

REVIEW

Proteomics, biomarkers, and HIV-1: A current perspective

Maire Rose Donnelly and Pawel Ciborowski

Department of Pharmacology and Experimental Neuroscience, University of Nebraska Medical Center, Omaha, NE 68198, USA

Despite more than three decades of extensive research, HIV-1 infection although well controlled with cART, remains incurable. Multifactorial complexity of the viral life-cycle poses great challenges in understanding molecular mechanisms underlying this infection and the development of biomarkers, which we hope will lead us to its eradication. For a more in-depth understanding of how the virus interacts with host target cells, T cells and macrophages, proteomic profiling techniques that offer strategies to investigate the proteome in its entirety were employed. Here, we review proteomic studies related to HIV-1 infection and discuss perspectives and limitations of proteomic and systems biology approaches in future studies.

Received: January 7, 2015

Revised: March 17, 2015

Accepted: May 27, 2015

Keywords:

Biomarker / Body fluids / HIV / Macrophage / Systems biology / T cell

1 Introduction

Efficient biomarkers are highly sought after for their ability to provide a window into physiological, developmental, and pathological changes in a biological system. Gaining measures of disease development and prognosis is only part of our quest for biomarkers; efficacy of treatment and measures of prevention extend this list. Desirable biomarkers should be sensitive to biological changes, predictive, and highly specific. Sensitivity of quantification will be easier to accomplish as analytical technologies and instrumentation improve and progress in single-cell proteomics is made. In contrast, achieving high specificity of biomarkers is more difficult and there are many reasons for this. Among diseased individuals there are broad responses to pathogenic factors and treatment inconsistencies due to genetic variability (i.e. SNPs) seem to be major contributors to difficulties in identifying specific biomarkers. Another factor is many diseases generate the same response, such as inflammation, which can be very sensitive to an ongoing pathological process, but not highly specific. We also need to consider the variability of response time between individuals in developing pathological

processes or responding to applied treatment, which also contributes to the difficulty of discovering universally identifiable biomarkers.

Proteomics as an experimental approach to global profiling has made a quantum leap in technological advancements in the past ten to 15 years. The field now offers a variety of methods based not only on polyacrylamide gel separations, but also on micro- and nano-flow LC fractionations. In addition, mass spectrometers have tremendously improved in analytical power, allowing protein identification at the attomolar level. This progress was accompanied by the development of sample preparation techniques, database search algorithms, and statistical models and bioinformatics tools. As a result, researchers are now equipped with a broad array of tools that can be used for systems biology approaches to simultaneously tackle multiple complex and dynamic processes, such as HIV-1 infection.

Treatment of human immunodeficiency virus-1 (HIV-1) would benefit greatly from advancements in biomarker discovery, which has proven to be difficult due to complexity of the HIV life cycle. To complete its life cycle, the virus requires multiple cellular elements that can be divided into four groups: (1) membrane proteins for entry, (2) cellular proteins for transport, assembly, and maturation, (3) nuclear factors for transcription, and (4) regulatory mechanisms associated with maintaining the viral life cycle. Such classification will help us understand that proteomic technologies are the most productive in generating new information that can be used for more effective treatments and ultimately eradication of viral infections. Moreover, each step of the viral life cycle is highly

Correspondence: Dr. Pawel Ciborowski, Department of Pharmacology and Experimental Neuroscience, University of Nebraska Medical Center, Omaha, NE 68198, USA

E-mail: pciborowski@unmc.edu

Abbreviations: ANAXA, Annexin; CyPA, cyclophilin A; HAD, HIV-1-associated dementia; HAND, HIV-associated neurocognitive disorder; HIV-1, human immunodeficiency virus-1; MDM, monocyte-derived macrophages; PIC, preintegration complex; STAT, signal transducers and activators of transcription; TMT, tandem mass tag; WCX2, weak cation exchange

Colour Online: See the article online to view Figs. 1 and 2 in colour.

complex and this is compounded by the fact that the host's response to infection will involve its own dynamic immune and cellular interaction with an invading pathogen.

In this review, we will guide the readers through accomplishments in this area over the past decade and will provide our perspective and guidance for future proteomics-based investigations. We also encourage readers to read reviews previously published by others and us in order to form a more complete understanding of the prior progress made in HIV proteomics [1–5].

2 Why systems biology approach to investigate HIV-1 infection?

Since the first publication of AIDS cases in Morbidity and Mortality Weekly Report in 1981 [6], thousands of studies have been conducted investigating various aspects of immune system deregulation caused by HIV-1 infection. Collectively, these studies generated more than a quarter million PubMed cited papers. As much as we have succeeded in controlling infection with cART, viral infection is not eradicated and thus it is not curable at this time. What also needs to be acknowledged is that HIV-1 infection is associated with the toxicity of antiretroviral drugs, numerous comorbidities, as well as other associated pathologies (i.e. use of illicit drugs). Therefore, HIV-1 infection, particularly in its chronic phase, cannot be investigated in a vacuum and other factors must be considered. However, considering multiple concurrent factors requires a specific experimental approach such as the systems biology approach, which meets the requirements to study complex systems with multiple variables.

The development of computational methods to process high throughput data, along with increased speed of data acquisition by new generations of mass spectrometers, has increased confidence in the interpretation of quantitative profiling experiments spanning across multiple variables. On the same token, dual comparisons of one experimental condition to a control is quite limiting in what we can learn about the function of biological systems. Alternatively, the comparison of multiple conditions poses an overwhelming challenge in understanding the significance and direction of dependability of variables representing intertwining conditions. In some instances, effects of disease and/or therapeutic intervention can be monitored continuously. Generally, samples for diagnosis are taken at certain time intervals, which may span from hours to weeks or months. The latter is more common with slowly developing diseases such as neurodegenerative disorders (i.e. Alzheimer's disease). A broad system, such as those involving whole cells with protein PTMs in all subcellular compartments, cross-talk signaling, regulatory, and metabolic pathways, and intracellular trafficking exposed to more than one stimulus, generates data with too many variables.

As cancer research leads in the number of studies applying proteomics and systems biology approaches, the first

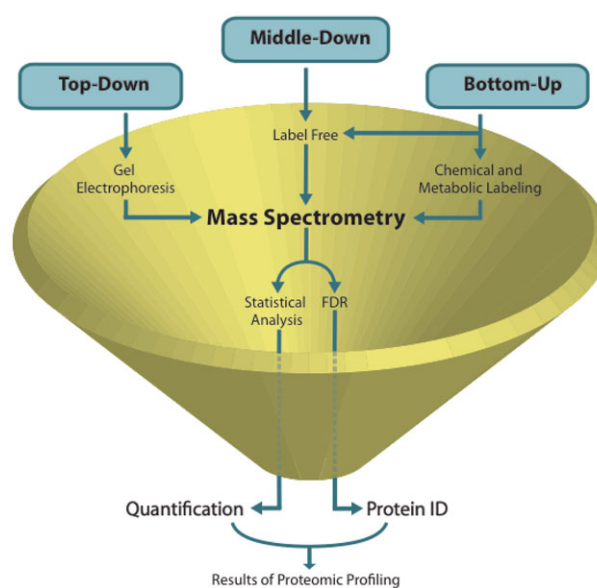


Figure 1. A summary of proteomic platforms. The three primary methods of proteomic analysis involving MS are top-down, middle-down, and bottom-up. While each requires different methods of sample preparation, they all have the same outcome; identification and quantification of samples, which leads to complete proteomic profiling.

such studies to understand the global effect of HIV-1 on the host's cells are emerging [7–9]. In a recent study by Jaworski et al. the authors used nanotrap particle technology to selectively capture specific HIV-1 molecules and to augment HIV-1 detection by concentrating viral proteins and infectious virions from infected samples [10]. Haverland et al. narrowed their studies to transcription factors and regulators [9]. In this study, the authors used SWATH-MS label-free quantitative profiling.

A compromise of narrowing the scope of investigations to subcellular compartments, subsets of proteins or particular signaling and metabolic pathways is making its way [11, 12]. It seems that for the near future, simplification of the scope of biological systems to be investigated is the best approach until the development of new and more powerful tools for statistical and bioinformatics analyses [13, 14]. However, this development will not replace critical thinking based on an in-depth understanding of the biological systems being investigated [14].

3 Proteomics platforms and HIV-1

3.1 An overview

Proteomics currently offers a variety of methods for quantitative profiling. Decisions regarding selection of which proteomic platform to be used in addressing any particular question is of utmost importance for downstream data interpretation and expectations of the output. A summary of proteomic platforms is presented in Fig. 1 and is divided into three

major groups: top-down, middle-down, and bottom-up. Methods representing top-down approaches are used for separation of intact proteins, including LC and gel-based techniques. Proteins can also be labeled with fluorescent dyes (i.e. Cy-dyes) or 2DE gels can be run without labeling. Methods based on the separation of intact proteins using LC can comprise single step or multistep fractionation approaches combining off-line and on-line configurations. Chemical or metabolic labeling, as well as label-free techniques, are widely used for bottom-up studies.

The majority of quantitative proteomic profiling techniques that have been utilized in HIV-1-related research and a comprehensive summary of these studies is presented in Table 1. Our intention here is to not only present past experience, but also provide guidance for the selection of proteomic platforms in future studies based on the questions being asked. The experimental design and sample processing should align with the type of material being investigated. Investigation of body fluid proteomes and culture supernatants or lysates of whole cells, tissues, or cellular compartments/organelles would all require use of a different platform, which can be determined from reviewing published studies. When designing experiments to study various aspects of HIV-1 infection or life cycle we need to acknowledge that this is a chronic disease and a relatively small percentage of cells are infected. This makes detection of differences much more challenging compared to, for example, exposure of cells to a drug where all cells are exposed at the same time and at the same level. Moreover, any proteomic profiling is almost always preceded by at least one-dimensional fractionation to reduce complexity of the samples. This can be accomplished by employing a wide variety of techniques including immunoaffinity removal of the most abundant proteins from plasma/CSF, organelle separation, isolation of specific groups of proteins such as phosphorylated or glycosylated proteins, ion exchange chromatography, etc. Extending pre-profiling fractionation may not always work toward our advantage, since every analytical step introduces some degree of variability. With improvements of sensitivity, resolution, mass accuracy, dynamic range, and speed of data acquisition of new generations of mass spectrometers, proteomic data are inevitably massive, and reduction of sample complexity is necessary.

3.2 SELDI-TOF

SELDI-TOF technology was developed in the early 1990s [15] and was commercialized as the ProteinChip system in 1997. SELDI-TOF was a simple TOF mass spectrometer used to detect differentially expressed proteins in various tissue samples as a tool for biomarker discovery. Although SELDI-TOF has not been used in HIV-1 research for some time, the MALDI-TOF-based approach is still used by many [16, 17]. The major advantage of SELDI-TOF is a set of chips coated with surfaces specifically absorbing proteins based on hydrophobic

properties, anion/cation exchange, and metal-binding surfaces. Another type of chip allows to covalently linking ligands (i.e. immunoglobulins) for affinity type reactions. The SELDI-TOF platform raised very high expectations for fast discovery of new biomarkers. However, problems with reproducibility and cross-validation soon limited enthusiasm for this method [18]. Despite these limitations, several HIV-related profiling studies have been conducted [19–24]. Because SELDI-TOF did not have the ability to fragment peptides, results were usually reported as features (presence or absence of peaks) characteristic of conditions under investigation [25]. A good example of such study is the investigation of a small number of CSF samples, matched sera that reported the presence of molecular features rather than specific proteins, which may correlate with the degree of HIV-1-associated dementia (HAD) [26]. However, none of these features were further identified as differentially expressed proteins. Wiederin et al. used the same chromatographic conditions in a preparative LC mode as they were used to bind proteins to the surface of a weak cation exchange (WCX2) SELDI chip to extend SELDI-MS profiling with protein identification. The same sample was absorbed to a column with exactly the same surface of WCX2 functional groups. Eluted proteins/peptides were subjected to fractionation by 1DE and then identified using MS/MS after in-gel trypsin digestion. This approach led to the identification of gelsolin, ceruloplasmin, and afamin as proteins differentially expressed between samples from HIV infected individuals with or without HAD. Interestingly, ceruloplasmin has been found to be downregulated in the CSF of subjects with HAD and upregulated in the plasma of these patients. This suggests that the CSF to plasma ratio of ceruloplasmin may be an important correlation of HIV-associated neurocognitive disorder (HAND) [27]. Similarly, Sun and coauthors identified secretory lysozyme as downregulated by macrophages isolated from HIV positive subjects with HAD [20]. Although none of the identified proteins became a diagnostic biomarker due to a lack of specificity, all differentially expressed proteins contributed to building upon our knowledge about global deregulation of homeostasis in HIV-infected individuals.

3.3 Gel electrophoresis

1D SDS-PAGE [28], followed by 2D SDS-PAGE and 2D-DIGE, were early methods for top-down, intact protein fractionations used in HIV research. These methods were used in several studies to discover biomarkers in body fluids of HIV-infected individuals [23, 29–33]. For the most part, these studies highlighted proteins representing inflammation, including components of complement cascade or lipid metabolism, such as apolipoproteins [29, 34–42]. Rozek et al. postulated that specific components and/or fragments of complement system might serve as a more specific biomarker than changes of complement in its entirety. They proposed that the Complement C3 α -40 chain fragment can be used as a potential

Table 1. Summary of proteomic platforms used in HIV-1 studies

Proteomic platform	Description of studies		References
	Experimental model	Major finding(s)	
SELDI	Immunodepleted sera from individuals with or without HIV Associated Dementia (HAD)	Differential expression of ceruloplasmin, afamin, and gelsolin.	[23]
	Cervical lavage fluid from HIV infected individuals	Detailed description of methodology, however no identification of specific biomarker candidates.	[24]
Gel Based	Monocyte supernatants from a cohort of HAD and non-HAD Thai volunteers prior to the initiation of ARV.	Profile of inflammatory proteins	[21]
	2DE-DIGE and iTRAQ platforms to identify differences between the proteomes of WT and nef-deleted viruses	Glucosidase II is enriched in WT virions, while Ezrin, ALG-2, CD81, and EHD4 are enriched in nef-deleted virions.	[34]
	Oral epithelial cells from HIV-infected HAART subjects and healthy controls analyzed by 2DE	Identification of 61 differentially expressed proteins. Down-regulated proteins are those associated with protein folding, pro- and anti-inflammatory responses, redox homeostasis, and detoxification. Up-regulated proteins are disulfide isomerases and those that expression was negatively regulated by Hsp90.	[35]
	2D-DIGE differential expression and identification of proteins in CD14+ monocytes from patients with HAD	Down-regulated proteins were ADP ribosylhydrolase, myeloperoxidase, thioredoxin, peroxiredoxin 3, NADPH, and GTPase activating protein.	[36]
	2DE-DIGE comparison of proteins from the frontal cortex of HAD and non-HAD patients	Identification of 76 differentially expressed proteins. Majority of these proteins were represented in the energy metabolic and signal transduction pathways.	[37]
iTRAQ	Cervicovaginal fluids from high-risk HIV-exposed seronegative individuals (HESN) and two control groups: low-risk HESN and HIV-positives	Serpin A5, a serine proteinase inhibitor, was upregulated and myeloblastin was down-regulated	[44]
	Plasma samples from HIV-1/HCV mono- and co-infected individuals	Apolipoproteins and complement proteins were identified as two major differentially regulated classes of proteins. APOA2, APOC2, APOE, C3, HRG proteins were upregulated in the plasma of all three HIV-1 mono-, HCV mono-, and co-infected patient samples compared to healthy control samples.	[47]
	Plasma samples from HIV-1-infected methamphetamine users	Among 390 proteins identified, 28 showed significant changes in expression. These proteins were involved in complement, coagulation pathways, and oxidative stress.	[27, 47]
	Plasma from rhesus monkeys before and after infection with simian immunodeficiency virus (SIV)	Afamin, a member of the albumin superfamily, was found to be significantly down-regulated after infection.	[34, 43–46]
	Profiling of changes in Treg macrophages	Reduction in virus release paralleled the upregulation of interferon-stimulated gene 15, a ubiquitin-like protein involved in interferon-mediated antiviral immunity.	[34, 43–46]
	CD4 lymphoblastoid SUP-T1 cells after infection with HIV-1 strain LAI.	This study found 266, 60 and 22 proteins differentially expressed at 4, 8, and 20 h postinfection. Functional analyses showed enrichment in several biological pathways including protein synthesis, cell proliferation, and T-cell activation.	[34, 43–46]

(Continued)

Table 1. Continued

Proteomic platform	Description of studies		References
	Experimental model	Major finding(s)	
TMT	Synaptosomes from the cortex of wild-type (wt) and HIV/gp120 transgenic mice	Out of 1301 proteins identified in both wt and transgenic mice, 107 proteins showed a significant change in expression. Bioinformatic analysis predicted deregulation of the PI3K/Akt pathway observed in human brains with HAND.	[98]
SILAC	Jurkat T-cells stably expressing HIV-1 Tat	Forty nine differentially expressed proteins indicated alterations in proteins collectively participating in ribosomal biogenesis, protein homeostasis, metabolic pathways including glycolytic, pentose phosphate, nucleotides, and amino acids biosynthetic pathways, stress response, T-cell signaling pathways, and genome integrity.	[48]
	U937 cells transduced with Vpr	Down-regulation of glutamate dehydrogenase 2 (GLUD2), adenylate kinase 2 (AK2), and transketolase (TKT) was shown.	[49]
	Plasma samples from HIV-1/HCV Mono- and Coinfected Individuals	Decrease of DNA packaging activities in parallel with virus-induced cytopathology and replication while increase of factors affecting antigen presentation and interferon-induced antiviral activity.	[50]
Label-Free	T cells culture supernatants and human plasma HIV Nucleoprotein Complexes	Glycoproteomic profiling Identification of 11055 proteins, including previously reported XRCC6, TFRC, and HSP70. The analysis identified 94 proteins unique in the infected fractions and 121 proteins unique to the control fractions. An additional 54 and 52 were classified as enriched in the infected and control samples, respectively, based on a threefold difference in total proteome discoverer probability score.	[84] [52]
	Bronchoalveolar lavage fluid in HIV infection	A total of 318 unique proteins in bronchoalveolar lavage fluid (BALF) of HIV- and HIV+ subjects were identified, of which 87 were differentially up- or downregulated between the two groups. Functional analysis of differentially regulated proteins implicated downregulation of immune responses in lungs of HIV+ patients and was significantly altered during HIV infection.	[85]
	Lavage fluid from 293 HIV-1-resistant, uninfected, and infected sex workers was analyzed using a 2D-LC-LTQ-FT-MS label-free platform	Identification of 360 unique proteins. Serpin A1, Serpin A3, Cystatin B, and an epithelial antiprotease A2ML1 were found to be significantly overabundant in HIV-1-resistant women. These mucosal antiprotease levels correlated with proinflammatory cytokine concentration, but independently of pro-inflammatory cytokine levels in HIV-1-resistant women including TNF-alpha, IL-1 alpha, IL-1 beta, IL-6, and IL-8.	[99]
	Interactions between affinity tagged HIV-1 and host proteins expressed in Jurkat and HEK293 cells	Identification of 497 high confidence interactions between HIV and host proteins, of which 40% were found in both cell types. Eleven host proteins were found to inhibit HIV-1 replication	[55]
SWATH-MS	Uninfected and HIV-1-infected MDM	Identification and quantification of 3608 proteins in uninfected and HIV-1-infected MDM, of which 420 were significantly altered upon HIV-1 infection. RNA binding and processing, as well as transcription regulation, were the most significant functional enrichments.	[9]

biomarker of HAD. The advantage of using this fragment is that regardless of how the precursor of complement C3 is processed and the abundance of intermediate forms, levels of the α -chain will indicate the final step of processing [29]. Although the authors did not further explore such possibility, it seems that the decrease of circulating α -40 chain of complement C3 is consistent among HIV-1 infected individuals with advanced neurocognitive symptoms. This, however, will require further investigation using larger cohorts of well-characterized samples.

3.4 Chemical isotope labeling and isobaric tags

iTRAQ and tandem mass tag (TMT) platforms are based on the same principal of generating chemically added tags that show the relative abundance of proteins after peptide fragmentation in MS/MS. These techniques emerged around 2003 and 2004 with high expectations for improved quantification of global proteomic profiling. iTRAQ offers two sets of labels (channels) for quantification, a 4-plex and 8-plex, allowing up to eight samples to be compared in one LC-MS/MS analysis while TMT offers a 6-plex system. Several studies used the iTRAQ platform to study cellular molecular mechanisms of HIV-1 infection [34, 43–46] and plasma proteomes for biomarker discovery [27, 47] (Table 1).

3.5 Metabolic labeling—SILAC

Although very few studies have been performed using this technique, it is worth being noted as a separate category. Jarboui et al. [48] cultivated stably transfected Jurkat T cells with Tat protein for five generations to achieve 95% incorporation of “heavy” amino acids. After mixing “heavy” and “light” samples, nucleoli were isolated, proteins were digested with trypsin and resulting peptides were analyzed by nano-LC-MS/MS. Out of 520 identified proteins, the authors found 49 to be differentially expressed using the IPI Database. Bioinformatic analysis of these data indicated alterations of transcription, RNA processing and to a lesser extent, cell-cycle processes and chromosome organization. Investigating macrophages infected with HIV-1 using the SILAC approach is much more challenging because macrophages are terminally differentiated and are nonproliferating cells. For SILAC experiments, monocytic U937 cells, which were differentiated using 100 ng/mL of phorbol myristate, were cultured for 5 days with daily change of media with “light” or “heavy” amino acids [49] followed by transfection with an Adeno-Vpr construct. Based on this study, the authors propose that HIV-1 steals the macrophage glucose metabolism pathway, inducing expression of hexokinase, glucose-6-phosphate dehydrogenase, and pyruvate kinase muscle type 2. In another study by Kraft-Terry et al. the authors used primary monocytes, which were first differentiated to macrophages using MCSF. SILAC media without MCSF was added on days 1, 3, or 5 following HIV-1 infection and cells were lysed 48 h following

SILAC medium addition. This approach allowed the authors to identify proteins synthesized *de novo* after viral infection. Data resulting from this experiment showed the complexity of the macrophages’ struggle to contain HIV-1 infection and develop antiretroviral mechanisms while the virus succeeds in acquiring protective mechanisms to support survival [50].

3.6 Label-free quantitative proteomics

The label-free approach appears to be very attractive because it can be used when metabolic labeling (SILAC) is not feasible or the number of analyzed samples is larger than molecular tags such as iTRAQ channels. Label-free quantification is based on the premise that although the data generated in MS experiments are not directly quantitative, samples can be compared directly using spectral count or using measurements of area under the peak, which is the classical approach in analytical chemistry quantification. Although it has been shown that peptide ion counts across experiments can be very reproducible [51], this method still raises reservations among many researchers. Nevertheless, Schweitzer et al. showed the viability of this approach [52]. First, the authors compared the number of high confidence peptide identifications per individual protein assignment, and then compared the number of individual protein assignments for each NCBI GI number across the seven biological replicates. This was followed by calculation and comparison of the coverage of each protein across biological replicates. To validate lack of bias, the authors compared the seven most abundant proteins common to both HIV-1 infected and control samples across all replicates. Lack of statistical differences between infected and control samples were used to conclude that there was no bias resulting from the comparison of proteins with high peptide counts and reinforced this methodology. Utilizing this analysis, several differences in proteins putatively constituting HIV-1 preintegration complex (PIC) have been identified and orthogonally validated [52]. In this study, PIC was partially purified using sucrose gradient ultracentrifugation and localization was measured by the activity of endogenous reverse transcriptase (EndoRT). As expected, fractions covering EndoRT activity contained more than two thousand unique protein IDs, but less than one hundred were unique to infected cells. The top ten proteins unique to infected samples were: XRCC6 (Ku70), FCRL6 Fc receptor-like 6, glutamate receptor, ionotropic, Nmethyl D-aspartate 2C, NONO non-POU domain containing, octamer-binding; p54 (nrb), MARCKS protein, MAX-like protein X, Golgi matrix protein 130 (GM130), FAM167B family with sequence similarity 167, member B, Serine/threonine kinase 33 (STK33), PDIA3 protein disulfide isomerase family A, member 3. The second group of proteins consisted of enriched proteins, which might be recruited during formation of PIC after viral entry. Annexin 6 (ANAXA6) was one of the proteins found in this group and validated by Western blot analysis. Interestingly, ANAXA 2 [53], another member of this family

of proteins, has been previously implicated in participation of HIV-1 particle formation in macrophages [54]. The role and biological validation of ANAXA 6 in HIV-1 infection of T cells requires further focused studies. In summary, this label-free proteomic study provided a panel of novel cellular proteins potentially associated with the viral life cycle in T cells.

Jager et al. published an elegant study of global landscape of HIV–host proteins interactions by pulling-down interacting partners from cells transfected with tagged viral proteins [55]. Although this study did not reveal specific biomarkers; it provides an indispensable source of information for future studies.

3.7 SWATH-MS

Recent advances in quantitative proteomics led to the development and commercialization of SWATH-MS, another label-free technique based on comparisons between the sums of areas under the peaks for a select number of transitions. Strength of this method is twofold. First, all information contained in analyzed samples is recorded; second, SWATH-MS provides a high dynamic range, thus increasing linearity of quantification. A limited number of applications have been published using this relatively new technique and only one was subjected to HIV-1 infection [9]. Nevertheless, this approach holds a good promise, particularly when multiple biological replicates across multiple conditions are compared. The authors of this study focused on identifying differential expression of transcription factors and regulators in primary human macrophages infected *in vitro* with HIV-1. This extensive study covered 3068 proteins and identified 420 as differentially expressed. Further analysis showed that HIV-1 infection leads to increased expression of YBOX1 and DHX15 proteins, which are predicted to be HIV-1 dependency factors and unique to the U11/U12 spliceosome. Other differentially expressed proteins, NCOR2 and HDAC2, have been implicated in the infectivity of HIV-1. NCOR1 is reported to be a positive mediator of HIV-1 latency. Two hundred forty-five proteins identified as differentially expressed by SWATH-MS, with no documentation to show association with HIV-1 infection, represent potentially novel discoveries.

3.8 What have we learned?

Every time we approach a profiling experiment, we hope to find a breakthrough biomarker or ideal target for new treatment. Instead we accumulate new information, which is collectively used to increase knowledge and advance our understanding about the course of pathological processes. It would be quite naive to believe that every proteomic profiling experiment will deliver “magic bullets,” yet stigmatizing these studies as a “fishing expedition” causes more damage than

creating a positive attitude. Yes, with proteomics, as with other types of studies, we have learned a lot over time. In fact, many studies could be designed better to increase the odds of delivering breakthrough discoveries. This, however, does not diminish the guidance value of existing studies. To support the latter claim, we present in Table 1 a summary of proteomic profiling studies published to date. Although the presented list contains major studies, it is by no means fully complete.

Considering the surprisingly low number of proteomic profiling studies aimed toward investigating aspects of HIV-1 infection, there has been an immense amount of new information gathered. Given this, do we have enough information to start drawing more general conclusions? In our view we do, but we have to consider several important limitations such as that each study represents a time snapshot of infection that has a relatively short acute phase and a much longer chronic phase. We also need to divide these studies by cell type: T cells and macrophages. There are two general methods of mimicking the course of infection, using either live virus or transfecting cells with viral proteins. The variety of proteomic platforms reviewed in this article, search engines, protein databases and bioinformatics tools used for protein identification and quantification will further contribute to the difficulty of direct data comparison and interpretation. Despite these limitations, the list of differentially expressed proteins is rapidly expanding and with increasing consistency from study to study. Such proteins should be used for comparison-based validation of profiling datasets. One such example is signal transducers and activators of transcription (STATs) pathways. Kanmogne’s reductionistic studies showed that HIV-1 activates proinflammatory and interferon- α -inducible genes that are linked to the Janus kinase/STAT pathway [56]. We found in our proteomic profiling-based systems biology study an increase in expression of STAT1. Although at this stage it is impossible to use the degree of STAT1 upregulation as a normalization factor and/or pattern, this fact can be used as a point for comparison-based validation across datasets. The fact that alterations in RNA processing pathways were identified in two otherwise distant studies, one using Jurkat T cells transfected with Tat protein [56] and the other using monocyte-derived macrophages (MDM) infected with live virus [9], further supports our view on how proteomic data should be validated.

In summary, the expectations of what proteomic profiling may deliver have not lived up to the results of actual studies. We should not expect discovery of “magic bullets” from just a few experiments with so much inherent variability. HIV-1 infection is a chronic disease that subsides with treatment but the virus is not eradicated and rebounds when antiretroviral therapy is interrupted. Confounding effects of toxicity of antiretroviral drugs, resistance, and the ability of viruses to hide in reservoirs requires the constant search for new ways of treating already infected individuals until a cure is found. Global profiling is and will be an indispensable tool in tackling such quest.

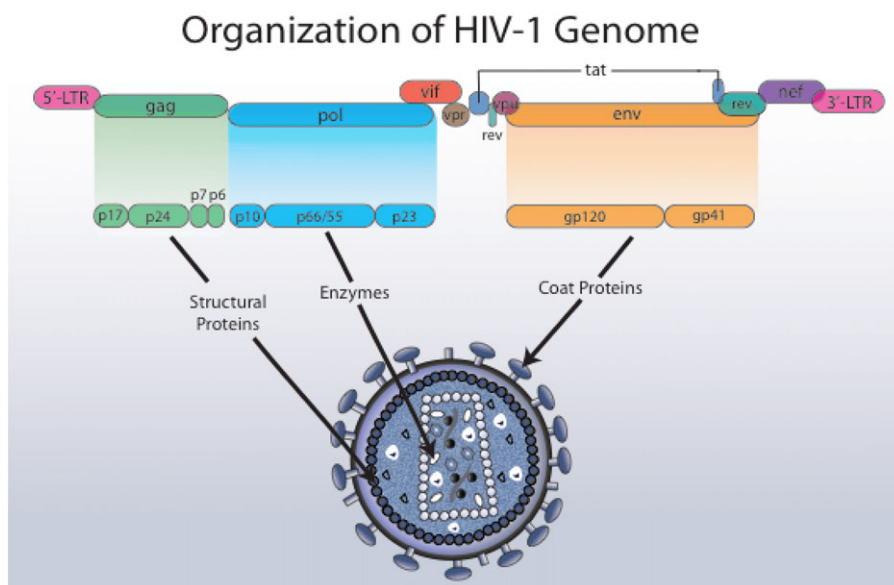


Figure 2. Organization of the HIV-1 genome. The expression levels of different HIV-1 proteins have been investigated as potential biomarkers with varying success.

4 HIV-1 proteins as biomarkers

4.1 Introduction

The most common HIV test is the antibody-screening test, which detects the presence of antibodies against HIV. Although such immunoassay test is generally accurate in providing a yes/no answer for initial screening, it measures response to infection rather than directly determining the levels of circulating viral proteins, or antigens. Since the first probable link between HIV and AIDS has been made, many studies have been conducted using viral proteins as specific, accurate, and sensitive biomarkers of various pathologies associated with viral infection, e.g. HAND. For the purpose of this review, we present in Fig. 2 the organization of the HIV-1 genome providing a list of viral proteins. These viral proteins have been investigated in a broad range of interests, also as causative factors of HIV-associated pathologies such as direct toxicity to neurons and the blood–brain barrier. Concurrently, their levels in body fluids have been explored as potential prognostic biomarkers for HIV-1-related pathologies, particularly HAND [57]. In this review, we focus on Tat and gp120, because numerous studies have shown their direct contribution to the pathobiology of HIV infection and because both proteins represent similar sets of issues associated with quantification of circulating amounts and linking this information to their predictive value as biomarkers.

4.2 Tat

There is overwhelming experimental evidence supporting multiple roles for Tat in the pathogenesis of HIV-1 infection and pathologies associated with this infection, mainly with HAND. Even though in vitro experiments in which cells

exposed to Tat show pathological effects, correlation of these findings to in vivo situations of infected individuals is much more complicated. Coming to meaningful conclusions about Tat roles is difficult because of the inability to efficiently quantify the protein. Hudson et al. used a highly sensitive ELISA assay (sensitivity of 2 ng Tat/mL of brain tissue) and was unable to detect Tat in extracts from the brains of individuals with HIV encephalitis, however Tat mRNA was detected in 4 out of 9 of these individuals [58]. Additionally, Tat can be differentially distributed between body compartments, such as blood, mucosa, and lymph node, which further complicates conclusions related to which site and corresponding concentration is most relevant from the perspective of prognosis. Also, Tat can be either trapped by other molecules, such as heparin sulfates becoming sequestered on extracellular matrices, or sequestered by anti-Tat antibody [59]. Nevertheless, Xiao et al. tested 80 anonymous HIV-1 patient sera from the National Institutes of Health Clinical Center and found that more than 1/3 of sera had antibody reactivity to Tat. Spotting of sera onto PVDF membranes and probing them with rabbit anti-Tat antibody showed positive detection in 5 out of 80 sera samples with concentrations ranging from 2 to 40 ng/mL [59]. These authors also suggest that comparison of such data between laboratories can be difficult due to a lack of standardized antibodies.

Alternative approach would be MS-based quantification; however, we were not able to identify this protein in any of our proteomic experiments, despite successfully detecting other proteins such as Gag and gp120 [30].

4.3 Gp120

Gp120 interaction with cell surface receptors, particularly CD4 on the surface of T-helper cells, is crucial for virus entry

into cells. As such, this protein was one of the first targets for vaccine development, thus making it a subject of extensive investigation. It has also been shown that viral gp120 contributes to neuronal death via various mechanisms [60,61] and was investigated as a potential biomarker for motor and cognitive disorder [62]. Importantly, gp120 is heavily N-glycosylated and contains highly variable regions. Therefore, it is represented by 182 272 unreviewed entries (TrEMBL), however only 185 entries (SwissProt) were reviewed as of December 2014 (<http://www.uniprot.org/uniprot/>). These database entries represent various clades and were collectively submitted over long periods of research. Similar to Tat, there is no general agreement regarding levels of gp120 circulating in blood. Although it has been hypothesized that gp120 contributes to immunological dysfunction during early stages of infection, it is not known what level of this protein is sufficient to cause such dysfunction therefore making it difficult to be measured as a biomarker. Rychert et al. showed that gp120 was detectable in plasma in only 33% of infected individuals, ranging between 0.5 and 15.6 ng/mL (4–130 pM). No correlation was found between acute and early stages of infection in this cohort of 109 individuals. A longitudinal follow-up of nine subjects showed stable levels of gp120 in blood and were independent from viral load. An early study performed by Gilbert et al. detected gp120 in only 11 out of 41 samples from AIDS patients [63] and the measurable range was 250–2000 pg/mL. Oh et al. detected gp120 in the majority of AIDS samples in a range of 0.1–0.8 nM [64]. A later study based on much more sensitive tests showed the range 2–20 pM in only a small portion of sera from p24-antigenemic AIDS and AIDS-related complex patients [63]. Although Klasse and Moore [65] stated that the cross-reactivity capabilities of the assay used by Oh et al. was questionable and therefore the ability to detect and quantify gp120 in plasma is problematic, they also admit that any estimation of how much gp120 was naturally present in the HIV-1 positive sera was clearly problematic. They further seem to agree with Parren et al. [66] that the majority of Env proteins (gp120 and gp41) are not secreted but rather released from dead cells as “viral debris.” Nevertheless, an overall conclusion is that plasma anti-gp120 antibodies may mask detection of the protein, thus skewing gp120 quantification results, similar to the case of Tat.

All of the studies reviewed here were based on ELISA assays in which various antibodies were used, thus adding an additional layer of difficulty in comparison and interpretation. We attempted to identify and characterize, but not quantify, circulating forms of gp120 using MS and found that identification of some viral proteins in the CSF or plasma of HIV-infected individuals [30]. However, in most instances only one unique peptide was detected, mostly from a conserved region putting a limitation on this type of assay as a reliable measure [29, 30]. Although we do not exclude the possibility of using mass spectrometric-based quantification of HIV proteins, it would require new approaches in sample preparation to facilitate this type of quantification. Based on

current knowledge, it still seems to be a stretch of the imagination to use proteomic profiling of viral proteins directly in ex vivo specimens as diagnostic biomarkers. Considering the presence of variable regions, immunoaffinity isolation would generate the same set of problems as ELISA assays used to date.

Since glycosylation can significantly alter protein conformation, as shown by deletion studies in the V1V2 regions of gp120, it can also alter biological properties of gp120, including immunogenicity and antigenicity [67,68]. For this reason, glycosylation of Env proteins drew significant attention. Studies investigating whether amino acid features of the V3 loop region can favor viral transmission from individual to individual did not correlate with the number of N-linked glycosylation sites [69]. We conclude that oligosaccharide moieties and/or the specific N-glycosylation site have a decisive role in altering biological activity of the viral gp120 protein.

It has also been postulated that variability in potential N-linked glycosylation sites as well as the number of glycosylated sites also result in increased viral fitness by altering sensitivity to potentially neutralizing antibodies. Therefore, a systematic glycoproteomic approach was required to advance our knowledge in this area. Desire and coworkers extensively characterized oligosaccharides associated with HIV gp120 protein. The initial goal of this research was to define glycosylation profiles of Env proteins and define specific carbohydrate moieties at each glycosylation site [67]. In the subsequent step, the authors attempted to determine which carbohydrate structures are involved in HIV-1 infectivity and are most suitable as a potential antigen to develop a protective immunological response to infection. The number of reports led not only to the refinement of mass spectrometric-based techniques used in glycoprotein analysis [68, 70, 71] but also provided a wide breath of new information about glycosylation of gp120 and gp41 [72–75]. Despite this effort, at this time, neither gp120, nor gp41 became suitable biomarker candidates. We need to further explore properties, either based on amino acid sequence or PTMs, of gp120 protein to uncover a useful biomarker.

4.4 Other viral proteins

Other HIV-1 proteins have been studied to a much lesser extent as possible direct biomarkers. Most of the work done has focused on immune responses to viral proteins. One such study performed by French et al. [76] found levels of IgG1 antibodies against HIV Gag (p18, p24, rp55) and Pol-encoded (p32, p51, p66) proteins higher in HIV controllers, a small group of infected individuals who control HIV replication in the absence of therapy. Yates et al. measured concentrations of IgG1 and IgG3 antibodies to eight different HIV-1 proteins, gp140 Env, gp120 Env, gp41 Env, p66 reverse transcriptase, p31 Integrase, Tat, Nef, and p55 Gag, during acute/recent HIV-1 infection [77]. The authors proposed that the decline in multiple HIV-specific IgG3 responses simultaneous with

persistent IgG1 responses during HIV-1 infection could serve as markers for detection of incident HIV infections.

4.5 Conclusions

It is disappointing that, with all the advances in research, viral proteins cannot be used for prediction or prognosis of other pathological conditions associated with this viral infection, even though they undoubtedly play a role in pathogenesis of HIV-1 infection. As much as we hope levels of viral protein circulating in body fluids or deposited in tissues would be more informative, it is evident that this is not the case due to the complex nature of HIV-1 infection and pathogenesis. Despite the presence of conserved regions, these proteins have not shown much promise as a target for anti-HIV vaccination due to its variable regions and shedding from the surface of mature virions [78]. The uneven degree of shedding results in circulation of highly variable amounts of viral proteins making it harder to establish criteria for their use as biomarkers.

5 Host proteins incorporated in virions

Although assembly of HIV-1 occurs at the plasma membrane of infected cells, many studies showed that virus also accumulates in intracellular compartments [79]. Pelchen-Matthews and coworkers demonstrated that the virus accumulates in late endosomes [79] supporting hypothesis that late endocytic membranes are where the bulk of infection viral particles reside. As virions mature, host proteins are encapsulated within viral envelope, raising the question of whether this is associated with recruitment of specific host proteins needed by the virus for early steps of the next life cycle. If this is correct, it would be interesting to determine which proteins are incorporated when virions are assembled at the plasma membrane and which are incorporated during virion assembly in late endosomes. Identifying which host proteins are purposely incorporated into virions will give us more information about complexity of the viral cell cycle and may also indicate new therapeutic targets. A major difficulty in the precise determination of proteome composition is dealing with the technical challenge of separating virions from copurified microvesicles, which have very similar densities. Linde and coworkers used iTRAQ-based quantification of HIV particles isolated from T cells and MDM. The authors found several proteins common in both cell types: actin isoforms, HLA proteins, chaperones, ERM proteins, the dynamin domain containing protein EH4, a phosphodiesterase, and cyclophilin A [80]. A major hurdle in interpretation of these results lies in the difficulty of differentiating between proteins necessary for completion of the viral life cycle from those that are interacting with viral proteins but are not crucial.

Misumi et al. suggested that two isoforms of cyclophilin A (CyPA) are inside the viral membrane, with a third isoform located outside the viral membrane. Very little is known about

the mechanism behind CyPA isoform distribution and their exact roles at each location. One suggestion is that CyPA may regulate the conformation of viral capsids [81].

As further proteomic profiling of cellular compartments is undertaken and more data is collected, it is imminent that this topic warrants a separate review. Therefore, for supplemental information we direct our readers to articles published by several groups [10, 80, 82, 83].

6 Proteomics of body fluids from HIV-1 infected individuals

Body fluids such as plasma/serum, cerebrospinal fluid, saliva, lung lavages, and cervical-vaginal fluid from HIV-1 positive individuals provide not only much needed diagnostic and prognostic biomarkers but also clues related to susceptibility to infection, efficacy of treatment, etc. Body fluids still remain as very informative material for proteomic profiling and every new approach to quantitative proteomics raises hopes for new discoveries. The inherent variability of proteins in body fluids and lack of good references for normalization are major obstacles in the way of faster and more efficient biomarker discovery. The concentration of proteins continually fluctuates depending on multiple factors. Even highly abundant serum albumin ranges from 34 to 54 g/L (reference serum values) and can fluctuate hourly, daily, weekly, or monthly, contributing to already existing high donor-to-donor variability. This might be very difficult to show for low abundant proteins such as those originating from tissue or cell leakage [84]. Even if highly abundant proteins are removed for proteomic profiling studies to lower the dynamic range of concentrations, the lasting portion will remain highly variable. Normalization of samples based on the amount of protein taken for analysis, independent of the volume of blood used to obtain this protein, seems to be the most common step in proteomic profiling of body fluids. Spiking in synthetic peptides aids absolute quantification of proteins in any given sample but will not help in overall normalization within populations of diseased and healthy individuals who remain highly variable. Similar issues apply to cerebrospinal fluid, lung lavages [85] as well as saliva [86] and cervical-vaginal fluid [44], with the latter two having been studied to a much lesser extent.

The necessity of pooling of body fluid samples to generate a sufficient amount of material for proteomic quantification is highly debated. For plasma/serum it is not, even if the remaining amount of proteins after immunodepletion is low because plasma usually is available in large quantities. In the case of CSF, it may be impossible to perform full proteomic analysis with samples from individual subjects without sample pooling [87], however, final outcome showed limited success.

The majority of biomarker discovery studies investigating various aspects of HIV-1 and SIV-1 infection have been performed using methods in which proteins and peptides were labeled. Initially, 2DE and 2D-DIGE were used [29, 30],

followed by the iTRAQ platform soon after (see Table 1). Since metabolic labeling such as SILAC cannot be applied to investigate human body fluids, chemical labeling, or label-free techniques became two popular choices. Protein arrays seem to be a feasible technique in the quest for biomarkers. However, its application for samples from HIV-1 infected individuals has not yet been shown. Superiority of any proteomic platform is difficult to assess because too few studies have been performed and none of the approaches have provided unique and highly specific biomarkers. The spectral counting approach sums the number of spectra identified for a given peptide in different samples and then integrates the results for all measured peptides of the protein(s) that are quantified. Label-free approaches in quantitative proteomics, either based on spectral count or precursor ion abundance, have been of interest in the proteomic community for some time [88, 89]. The latter requires high-resolution peptide signals on the MS level but separates the quantification from the identification process in MS/MS mode. It has been postulated that the label-free methods provide a higher dynamic range of quantification [89, 90]. Nevertheless, spectral count as a quantitative measure for global profiling is not commonly used and is criticized by some for issues such as linearity and accuracy of proteins represented by low number of spectra. Some studies showed successful use of this approach in studying proteomes of saliva [91]; however it has not been applied to HIV-1 related biomarker discovery in body fluids. While the recent developments of data independent acquisition and the commercialization of SWATH-MS technology [9, 92, 93] has opened new avenues in label-free quantification, formal studies of body fluids have not yet been published.

Careful preparation is critical for the proteomic profiling of samples such as plasma/serum, CSF and ascites fluid due to very high complexity and a wide dynamic range of concentrations. Based on findings from the Human Plasma Proteome Project, the Core Dataset contains 9504 IPI proteins identified with one or more peptides and 3020 proteins identified with two or more peptides in human blood. The 12 most abundant proteins constitute approximately 96% of the protein mass. Various methods involving depletion of the most abundant proteins have been proposed. Among the immunodepletion methods, use of IgY columns seems to be the most efficient [27, 94, 95] as well as the ProteoMiner kit from BioRad (Hercules, CA), which depletes serum albumin in a cleaned-up sample [47]. Other methods of plasma sample enrichment were also used. Yang et al. enriched glycopeptides and performed a search with PNGase F-catalyzed conversion of Asn to Asp. This method allowed for the identification of 829 unique glycopeptides from 411 unique glycoproteins with N-deamidation within the consensus N-linked glycosylation motif. Further analysis using the MaxQuant software package with Andromeda algorithm provided identification of 671 glycopeptides from 375 glycoproteins. After normalization, 326 proteins were used for quantification based on a fold-change of 3 and a *p*-value < 0.05. Of these, 59

glycoproteins were identified as being significantly differentially expressed.

A methodologically thorough study of CSF proteomes from 12 HIV-1-infected subjects, 12 controls, and 14 HIV-1-infected individuals before and after initiation of cART identified 2333 peptides covering 307 proteins and showed that protein abundances differed between and within subjects, thus differentiating those with and without HIV infection [96]. This would be expected because of the multilevel effect retroviral infection has on the overall function of the human body. Bioinformatic analysis showed alteration of the complement cascade, cell adhesion, acute phase signaling, immune responses, and wound healing, all of which has previously been reported and is not HIV specific. The authors tried to correlate neuronal injury with the levels of neopterin, which is a catabolic product of guanosine triphosphate. Because neopterin is produced by INF- α activated macrophages, indicating a proinflammatory response, its levels are only a reflection of immune system activation.

Besides viral proteins, what would we expect a biomarker to represent and how large and persistent would the quantitative change be? First, we must realize that biomarkers of chronic and slowly progressing diseases will be different in nature than those with a relatively short but highly acute phase. In the case of HIV-1 infection, which can be controlled but not eradicated, there is an acute phase first followed by chronic infection, which is controlled by lifelong antiretroviral treatment. Wiederin et al. compared plasma proteomes of SIV-infected monkeys from acute to chronic states [94]. Because SIV infection of the monkeys was fully controlled, the results were disconnected from coinfections, drug toxicity, multidrug abuse, and other confounding factors, which is very difficult if not impossible to avoid when using samples from patients enrolled in clinical studies. The authors made some interesting observations. First, despite immunodepletion of the 12 most abundant proteins (albumin, IgG, fibrinogen, transferrin, IgA, IgM, haptoglobin, apolipoprotein A-I, apolipoprotein A-II, α 1-antitrypsin, α 1-acid glycoprotein, and α 2-macroglobulin) and 2D fractionation (SCX and RP chromatography) only medium to highly abundant proteins were identified. Out of them, 64 proteins showed statistically significant differences in expression and represented inflammatory responses to microbial infections. In the subsequent study [95] in which the effect of opiate administration on the course of acute and chronic SIV infection of monkeys was investigated, a relatively small subset of proteins were identified and quantified and the majority of proteins also represented inflammatory responses to microbial infections. Despite using a well-controlled animal model and reduced technical variability as much as possible, the authors still observed highly variable changes in the plasma proteome, which they linked to differences in genetic background and individual responses to exogenous conditions—an opiate administration. When all was taken under consideration, the authors concluded that in chronic conditions, biological variance,

rather than technical, is a major obstacle in biomarker discovery.

It is important to note that despite thousands of proteins being identified in human plasma, biomarker discovery studies identify relatively low numbers of proteins, usually not exceeding a few hundred. Shetty et al., using the iTRAQ platform, identified altogether 70 differentially expressed proteins in three sets of samples [47]. Many of these proteins represent complement components changed during HIV infection or HIV/HCV coinfection. Pottiez and coworkers compared profiles of plasma samples from HIV-1 positive individuals actively using illicit drugs to those who successfully entered a treatment program and abstain from drug abuse for relatively short periods of time or longer [27]. Interestingly, the authors showed that differences induced by use of illicit drugs are short lived and a few weeks void of drugs brings levels of proteins to those who were HIV negative and not using drugs of abuse. Change in levels of apolipoprotein A-1 (ApoA1), apolipoprotein B (ApoB), histine-rich glycoprotein, alpha-1 B glycoprotein, and orosomucoid 2 in individuals treated with cART and showing symptoms of fatigue may represent a broad range of altered physiological and molecular mechanisms [97].

An interesting observation has been reported by Rozek et al. [29], who analyzed CSF samples using 2D-DIGE and found that in samples from individuals with HAD, the α -chain of complement C3 is consistently downregulated compared to those without dementia. This would be an indirect indication of alterations in immune system function. However, a more thorough quantitative analysis of all components derived from complement C3 precursor need to be performed.

7 Overall conclusions

Throughout years of research, we have accumulated ample amount of new knowledge regarding HIV-1–host interactions during infection and the significant role of host cellular machinery in successful completion of the viral life cycle. Some of the host cell machinery involves changes in several proteins associated with specific regulatory mechanisms. One example is NCOR1, NCOR2, and HDAC2 as a part of transcription regulation complexes [9]. Another two regulatory factors are YBOX1 and DHX15, which regulate RNA splicing [9]. Nevertheless, much more work needs to be done in validating findings of proteomic profiling using less laborious approaches. This can only be accomplished by working through many of the aforementioned obstacles and challenges.

Despite technological advances, investigation of proteomes in cells infected by HIV-1 poses a challenge in experimental design. After in vitro infection of macrophages with HIV-1, a low percentage of cells are infected within the first several hours and the remaining cells are infected within 3–4 days from secondary infections. Cells that have been already infected become activated and start secreting

proinflammatory cytokines, which in turn leads to activation of uninfected cells. The overall effect is a mosaic of cells at various stages of infection and activation, making interpretation of proteomic data taken as a time-point snap shot more difficult. Infection of T cells can be made more uniform by utilizing the spinoculation technique [52]. With this method, not all cells are infected at the same time but can be activated simultaneously. Timing of infection and activation needs to be considered when experiments are designed and data are interpreted.

Challenges involved in biomarker discovery and fully understanding molecular mechanisms of HIV-1 infection are very similar to those associated with other diseases. Continuous development of proteomic methods, protocols, and technology platforms, along with improvement of sensitivity and specificity, has generated more and more massive datasets. Undoubtedly, these datasets contain valuable information, however the major hindrance is data analysis when multiple factors change concurrently and in opposite directions. Almost all models of HIV-1 infection have been explored: primary cells and cell lines directly infected with the virus, cells stably transfected with viral proteins, ex vivo and post-mortem human clinical material, humanized mice and SIV with nonhuman primates. With these models, almost all proteomic platforms have been employed. To further our knowledge in this area, we need better tools for statistical analysis and more refined tools for bioinformatics analyses. With the amount of recent progress made in proteomics, there is no doubt that our collective experiences provide enough information to design more efficient and informative experiments. It is our hope that with better design of experiments, more solid conclusions about HIV-1 infection can be made.

Financial support was provided by National Institutes of Health grants R01 DA030962, P30 MH062261. We would like to thank Ms. Robin Taylor for help in editing this manuscript.

The authors have declared no conflict of interest.

8 References

- [1] Ciborowski, P., Biomarkers of HIV-1-associated neurocognitive disorders: challenges of proteomic approaches. *Biomark. Med.* 2009, 3, 771–785.
- [2] Calabria, A. R., Shusta, E. V., Blood-brain barrier genomics and proteomics: elucidating phenotype, identifying disease targets and enabling brain drug delivery. *Drug Discov. Today* 2006, 11, 792–799.
- [3] Simpson, R. J., Lim, J. W., Moritz, R. L., Mathivanan, S., Exosomes: proteomic insights and diagnostic potential. *Expert. Rev. Proteomics* 2009, 6, 267–283.
- [4] Ciborowski, P., Gendelman, H. E., Human immunodeficiency virus-mononuclear phagocyte interactions: emerging avenues of biomarker discovery, modes of viral persistence and disease pathogenesis. *Curr. HIV Res.* 2006, 4, 279–291.

- [5] Zhang, L., Zhang, X., Ma, Q., Zhou, H., Host proteome research in HIV infection. *Genomics Proteomics Bioinformatics* 2010, 8, 1–9.
- [6] Center for Disease Control. Pneumocystis Pneumonia – Los Angeles. *MMWR* 1981, 30, 250–252.
- [7] Stein, D. R., Burgener, A., Ball, T. B., Proteomics as a novel HIV immune monitoring tool. *Curr. Opin. HIV AIDS* 2013, 8, 140–146.
- [8] Zak, D. E., Aderem, A., Overcoming limitations in the systems vaccinology approach: a pathway for accelerated HIV vaccine development. *Curr. Opin. HIV AIDS* 2012, 7, 58–63.
- [9] Haverland, N. A., Fox, H. S., Ciborowski, P., Quantitative proteomics by SWATH-MS reveals altered expression of nucleic acid binding and regulatory proteins in HIV-1-infected macrophages. *J. Proteome Res.* 2014, 13, 2109–2119.
- [10] Jaworski, E., Saifuddin, M., Sampey, G., Shafagati, N. et al., The use of Nanotrap particles technology in capturing HIV-1 virions and viral proteins from infected cells. *PLoS One* 2014, 9, e96778.
- [11] Bhat, A., Dakna, M., Mischak, H., Integrating proteomics profiling data sets: a network perspective. *Methods Mol. Biol.* 2015, 1243, 237–253.
- [12] Liu, Y., Chance, M. R., Integrating phosphoproteomics in systems biology. *Comput. Struct. Biotechnol. J.* 2014, 10, 90–97.
- [13] Philp, A., Rowland, T., Perez-Schindler, J., Schenk, S., Understanding the acetyome: translating targeted proteomics into meaningful physiology. *Am. J. Physiol. Cell Physiol.* 2014, 307, C763–C773.
- [14] Hoopmann, M. R., Moritz, R. L., Current algorithmic solutions for peptide-based proteomics data generation and identification. *Curr. Opin. Biotechnol.* 2013, 24, 31–38.
- [15] Hutchens, T. W. a. Y., T.-T., New desorption strategies for the mass spectrometric analysis of macromolecules. *Rapid Commun. Mass Spectrom.* 1993, 7, 576–580.
- [16] Moon, Y. S., Park, E. S., Kim, T. O., Lee, H. S., Lee, S. E., SELDI-TOF MS-based discovery of a biomarker in *Cucumis sativus* seeds exposed to CuO nanoparticles. *Environ. Toxicol. Pharmacol.* 2014, 38, 922–931.
- [17] von Saltzwedel, H., Maraun, M., Scheu, S., Schaefer, I., Evidence for frozen-niche variation in a cosmopolitan parthenogenetic soil mite species (acari, oribatida). *PLoS One* 2014, 9, e113268.
- [18] Baggerly, K. A., Morris, J. S., Coombes, K. R., Reproducibility of SELDI-TOF protein patterns in serum: comparing datasets from different experiments. *Bioinformatics* 2004, 20, 777–785.
- [19] Luo, X., Carlson, K. A., Wojna, V., Mayo, R. et al., Macrophage proteomic fingerprinting predicts HIV-1-associated cognitive impairment. *Neurology* 2003, 60, 1931–1937.
- [20] Sun, B., Rempel, H. C., Pulliam, L., Loss of macrophage-secreted lysozyme in HIV-1-associated dementia detected by SELDI-TOF mass spectrometry. *Aids* 2004, 18, 1009–1012.
- [21] Ratto-Kim, S., Chuenchitra, T., Pulliam, L., Paris, R. et al., Expression of monocyte markers in HIV-1 infected individuals with or without HIV associated dementia and normal controls in Bangkok Thailand. *J. Neuroimmunol.* 2008, 195, 100–107.
- [22] Toro-Nieves, D. M., Rodriguez, Y., Plaud, M., Ciborowski, P. et al., Proteomic analyses of monocyte-derived macrophages infected with human immunodeficiency virus type 1 primary isolates from Hispanic women with and without cognitive impairment. *J. Neurovirol.* 2009, 15, 36–50.
- [23] Wiederin, J., Rozek, W., Duan, F., Ciborowski, P., Biomarkers of HIV-1 associated dementia: proteomic investigation of sera. *Proteome Sci.* 2009, 7, 8.
- [24] Burgener, A., Profiling cervical lavage fluid by SELDI-TOF mass spectrometry. *Methods Mol. Biol.* 2012, 818, 143–152.
- [25] Laspiur, J. P., Anderson, E. R., Ciborowski, P., Wojna, V. et al., CSF proteomic fingerprints for HIV-associated cognitive impairment. *J. Neuroimmunol.* 2007, 192, 157–170.
- [26] Berger, J. R., Avison, M., Mootoor, Y., Beach, C., Cerebrospinal fluid proteomics and human immunodeficiency virus dementia: preliminary observations. *J. Neurovirol.* 2005, 11, 557–562.
- [27] Pottiez, G., Jagadish, T., Yu, F., Letendre, S. et al., Plasma proteomic profiling in HIV-1 infected methamphetamine abusers. *PLoS One* 2012, 7, e31031.
- [28] Bosley, A., Marshall, H. N., Badralmaa, Y., Natarajan, V., A method of HIV-1 inactivation compatible with antibody-based depletion of abundant proteins from plasma. *Proteomics Clin. Appl.* 2008, 2, 904–907.
- [29] Rozek, W., Ricardo-Dukelow, M., Holloway, S., Gendelman, H. E. et al., Cerebrospinal fluid proteomic profiling of HIV-1-infected patients with cognitive impairment. *J. Proteome Res.* 2007, 6, 4189–4199.
- [30] Rozek, W., Horning, J., Anderson, J., Ciborowski, P., Sera proteomic biomarker profiling in HIV-1 infected subjects with cognitive impairment. *Proteomics Clin. Appl.* 2008, 2, 1498–1507.
- [31] Kim, S. S., Kim, M. H., Shin, B. K., Na, H. J. et al., Different isoforms of apolipoprotein AI present heterologous post-translational expression in HIV infected patients. *J. Proteome Res.* 2007, 6, 180–184.
- [32] Zhang, L., Jia, X., Zhang, X., Cao, J. et al., Alpha-1 antitrypsin variants in plasma from HIV-infected patients revealed by proteomic and glycoproteomic analysis. *Electrophoresis* 2010, 31, 3437–3445.
- [33] Ciborowski, P., Kadiu, I., Rozek, W., Smith, L. et al., Investigating the human immunodeficiency virus type 1-infected monocyte-derived macrophage secretome. *Virology* 2007, 363, 198–209.
- [34] Bregnard, C., Zamborlini, A., Leduc, M., Chafey, P. et al., Comparative proteomic analysis of HIV-1 particles reveals a role for Ezrin and EHD4 in the Nef-dependent increase of virus infectivity. *J. Virol.* 2013, 87, 3729–3740.
- [35] Yohannes, E., Ghosh, S. K., Jiang, B., McCormick, T. S. et al., Proteomic signatures of human oral epithelial cells in HIV-infected subjects. *PLoS One* 2011, 6, e27816.
- [36] Kraft-Terry, S., Gerena, Y., Wojna, V., Plaud-Valentin, M. et al., Proteomic analyses of monocytes obtained from Hispanic

- women with HIV-associated dementia show depressed antioxidants. *Proteomics Clin. Appl.* 2010, 4, 706–714.
- [37] Zhou, L., Diefenbach, E., Crossett, B., Tran, S. L. et al., First evidence of overlaps between HIV-Associated Dementia (HAD) and non-viral neurodegenerative diseases: proteomic analysis of the frontal cortex from HIV+ patients with and without dementia. *Mol. Neurodegener.* 2010, 5, 27.
- [38] Kadiu, I., Wang, T., Schlautman, J. D., Dubrovsky, L. et al., HIV-1 transforms the monocyte plasma membrane proteome. *Cell. Immunol.* 2009, 258, 44–58.
- [39] Reynolds, J. L., Mahajan, S. D., Sykes, D. E., Schwartz, S. A., Nair, M. P., Proteomic analyses of methamphetamine (METH)-induced differential protein expression by immature dendritic cells (IDC). *Biochim. Biophys. Acta* 2007, 1774, 433–442.
- [40] Reynolds, J. L., Mahajan, S. D., Sykes, D., Nair, M. P., Heroin induces differential protein expression by normal human astrocytes (NHA). *Am. J. Infectious Dis.* 2006, 2, 49–57.
- [41] Reynolds, J. L., Mahajan, S. D., Bindukumar, B., Sykes, D. et al., Proteomic analysis of the effects of cocaine on the enhancement of HIV-1 replication in normal human astrocytes (NHA). *Brain Res.* 2006, 1123, 226–236.
- [42] Coiras, M., Camafeita, E., Urena, T., Lopez, J. A. et al., Modifications in the human T cell proteome induced by intracellular HIV-1 Tat protein expression. *Proteomics* 2006, 6 (Suppl 1), S63–S73.
- [43] Pendyala, G., Trauger, S. A., Siuzdak, G., Fox, H. S., Quantitative plasma proteomic profiling identifies the vitamin E binding protein afamin as a potential pathogenic factor in SIV induced CNS disease. *J. Proteome Res.* 2010, 9, 352–358.
- [44] Van Raemdonck, G., Zegels, G., Coen, E., Vuylsteke, B. et al., Increased Serpin A5 levels in the cervicovaginal fluid of HIV-1 exposed seronegatives suggest that a subtle balance between serine proteases and their inhibitors may determine susceptibility to HIV-1 infection. *Virology* 2014, 458–459, 11–21.
- [45] Huang, X., Stone, D. K., Yu, F., Zeng, Y., Gendelman, H. E., Functional proteomic analysis for regulatory T cell surveillance of the HIV-1-infected macrophage. *J. Proteome Res.* 2010, 9, 6759–6773.
- [46] Navare, A. T., Sova, P., Purdy, D. E., Weiss, J. M. et al., Quantitative proteomic analysis of HIV-1 infected CD4+ T cells reveals an early host response in important biological pathways: protein synthesis, cell proliferation, and T-cell activation. *Virology* 2012, 429, 37–46.
- [47] Shetty, V., Jain, P., Nickens, Z., Sinnathamby, G. et al., Investigation of plasma biomarkers in HIV-1/HCV mono- and coinfecting individuals by multiplex iTRAQ quantitative proteomics. *Omics* 2011, 15, 705–717.
- [48] Jarboui, M. A., Bidoia, C., Woods, E., Roe, B. et al., Nucleolar protein trafficking in response to HIV-1 Tat: rewiring the nucleolus. *PLoS One* 2012, 7, e48702.
- [49] Barrero, C. A., Datta, P. K., Sen, S., Deshmane, S. et al., HIV-1 Vpr modulates macrophage metabolic pathways: a SILAC-based quantitative analysis. *PLoS One* 2013, 8, e68376.
- [50] Kraft-Terry, S. D., Engebretsen, I. L., Bastola, D. K., Fox, H. S. et al., Pulsed stable isotope labeling of amino acids in cell culture uncovers the dynamic interactions between HIV-1 and the monocyte-derived macrophage. *J. Proteome Res.* 2011, 10, 2852–2862.
- [51] Wang, W., Zhou, H., Lin, H., Roy, S. et al., Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal. Chem.* 2003, 75, 4818–4826.
- [52] Schweitzer, C. J., Jagadish, T., Haverland, N., Ciborowski, P., Belshan, M., Proteomic analysis of early HIV-1 nucleoprotein complexes. *J. Proteome Res.* 2013, 12, 559–572.
- [53] Ryzhova, E. V., Vos, R. M., Albright, A. V., Harrist, A. V. et al., Annexin 2: a novel human immunodeficiency virus type 1 Gag binding protein involved in replication in monocyte-derived macrophages. *J. Virol.* 2006, 80, 2694–2704.
- [54] Enrich, C., Rentero, C., deMuga, S. V., Reverter, M. et al., Annexin A6-linking Ca(2+) signaling with cholesterol transport. *Biochim. Biophys. Acta* 2011, 1813, 935–947.
- [55] Jager, S., Cimermancic, P., Gulbahce, N., Johnson, J. R. et al., Global landscape of HIV-human protein complexes. *Nature* 2012, 481, 365–370.
- [56] Chaudhuri, A., Yang, B., Gendelman, H. E., Persidsky, Y., Kanmogne, G. D., STAT1 signaling modulates HIV-1-induced inflammatory responses and leukocyte transmigration across the blood-brain barrier. *Blood* 2008, 111, 2062–2072.
- [57] Pratt, R. D., Nichols, S., McKinney, N., Kwok, S. et al., Virologic markers of human immunodeficiency virus type 1 in cerebrospinal fluid of infected children. *J. Infectious Dis.* 1996, 174, 288–293.
- [58] Hudson, L., Liu, J., Nath, A., Jones, M. et al., Detection of the human immunodeficiency virus regulatory protein tat in CNS tissues. *J. Neurovirol.* 2000, 6, 145–155.
- [59] Xiao, H., Neuveut, C., Tiffany, H. L., Benkirane, M. et al., Selective CXCR4 antagonism by Tat: implications for in vivo expansion of coreceptor use by HIV-1. *Proc. Natl. Acad. Sci. USA* 2000, 97, 11466–11471.
- [60] Kanmogne, G. D., Schall, K., Leibhart, J., Knipe, B. et al., HIV-1 gp120 compromises blood-brain barrier integrity and enhances monocyte migration across blood-brain barrier: implication for viral neuropathogenesis. *J. Cerebral Blood Flow Metab.* 2007, 27, 123–134.
- [61] Corasaniti, M. T., Piccirilli, S., Paoletti, A., Nistico, R. et al., Evidence that the HIV-1 coat protein gp120 causes neuronal apoptosis in the neocortex of rat via a mechanism involving CXCR4 chemokine receptor. *Neurosci. Lett.* 2001, 312, 67–70.
- [62] Jurado, A., Rahimi-Moghaddam, P., Bar-Jurado, S., Richardson, J. S. et al., Genetic markers on HIV-1 gp120 C2-V3 region associated with the expression or absence of cognitive motor complex in HIV/AIDS. *J. Neuro AIDS* 1999, 2, 15–28.
- [63] Gilbert, M., Kirihara, J., Mills, J., Enzyme-linked immunoassay for human immunodeficiency virus type 1 envelope glycoprotein 120. *J. Clin. Microbiol.* 1991, 29, 142–147.
- [64] Oh, S. K., Cruikshank, W. W., Raina, J., Blanchard, G. C. et al., Identification of HIV-1 envelope glycoprotein in the serum of AIDS and ARC patients. *J. Acquir. Immune Defic. Syndr.* 1992, 5, 251–256.

- [65] Klasse, P. J., Moore, J. P., Is there enough gp120 in the body fluids of HIV-1-infected individuals to have biologically significant effects? *Virology* 2004, *323*, 1–8.
- [66] Parren, P. W., Burton, D. R., Sattentau, Q. J., HIV-1 antibody—debris or virion? *Nat. Med.* 1997, *3*, 366–367.
- [67] Go, E. P., Irungu, J., Zhang, Y., Dalpathado, D. S. et al., Glycosylation site-specific analysis of HIV envelope proteins (JR-FL and CON-S) reveals major differences in glycosylation site occupancy, glycoform profiles, and antigenic epitopes' accessibility. *J. Proteome Res.* 2008, *7*, 1660–1674.
- [68] Irungu, J., Go, E. P., Zhang, Y., Dalpathado, D. S. et al., Comparison of HPLC/ESI-FTICR MS versus MALDI-TOF/TOF MS for glycopeptide analysis of a highly glycosylated HIV envelope glycoprotein. *J. Am. Soc. Mass Spectrom.* 2008, *19*, 1209–1220.
- [69] Sagar, M., Laeyendecker, O., Lee, S., Gamiel, J. et al., Selection of HIV variants with signature genotypic characteristics during heterosexual transmission. *J. Infectious Dis.* 2009, *199*, 580–589.
- [70] Go, E. P., Hewawasam, G., Liao, H. X., Chen, H. et al., Characterization of glycosylation profiles of HIV-1 transmitted/founder envelopes by mass spectrometry. *J. Virol.* 2011, *85*, 8270–8284.
- [71] Zhu, Z., Su, X., Clark, D. F., Go, E. P., Desaire, H., Characterizing O-linked glycopeptides by electron transfer dissociation: fragmentation rules and applications in data analysis. *Anal. Chem.* 2013, *85*, 8403–8411.
- [72] Go, E. P., Chang, Q., Liao, H. X., Sutherland, L. L. et al., Glycosylation site-specific analysis of clade C HIV-1 envelope proteins. *J. Proteome Res.* 2009, *8*, 4231–4242.
- [73] Go, E. P., Zhang, Y., Menon, S., Desaire, H., Analysis of the disulfide bond arrangement of the HIV-1 envelope protein CON-S gp140 DeltaCFI shows variability in the V1 and V2 regions. *J. Proteome Res.* 2011, *10*, 578–591.
- [74] Ma, B. J., Alam, S. M., Go, E. P., Lu, X. et al., Envelope deglycosylation enhances antigenicity of HIV-1 gp41 epitopes for both broad neutralizing antibodies and their unmutated ancestor antibodies. *PLoS Pathogens* 2011, *7*, e1002200.
- [75] Go, E. P., Liao, H. X., Alam, S. M., Hua, D. et al., Characterization of host-cell line specific glycosylation profiles of early transmitted/founder HIV-1 gp120 envelope proteins. *J. Proteome Res.* 2013, *12*, 1223–1234.
- [76] French, M. A., Center, R. J., Wilson, K. M., Fleyfel, I. et al., Isotype-switched immunoglobulin G antibodies to HIV Gag proteins may provide alternative or additional immune responses to 'protective' human leukocyte antigen-B alleles in HIV controllers. *Aids* 2013, *27*, 519–528.
- [77] Yates, N. L., Lucas, J. T., Nolen, T. L., Vandergrift, N. A. et al., Multiple HIV-1-specific IgG3 responses decline during acute HIV-1: implications for detection of incident HIV infection. *Aids* 2011, *25*, 2089–2097.
- [78] Chertova, E., Bess, J. W., Jr., Crise, B. J., Sowder, I. R. et al., Envelope glycoprotein incorporation, not shedding of surface envelope glycoprotein (gp120/SU), is the primary determinant of SU content of purified human immunodeficiency virus type 1 and simian immunodeficiency virus. *J. Virol.* 2002, *76*, 5315–5325.
- [79] Pelchen-Matthews, A., Kramer, B., Marsh, M., Infectious HIV-1 assembles in late endosomes in primary macrophages. *J. Cell Biol.* 2003, *162*, 443–455.
- [80] Linde, M. E., Colquhoun, D. R., Ubaida Mohien, C., Kole, T. et al., The conserved set of host proteins incorporated into HIV-1 virions suggests a common egress pathway in multiple cell types. *J. Proteome Res.* 2013, *12*, 2045–2054.
- [81] Misumi, S., Fuchigami, T., Takamune, N., Takahashi, I. et al., Three isoforms of cyclophilin A associated with human immunodeficiency virus type 1 were found by proteomics by using two-dimensional gel electrophoresis and matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J. Virol.* 2002, *76*, 10000–10008.
- [82] Mascarenhas, A. P., Musier-Forsyth, K., The capsid protein of human immunodeficiency virus: interactions of HIV-1 capsid with host protein factors. *FEBS J.* 2009, *276*, 6118–6127.
- [83] Saphire, A. C., Galloway, P. A., Bark, S. J., Proteomic analysis of human immunodeficiency virus using liquid chromatography/tandem mass spectrometry effectively distinguishes specific incorporated host proteins. *J. Proteome Res.* 2006, *5*, 530–538.
- [84] Yang, W., Zhou, J. Y., Chen, L., Ao, M. et al., Glycoproteomic analysis identifies human glycoproteins secreted from HIV latently infected T cells and reveals their presence in HIV+ plasma. *Clin. Proteomics* 2014, *11*, 9.
- [85] Nguyen, E. V., Gharib, S. A., Crothers, K., Chow, Y. H. et al., Proteomic landscape of bronchoalveolar lavage fluid in human immunodeficiency virus infection. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 2014, *306*, L35–L42.
- [86] Cuevas-Cordoba, B., Santiago-Garcia, J., Saliva: a fluid of study for OMICS. *Omic*s 2014, *18*, 87–97.
- [87] Bora, A., Ubaida Mohien, C., Chaerkady, R., Chang, L. et al., Identification of putative biomarkers for HIV-associated neurocognitive impairment in the CSF of HIV-infected patients under cART therapy determined by mass spectrometry. *J. Neurovirol.* 2014, *20*, 457–465.
- [88] Mann, M., Comparative analysis to guide quality improvements in proteomics. *Nat. Methods* 2009, *6*, 717–719.
- [89] Patel, V. J., Thalassinou, K., Slade, S. E., Connolly, J. B. et al., A comparison of labeling and label-free mass spectrometry-based proteomics approaches. *J. Proteome Res.* 2009, *8*, 3752–3759.
- [90] Bantscheff, M., Lemeer, S., Savitski, M. M., Kuster, B., Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal. Bioanal. Chem.* 2012, *404*, 939–965.
- [91] Zhang, N., Zhang, Z., Feng, S., Wang, Q. et al., Quantitative analysis of differentially expressed saliva proteins in human immunodeficiency virus type 1 (HIV-1) infected individuals. *Anal. Chim. Acta* 2013, *774*, 61–66.
- [92] Gillet, L. C., Navarro, P., Tate, S., Rost, H. et al., Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* 2012, *11*, O111 016717.
- [93] Villeneuve, L. M., Stauch, K. L., Fox, H. S., Proteomic analysis of the mitochondria from embryonic and postnatal rat

- brains reveals response to developmental changes in energy demands. *J. Proteomics* 2014, 109C, 228–239.
- [94] Wiederin, J. L., Donahoe, R. M., Anderson, J. R., Yu, F. et al., Plasma proteomic analysis of simian immunodeficiency virus infection of rhesus macaques. *J. Proteome Res.* 2010, 9, 4721–4731.
- [95] Wiederin, J. L., Yu, F., Donahoe, R. M., Fox, H. S. et al., Changes in the plasma proteome follows chronic opiate administration in simian immunodeficiency virus infected rhesus macaques. *Drug Alcohol Depend.* 2012, 120, 105–112.
- [96] Angel, T. E., Jacobs, J. M., Spudich, S. S., Gritsenko, M. A. et al., The cerebrospinal fluid proteome in HIV infection: change associated with disease severity. *Clin. Proteomics* 2012, 9, 3.
- [97] Jensen, K., Goo, Y. A., Yahiaoui, A., Bajwa, S. et al., Identification of fatigue biomarkers in treated and treatment-naive HIV patients: preliminary results. *Biol. Res. Nurs.* 2013, 16, 278–287.
- [98] Banerjee, S., Liao, L., Russo, R., Nakamura, T. et al., Iso-baric tagging-based quantification by mass spectrometry of differentially regulated proteins in synaptosomes of HIV/gp120 transgenic mice: implications for HIV-associated neurodegeneration. *Exp. Neurol.* 2012, 236, 298–306.
- [99] Burgener, A., Boutilier, J., Wachihi, C., Kimani, J. et al., Identification of differentially expressed proteins in the cervical mucosa of HIV-1-resistant sex workers. *J. Proteome Res.* 2008, 7, 4446–4454.