed by nanocapillary reversed phase high-performance liquid chromatography (RP-HPLC) tryptic peptide separation prior to MS/MS analysis to detect low-abundance proteins in human plasma and serum.

Görg and colleagues ***(53)*** developed a solid-phase prefractionation IEF using granulated gels. In brief, a Sephadex slurry is made with Sephadex G-200 superfine and a solution containing urea, 3-[(3-cholamidopropyl)dimethylamino]-1-propanesulfonate (CHAPS), dithiothreitol (DTT), and carrier ampholytes. This slurry is mixed with the sample solution, and the mixture is pipeted into the trough of the template inserted into the IPG DryStrip kit for IEF. After IEF, individual Sephadex fractions are removed with a spatula and applied onto rehydrated, narrow-pH-range IPG strips. When IPG-IEF is performed, prefractionated proteins in the Sephadex fraction are electrophoretically transferred to IPG strips and focused. This method does not require special equipment and is relatively free from protein dilution and loss, which may occur in liquid-phase IEF.

2.3.2.1.2. Free-Flow Electrophoresis and Chromatofocusing. Although the prefractionation-IEF methods introduced above are specifically devised for 2-DE, FFE, and chromatofocusing, they provide better results in combination with liquid chromatography; they can also be used with gel-based technologies. Both FFE and chromatofocusing are liquid-phase IEF, and, as their names indicate, the basic principles of FFE and chromatofocusing are based on electrophoresis and chromatography, respectively.

In FFE, samples are continuously injected into a carrier ampholyte solution flowing as a thin film between two parallel plates. By introducing an electric field perpendicular to the flow direction, proteins are separated by IEF according to their different p*I* values and collected ***(29,54)***. Key advantages of this method are improved sample recovery (owing to the absence of solid membrane

supports) and sample loading capacity. (Sample loading is continuous during FFE and hence not rate limiting ***[29]***).

Although FFE can be coupled off-line to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) ***(54)***, the restricted separation capacity of SDS-PAGE presents limitations in resolution, recovery of low-M_r proteins, and sample loadability. In contrast, when FFE is coupled off-line to RP-HPLC, the high resolving power produced in the first-dimension IEF step, in which very narrow-range pH gradients can easily be generated, coupled to the high resolution of modern RP-HPLC stationary phases, extends the resolving power of this 2D protein separation system over other previously described 2D systems based solely on coupled HPLC columns ***(29)***. For the fundamental principles and experimental protocols of FFE, including the introduction of commercial instrumentation, *see* Krivankova and Bocek ***(55)*** and Weber et al. ***(56)***. For a detailed protocol for the application of FFE for proteins and peptides, *see* Moritz and Simpson ***(57)***.

In chromatofocusing, usually a weak anion exchanger is used as the matrix in which the functional groups are amines, and the eluent is a buffer containing a large number of buffering species, which together give a uniform buffering capacity over a broad pH range. Unlike ion-exchange chromatography, in which a pH gradient is normally formed using a gradient mixer, chromatofocusing takes advantage of the buffering action of the charged functional groups on the matrix, and the pH gradient is formed automatically as the eluting buffer titrates the functional groups on the matrix. As elution progresses, the pH at each point in the column is gradually lowered, and proteins with different p*I* values will migrate at different distances on the column before binding. In this way, proteins elute in the order of their p*I*s ***(58)***. Yan and colleagues ***(59)*** used chromatofocusing coupled to nonporous (NPS) RP-HPLC for fractionating and comparing protein expression using a drug-treated cell line vs the same untreated cell line. This method provides a 2D map based on p*I* values and hydrophobicity and has been shown to be highly reproducible for quantitative differential expression analysis. Soldi and colleagues ***(60)*** used a commercial platform combining chromatofocusing and NPS RP-HPLC for protein profiling of human urine and showed that this method could be a complementary system to 2-DE in body fluid research.

2.3.2.2. 2D Electrophoresis

2-DE has been the most commonly used technique in the field of proteomics since its development in 1975 by O'Farrell ***(26)*** and Klose ***(27)***. This technique couples IEF in the first dimension with SDS-PAGE in the second dimension and allows the separation of complex mixtures of proteins according to their respective p*I* and M_r values. Depending on the gel size and pH gradient used, 2-DE can resolve more than 5000 proteins simultaneously (more than 2000

proteins routinely) and can detect less than 1 ng of protein per spot ***(46)***. Since 2-DE has suffered from problems such as reproducibility, resolution, proteins with extremes of p*I*, and recovery of hydrophobic proteins ***(61)***, it is now seriously challenged by other non-gel-based approaches. However, as Rabilloud ***(31)*** pointed out, if the goal of the proteomic experiment is to look for quantitative changes, 2-DE will remain unrivalled for some time.

Body fluid research is closely related to finding disease biomarkers, and therefore, quantitative analysis of differentially expressed proteins in normal and disease groups is important. Although 2-DE is a good quantitative tool, its ability has been hampered by important limitations. First, the predominant protein staining methods are either not sensitive enough (Coomassie brilliant blue) or have a limited linearity (silver staining) ***(62)***. The application of radioactive labeling or fluorescent stains can alleviate these problems, but only partially. Second, the intrinsic gel-to-gel variation of 2-DE masks the biological difference between the samples and compromises any quantitative comparison of protein expression levels ***(63)***. DIGE ***(64)*** circumvents many of the issues associated with traditional 2-DE, such as reproducibility and limited dynamic range, and allows for more accurate and sensitive quantitative analysis ***(65)***.

In DIGE, two samples are labeled in vitro using two different fluorescent cyanine dyes (CyDyes; Amersham Biosciences) differing in their excitation and emission wavelengths, then mixed before IEF, and separated on a single 2D gel. After consecutive excitation with both wavelengths, the images are overlaid and subtracted (normalized), whereby especially differences (e.g., up- or down-regulated, and/or posttranslationally modified proteins) between the two samples can be visualized ***(46)***. This multiplex approach instead of the "one gel one sample" approach solves most of the problems associated with gel-to-gel variation, spot matching, and normalization. In addition, CyDye has a detection limit of 150 to 500 pg for a single protein with a linear response in protein concentration over 5 orders of magnitude, whereas silver staining has a protein detection limit of approx 1 ng with a dynamic range of less than 2 orders of magnitude ***(65)***. At least five replicate gels should be run per sample for quantitative analysis in traditional 2-DE, and owing to the high variability from sample comparisons run in different gels, the threshold for accepting a meaningful variation is set at a factor of 2.0 (100% variation) ***(66)***. In contrast, DIGE can detect quantitative changes as low as 10% with 95% confidence, and the use of an internal standard helps to minimize false positives and negatives ***(25)***.

For biomarker research using body fluids, protein profiling of a large set of samples is essential. In this case, the primary benefit of sample multiplexing is that a pooled standard can be included on each gel, which comprises equal amounts of each sample and represents the average of all the samples being analyzed. The pooled standard approach is used to normalize protein abundance

Table 2
Commonly Used Liquid Chromatography Methods in Proteomics

Principle of separation	Type of chromatography
Size and shape	Size-exclusion chromatography (gel-filtration or gel-permeation chromatography)
Net charge	Ion-exchange chromatography
Hydrophobicity	Hydrophobic interaction chromatography
	Reversed-phase high-performance liquid chromatography
Antigen-antibody interaction	Immunoaffinity chromatography
Isoelectric point (p*I*)	Chromatofocusing
Metal binding	Immobilized metal ion affinity chromatography

measurements across multiple gels in an experiment, making it possible to compare more than two samples accurately ***(67)***. A more detailed review of multiplexed dye technologies is presented by Patton in chapter 4.

2.3.2.3. Liquid Chromatography

Chromatography is a widely used technique for separating the components of a mixture by allowing the sample (the analyte) to distribute between the stationary and mobile phases. Stationary phases, the key elements of LC, are made of the support matrix chemically coated with a bonded phase containing functional groups that provide the desired specific binding interaction ***(68)***. **Table 2** summarizes commonly used chromatographic methods and their principles of separation. Chromatography can be used for the enrichment of low-abundance proteins as well as for multidimensional analysis of body fluids. For a more detailed review of chromatographic methods for separating proteins and peptides, *see* Simpson ***(32)*** and Mant and Hodges (***44***).

2.3.2.3.1. Chromatographic Prefractionation for 2-DE. As mentioned above, immunoaffinity chromatography is the most commonly used prefractionation tool in body fluid research. However, other chromatographic methods such as ion-exchange chromatography and RP-HPLC can also be considered for prefractionation. Combining a chromatographic step with 2-DE provides a third orthogonal dimension for protein separation. For example, if we use RP-HPLC before 2-DE, we are separating proteins based on a combination of their hydrophobicity, p*I,* and M_r. For an overview of chromatographic prefractionation prior to 2-DE, *see* Lescuyer et al. ***(69)***.

2.3.2.3.2. Multidimensional Analysis Using Chromatography. As we discussed in chromatofocusing, multiple chromatographic methods can be coupled for separating the components of a mixture (e.g., combination of chromatofocusing

Table 3
MDLC of Intact Proteins versus MudPIT

	MDLC	MudPIT
Peptide fragment correlation	Good correlation between peptides and their original protein	Poor correlation owing to digestion prior to fractionation
Mixture complexity	Still complex	About 50-fold more complex
Computational requirements	Less challenging; MS/MS data search can be supported by elution profiles and protein properties.	One of the biggest challenges
Solubility	Problematic	A significant advantage over MDLC

Abbreviations: MDLC, multidimensional high-performance liquid chromatography; MS, mass spectrometry; MudPIT, multidimensional protein identification Technology.

and RP-HPLC). This approach, termed multidimensional HPLC (MDLC), can be fully automated to join the various separation steps into a single seamless procedure and can also interface protein and peptide separations directly to mass spectrometers. In MDLC of intact proteins, the protein complex is fractionated and digested to peptides for subsequent mass spectrometric analysis. In an alternative MDLC, termed multidimensional protein identification technology (MudPIT), complex protein samples are enzymatically digested to produce extremely complex peptide mixtures, which are then subjected to multidimensional chromatographic separations and mass spectrometric analysis. For detailed reviews of MDLC of intact proteins and MudPIT, see Apffel ***(70)*** and Wolters et al. ***(71)***. **Table 3** presents a summary of comparisons between MDLC of intact proteins and MudPIT. This table is based on Apffel's review ***(70)***.

2.4. Mass Spectrometry

MS has become the method of choice for the identification and characterization of proteins in complex mixtures, largely as a result of the development of soft ionization methods for proteins and the availability of gene and genome sequence databases ***(11)***. In expression proteomics of body fluids, the main applications of MS are determination of primary structure of peptides, quantitative analysis, and characterization of posttranslational modifications. For a review of MS-based proteomics, see Aebersold and Mann ***(72)***, and for a "hands-on" description of current MS-based proteomics methods, see the proteomics laboratory manual of Simpson ***(43)***. Glycosylation, an important posttranslational modification in body fluids, will be discussed by Bunkenborg and colleagues in chapter 5.

Table 4
Current Approaches to Protein Identification Using Mass Spectrometry (MS)

MS approach	Features
Top-down	Analysis of the fragmentation pathway of intact proteins Complete sequence coverage (useful for examining site-specific mutations and posttranslational modifications) ***(96)*** Requirement of specialized equipment Early stage of development
Bottom-up	Digestion of proteins to peptides prior to MS Limited sequence coverage, but improved sequencing properties and detection efficiencies of peptide ***(97)***
PMF	Comparison of peptide mass fingerprints with virtual fingerprints obtained by theoretical cleavage of protein sequences in databases
MS/MS	Peptide sequencing by analyzing the fragmentation patterns of peptides More sensitive and specific than PMF

Abbreviations: PMF, peptide mass fingerprinting; MS/MS, tandem mass spectrometry

2.4.1. Protein Identification

2.4.1.1. Protein Identification Using Mass Spectrometry

Currently, the bottom-up approach using MS/MS is most widely used for protein identification. In MS/MS, peptide ions are isolated, fragmented, and analyzed to produce MS/MS spectra. Then these experimental MS/MS spectra are compared with the theoretical MS/MS spectra generated from protein sequence databases using search algorithms, which assign scores indicating the degree of similarity between the experimental and theoretical MS/MS spectra ***(73)*** **(Table 4)**.

However, it has been increasingly realized that the protein inference problem, i.e., the task of assembling the sequences of identified peptides to infer the proteins of their origin, is far from being trivial and requires special attention. Protein digestion makes peptides, not proteins, the currency of MS/MS, and the connectivity between peptides and proteins is lost at the digestion stage, which complicates computational analysis and biological interpretation of the data, especially in the case of higher eukaryotic organisms in which the same peptide sequence can be present in multiple different proteins ***(74)***.

2.4.1.2. MS/MS Search Algorithms for Protein Identification

Currently MS/MS search algorithms scoring functions can essentially be classified into two categories: heuristic and probabilistic algorithms. Heuristic

algorithms, such as SEQUEST, Spectrum Mill, X!Tandem, and Sonar, correlate the experimental MS/MS spectrum with a theoretical spectrum and calculate a score based on the similarity between the two. On the other hand, probabilistic algorithms, such as MASCOT, model to some extent the peptide fragmentation process (e.g., ladders of sequence ions) and calculate the probability that a particular peptide sequence produced the observed spectrum by chance. Important considerations when one is carrying out MS/MS database searches are the specified search parameters (i.e., mass tolerance, which is dependent on the instrument and calibration), search strategy (i.e., semitryptic vs tryptic), chosen protein sequence database to query (i.e., IPI vs NCBI NR, which is dependent on the particular experiment), and chosen search engine. In addition, it is recommended to use an algorithm that demonstrates high sensitivity in conjunction with an algorithm that demonstrates high specificity ***(75)***. For more detailed discussion, see Kapp et al. ***(75)*** and Sadygov et al. ***(76)***. For publication guidelines for peptide and protein identification data, *see* Carr et al. ***(77)***.

2.4.2. Quantitative Analysis Using Mass Spectrometry

MS has been used successfully to identify and characterize proteins in complex mixtures (especially suborganellar proteomes such as mitochondria, phagosomes, Golgi bodies, and so on; for reviews, *see* Taylor et al. ***[78]*** and Brunet et al. ***[79]***). MS research so far has been mainly qualitative, yet the recent advent of new methodologies provides the opportunity to obtain quantitative proteomics data sets. There are two approaches in MS-based quantitative proteomics: *stable isotope labeling*, which permits direct comparison of two proteome states in the same analysis, and *ion current-based quantitation* (label-free methods), which compares the ion currents of the same peptides in different experiments ***(80)***.

Two main approaches, based on stable isotope methods, are currently used for relative quantitation using MS: metabolic labeling and chemical tagging. In metabolic labeling, stable isotope-labeled atoms are metabolically incorporated into newly synthesized proteins in vivo. These labeled cells (or their lysates) are then added as an internal standard to cells grown in material with natural abundance isotopes at the beginning of the experiment to account for errors accrued during sample preparation and measurement ***(81)***. This approach can be used for model organisms, as well as cell lines in culture. In stable isotope labeling in cell culture (SILAC), a prototype approach of metabolic labeling in cell culture, mammalian cell lines are grown in a defined medium containing isotope-labeled amino acids. Samples grown in the presence of the natural and heavy isotopes can be pooled and analyzed together. Then, the signal intensities of the light and heavy versions of the same peptides are measured, which allows their relative quantitation ***(82)***. In chemical tagging, protein samples are labeled with chemical tags of light and heavy formats. After labeling, samples are pooled and analyzed together for the

same purpose just explained for metabolic labeling. In isotope-coded affinity tag (ICAT), a prototype approach of chemical tagging, cysteine residues of proteins samples are labeled with biotinylated tags of light and heavy formats. After labeling, samples are pooled, digested, and analyzed. As in SILAC, the signal intensities of the light and heavy versions of the same peptide are measured for relative quantitation ***(83)***. Although metabolic labeling has only a limited value in human body fluid research, chemical tagging is fully applicable to it. There are other derivatives of chemical tagging methods such as iTRAQ™, or ^{18}O labeling ***(84,85)***. For a more detailed review of stable isotope methods and metabolic labeling of proteins, *see* Schneider and Hall ***(86)*** and Beynon and Pratt ***(87)***.

Recently, Pan and colleagues ***(88)*** reported a new approach for the detection and quantification of targeted proteins in a complex mixture. In this method, proteotypic peptides that uniquely represent proteins are selected from databases and used as reference peptides. The reference peptides are generated by chemical synthesis and contain heavy stable isotopes. Protein samples are digested and combined with a mixture of defined amounts of isotopically labeled reference peptides. Then the peptide mixture is separated by capillary RP-HPLC and deposited on a matrix-assisted laser desorption ionization (MALDI) plate to be analyzed using a MALDI tandem mass spectrometer. The identification and quantification of targeted proteins is based on searching and identifying the corresponding signature peptide pairs directly ***(88)***. This method holds the promise that it can improve the throughput and confidence of protein identification as well as allowing absolute quantitation. For a more detailed discussion of the proteotypic peptide approach, *see* Kuster et al. ***(89)***.

The second approach of MS-based quantitation is comparing the ion currents of peptides. Quantitation of small molecules by integration of LC-MS-extracted ion currents (XIC) has a long history in analytical chemistry, and similar quantitation techniques have been applied to proteolytic digests of protein mixtures ***(90)***. Obvious advantages of XIC-based quantification are that no labeling is used and it can be performed with any type of sample, whereas clear disadvantages are the multiple occasions for quantitation error to occur during sample processing and LC-MS analysis, as well as the presence of interfering substances in one of the states to be compared (i.e., extremely reproducible conditions are required) ***(80)***. According to Higgs and colleagues ***(90)***, it appears that relatively small (20%) changes in protein relative levels between different biological sample sets are discoverable with a fully automated sample processing and analysis system, which is implemented using a high-throughput computational environment.

3. Future Perspectives

As previously mentioned, one of the main goals of body fluid proteomics is to find protein fingerprints or biomarkers, which may reflect various disease states,

and it is becoming more evident that a single biomarker cannot represent all the complex mechanisms behind diseases including cancer. According to Diamandis ***(91)***, the prevailing view in cancer biomarker research is that the most powerful single cancer biomarkers may have already been discovered. Probably, we are now bound to discover biomarkers that could be used in panels with improved sensitivity and specificity (i.e., a multivariate approach). In other words, we may need to construct more detailed patterns to detect a certain phenotype. Then the next challenge will be to understand the relationships between the components comprising disease-specific patterns. Although this approach, referred to as *integrative systems biology*, has been increasingly applied to the study of animal models or single cells in which informative pathway information can be gained at an early stage of analysis, its application has been limited in body fluids in which the relationships between the components may not be revealed without further study ***(92)***. Recently, Davidov and colleagues ***(93)*** reported methods for the differential integrative analysis of plasma. Solely based on a body fluid analysis, their effort represents the first attempt to explain the relationships between molecular phenotypic fingerprints by combining quantitative proteomic and metabolomic data. In the future, it will be possible to enhance this approach by including the genomic component in the form of differential transcription analysis of multiple tissues and make it truly global with respect to understanding pleiotropic effects of gene perturbation on body fluids ***(93)***.

Acknowledgments

This work was supported by Australia-Asia Awards (to S.M.A.) and a National Health and Medical Research Council program grant (no. 280912) (to R.J.S.).

References

1. Skandarajah A, Moritz RL, Tjandra JJ, Simpson RJ. Proteomic analysis of colorectal cancer: discovering novel biomarkers. *Expert Rev Proteomics* 2005;2: 681–692.
2. Robin ED. Limits of the internal environment. In: Robin ED, ed. *Claude Bernard and the Internal Environment: A Memorial Symposium.* New York: M. Dekker, 1979:257–267.
3. Guyton AC, Hall JE. The body fluid compartments: extracellular and intracellular fluids, interstitiatl fluid and edema. In: Guyton AC, Hall JE, eds. *Textbook of Medical Physiology*, 10th ed. Philadelphia: W.B. Saunders, 2000:264–294.
4. Liotta LA, Kohn EC. The microenvironment of the tumour-host interface. *Nature* 2001;411:375–379.
5. Liotta LA, Ferrari M, Petricoin E. Clinical proteomics: written in blood. *Nature* 2003;425:905.
6. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 2002;1:845–867.

7. Celis JE, Gromov P, Cabezon T, et al. Proteomic characterization of the interstitial fluid perfusing the breast tumor microenvironment: a novel resource for biomarker and therapeutic target discovery. *Mol Cell Proteomics* 2004;3: 327–344.
8. Leak LV, Liotta LA, Krutzsch H, et al. Proteomic analysis of lymph. *Proteomics* 2004;4:753–765.
9. Sedlaczek P, Frydecka I, Gabrys M, Van Dalen A, Einarsson R, Harlozinska A. Comparative analysis of CA125, tissue polypeptide specific antigen, and soluble interleukin-2 receptor alpha levels in sera, cyst, and ascitic fluids from patients with ovarian carcinoma. *Cancer* 2002;95:1886–1893.
10. Trape J, Molina R, Sant F. Clinical evaluation of the simultaneous determination of tumor markers in fluid and serum and their ratio in the differential diagnosis of serous effusions. *Tumour Biol* 2004;25:276–281.
11. Simpson RJ. Role of separation science in proteomics. In: Simpson RJ, ed. *Purifying Proteins for Proteomics: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2004:1–15.
12. Dufour R. Sources and control of preanalytical variation. In: Kaplan LA, Pesce AJ, Kazmierczak SC, eds. *Clinical Chemistry: Theory, Analysis, Correlation*, 4th ed. St. Louis, MO; Mosby, 2003:64–82.
13. Omenn GS. The Human Proteome Organization Plasma Proteome Project pilot phase: reference specimens, technology platform comparisons, and standardized data submissions and analyses. *Proteomics* 2004;4:1235–1240.
14. Rand MJ, Murray RK. Plasma proteins, immunoglobulins, and blood coagulation. In: Murray RK, Granner DK, Mayes PA, Rodwell VW, eds. *Harper's Biochemistry,* 25th ed. New York: McGraw-Hill, 2000:737–762.
15. Rai AJ, Gelfand CA, Haywood BC, et al. HUPO Plasma Proteome Project specimen collection and handling: towards the standardization of parameters for plasma proteome samples. *Proteomics* 2005;5:3262–3277.
16. Muller N, Danckworth HP. [Coagulation properties of the extravascular fluid. I. Coagulation factors in thoracic-duct lymph]. *Z Lymphol* 1980;4:11–17.
17. Chang P, Aronson DL, Borenstein DG, Kessler CM. Coagulant proteins and thrombin generation in synovial fluid: a model for extravascular coagulation. *Am J Hematol* 1995;50:79–83.
18. Hulmes JD, Bethea B, Ho K, et al. An investigation of plasma collection, stabilization, and storage procedures for proteomic analysis of clinical samples. *Clin Proteomics* 2004;1:17–32.
19. Young DS, Bermes EW. Specimen collection and other preanalytical variables. In: Burtis CA, Ashwood ER, eds. *Tietz Fundamentals of Clinical Chemistry*, 5th ed. Philadelphia: Saunders, 2001:30–54.
20. Durham BH, Robinson J, Fraser WD. Differences in the stability of intact osteocalcin in serum, lithium heparin plasma and EDTA plasma. *Ann Clin Biochem* 1995;32:422–423.
21. Chan BY, Buckley KA, Durham BH, Gallagher JA, Fraser WD. Effect of anticoagulants and storage temperature on the stability of receptor activator for nuclear

factor-kappa B ligand and osteoprotegerin in plasma and serum. *Clin Chem* 2003;49:2083–2085.
22. Drake SK, Bowen RA, Remaley AT, Hortin GL. Potential interferences from blood collection tubes in mass spectrometric analyses of serum polypeptides. *Clin Chem* 2004;50:2398–2401.
23. Lewis MR, Callas PW, Jenny NS, Tracy RP. Longitudinal stability of coagulation, fibrinolysis, and inflammation factors in stored plasma samples. *Thromb Haemost* 2001;86:1495–1500.
24. Rouy D, Ernens I, Jeanty C, Wagner DR. Plasma storage at −80 degrees C does not protect matrix metalloproteinase-9 from degradation. *Anal Biochem* 2005; 338:294–298.
25. Knowles MR, Cervino S, Skynner HA, et al. Multiplex proteomic analysis by two-dimensional differential in-gel electrophoresis. *Proteomics* 2003;3:1162–1171.
26. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 1975;250:4007–4021.
27. Klose J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* 1975;26:231–243.
28. Gorg A, Drews O, Weiss W. Separation of proteins using two-dimensional gel electrophoresis. In: Simpson RJ, ed. *Purifying Proteins for Proteomics: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2004:391–430.
29. Moritz RL, Ji H, Schutz F, et al. A proteome strategy for fractionating proteins and peptides using continuous free-flow electrophoresis coupled off-line to reversed-phase high-performance liquid chromatography. *Anal Chem* 2004;76: 4811–4824.
30. Corthals GL, Wasinger VC, Hochstrasser DF, Sanchez JC. The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* 2000;21: 1104–1115.
31. Rabilloud T. Two-dimensional gel electrophoresis in proteomics: old, old fashioned, but it still climbs up the mountains. *Proteomics* 2002;2:3–10.
32. Simpson RJ. *Purifying Proteins for Proteomics: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2004.
33. Righetti PG, Castagna A, Antonioli P, Boschetti E. Prefractionation techniques in proteome analysis: the mining tools of the third millennium. *Electrophoresis* 2005;26:297–319.
34. Tirumalai RS, Chan KC, Prieto DA, Issaq HJ, Conrads TP, Veenstra TD. Characterization of the low molecular weight human serum proteome. *Mol Cell Proteomics* 2003;2:1096–1103.
35. Pieper R, Su Q, Gatlin CL, Huang ST, Anderson NL, Steiner S. Multi-component immunoaffinity subtraction chromatography: an innovative step towards a comprehensive survey of the human plasma proteome. *Proteomics* 2003;3:422–432.
36. Zolg JW, Langen H. How industry is approaching the search for new diagnostic markers and biomarkers. *Mol Cell Proteomics* 2004;3:345–354.

37. Bjorhall K, Miliotis T, Davidsson P. Comparison of different depletion strategies for improved resolution in proteomic analysis of human serum samples. *Proteomics* 2005;5:307–317.
38. Fountoulakis M, Juranville JF, Jiang L, et al. Depletion of the high-abundance plasma proteins. *Amino Acids* 2004;27:249–259.
39. Mehta AI, Ross S, Lowenthal MS, et al. Biomarker amplification by serum carrier protein binding. *Dis Markers* 2003;19:1–10.
40. Zhang H, Li XJ, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* 2003;21:660–666.
41. Kaji H, Saito H, Yamauchi Y, et al. Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat Biotechnol* 2003; 21:667–672.
42. Westermeier R, Gronau S. *Electrophoresis in Practice: A Guide to Methods and Applications of DNA and Protein Separations*, 4th ed. Weinheim: Wiley-VCH, 2005.
43. Simpson RJ. *Proteins and Proteomics: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2003.
44. Mant CT, Hodges RS. *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis, and Conformation*. Boca Raton; CRC Press, 1991.
45. Stochaj W, Berkelman T, Laird N. Preparative two-dimensional gel electrophoresis with immobilized pH gradients. In: Simpson RJ, ed. *Proteins and Proteomics: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2003:143–218.
46. Gorg A, Weiss W, Dunn MJ. Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 2005;5:826–827.
47. Herbert BR, Righetti PG, McCarthy J, et al. Sample preparation for high-resolution two-dimensional electrophoresis by isoelectric fractionation in an MCE. In: Simpson RJ, ed. *Purifying Proteins for Proteomics: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2004:431–442.
48. Bier M. Recycling isoelectric focusing and isotachophoresis. *Electrophoresis* 1998;19:1057–1063.
49. Herbert BR, Righetti PG, McCarthy J, et al. Sample preparation for high-resolution two-dimensional electrophoresis by isoelectric fractionation in an MCE. In: Simpson RJ, ed. *Purifying Proteins for Proteomics: A Laboratory Manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press, 2004:431–442.
50. Zuo X, Speicher DW. A method for global analysis of complex proteomes using sample prefractionation by solution isoelectrofocusing prior to two-dimensional electrophoresis. *Anal Biochem* 2000;284:266–278.
51. Zuo X, Speicher DW. Microscale solution isoelectrofocusing: a sample prefractionation method for comprehensive proteome analysis. *Methods Mol Biol* 2004;244:361–375.
52. Tang HY, Ali-Khan N, Echan LA, Levenkova N, Rux JJ, Speicher DW. A novel four-dimensional strategy combining protein and peptide separation methods

enables detection of low-abundance proteins in human plasma and serum proteomes. *Proteomics* 2005;5:3329–3342.
53. Gorg A, Boguth G, Kopf A, Reil G, Parlar H, Weiss W. Sample prefractionation with Sephadex isoelectric focusing prior to narrow pH range two-dimensional gels. *Proteomics* 2002;2:1652–1657.
54. Hoffmann P, Ji H, Moritz RL, et al. Continuous free-flow electrophoresis separation of cytosolic proteins from the human colon carcinoma cell line LIM 1215: a non two-dimensional gel electrophoresis-based proteome analysis strategy. *Proteomics* 2001;1:807–818.
55. Krivankova L, Bocek P. Continuous free-flow electrophoresis. *Electrophoresis* 1998;19:1064–1074.
56. Weber PJA, Weber G, Eckerskorn C. Protein purification using free-flow electrophoresis. In: Simpson RJ, ed. *Purifying Proteins for Proteomics: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2004:463–478.
57. Moritz RL, Simpson RJ. Liquid-based free-flow electrophoresis-reversed-phase HPLC: a proteomic tool. *Nat Methods* 2005;2:863–873.
58. Mohammad J. Chromatofocusing. In: Simpson RJ, ed. *Purifying Proteins for Proteomics: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2004:355–379.
59. Yan F, Subramanian B, Nakeff A, Barder TJ, Parus SJ, Lubman DM. A comparison of drug-treated and untreated HCT-116 human colon adenocarcinoma cells using a 2-D liquid separation mapping method based upon chromatofocusing PI fractionation. *Anal Chem* 2003;75:2299–2308.
60. Soldi M, Sarto C, Valsecchi C, et al. Proteome profile of human urine with two-dimensional liquid phase fractionation. *Proteomics* 2005;5:2641–2647.
61. Lilley KS, Razzaq A, Dupree P. Two-dimensional gel electrophoresis: recent advances in sample preparation, detection and quantitation. *Curr Opin Chem Biol* 2002;6:46–50.
62. Westermeier R, Marouga R. Protein detection methods in proteomics research. *Biosci Rep* 2005;25:19–32.
63. Van den Bergh G, Arckens L. Fluorescent two-dimensional difference gel electrophoresis unveils the potential of gel-based proteomics. *Curr Opin Biotechnol* 2004;15:38–43.
64. Unlu M, Morgan ME, Minden JS. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 1997;18:2071–2077.
65. Lilley KS, Friedman DB. All about DIGE: quantification technology for differential-display 2D-gel proteomics. *Expert Rev Proteomics* 2004;1:401–409.
66. Marengo E, Robotti E, Antonucci F, Cecconi D, Campostrini N, Righetti PG. Numerical approaches for quantitative analysis of two-dimensional maps: a review of commercial software and home-made systems. *Proteomics* 2005;5:654–666.

67. Alban A, David SO, Bjorkesten L, et al. A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics* 2003;3:36–44.
68. Simpson RJ. Introduction to chromatographic methods for protein and peptide purification. In: Simpson RJ, ed. *Purifying Proteins for Proteomics: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2004: 41–74.
69. Lescuyer P, Hochstrasser DF, Sanchez JC. Comprehensive proteome analysis by chromatographic protein prefractionation. *Electrophoresis* 2004;25:1125–1135.
70. Apffel A. Multidimensional chromatography of intact proteins. In: Simpson RJ, ed. *Purifying Proteins for Proteomics: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2004:75–100.
71. Wolters DA, Washburn MP, Yates JR 3rd. An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* 2001;73:5683–5690.
72. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003;422: 198–207.
73. Weatherly DB, Atwood JA 3rd, Minning TA, Cavola C, Tarleton RL, Orlando R. A heuristic method for assigning a false discovery rate for protein identifications from mascot database search results. *Mol Cell Proteomics* 2005.
74. Nesvizhskii AI, Aebersold R. Interpretation of shotgun proteomics data: The protein inference problem. *Mol Cell Proteomics* 2005;4:1419–1440.
75. Kapp EA, Schutz F, Connolly LM, et al. An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: sensitivity and specificity analysis. *Proteomics* 2005;5:3475–3490.
76. Sadygov RG, Cociorva D, Yates JR 3rd. Large-scale database searching using tandem mass spectra: looking up the answer in the back of the book. *Nat Methods* 2004;1:195–202.
77. Carr S, Aebersold R, Baldwin M, Burlingame A, Clauser K, Nesvizhskii A. The need for guidelines in publication of peptide and protein identification data: Working Group on Publication Guidelines for Peptide and Protein Identification Data. *Mol Cell Proteomics* 2004;3:531–533.
78. Taylor SW, Fahy E, Ghosh SS. Global organellar proteomics. *Trends Biotechnol* 2003;21:82–88.
79. Brunet S, Thibault P, Gagnon E, Kearney P, Bergeron JJ, Desjardins M. Organelle proteomics: looking at less to see more. *Trends Cell Biol* 2003;13:629–638.
80. Ong SE, Mann M. Mass-spectrometry-based proteomics turns quantitative. *Nat Chem Biol* 2005;1:252–262.
81. Wu CC, MacCoss MJ, Howell KE, Matthews DE, Yates JR 3rd. Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal Chem* 2004;76:4951–4959.
82. Ong SE, Blagoev B, Kratchmarova I, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002;1:376–386.

83. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 1999;17:994–999.
84. Ross PL, Huang YN, Marchese JN, et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 2004;3:1154–1169.
85. Stewart II, Thomson T, Figeys D. 18O labeling: a tool for proteomics. *Rapid Commun Mass Spectrom* 2001;15:2456–2465.
86. Schneider LV, Hall MP. Stable isotope methods for high-precision proteomics. *Drug Discov Today* 2005;10:353–363.
87. Beynon RJ, Pratt JM. Metabolic labeling of proteins for proteomics. *Mol Cell Proteomics* 2005;4:857–872.
88. Pan S, Zhang H, Rush J, et al. High throughput proteome screening for biomarker detection. *Mol Cell Proteomics* 2005;4:182–190.
89. Kuster B, Schirle M, Mallick P, Aebersold R. Scoring proteomes with proteotypic peptide probes. *Nat Rev Mol Cell Biol* 2005;6:577–583.
90. Higgs RE, Knierman MD, Gelfanova V, Butler JP, Hale JE. Comprehensive label-free method for the relative quantification of proteins from biological samples. *J Proteome Res* 2005;4:1442–1450.
91. Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. *Mol Cell Proteomics* 2004;3: 367–378.
92. van der Greef J, Stroobant P, van der Heijden R. The role of analytical sciences in medical systems biology. *Curr Opin Chem Biol* 2004;8:559–565.
93. Davidov E, Clish CB, Oresic M, et al. Methods for the differential integrative omic analysis of plasma from a transgenic disease animal model. *Omics* 2004;8: 267–288.
94. Guyton AC, Hall JE. The microcirculation and the lymphatic system: capillary fluid exchange, interstitial fluid, and lymph flow. In: Guyton AC, Hall JE, eds. *Textbook of Medical Physiology*, 10th ed. Philadelphia: WB Saunders, 2000: 162–174.
95. Guyton AC, Hall JE. Cerebral blood flow; the cerebrospinal fluid; and brain metabolism. In: Guyton AC, Hall JE, eds. *Textbook of Medical Physiology*, 10th ed. Philadelphia: WB Saunders, 2000:709–715.
96. Bogdanov B, Smith RD. Proteomics by FTICR mass spectrometry: top down and bottom up. *Mass Spectrom Rev* 2005;24:168–200.
97. Steen H, Mann M. The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol* 2004;5:699–711.