REVIEW

For reprint orders, please contact: reprints@futuremedicine.com

Mass spectrometry-based proteomic approaches for discovery of HIV–host interactions

Yang Luo^{*1} & Mark A Muesing¹

ABSTRACT A molecular understanding of viral infection requires a multi-disciplinary approach. Mass spectrometry has emerged as an indispensable tool to investigate the complex and dynamic interactions between HIV-1 and its host. It has been employed to study protein associations, changes in protein abundance and post-translational modifications occurring after viral infection. Here, we review and provide examples of mass spectrometry-based proteomic approaches currently used to explore virus–host interaction. Efforts in this area are certain to accelerate the discovery of the unique molecular strategies utilized by the virus to commandeer the cell as well as mechanisms of host defense.

Viruses, as intracellular parasites, rely heavily on host factors to complete their life cycle. They have evolved to manipulate the cellular machinery, evade the host immune response and hijack signaling pathways for their own benefit. Identification of those requisite cellular factors is invaluable to gain an intimate understanding of viral invasion and host defense strategies, as exemplified by recent proteomic studies of influenza A [1], hepatitis C virus [2] and herpes simplex virus type 1 [3]. It would also provide potential targets for antiviral intervention, particularly for viruses causing pandemic diseases such as HIV-1, the major focus of this review.

HIV-1 is a retrovirus with an RNA genome of approximately 9200 nucleotides. It encodes nine open reading frames, capable of producing 15 mature viral proteins after proteolytic cleavage. Given the limited genetic capacity of the virus and the complex molecular events occurring during infection, it is likely that the virus interacts with an extensive array of cellular factors to sustain its passage through the host [4]. Despite the expanding knowledge and extensive efforts in the field of HIV biology, most of the elaborate interplay between the virus and its host remain obscured from scientific investigation. Various molecular and biochemical techniques have been employed to identify cellular genes or gene products essential for retroviral infection [5–9]. However, each screening technique comes with its own inherent drawback(s) and many limit the study to individual host factors instead of multicomponent complexes. Chemical mutagenesis induces mutations in cellular coding regions to create cell lines resistant to viral replication [10,11]. However, the relevant mutant gene cannot be readily identified. This fact, coupled with the possibility that several genes might be altered in parallel, makes it extremely difficult to identify the mutations responsible for the associated phenotype. Insertional mutagenesis randomly inactivates host genes by retroviral integration [12,13], whereas cDNA complementation with cellular genes has led to discoveries of host receptors or co-receptors required for viral entry [14] as well as restriction factors against viral infection [15–17]. Nevertheless, both approaches will succeed only when one gene product is required for either its positive or negative effect on viral infection. Furthermore, insertional mutagenesis

KEYWORDS

• affinity purification • HIV-1
• MS-based proteomics
• post-translational modification • quantitative MS • virus–host interactions

¹Aaron Diamond AIDS Research Center, Affiliate of The Rockefeller University, 455 First Avenue 7th Floor, New York, NY 10016, USA

^{*}Author for correspondence: yloeo@adarc.org

inactivates only one homolog and therefore can identify only those genes whose corresponding gene products are amenable to trans-dominant interference or can be detected by their functional haplo-insufficiency. Direct binary interactions between host and viral proteins have also been identified by the yeast two-hybrid system (Y2H) [18,19]. However, Y2H requires tight interactions that take place in the yeast fungi nucleus, and is notorious for producing false positives. Moreover, this approach is not optimized for the study of most viral proteins in their natural mammalian intracellular environment. Gene down-modulation strategies using antisense or siRNA have been employed for functional identification of genes required for infection. Apart from off-target cross-reactivity and variable efficiencies of gene knockdown, studies of HIV-1 infection have resulted in incongruent sets of host proteins required for viral replication [8]. It is certain that the scope of the host machinery usurped by HIV-1 remains largely unknown and new approaches are necessary to identify those interactions acting at the molecular interface between the virus and its host. Furthermore, systematic approaches may uncover interaction networks and therefore help to decipher the interconnection between different viral and cellular components, providing a dynamic picture of the basic biological strategies employed by viruses to gain control of the host cell.

In recent years, mass spectrometry (MS)-based proteomic technologies have advanced immensely to facilitate the identification and characterization of host factors required for viral infection. Host proteins have been found to be involved in different aspects of viral replication ranging from viral entry, chemokine receptor-mediated signal transduction, intracellular trafficking, transcription, chromatin remodeling and mRNA export and assembly. Here we describe the MS-based approaches developed for studying viral and cellular proteins that are functionally, structurally and dynamically associated during HIV-1 infection, demonstrating the exceptional technical capabilities in exploring individual interactions and quantitating intracellular changes. A typical workflow for protein MS analysis is illustrated in **Figure 1**. The key to successful MS-based proteomic study is the combination of good experimental design, proper sample preparation, statistical analysis of the MS data and, ultimately, functional validation of identified candidates. We summarize

recent advances in areas where MS-based proteomics can make unique contributions in probing virus–host interactions. Their application to the study of HIV-1 will be discussed.

Affinity purification MS approaches to identify HIV–host protein interactions

In recent years, MS coupled to diverse affinity-based purification steps has emerged as a powerful tool to elucidate the protein interactions between virus and host [20–22]. Originally carried out in large-scale studies of protein complexes in yeast [23], affinity purification (AP) followed by MS (AP-MS) has been used extensively to identify interaction partners in complex with a given protein of interest (POI). Based on the macromolecular interactions between an ectopically expressed viral protein and its reactive host cell protein(s), these methods have the capability to investigate the protein interaction as a potentially physiologically relevant complex or network. In this type of study, it is essential that the target protein and associated complexes can be efficiently purified. When potent antibodies specifically targeting the POI are available, endogenous protein immunoprecipitation can be performed to map its interactions [24]. However, as in the case of HIV-1, suitable antibodies raised against an individual viral protein are not always available. As a common alternative, a potent affinity tag can be engineered as a fusion with either the N- or C-terminus of the target viral protein. Earlier approaches identified cellular factors by AP followed by immunoblotting [25,26]. Being hypothesis driven, however, this method relies heavily on prior knowledge of target protein function. During the last decade, advances in AP-MS have allowed unbiased assignment of viral-binding proteins and led to surprising discoveries of previously unknown cellular functions and pathways commandeered by the virus. A number of cellular factors have been identified by this conventional and straightforward means. For instance, LEDGF/p75, a protein implicated in regulation of cellular gene expression and stress response, was identified by liquid chromatography (LC)-MS/MS in complex with HIV-1 integrase (IN) in nuclear extracts of cells expressing C-terminal FLAG-tagged IN [27]. Characterized as a host factor responsible for HIV-1 integration site selection, the efficient knockdown of LEDGF severely impairs integration of the viral DNA into the host chromosome and viral replication [28].

Figure 1. Schematic representation of a generalized mass spectrometry-based proteomic workflow.

Currently, the interaction between IN and LEDGF is a validated target for anti-HIV drug discovery [29,30].

In a very recent study, six independent AP-MS screens were performed to identify potential cellular interaction partners of HIV-1 p55 precursor polyprotein Gag, using as bait either a C-terminal green fluorescent protein-tagged Gag or Gag tagged with internal green fluorescent protein insertion between its matrix and capsid domains and expressed in the presence of additional HIV-1 proteins [31]. Although the position and identity of the tag was shown to influence the composition of the recovered interactors, a common subset of host proteins was identified. This analysis showed an enrichment of distinct protein motifs and molecular pathways associated with Gag, including cytoskeleton, SR proteins and other RNA-binding proteins, providing a previously unknown source of cellular interaction partners of HIV-1 Gag.

As a variation of the AP-MS method, tandem AP (TAP), in which two contiguous epitope tags are incorporated into the POI followed by two-stage purification, offers an effective and highly specific means to isolate the target protein and its associated complexes [32,33]. The original TAP tag consists of calmodulin-binding peptide followed by a tobacco etch virus (TEV) protease cleavage site and Protein A. As this technique has been widely exploited, a number of alternative epitope-tag combinations such as FLAG-hemagglutinin (HA) have also been developed.

Successive AP steps greatly reduce the amount of nonspecific binding contamination, benefiting downstream MS identification. Using tandem streptavidin calmodulin binding peptide tagging at the N-terminus of HIV-1 Vpr, damage-specific DNA-binding protein 1 was identified by MS as the predominant cellular protein in a Vpr complex [34]. Studies showed that interaction with DNA-binding protein 1 mediates Vpr-induced apoptosis and G2 arrest [34,35]. In addition, this method has been extended to explore cellular proteins bound to particular domains of viral proteins. For example, to elucidate the functions of the highly conserved 151 amino acids (aa) C-terminal cytoplasmic domain (envelope [Env]-CT) of the HIV-1 glycoprotein, TAP-MS led to the identification of the Phb1/2 heterodimer as an interaction partner [36]. The finding pointed to the importance of Phb1–Phb2 complex binding to Env-CT for the infection of nonpermissive cell types, and possibly sheds light on an unidentified function essential for HIV-1 replication.

An interesting twist to the AP-MS approach can be applied to uncover antiviral factors in the absence of their cognate viral antagonist. SAMHD1, a host protein that is partially responsible for inefficient HIV-1 infection of dendritic cells and macrophages, was identified using TAP-MS approaches [37,38]. Unlike HIV-1, both HIV-2 and related simian immunodeficiency viruses can efficiently transduce those cell types owing to their virion-associated

Vpx accessory protein. Given that Vpx might overcome the host's antiviral mechanism by recruiting cellular target(s) toward degradation pathways, proteomic screens were used to search for cellular protein(s) that associate with tagged Vpx. Furthermore, factors were identified that bind to the CRL4^{DCAF1} ubiquitin ligase complex only in the presence of Vpx [37]. Both studies identified SAMHD1, a dGTP-dependent deoxynucleoside triphosphohydrolase, as the antiviral protein. Vpx protein mediates the degradation of SAMHD1 by recruiting it to the CRL4^{DCAF1} complex. To elucidate its mechanism of action, MS mapped the phosphorylation sites within SAMHD1 [39]. Mutational analysis of one of the identified sites (T592) indicates that SAMHD1 phosphorylation might be a negative regulator of its antiviral activity. To further explore SAMHD1 related complexes, AP-MS was performed using HA-tagged human SAMHD1 to identify two additional host proteins, cyclin-dependent kinase 2 and S-phase kinase-associated protein 2 [40]. Together, these studies revealed cyclin-dependent kinase 2 as a cofactor responsible for T592 phosphorylation thereby regulating the function of SAMHD1-mediated HIV-1 restriction. As for why HIV-1 does not counteract SAMHD1, a possible explanation is that Vpx deficiency could be beneficial for HIV-1 by avoiding the triggering of a strong 'cryptic sensor' immune response in dendritic cells [41]. Such discoveries, assisted by MS, will lead to deeper understanding of the sophisticated interplay existing between the virus and host.

In a more comprehensive proteomic study, AP-MS was used to systematically identify the binding partners of all 15 HIV-1 proteins and three precursor polyproteins (Gag, Gag-Pol and Env gp160) by overexpressing affinity-tagged versions of each protein individually in two different mammalian cells including a human T-cell line [22]. Although performed in the absence of viral infection, this study presents an interaction landscape of these proteins and their corresponding host protein complex(es). This study provided a broad identification of potential host factors that may impact HIV-1 replication.

In AP-based approaches, the position of the tag might affect the conformation or mobility of the POI, potentially interfering with normal protein folding or function. Therefore whenever possible, the proper enzymatic function and subcellular localization of POI after tag insertion should be confirmed. Furthermore, it is preferable to

maintain the natural level of protein expression as overexpression can lead to altered intracellular localization and has the potential to result in the formation of nonphysiological protein contacts [42]. To address this concern for cellular proteins, transgenes have been cloned using bacterial artificial chromosomes along with their endogenous regulatory sequences to ensure an expression level that closely matches that of the endogenous gene [43]. However, existing strategies do not recapitulate the differential expression levels, sometimes with temporal variation, of the viral proteins at various stages during HIV-1 infection. In addition, the late genes of HIV-1 are conditionally expressed from intron-containing mRNAs and depend on the Rev protein for their nuclear export and translation. Therefore codon optimization is necessary to remove inhibitory sequences to achieve a suitable level of expression in mammalian cells. An interesting alternative is to insert an epitope tag in the context of a full-length infectious virus, and isolate the complexes directly during active infection. This strategy ensures the isolation of physiologically relevant virus–host complexes as its expression is under native regulatory control. As a notable example, an elegant study characterized cellular factors that interact with HIV-1 accessory protein Vif. Vif is essential for viral evasion of the host antiviral factor APOBEC3G (A3G), a potent mediator of anti-HIV-1 replication [15] which induces hypermutation in newly synthesized minus-strand viral DNA converting cytosines to uracils and ultimately damaging viral genomic integrity [44]. An infectious HIV-1 clone was constructed in which the C-terminus of Vif was fused to an HA tag to efficiently isolate Vif from infected cells. Associated cellular proteins were subsequently identified by AP-MS. The results show that Vif protects the viral genome by recruiting A3G to a Cul5–ElonginB/C complex and inducing its ubiquitination and degradation, thereby blocking incorporation of A3G into egressing viral particles [45]. Further study using the same system identified another Vif interacting protein, CBF β [46]. Interestingly, CBF β was independently isolated as a Vif-binding factor in the global landscape study described above [22,47]. These studies demonstrate mechanisms of HIV-1 evasion of host antiviral defenses.

One important prerequisite for successful implementation of this AP strategy is the generation of infectious virus carrying an exogenous epitope tag. Maintenance of replication

competency upon tag insertion is crucial since the location of the tag can disrupt normal protein function and obscure physiologically relevant complex formation with the native viral protein, likely resulting in defective virus. Despite the few exceptions whereby tags can be accommodated at the termini of a viral protein without loss of replication competency, HIV-1 notoriously cannot tolerate foreign insertion and becomes unstable with deleterious effect. This is because HIV-1 genome is exceptionally compact and all gene products are essential. Genomic hypervariation and reverse transcription (RT)-mediated recombination can further exacerbate the situation and result in the deletion of exogenous sequences during subsequent viral passage. Moreover, tagging the termini of the components generated from either Gag or Gag-Pol polyprotein processing may disrupt the pattern of viral protease cleavage. Last, the existence of several overlapping reading frames (e.g., *pol* overlaps with *p6* and *vif*, *env* overlaps with *rev* and *tat*) poses a particular challenge as tag insertion in one viral protein might disrupt the expression of another.

In an endeavor to systematically identify factors that directly interact with the HIV-1 machinery during viral replication, we developed a unique strategy to engineer HIV-1 to incorporate a potent epitope tag but remain robustly infectious. Transposon-based saturation linker scanning mutagenesis [48] was performed on individual gene segments of the HIV-1 genome. This creates a large library of viral clones, each uniquely marked with a 5-aa in-frame insertion including a unique restriction endonuclease site that can be further accessed to insert an exogenous epitope tag at random locations. Viruses that persist through multiple rounds of passage, a stringent selection step for replication competency, are isolated as they can accommodate a tag without loss of biological ‘fitness’. In turn, replication-competent viruses harboring a 3x FLAG tag were generated. This allows for the large-scale affinity capture of viral–host macromolecular complexes formed during infection for subsequent MS analysis. Since the modified viruses are replication competent, each tagged protein is engaged with the authentic repertoire of cellular factors required during natural infection, facilitating a comprehensive definition of the HIV–host interactome. Most importantly, because the complexes are isolated directly from infection, this approach makes it possible for the

first time to directly explore the virus–host interactions occurring at various stages throughout the viral life cycle. We believe this systematic tagging strategy can be extended to other viruses and pathogens, benefiting the proteomic study of pathogen–host interaction dynamics and our understanding of the molecular details of pathogen establishment and persistence.

Extended AP-MS-based methodologies

In addition to protein–protein complexes, several viral proteins (i.e., IN, RT, NC, Rev and Tat) interact with viral nucleic acids and presumably form nucleoprotein complexes with many host proteins. For example, integrated viral DNA is regulated by an intricate interplay among viral regulatory proteins and host cellular transcription factors acting at the viral long terminal repeat. Such DNA-binding proteins, often in low abundance, can be captured via oligonucleotide probes immobilized on a chromatographic support and identified by MS. A one-step DNA-affinity capture MS method has been described to identify many cellular transcriptional regulators in association with the HIV-1 5'-long terminal repeat [49]. Among them, the homeodomain-containing transcription factor, myeloid ectopic integration site, was confirmed to functionally downregulate HIV-1 expression *in vitro*. This nucleic acid affinity MS approach can be an efficient screening tool to identify proteins that associate with a particular DNA and/or RNA sequence of interest. Using a similar approach, catalytically active preintegration complexes (PICs) from HIV-1 infected CD4⁺ cells were isolated using a biotinylated target DNA, and the copurified proteins identified by MS [50]. The study identified 19 host proteins associated with PICs, whose function ranged from chromatin organization to protein transport. Validation and detailed characterization of these proteins could provide insights into the early steps of infection.

Viral and host proteins form not only stable but also transient macromolecular complexes with essential roles, such as enzyme–substrate binding, signaling cascades and regulatory molecular switches, all contributing significantly to the dynamic control of protein function. In conventional AP-MS studies, these types of interactions often fall outside of the dynamic range of MS measurement and remain undetectable. Chemical cross-linking has been employed to capture transient or weak interactions. Although

successfully incorporated into MS analysis with possibility to elucidate the spatial and temporal regulation of disease-related mechanisms [51], cross-linking currently has been underutilized in the study of transient HIV–host interactions. Alternatively, weak or transient interactions can be selectively enriched from the relevant subcellular compartment. For example, to investigate cofactors that bind to the HIV-1 RT complex (RTC), MS was used to identify cellular proteins from the fractionation of cell lysates that stimulate late RT *in vitro* [52]. Among the 25 proteins identified from four replicate experiments, both the eEF1A and eEF1G subunits of eEF1 were identified as functional RT cofactors, providing a basis for studying RT and trafficking of the RTC to the nucleus. In a similar study, proteomic analysis was carried out in RTCs and PICs partially purified by gradient centrifugation fractionation [53]. LC-MS/MS analysis of seven replicates identified 94 cellular proteins unique to the infected fractions. This study may contribute additional candidates involved in viral replication.

Cellular proteins found in virions may also contribute to viral pathogenesis. However, these analyses are complicated by contamination of the sample with microvesicles, exosomes or other nonvirion protein-containing particulates. Therefore, cellular protein profiling of virions requires well-designed separation methods to efficiently remove these contaminants. Proteomic composition of HIV-1 virions produced in lymphocytes, epithelial cell model systems [54] or monocyte-derived macrophages [55] have been subjected to MS analysis. These studies revealed large numbers of cellular proteins not previously described to be in association with HIV-1, providing important leads for further investigation. In particular, MS identified clathrin, a cytosolic protein functioning in vesicle genesis and transport, not only as a highly abundant protein within the virion but also recruited with high specificity [56,57]. These studies implied that the virus might employ clathrin to facilitate accurate morphogenesis of infectious particles, possibly preventing premature proteolytic processing of the virion polyproteins during assembly.

MS-based methods for mapping post-translational modification

In addition to protein identification, MS-based proteomic methodologies have also been utilized

in the characterization of post-translational modification (PTM). Covalent alteration of specific amino acid side chains, such as phosphorylation, acetylation, ubiquitination, methylation and glycosylation, are commonly involved in regulatory pathways and represents a diversity of protein isoforms. Advances in MS along with efficient enrichment and separation methodologies have greatly improved the detection of PTMs during infection. For example, MS identified phosphorylation of several serine and threonine residues in p6, the C-terminal domain of HIV-1 Gag with titanium dioxide for phosphopeptide enrichment and LC-MS/MS. The phosphorylation profile guided further p6 mutagenic studies [58]. To test if T-cell receptor signaling induces critical PTMs enhancing interactions between P-TEF β and the HIV-1 transactivator protein Tat, AP-MS/MS was performed to define key PTMs on P-TEF β subunits in response to T-cell receptor signaling [59]. The results showed that phosphorylation of CDK9 at S175 played a critical role in altering the competitive binding of Tat and bromodomain protein BRD4 to P-TEF β .

As a regulatory PTM in cell signaling cascades, histones and nonhistone proteins can serve as substrates for different histone acetyltransferase enzymes, leading to acetylation. MALDI-MS was used to map the sites of p300-mediated acetylation of Tat *in vitro* [60]. Interestingly, in addition to the Tat K50 acetylation site in its RNA-binding region, the study showed that the cysteine-rich region, necessary for Tat transactivation activity, is also acetylated at multiple cysteine residues. Another study using matrix-assisted laser desorption/ionization (MALDI)-MS mapped the acetylation sites of NF- κ B to its DNA-binding domain, and reported Tat enhanced the *in vitro* acetylation of the NF- κ B p50 subunit, increasing p50 DNA-binding affinity [61]. Taken together, these studies reveal new insights into Tat regulation of viral transcriptional activity.

Ubiquitination affects cellular processes by regulating proteasomal degradation, cellular localization and transcriptional regulation. In a follow-up study of HIV-1 Vif interaction described above, ubiquitin (Ub) remnant profiling, using an antibody recognizing the K-GG motif of trypsinized Ub peptides to enrich for ubiquitinated proteins prior to MS analysis [62], was employed to identify Ub acceptor sites in A3G and A3F targeted by HIV-1 Vif [63]. This unbiased proteomic approach identified

dispersed internal lysines as the dominant polyUb acceptor sites, facilitating understanding of APOBEC3 binding to the Vif–Ub ligase complex.

Methylation, has been implicated in transcriptional regulation, epigenetics, DNA repair, mRNA splicing and signal transduction. For example, arginine methylation, catalyzed by PRMT6, has been documented to interfere with the functions of HIV-1 Tat, Rev and nucleocapsid and inhibit HIV-1 replication in culture. Using purified recombinant PRMT6 in an *in vitro* methylation assay, LC-MS/MS identified R35 as the target residue for automethylation in PRMT6. Mutagenesis has shown that automethylation of PRMT6 at R35 regulates its stability and its ability to restrict HIV-1 replication during viral production as well as subsequent rounds of viral growth [64].

HIV-1 Env spike, a trimer of gp120 and gp41 subunits, is one of the most highly N-glycosylated structures found in nature, with 50% of its mass comprised of carbohydrate [65]. The analysis of Env carbohydrates is critical for understanding HIV-1 transmission and immune escape. Glycopeptide-based mass mapping approaches have been employed to categorize the glycan species on the native Envs from transmitted/founder viruses [66] and primary isolates of different viral clades [67]. In contrast to recombinant monomeric gp120 produced by transfection, the oligomannose profile of purified virions was drastically different, highly conserved among the clades and almost devoid of complex-type glycans. Interestingly, one study also noted that specific glycans recognized by the broadly neutralizing antibody, 2G12, were significantly more abundant on the native Env than on the recombinant monomer. Such studies extend the understanding of the HIV-1 glycan shield and its roles in infection and immune evasion with implications for vaccine design.

Quantitative proteomic approaches for virus–host interactions

A systematic analysis of temporal and spatial changes in host and viral protein throughout the course of a productive infection would provide dynamic insights into critical virus–host interactions. Although various molecular approaches have been used to probe alterations induced by infection, changes in mRNA level do not always correlate with changes at the protein level. Proteomics, therefore, can provide a

more extensive and complementary description of the cellular mechanisms altered during the infectious cycle. In the last decade, MS-based approaches have advanced significantly to allow accurate and sensitive quantitation of proteins and peptides in complex biological systems. These quantitative measures use label-free, chemical tagging and stable isotopic labeling to detect small changes in peptide and protein abundance. Various methodologies, such as differential in gel electrophoresis [68] and LC-MS/MS coupled with quantitative proteomics and bioinformatics [69], have been employed to detect cellular changes that occur during viral infection as well as to study the interactome of viral proteins with the host cell.

Label-free quantification via spectral counting and signal intensity measurement of peptide peaks are the most intuitive means of obtaining quantitative information. It can provide robust analysis of comparative protein expression if performed with sufficient replicates with minimal analytical variation. Based on the observation that more abundant proteins will produce more MS/MS spectra and abundant peptides triggering more fragment ion scans than ones of lower abundance, spectral counting quantifies proteins by comparing the number of spectra originating from identical proteins in different samples. Similarly, MS peptide ion intensity can be used to quantify peptide abundance by integrating the peak areas of the same ion species from different samples. The major advantages of label-free techniques are that they are fast and easy to perform. For example, shotgun LC-MS/MS analysis of *ex vivo* infection of human primary CD4 cells revealed two distinct profiles at early and late stages of HIV-1 infection [70]. The detection specificity was enhanced by multiplexed comparison of protein abundance changes over time, as well as in the presence of an antiviral RT inhibitor included as a negative control. This signal intensity based label-free strategy highlighted the temporal effects of HIV-1 infection on cellular protein profiles and pinpointed the specific rerouting of cellular pathways dependent upon infection. In another study, significantly elevated expression of certain factors that influence the success of HIV-1 infection was detected in the activated T cells [71].

Although label-free methods can be easily incorporated into experiments, they cannot reliably detect small changes in quantity or proteins of low abundance. Techniques

incorporating chemical or metabolic labeling have been introduced to address the necessity for more accurate quantification. Stable isotopic labeling techniques are based on the introduction of a differential mass label without affecting the chemical properties of the protein. Chemical labeling such as isotope-coded affinity tag and isobaric tag for relative and absolute quantitation (iTRAQ) introduce a covalent label to proteins or peptides. In the isotope-coded affinity tag approach, isotopically light or heavy marked linkers specifically bind to cysteine residues [72]. However, it requires cysteine-containing peptides and reduces the ability to detect PTMs. iTRAQ, a multiplex isotopic method instead labels the N-terminus and lysine residues of a given peptide or protein [73]. It introduces mass-balanced labels at the level of tryptic peptides, resulting in labeled peptides of the same total mass that differentiate upon peptide fragmentation during MS/MS. Despite variability in labeling efficiencies and protein digestion, iTRAQ can greatly facilitate peptide identification due to the enhanced intensities of the parental and fragment ions, allowing quantification of two to eight parallel samples. It has been applied as an investigative tool to quantify the differential host proteome composition in HIV-1 virions purified from multiple cell types [74] and in the characterization of the early host cell responses in infected CD4⁺ lymphoblastoid cells [75]. However, due to its sample complexity, iTRAQ is often considered to be limited in dynamic range and may have accuracy or precision issues when compared with other methods [76].

Despite its relative high cost and requirement of metabolically active culture, metabolic stable isotopic labeling provides ubiquitous labeling of all expressed proteins therefore minimizing experimental variation and improving the accuracy for relative quantification as well as the detection of PTMs. First utilized in a study to quantify modifications at specific sites in individual proteins [77], stable isotope labeling of aa in cell culture (SILAC) utilizes the cellular metabolic machinery to achieve the incorporation of heavy isotopic aa (usually ¹³C₆ lysine, ¹³C₆ and/or ¹⁵N₇ arginine) to differentially label all proteins in cell culture [78]. Several cycles of cell division ensures that all tryptic peptides contain at least one heavy-isotope labeled amino acid that can be distinguished from their sibling light peptide (i.e., ¹²C₆ lysine and ¹²C₆-¹⁴N₇ arginine) derived from a differential experimental condition. A

SILAC-based quantitative analysis was carried out to characterize the Vpr response in macrophages [79]. The study identified 136 cellular proteins that are significantly altered upon Vpr overexpression, indicating Vpr's ability to modulate macrophage metabolic pathways. In another study, SILAC quantitative phosphoproteomics was used to examine the CXCL12/CXCR4 signaling pathway, implicated in HIV-1 infection, in the human lymphoblastic CEM cell line in an unbiased fashion [80]. The findings can potentially expand the understanding of signaling pathways involving CXCL12 signaling and further our understanding of chemokine signaling regulation by phosphorylation. To examine the signaling events induced during HIV-1 entry into primary human CD4⁺ cells, a similar quantitative screen established that the phosphorylation profiles of 175 gene products changed significantly within 1 min after exposure to HIV-1, including several proteins in pathways known to be affected by HIV-receptor binding as well as proteins of the alternative splicing machinery [81].

Unequivocal assignment of a given intracellular interaction using affinity-based extraction of protein complexes is frequently compromised by the presence of nonspecifically interacting proteins that co-enrich during purification. A particularly advantageous application of metabolic labeling, termed isotopic differentiation of interactions as random or targeted (I-DIRT), has been used to unambiguously assign authentic interactions and discriminate against those contaminants that bind nonspecifically to isolated protein complexes [82]. We have employed I-DIRT to probe the virus–host interactome during HIV-1 infections in our studies described above. Cells infected with replication competent 3x FLAG-tagged virus were harvested in media supplemented with ¹³C₆ (heavy [H]) lysine and arginine while the untagged parental virus was expanded in standard ¹²C₆ (light [L]) media. An equal mass of cells from each infection providing an equal H/L isotopic ratio was mixed prior to cell lysis and AP. Stable specific association with the tagged viral protein exhibits high H/L ratio in MS analysis, while nonspecific interactions are revealed by approximately 1:1 H/L ratios. The discrimination of specificity afforded by I-DIRT was further confirmed by reversing the isotope labels in growth media. This reverse I-DIRT experiment mitigates false quantitation ratios and provides an additional biological

replicate. This methodology is particularly suitable to identify specific stable interactions that do not easily disassemble over time. However, I-DIRT cannot distinguish authentic binding partners with rapid exchange rates from those proteins binding nonspecifically. Therefore, specific but rapidly exchanging interactions can register as false negatives by this method, requiring independent experimental confirmation of their *bona fide* association within the cell.

Validation of MS-based proteomic data

Despite its exceptional capabilities, MS-based proteomics is not without its limitations. The large number of identified candidates in most studies, confounded by the high sensitivity of MS measurement, presents challenges for determining biological relevancy. First, it remains difficult to distinguish true interactions from contaminants in AP-MS. In addition to the I-DIRT strategy described above, the most common solution is to carry out a parallel untagged or an unrelated tagged protein control purification to define the specificity for each interaction. Proteins present in the control AP are considered to be nonspecific if they are of equal abundance in both tagged and control isolation(s), assuming reproducible chromatography and consistent MS accuracy. Computational algorithms based on label-free quantitation has been developed to distinguish true interactors from contaminants [83]. For example, significance analysis of interactome (SAINT) computes confidence scores for protein–protein interaction data generated by AP-MS and derives the probability of *bona fide* interactions by modeling the distribution for true and false interactions [84]. Perhaps currently the only known algorithm designed for the study of HIV–host interactions, MS interaction statistics analyzes interaction specificity using the weighed sum of protein abundance measured by peak intensity, reproducibility across replicates and uniqueness of a particular interaction in all AP-MS performed using different tagged HIV-1 proteins [22]. In addition to computer algorithms, protein interaction databases can help in the assessment of interaction specificity. The contaminant repository for AP (CRAPome) [85] characterizes contaminants associated with particular experimental protocols by aggregating negative control data from multiple AP-MS studies [86]. In addition, the Virus Molecular Interaction (VirusMINT) database [87] aims to collate all protein interactions reported in the

literature between hosts and viruses, including HIV-1 [88].

The frequently complex data set of proteomic analysis presents a challenge for deducing biologically significant targets and pathways. By retrieving information from a variety of databases, computational tools can facilitate the visualization and functional analysis of protein networks, and in turn, infer the biological significance of proteomic findings. In addition, clusters visualized in the functional network may also serve as a confirmation on the biological relevance of the data sets. The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database [89], for example, creates interactive visualization for predicted functional associations between proteins based on experimental data and literature searches [90]. Ingenuity Pathway Analysis database [91], constructing physical and functional links among cellular proteins, has been successfully employed to predict disruption of cellular function pathway(s) during viral infection [92]. Functional analysis of clustered protein association based on gene ontology (GO) can facilitate the interpretation of large interactome data sets by categorizing interaction into biological processes, cellular components and molecular function. ClueGO, a GO annotation database, creates a functionally organized GO/pathway network, providing an intuitive representation of the GO classification [93].

Although statistical analysis has significantly advanced proteomic studies, the intrinsic complexity of protein isolation and MS analysis poses limitations to the authenticity of the interactions identified. Therefore, independent assessments are required to functionally validate true intracellular interactions. These approaches include immunofluorescence staining, fluorescence resonance energy transfer and bimolecular fluorescence complementation for protein relocation and co-localization, mutagenic profiling, DNA/RNA microarrays, target gene knockdown and biochemical characterization. For example, as illustrated in some of the studies described here, the specificity of Phb1–Phb2 complex for the Env-CT domain was confirmed by co-immunoprecipitation, and the interaction between Phb1–Phb2 and the full length Env was also verified [36]. In addition to co-immunoprecipitation [22], live cell imaging and fluorescence resonance energy transfer [46] confirmed the CBF β –Vif

interaction. Moreover, knockdown of endogenous CBF β expression demonstrated that depletion of the protein impaired the ability of Vif to coordinate the ubiquitination of A3G [46,47].

Conclusion

Advances in MS-based proteomic methodologies have revolutionized our ability to decipher virus–host interactomes, allowing the discovery of previously unexplored processes and pathways, providing insights into how HIV-1 commandeers the host machinery, evades the

human immune system and persists throughout the lifetime of the infected individual (Figure 2). Indeed, the rapid advances in MS-based proteomic strategies have already aided in the identification and characterization of several critical interactions established during the viral life cycle. However, current challenges arise when the ever-expanding repertoire of HIV–host interactions remain descriptive but functionally unverified. To elucidate the pathogenic biological networks, further development of computational tools and high-throughput technologies as well as downstream functional

Figure 2. Summary of current mass spectrometry-based approaches for the study of HIV–host interactions.

AP: Affinity purification; ICAT: Isotope-coded affinity tag; I-DIRT: Isotopic differentiation of interactions as random or targeted; iTRAQ: Isobaric tag for relative and absolute quantitation; PTM: Post-translational modification; TAP: Tandem affinity purification.

validation is required. We believe that the next generation of proteomic methods will expand our ability to study the temporal and spatial dynamics of virus–host interactions. Such a comprehensive view of the viral interactome will afford the discovery of new cellular interactors, providing the impetus for unique strategies of antiviral intervention against those factors in part dispensable in the host but required for viral pathogenesis.

Future perspective

MS-based proteomics strategies will lead to more comprehensive examination of the temporal and spatial dynamics of virus–host interactions. We believe that such an integrative

view of the viral interactome will afford the discovery of new cellular interactors, providing the impetus for unique strategies of antiviral intervention.

Financial & competing interests disclosure

Support is acknowledged from the NIH (R01 AI081615 and R21 AI097233, both awarded to MA Muesing) to help defray costs associated with the preparation and publication of this manuscript. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

EXECUTIVE SUMMARY

Affinity purification mass spectrometry approaches to identify HIV–host protein interactions

- Ectopically expressed viral proteins have been used for affinity purification (AP)-mass spectrometry (MS) to identify associated protein complex(es).
- AP in the context of viral infection ensures the isolation of physiologically relevant virus–host complexes.

Extended AP-MS-based methodologies

- Nucleic acid affinity MS approaches can be efficient screening tool to identify proteins that associate with a particular DNA and/or RNA sequence.
- Selective sample enrichment has been utilized for identification of transient interactions and cellular protein profiling of virions.

MS-based methods for mapping post-translational modification

- Advances in MS along with efficient enrichment and separation methodologies have greatly improved the detection of post-translational modifications during infection.

Quantitative proteomic approaches for virus–host interactions

- Quantitative MS uses label-free, chemical tagging and stable isotopic labeling to detect alterations in protein abundance, composition and post-translational modifications related to infection.
- I-DIRT can be used to unambiguously assign specificity to protein interactions.

Validation of MS-based proteomic data

- Computational algorithms can facilitate the determination of interaction specificity and deduce biologically significant targets and pathways.
- Functional validation is essential to confirm suspected roles of MS identified proteins.

References

Papers of special note have been highlighted as:

• of interest; •• of considerable interest

- 1 Lietzen N, Ohman T, Rintahaka J *et al.* Quantitative subcellular proteome and secretome profiling of influenza A virus-infected human primary macrophages. *PLoS Pathog.* 7(5), e1001340 (2011).
- 2 Germain MA, Chatel-Chaix L, Gagne B *et al.* Elucidating novel hepatitis C virus–host interactions using combined mass spectrometry and functional genomics approaches. *Mol. Cell. Proteomics* 13(1), 184–203 (2014).
- 3 Santamaria E, Sanchez-Quiles V, Fernandez-Irigoyen J, Corrales F. Contribution of MS-based proteomics to the understanding of herpes simplex virus type 1 interaction with host cells. *Front. Microbiol.* 3, 107 (2012).
- 4 Sorin M, Kalpana GV. Dynamics of virus–host interplay in HIV-1 replication. *Curr. HIV Res.* 4(2), 117–130 (2006).
- 5 Blanco-Melo D, Venkatesh S, Bieniasz PD. Intrinsic cellular defenses against human immunodeficiency viruses. *Immunity* 37(3), 399–411 (2012).

- 6 Friedrich BM, Dziuba N, Li G, Endsley MA, Murray JL, Ferguson MR. Host factors mediating HIV-1 replication. *Virus Res.* 161(2), 101–114 (2011).
- 7 Goff SP. Host factors exploited by retroviruses. *Nat. Rev. Microbiol.* 5(4), 253–263 (2007).
- 8 Goff SP. Knockdown screens to knockout HIV-1. *Cell* 135(3), 417–420 (2008).
- 9 Van Manen D, Van 'T Wout AB, Schuitemaker H. Genome-wide association studies on HIV susceptibility, pathogenesis and pharmacogenomics. *Retrovirology* 9, 70 (2012).
- 10 Gao G, Goff SP. Somatic cell mutants resistant to retrovirus replication: intracellular blocks during the early stages of infection. *Mol. Biol. Cell* 10(6), 1705–1717 (1999).
- 11 Bruce JW, Bradley KA, Ahlquist P, Young JA. Isolation of cell lines that show novel, murine leukemia virus-specific blocks to early steps of retroviral replication. *J. Virol.* 79(20), 12969–12978 (2005).
- 12 Zheng XH, Hughes SH. An avian sarcoma/leukosis virus-based gene trap vector for mammalian cells. *J. Virol.* 73(8), 6946–6952 (1999).
- 13 Dziuba N, Ferguson MR, O'brien WA *et al.* Identification of cellular proteins required for replication of human immunodeficiency virus type 1. *AIDS Res. Hum. Retroviruses* 28(10), 1329–1339 (2012).
- 14 Deng HK, Unutmaz D, Kewalramani VN, Littman DR. Expression cloning of new receptors used by simian and human immunodeficiency viruses. *Nature* 388(6639), 296–300 (1997).
- 15 Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418(6898), 646–650 (2002).
- 16 Stremlau M, Owens CM, Perron MJ, Kiessling M, Autissier P, Sodroski J. The cytoplasmic body component TRIM5 α restricts HIV-1 infection in Old World monkeys. *Nature* 427(6977), 848–853 (2004).
- 17 Neil SJ, Zang T, Bieniasz PD. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 451(7177), 425–430 (2008).
- 18 Verplank L, Bouamr F, Lagrassa TJ *et al.* Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). *Proc. Natl Acad. Sci. USA* 98(14), 7724–7729 (2001).
- 19 Kalpana GV, Marmon S, Wang W, Crabtree GR, Goff SP. Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. *Science* 266(5193), 2002–2006 (1994).
- 20 Cristea IM, Moorman NJ, Terhune SS *et al.* Human cytomegalovirus pUL83 stimulates activity of the viral immediate-early promoter through its interaction with the cellular IFI16 protein. *J. Virol.* 84(15), 7803–7814 (2010).
- 21 Terhune SS, Moorman NJ, Cristea IM *et al.* Human cytomegalovirus UL29/28 protein interacts with components of the NuRD complex which promote accumulation of immediate-early RNA. *PLoS Pathog.* 6(6), e1000965 (2010).
- 22 Jager S, Cimermancic P, Gulbahce N *et al.* Global landscape of HIV-human protein complexes. *Nature* 481(7381), 365–370 (2012).
- **A comprehensive AP-MS study using individually overexpressed viral proteins to systematically identify the binding partners of each HIV-1 protein**
- 23 Gavin AC, Aloy P, Grandi P *et al.* Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440(7084), 631–636 (2006).
- 24 Malovannaya A, Lanz RB, Jung SY *et al.* Analysis of the human endogenous coregulator complexome. *Cell* 145(5), 787–799 (2011).
- 25 Mulder LC, Chakrabarti LA, Muesing MA. Interaction of HIV-1 integrase with DNA repair protein hRad18. *J. Biol. Chem.* 277(30), 27489–27493 (2002).
- 26 Batonick M, Favre M, Boge M, Spearman P, Honing S, Thali M. Interaction of HIV-1 Gag with the clathrin-associated adaptor AP-2. *Virology* 342(2), 190–200 (2005).
- 27 Cherepanov P, Maertens G, Proost P *et al.* HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J. Biol. Chem.* 278(1), 372–381 (2003).
- 28 Llano M, Saenz DT, Meehan A *et al.* An essential role for LEDGF/p75 in HIV integration. *Science* 314(5798), 461–464 (2006).
- 29 Christ F, Voet A, Marchand A *et al.* Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat. Chem. Biol.* 6(6), 442–448 (2010).
- 30 Luo Y, Muesing MA. Prospective strategies for targeting HIV-1 integrase function. *Future Med. Chem.* 2(7), 1055–1060 (2010).
- 31 Engeland CE, Brown NP, Borner K *et al.* Proteome analysis of the HIV-1 Gag interactome. *Virology* 460–461, 194–206 (2014).
- 32 Puig O, Caspary F, Rigaut G *et al.* The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* 24(3), 218–229 (2001).
- **Describes tandem affinity purification as a tool to allow rapid purification under native conditions of complexes.**
- 33 Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B. A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17(10), 1030–1032 (1999).
- 34 Schrofelbauer B, Hakata Y, Landau NR. HIV-1 Vpr function is mediated by interaction with the damage-specific DNA-binding protein DDB1. *Proc. Natl Acad. Sci. USA* 104(10), 4130–4135 (2007).
- 35 Belzile JP, Duisit G, Rougeau N, Mercier J, Finzi A, Cohen EA. HIV-1 Vpr-mediated G2 arrest involves the DDB1-CUL4AVPRBP E3 ubiquitin ligase. *PLoS Pathog.* 3(7), e85 (2007).
- 36 Emerson V, Holtkotte D, Pfeiffer T *et al.* Identification of the cellular prohibitin 1/prohibitin 2 heterodimer as an interaction partner of the C-terminal cytoplasmic domain of the HIV-1 glycoprotein. *J. Virol.* 84(3), 1355–1365 (2010).
- 37 Hrecka K, Hao C, Gierszewska M *et al.* Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* 474(7353), 658–661 (2011).
- 38 Laguette N, Sobhian B, Casartelli N *et al.* SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* 474(7353), 654–657 (2011).
- 39 Welbourn S, Dutta SM, Semmes OJ, Strebel K. Restriction of virus infection but not catalytic dNTPase activity is regulated by phosphorylation of SAMHD1. *J. Virol.* 87(21), 11516–11524 (2013).
- 40 St Gelais C, De Silva S, Hach JC *et al.* Identification of cellular proteins interacting with the retroviral restriction factor SAMHD1. *J. Virol.* 88(13), 7689 (2014).
- 41 Manel N, Hogstad B, Wang Y, Levy DE, Unutmaz D, Littman DR. A cryptic sensor for HIV-1 activates antiviral innate immunity in dendritic cells. *Nature* 467(7312), 214–217 (2010).
- 42 Berggard T, Linse S, James P. Methods for the detection and analysis of protein-protein interactions. *Proteomics* 7(16), 2833–2842 (2007).
- 43 Poser I, Sarov M, Hutchins JR *et al.* BAC TransgeneOmics: a high-throughput method for exploration of protein function in

- mammals. *Nat. Methods* 5(5), 409–415 (2008).
- 44 Harris RS, Bishop KN, Sheehy AM *et al.* DNA deamination mediates innate immunity to retroviral infection. *Cell* 113(6), 803–809 (2003).
- 45 Yu X, Yu Y, Liu B *et al.* Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* 302(5647), 1056–1060 (2003).
- This notable study is the first to characterize cellular factors that interact with HIV-1 accessory protein Vif directly during active infection by addition of a terminal epitope tag and isolation in the context of a full-length infectious virus.
- 46 Zhang W, Du J, Evans SL, Yu Y, Yu XF. T-cell differentiation factor CBF- β regulates HIV-1 Vif-mediated evasion of host restriction. *Nature* 481(7381), 376–379 (2012).
- 47 Jager S, Kim DY, Hultquist JF *et al.* Vif hijacks CBF- β to degrade APOBEC3G and promote HIV-1 infection. *Nature* 481(7381), 371–375 (2012).
- 48 Biery MC, Stewart FJ, Stellwagen AE, Raleigh EA, Craig NL. A simple *in vitro* Tn7-based transposition system with low target site selectivity for genome and gene analysis. *Nucleic Acids Res.* 28(5), 1067–1077 (2000).
- 49 Tachenay A, Michel S, Dieu M, Payen L, Arnould T, Renard P. Unbiased proteomic analysis of proteins interacting with the HIV-1 5'LTR sequence: role of the transcription factor Meis. *Nucleic Acids Res.* 40(21), e168 (2012).
- 50 Raghavendra NK, Shkriabai N, Graham R, Hess S, Kvaratskhelia M, Wu L. Identification of host proteins associated with HIV-1 preintegration complexes isolated from infected CD4⁺ cells. *Retrovirology* 7, 66 (2010).
- 51 Leitner A, Walzthoeni T, Kahraman A *et al.* Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics* 9(8), 1634–1649 (2010).
- 52 Warren K, Wei T, Li D *et al.* Eukaryotic elongation factor 1 complex subunits are critical HIV-1 reverse transcription cofactors. *Proc. Natl Acad. Sci. USA* 109(24), 9587–9592 (2012).
- 53 Schweitzer CJ, Jagadish T, Haverland N, Ciborowski P, Belshan M. Proteomic analysis of early HIV-1 nucleoprotein complexes. *J. Proteome Res.* 12(2), 559–572 (2013).
- 54 Cantin R, Methot S, Tremblay MJ. Plunder and stowaways: incorporation of cellular proteins by enveloped viruses. *J. Virol.* 79(11), 6577–6587 (2005).
- 55 Chertova E, Chertov O, Coren LV *et al.* Proteomic and biochemical analysis of purified human immunodeficiency virus type 1 produced from infected monocyte-derived macrophages. *J. Virol.* 80(18), 9039–9052 (2006).
- 56 Zhang F, Zang T, Wilson SJ, Johnson MC, Bieniasz PD. Clathrin facilitates the morphogenesis of retrovirus particles. *PLoS Pathog.* 7(6), e1002119 (2011).
- 57 Popov S, Strack B, Sanchez-Merino V, Popova E, Rosin H, Gottlinger HG. Human immunodeficiency virus type 1 and related primate lentiviruses engage clathrin through Gag-Pol or Gag. *J. Virol.* 85(8), 3792–3801 (2011).
- 58 Radestock B, Morales I, Rahman SA *et al.* Comprehensive mutational analysis reveals p6Gag phosphorylation to be dispensable for HIV-1 morphogenesis and replication. *J. Virol.* 87(2), 724–734 (2013).
- 59 Mbonye UR, Gokulrangan G, Datt M *et al.* Phosphorylation of CDK9 at Ser175 enhances HIV transcription and is a marker of activated P-TEFb in CD4⁺ T lymphocytes. *PLoS Pathog.* 9(5), e1003338 (2013).
- 60 Dormeyer W, Dorr A, Ott M, Schnolzer M. Acetylation of the HIV-1 Tat protein: an *in vitro* study. *Anal. Bioanal. Chem.* 376(7), 994–1005 (2003).
- 61 Furia B, Deng L, Wu K *et al.* Enhancement of nuclear factor- κ B acetylation by coactivator p300 and HIV-1 Tat proteins. *J. Biol. Chem.* 277(7), 4973–4980 (2002).
- 62 Xu G, Paige JS, Jaffrey SR. Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nat. Biotechnol.* 28(8), 868–873 (2010).
- 63 Albin JS, Anderson JS, Johnson JR *et al.* Dispersed sites of HIV Vif-dependent polyubiquitination in the DNA deaminase APOBEC3F. *J. Mol. Biol.* 425(7), 1172–1182 (2013).
- 64 Singhroy DN, Mesplede T, Sabbah A, Quashie PK, Falgout JP, Wainberg MA. Automethylation of protein arginine methyltransferase 6 (PRMT6) regulates its stability and its anti-HIV-1 activity. *Retrovirology* 10, 73 (2013).
- 65 Zhang M, Gaschen B, Blay W *et al.* Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology* 14(12), 1229–1246 (2004).
- 66 Go EP, Hewawasam G, Liao HX *et al.* Characterization of glycosylation profiles of HIV-1 transmitted/founder envelopes by mass spectrometry. *J. Virol.* 85(16), 8270–8284 (2011).
- A glycopeptide-based mass mapping approach to categorize the glycan species on the native Envs from transmitted/founder viruses, an example of complex post-translational modification for a viral protein
- 67 DooRes. KJ, Bonomelli C, Harvey DJ *et al.* Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. *Proc. Natl Acad. Sci. USA* 107(31), 13800–13805 (2010).
- 68 Ringrose JH, Jeeninga RE, Berkhout B, Speijer D. Proteomic studies reveal coordinated changes in T-cell expression patterns upon infection with human immunodeficiency virus type 1. *J. Virol.* 82(9), 4320–4330 (2008).
- 69 Munday DC, Surtees R, Emmott E *et al.* Using SILAC and quantitative proteomics to investigate the interactions between viral and host proteomes. *Proteomics* 12(4–5), 666–672 (2012).
- This review summarizes quantitative proteomics and bioinformatics for the detection of cellular changes that occur during viral infection
- 70 Chan EY, Sutton JN, Jacobs JM, Bondarenko A, Smith RD, Katze MG. Dynamic host energetics and cytoskeletal proteomes in human immunodeficiency virus type 1-infected human primary CD4 cells: analysis by multiplexed label-free mass spectrometry. *J. Virol.* 83(18), 9283–9295 (2009).
- 71 Raposo RA, Trudgian DC, Thomas B, Van Wilgenburg B, Cowley SA, James W. Protein kinase C and NF- κ B-dependent CD4 downregulation in macrophages induced by T cell-derived soluble factors: consequences for HIV-1 infection. *J. Immunol.* 187(2), 748–759 (2011).
- 72 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.* 17(10), 994–999 (1999).
- 73 Ross PL, Huang YN, Marchese JN *et al.* Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* 3(12), 1154–1169 (2004).
- 74 Linde ME, Colquhoun DR, Ubaida Mohien C *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.* 12(5), 2045–2054 (2013).

- 75 Navare AT, Sova P, Purdy DE *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* 429(1), 37–46 (2012).
- 76 Karp NA, Huber W, Sadowski PG, Charles PD, Hester SV, Lilley KS. Addressing accuracy and precision issues in iTRAQ quantitation. *Mol. Cell. Proteomics* 9(9), 1885–1897 (2010).
- 77 Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl Acad. Sci. USA* 96(12), 6591–6596 (1999).
- 78 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* 1(5), 376–386 (2002).
- 79 Barrero CA, Datta PK, Sen S *et al.* HIV-1 Vpr modulates macrophage metabolic pathways: a SILAC-based quantitative analysis. *PLoS ONE* 8(7), e68376 (2013).
- 80 Wojcechowskyj JA, Lee JY, Seeholzer SH, Doms RW. Quantitative phosphoproteomics of CXCL12 (SDF-1) signaling. *PLoS ONE* 6(9), e24918 (2011).
- 81 Wojcechowskyj JA, Didigu CA, Lee JY *et al.* Quantitative phosphoproteomics reveals extensive cellular reprogramming during HIV-1 entry. *Cell Host Microbe* 13(5), 613–623 (2013).
- 82 Tackett AJ, Degrasse JA, Sekedat MD, Oeffinger M, Rout MP, Chait BT. I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. *J. Proteome Res.* 4(5), 1752–1756 (2005).
- Describes a metabolic labeling method to unambiguously assign authentic interactions and discriminate against those contaminants that bind nonspecifically to isolated protein complexes.
- 83 Nesvizhskii AI. Computational and informatics strategies for identification of specific protein interaction partners in affinity purification mass spectrometry experiments. *Proteomics* 12(10), 1639–1655 (2012).
- This review summarizes computational algorithms and informatics strategies for detecting specific protein interaction partners in AP/MS experiments.
- 84 Choi H, Larsen B, Lin ZY *et al.* SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nat. Methods* 8(1), 70–73 (2011).
- 85 CRAPome database. www.crapome.org
- 86 Mellacheruvu D, Wright Z, Couzens AL *et al.* The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods* 10(8), 730–736 (2013).
- 87 VirusMINT database. <http://mint.bio.uniroma2.it/virusmint/>
- 88 Chatr-Aryamontri A, Ceol A, Peluso D *et al.* VirusMINT: a viral protein interaction database. *Nucleic Acids Res.* 37(Database issue), D669–D673 (2009).
- 89 STRING database. www.bork.embl-heidelberg.de/STRING/
- 90 Von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res.* 31(1), 258–261 (2003).
- 91 Ingenuity pathway analysis. www.ingenuity.com/
- 92 Munday DC, Emmott E, Surtees R *et al.* Quantitative proteomic analysis of A549 cells infected with human respiratory syncytial virus. *Mol. Cell. Proteomics* 9(11), 2438–2459 (2010).
- 93 Bindea G, Mlecnik B, Hackl H *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25(8), 1091–1093 (2009).