Proteomics Tracing the Footsteps of Infectious Disease*

Todd M. Greco‡ and Ileana M. Cristea‡§

Every year, a major cause of human disease and death worldwide is infection with the various pathogens—viruses, bacteria, fungi, and protozoa—that are intrinsic to our ecosystem. In efforts to control the prevalence of infectious disease and develop improved therapies, the scientific community has focused on building a molecular picture of pathogen infection and spread. These studies have been aimed at defining the cellular mechanisms that allow pathogen entry into hosts cells, their replication and transmission, as well as the core mechanisms of host defense against pathogens. The past two decades have demonstrated the valuable implementation of proteomic methods in all these areas of infectious disease research. Here, we provide a perspective on the contributions of mass spectrometry and other proteomics approaches to understanding the molecular details of pathogen infection. Specifically, we highlight methods used for defining the composition of viral and bacterial pathogens and the dynamic interaction with their hosts in space and time. We discuss the promise of MS-based proteomics in supporting the development of diagnostics and therapies, and the growing need for multi-omics strategies for gaining a systems view of pathogen infection.

Molecular & Cellular Proteomics 16: 10.1074/mcp.O116.066001, S5–S14, 2017.

Infectious diseases, including HIV/AIDS, lower respiratory bacterial infections, and malaria, account for almost 25% of deaths worldwide (1). The causative agents of infectious diseases are microorganisms—viruses, bacteria, fungi, and protozoa—that exist in water, air, and soil. For many of these pathogenic agents, their evolution is linked to the organisms they infect, i.e. their hosts. Depending on the agent, the infection can produce a wide-range of symptoms. In humans, this can range from the seemingly innocuous runny nose, typical of rhinoviruses, to uncontrolled bleeding and multiple organ failure seen in emerging viruses such as Ebola. The ability of microorganisms to cause severe illness contrasts with their relatively simple genomes and proteomes when compared with their hosts. Indeed, microorganisms are intimate members of our ecosystem and have expertly evolved mechanisms to bypass and/or inhibit host immune surveillance programs.

In efforts to control the spread and prevalence of infectious disease, scientific inquiry has largely focused on three questions: (1) what host cell conditions facilitate permissive infections, (2) how the host immune system responds to the infection, and (3) what are the mechanisms of pathogen replication and transmission? The answers to these questions paint a picture of the pathogen’s life cycle and can aid in the development of precision therapies. Yet, the ability of many pathogens, such as influenza virus or Staphylococcus aureus bacteria, to rapidly adapt to environmental pressures and develop treatment-resistant strains and increased pathogenicity is a serious concern (2). Therefore, it is critical that our understanding of these pathogens and the development of more effective antiviral and antibacterial strategies stays ahead of their ability to adapt.

In the last century, the field of microbiology has benefited from continued technological advances in genetic, analytical, molecular, and imaging techniques. In particular, in the last two decades, mass spectrometry and proteomics-based technologies have been increasingly used to characterize the molecular details of pathogen-host relationships and provide insights into the biological basis of infectious diseases (3). Here, we provide a perspective on the contributions of mass spectrometry (MS)-based proteomics in defining the molecular structure and composition of viral and bacterial pathogens, as well as their interaction with the host in space and time (Fig. 1). We highlight several areas where MS-based proteomics is supporting the development of diagnostics and therapies and the emerging role of multi-omics strategies in providing a systems biology view of pathogen-host relationships.

Characterizing the Composition of Pathogens—The earliest applications of protein-based mass spectrometry to microbiological research began with the analysis of intact bacteria by particle beam mass spectrometry (4). However, widespread use of mass spectrometry for biological applications did not occur until the development of matrix-assisted laser desorption/ionization (MALDI)1 and electrospray ionization (ESI)

From the ‡Department of Molecular Biology, Princeton University, Princeton, New Jersey, 08544
Received November 30, 2016, and in revised form, January 25, 2017
Published, MCP Papers in Press, February 5, 2017, DOI 10.1074/mcp.O116.066001
Author contributions: T.M.G. and I.M.C. wrote the paper.

1 The abbreviations used are: MALDI, matrix-assisted laser desorption/ionization; ESI, electrospray ionization; TrV, Triatoma virus; IP-MS, immunoaffinity purification-mass spectrometry; DDA, data-dependent acquisition; Dia, data-independent acquisitions; hRSV, human respiratory syncytial virus; HCMV, human cytomegalovirus; PLRV, Potato leafroll virus.
techniques, which allowed the ionization of larger biomolecules (5–7). Using these soft ionization techniques, the direct analysis of bacteria generated highly complex mass spectra, but which nonetheless provided a mass "fingerprint" that could discriminate different species and strains, such as *Escherichia coli*, *Shingella flexneri* (8, 9). In efforts to reduce complexity, Liu and colleagues developed an online microdialysis path coupled to electrospray ionization-mass spectrometry (10). This design allowed selective isolation and mass measurement of lower molecular weight proteins directly from crude microbial cell lysates.

These studies provided an initial glimpse into bacterial proteomes, but also highlighted the analytical challenges associated with intact protein measurements. To address some of these challenges, complementary strategies were implemented that digested microbial proteins into peptides, commonly with the protease trypsin, followed by reverse-phase liquid chromatography to reduce complexity prior to mass spectrometry analysis. Despite sacrificing proteoform information, these approaches, collectively referred to as “bottom-up” proteomics (11), currently predominate because of advantages in overall sequence coverage and depth of analysis. These approaches proved effective at characterizing the composition of viral and bacterial pathogens. For instance, the combination of “bottom-up” proteomics and ESI-tandem mass spectrometry has helped define the composition of infectious viral particles for DNA viruses (e.g. herpes simplex virus 1 and human cytomegalovirus) and RNA viruses (e.g. influenza virus and filoviruses) (12–17). Proteomics has also been used to discover differences between related viral strains and determine virion components (see Fig. 2A, left) necessary for viral replication and spread. Using metabolic labeling and mass spectrometry, comparison of virion compositions from several herpesvirus mutant strains revealed that deletion of just a single protein, such as the viral glycoprotein E, can trigger changes in the incorporation levels of several viral structural proteins (pUL46, pUL48, pUL49) and host proteins (18). As many of mutant virus strains have known defects in virus replication, this study highlights the ability of proteomic analysis to link proteome composition to phenotypes. Overall, by understanding the molecular composition of microorganisms, including host protein incorporation into viruses, proteomics will inform on the molecular basis of pathogenicity and virulence.

Similarly, for bacterial pathogens, methodological and technological advancements in proteomics have made it possible to achieve near-complete proteome depth in less than a day (19), due in part to their reduced complexity compared with animal proteomes. For example, *Mycobacterium tuberculosis* lacks membrane-bound organelles, and its genomic DNA encodes for only about 4000 proteins, though the bacterium can take up foreign plasmid DNA (Fig. 2A, right), increasing its protein-coding abilities and cellular functions. Given these unique aspects of bacteria, proteomics has been an invaluable technique for studying bacterial strains cultured under various environment conditions, such as nutrient deprivation, hypoxia, pH, and drug treatment (20–22). In addition, bacteria possess several protein secretion systems that are responsible for exporting factors that determine the pathogen’s virulence (23). Therefore, bacterial secretory pathways represent attractive targets for MS-based proteomic analysis (24). Recent advances in bacterial model systems have further extended bacterial proteome analyses to the *in vivo* state. For example, using a guinea pig model of tuberculosis, the *Mycobacterium tuberculosis* proteome was characterized in lungs from early and chronic disease stages (25). Despite the overwhelming abundance and heterogeneity of the host tissue proteome, over 500 bacterial proteins were identified, including annotated functions in cell wall and intermediary metabolism. Undoubtedly, with the continued development of
bacterial and viral infection model systems and improvements in overall MS sensitivity, proteomics will reveal additional aspects of the pathogen proteome and interactions with hosts.

Although the success of bottom-up proteomics is undeniable, the increasing resolution of FT-ICR and Orbitrap-based instruments has taken intact mass analysis to the next level. As another way to define pathogen composition, native MS, paired with ion mobility separation, is emerging as a complementary tool for structural microbiologists to examine macromolecular viral protein assemblies (26) and intact microbial particles (27). Most viruses protect their genetic material in a proteinaceous capsid, with masses exceeding one megadalton (Fig. 2A). Biophysical characterization of these intact viral protein assemblies, i.e. their composition, structure, and stability, provides insight into how mature viral particles are constructed for release and spread of the infection. Notably, the role of viral genomes in this process is an emerging topic that native mass spectrometry is well positioned to investigate. In a significant technological achievement, Snijder et al., analyzed viral capsids from the picorna-like Triatoma virus (TrV) using ion mobility separation-native mass spectrometry to determine the masses of viral assemblies, and atomic force microscopy to visualize the topology of individual particles (28). TrV is related to the human pathogen, poliovirus, but it infects the insect Triatoma infestans, which is itself a vector for Chagas disease, a neglected tropical disease in humans. In this context, the adage “the enemy of my enemy is my friend” is quite appropriate, as TrV is being studied for its potential as a biopesticide (29). Overall, the importance of understanding viral capsid structure and function cannot be understated. Proper capsid assembly containing a viral genome is essential for the generation of an infectious particle, and understanding the key players in this process has the potential to guide targeted anti-viral therapies. As the structural features of more viruses are cataloged using proteomic...
approaches, it may be possible to infer cell and host tropisms of poorly characterized viruses based on conserved structural features from well-studied viruses.

**Intracellular Host-Pathogen Interactions**—During the last decade, proteomic techniques have made significant contributions to understanding host-pathogen interactions during the cellular life cycle of a pathogen. Specifically, the hyphenation of traditional analytical and biochemical techniques with mass spectrometry has generated proteomic approaches that probe different facets of the host-pathogen relationship. In particular, coupling of immunoaffinity purification techniques to mass spectrometry (IP-MS) has accelerated the identification of critical proteins involved in host interactions with pathogen proteins and their genetic material and in propagation of immune defense signaling.

Following entry of the pathogen into the host cell (Fig. 2B, circles 1A and 1B), a surge of anti and propathogenic cellular responses occurs, with each arm relying on stimulus-induced changes in signaling. These signals cause global proteome abundance changes and also form or dissolve protein-protein interactions. As a consequence, host proteins can often assume additional functions beyond their basic cellular housekeeping roles. IP-MS has emerged as a powerful technique to discover previously undocumented protein functions and activities in the context of infection (30). The implementation of IP-MS in virus-host protein interaction studies started about ten years ago, demonstrating the value of mass spectrometry for understanding the biology of RNA and DNA viruses, i.e. Sindbis and HSV-1 viruses (31, 32). Since then, the biological insights gained from IP-MS approaches have included mechanisms of viral replication and maturation (Fig. 2B, circles 2–4A) (33–35) and host cellular defense (36), and have driven the expansion of the scopes of such studies to capturing larger protein interactomes during viral infection (37, 38). Overall, these studies have improved the molecular picture of the host response to infection, both from the host and pathogen perspective.

In general, pathogenic protein interactions are ideally determined during the course of natural infections (31, 32, 34, 39). When pathogenic model systems, such as recombinant viral strains with genome-encoded affinity tagged viral proteins, are not available for proteomic studies, the overexpression of individual pathogenic proteins can also provide high quality protein interactions. Although overall interactome coverage may be lower in the absence of other pathogen proteins, this approach has proven valuable for studying viruses difficult to tag, as was shown for the HIV-1 interactome (40) and the Ebola virus matrix protein, VP40 (41). Toward the goal of generating infection model systems for proteomic analysis of virus-host interactions, recent work from Luo and co-workers developed a strategy for generating epitope-tagged HIV-1 strains (39). Transposon-mediated saturation linker scanning mutagenesis was used to find sites within the viral gene coding region that were permissive to insertion of an epitope tag, i.e. sites that would not inhibit viral replication. Infection of human CEM-T with heterogeneous HIV-1 virus stocks (made from at least 100,000 clones) led to the survival of only the replication-competent mutants, which were then sequenced to determine the specific gene loci with the mutational marker. Indeed, these positions were shown to allow for replication competent, 3XFLAG-tagged HIV-1 strains, which were then used to profile the viral-host interactions of selected HIV-1 proteins in infected cells. As there will still be cases when virus strains or high-affinity antibodies are not available, homology-based modeling of protein interactions can be informative, generating resources of predicted interactions, as was performed for viral-viral interactions in HSV-1 (42). Overall, IP-MS studies, in conjunction with biological follow-up analyses, have demonstrated the fundamental contribution of proteomics to understanding pathogen infections. However, knowledge of viral-host interactions still remains sparse for numerous viruses, including those well-established as major human pathogens and emerging viruses. Additional studies are needed to define virus-host interactions at different stages of infection, which can help identify new targets for antiviral therapies.

Similar to viruses, bacterial infections depend on host cell protein machineries for entry and vesicular trafficking (Fig. 2B, circle 1B). Bacteria can also modulate host cell immune defense programs, which can result in opposing outcomes of pathogen degradation (Fig. 2B, circle 2B) or immune evasion followed by bacterial replication and maturation (Fig. 2B, circle 3B). The full complement of molecular players underlying these events remains to be understood, with the majority of studies relying on functional knockdowns to identify critical proteins. Although IP-MS has been used to study protein complexes within bacteria (43), it has not been widely used to study in vivo pathogen-host cell protein interactions and their dynamic regulation during infection. One exception to this is recent work from Mirrashidi et al., which examined host-pathogen interactions for the secreted inclusion membrane of *Chlamydia trachomatis* during infection of human cells (44). Although bacterial systems in general lend themselves to genetic manipulation, outside of well-established bacterial models, such as *Escherichia coli* and *Bacillus subtilis*, many bacterial species have limited genetic tools and are difficult to culture.

Complementing IP-MS approaches, proteomic approaches based on protein microarrays show exciting possibilities for high-throughput screening of pathogen interactions. In particular, the Nucleic Acid Programmable Protein Array (NAPPA) technology has circumvented the common pitfalls affecting microarrays imprinted with purified proteins, by instead spotting protein-encoding plasmid DNA, which is then translated using cell-free expression systems just prior to application of the sample (45). This technology has been used to determine the direct interactions of *Legionella pneumophila* proteins (SidM and LidA) with 10,000 human proteins imprinted by
In addition to defining protein–protein interactions, proteomic methods have provided insights into associations between host proteins and pathogenic nucleic acids. At the onset of infection, the host cell relies on intracellular immune surveillance mechanisms, often through recognition of foreign nucleic acids by sensor proteins (48). Alternatively, binding of host proteins to foreign genomes may be pathogenetic, recruited to their genomes to facilitate pathogen replication (49). Given these possibilities, unbiased identification of host-genome interactions using nucleic acid affinity chromatography paired with functional knockdown studies have proven useful to rapidly screen for candidate host factors, particularly during replication of RNA viruses, such as norovirus (50), poliovirus (51), dengue virus (52, 53), and hepatitis C virus (54). One of the main challenges in identifying RNA-protein interactions during viral infection is achieving adequate efficiency of viral RNA (vRNA) isolation, while maintaining in vivo protein interactions and minimizing nonspecific contaminants. To address these issues, infected cells are treated with long-wave UV light to cross-link proteins to vRNA in which a nucleoside analog, 4-thiouridine, has been incorporated (51). This uridine base analog permits cross-linking of proteins only in close apposition to the vRNA and for the isolation to proceed under denaturing conditions using oligoT-based capture. For RNA viruses that require host cell machinery for replication, affinity purification of vRNA can be performed with a biotinylated antisense oligonucleotide (PNA) moiety can also be used, which retains the complementary DNA bases that recognize the vRNA, but replaces the sugar-phosphate backbone with peptide bonds, providing high affinity and specificity for viral RNA capture (54). Of note, similar unbiased proteomic studies have not been reported for interaction of host protein with the genomes of DNA viruses. Yet, given the increasing use of MS-based proteomics to unravel the complex dynamics of chromatin states (55, 56), future efforts will undoubtedly reveal additional host and viral factors that are essential for the transcription of pathogen mRNAs and for replication of DNA genomes. In summary, as additional virus strains, model systems, and proteomic techniques are developed, the study of host-pathogen protein–protein and protein–nucleic acid interactions in their natural state, such as in tissues obtained from infected hosts, will likely become more routine, and thus help to assemble a systems biology perspective of infectious disease states.

Host Cell Proteomes under Infection—Pathogen infection triggers diverse alterations in host cellular signaling, which reflects the biochemical and physical demands of the pathogen’s life cycle. Understanding these signature infection-driven changes in host proteomes can provide critical insights into pathways required for the replication of the pathogen or for host defense. Recent advances in mass spectrometry technologies have produced instruments with increased sequencing speed, resolution, and mass accuracy. Combined with intelligent data acquisition strategies, such as data-dependent (DDA) and data-independent acquisitions (DIA), the depth and coverage of the proteome has grown remarkably. In the past decade, proteomic platforms have increased the proteome coverage from about 1000 protein groups to almost 10,000 in a single cell type (57). This has allowed the identification of lower abundance proteins, such as transcription factors and organelle-specific proteins. In the context of infection, using MS-based approaches to maximize the proteome coverage of infected cells or tissues provides an increasingly comprehensive picture of differentially regulated proteins. Moreover, global proteomic approaches that employ DDA inherently lack a pathway bias, because specific cellular pathways are not targeted. In a notable example, this approach allowed the identification of a host pathway, Sec61-mediated cotranslational translocation, that when disrupted, was broadly inhibitory to influenza, HIV, and dengue virus growth and infectivity (58). Another example of expertly leveraging global whole cell proteomics and in-solution protein isoelectric focusing was a study from Dave and colleagues, in which the epithelial host cell response to human respiratory syncytial virus (hRSV) infection was investigated (59). The depth of analysis facilitated the quantification of proteins in the tumor necrosis factor-alpha and nuclear factor kappa B (NF-κB)-mediated cellular signaling pathways. The use of insoluble protein-level fractionation by isoelectric point prior to bottom-up proteomic analysis, which was further expanded in a follow-up study from the same group (60), allowed separation of proteoforms. Therefore, this strategy identified protein-level regulation stimulated by hRSV that would be difficult to detect by “bottom-up” approaches only. For example, several isoforms and truncations of signal transducer and activator of transcription 1 (STAT1) were found only in infected cells. Top-down proteomic approaches also have the potential to contribute to understanding proteolytic events and changes in post-translational modification (PTM) status induced during infection.

It is known that the regulation of host and pathogen PTM signaling is critical during infection, initiating changes in protein interactions, activity, and subcellular localization, and ultimately impacting virus replication (61). This is particularly relevant for phosphorylation as many pathogens encode their own kinases and effectively manipulate host kinase signaling pathways. Mass spectrometry has been instrumental in identifying specific sites of modification, both on host and viral
Proteomics in Infectious Diseases

proteins (36, 62–66). However, to date only few studies have employed global PTM profiling (phosphorylation, acetylation, glycosylation, among others) during viral infection (67, 68). The continued application and expansion of these techniques to simultaneously identify multiple modifications will likely uncover novel pathogen-specific PTM signaling cascades that could serve as the basis for targeted therapeutics.

Access to lower abundance proteins and specialized pathways manipulated by pathogens has also been accomplished using subcellular fractionation prior to proteomic analysis. A type of host proteome change that has been of particular interest is alteration of the cellular periphery, i.e. plasma membrane. The identification of host cell surface receptors that bind and facilitate pathogen entry into cells (see Fig. 2B, circles 1A and 1B), as well as other plasma membrane-resident proteins that contribute to viral replication has provided attractive therapeutic targets. In fact, about two-thirds of existing drugs target cell surface proteins (69). Pathogen-specific cell surface targets have been identified directly from plasma membrane enriched fractions (70), or following affinity isolation, such as the isolation of pathogen-specific cell surface proteins (71), or cell surface labeling and capture of glycosylated proteins (72, 73) or peptides (74). Affinity capture techniques at the peptide level often have advantages in specificity and sensitivity for detecting low abundance receptors, as well as the ability to detect site specific changes in N-glycosylation abundance. The application of these techniques to a greater number of infection model systems will expand the repertoire of known host cell pathogen receptors that facilitate pathogen entry and of host cell surface responses to infection that could participate in intercellular immune signaling mechanisms.

The challenge of studying biological systems, particularly in the realm of pathogen infection, is the critical dependence on the temporality of cellular signaling responses. A prime example is studying the progression of herpesvirus infections in cellular models, which from initial cell entry to release of progeny virus, can take days to weeks. One such herpesvirus is human cytomegalovirus (HCMV), which has a lifecycle of 4–5 days, during which it has a temporally coordinated cascade of viral transcription and translation (75, 76) (Fig. 2B, circle 3A). HCMV has a relatively large proteome with the capacity to express nearly 200 proteins with diverse functionality during either active or latent infections. Like many DNA viruses, HCMV also possesses several proteins with unique abilities to evade host immune defenses (77). Therefore, HCMV has many molecular strategies to influence the host cell proteome on several regulatory levels at different stages of infection. To gain a richer understanding of the HCMV lifecycle, Weekes and co-workers conducted temporal profiling of both viral and fibroblast host proteome abundances at critical time points throughout infection (78). The ability of this study to achieve temporal resolution of the infection, maintain a feasible experimental design, while still obtaining proteome depth was accomplished using quantitative multiplexed MS-based proteomics combined with MS instrumentation employing synchronous precursor selection (79). In multiplexed proteomic approaches, a key innovation was the development of isobaric, isotope-coded tagging reagents (80, 81), which currently can encode for up to 10 experimental parameters (e.g. time, space, drug treatment) into the mass (m/z) domain of tandem mass spectra that can then be decoded offline into protein abundances. In the study of proteome dynamics during HCMV infection, Weekes et al. achieved extensive host and viral proteome coverage, revealing multiple dysregulated pathways, including down- and up-regulation of known IFN signaling and fatty acid metabolic pathways, respectively, as well as a previously unknown down-regulation of the wnt signaling pathway (78). In addition, the analysis of plasma membrane-enriched fractions revealed modulation of several natural killer cell ligands during infection, which may ultimately prove to be targets of HCMV anti-immune responses.

Proteome Organization in Space and Time during Infection—The studies above highlight the strength of proteome profiling and quantification of protein abundances in pathogen infections. Yet, cellular responses have additional layers of complexity; independent of protein abundance, they can be fine-tuned by changing protein localization/interaction, activity, and posttranslational modification status (82). These themes are even more salient when put in the context of pathogen infection. Indeed, understanding proteome organization and host cell responses during infection is a research area of growing interest. Several MS-based approaches can be used to gain spatial and temporal information during infection. For example, one early implementation of MS involved the use of fluorescently-labeled viruses as both visualization and affinity tools to define virus-host protein interactions in space and time during Sindbis virus infection (32). Building a more comprehensive picture of intracellular protein dynamics during infection will also require understanding the precise connectivity among proteins in their native cellular state. Toward this goal, the continued development of in vivo chemical labeling techniques paired with mass spectrometry will increase the specificity of identifying interactions in infected cells and provide the direct points of host-pathogen contact. For instance, an in vivo chemical cross-linking strategy paired with protein interaction reporter technology and mass spectrometry analysis has been used to identify protein interaction interfaces in cultured cells (83), viruses (84), and bacteria (85). The first application of this technology to viral infection characterized the protein interaction topologies of the plant pathogen Potato leafroll virus (PLRV) (84). Although this technology has not yet been applied in human infections, the study found several instances of interaction topologies that could reflect preserved functions in unrelated human viruses. For example, a PLRV-host interaction was identified that was analogous to the known structural bind-
ing of the animal zinc binding protein BS69 to a conserved PXLXP motif in viral proteins from the Herpesviridae and Adenoviridae families (84). Overall, this technology has enormous potential in pathogen proteomics, with the ability to identify in a single experiment both the interacting protein pairs and the higher order structures that mediate the interaction.

Another important aspect of proteome organization during infection is the virus-induced alteration in subcellular organelles. Here again, HCMV infection provides a quintessential illustration, as the progression of infection depends on the reorganization of numerous organelles, such as mitochondria for energy production and Golgi and the ER for formation of the virion assembly complex (Fig. 2B, circle 4A). Demonstrating the value of MS approaches in the context of infection, Jean Beltran and colleagues used both label-free and multiplexed labeling quantitative MS to probe organelle composition during the progression of HCMV-infection (86). Specifically, the study used density gradient fractionation, label-free and the TMT-based Localization of Organelle Proteins by Isotope Tagging technique (87), and supervised machine learning to assign localization of viral and host proteins to organelles and discover host protein redistribution and viral protein trafficking in space and time. Reorganization of several secretory pathways was observed, including the division of lysosomal proteins into two functionally- and localization-distinct populations. Many host factors were also found translocated from one organelle to another in an infection-dependent manner. One interesting example was the unconventional myosin, MYO18A, which relocated from the plasma membrane to the viral assembly complex and was implicated in virus release and spread of infection. Complementary to organellar temporal-spatial profiling, proximity-based biotinylation can also be advantageous for labeling proteins in specific subcellular compartments directly in cultured cells (88) or in tissues (89). In this approach, organelle-specific local protein biotinylation is induced through expression of an engineered peroxidase, followed by streptavidin-affinity isolation and mass spectrometry analysis. Although not yet employed for studying pathogen infection, this technology could enable in depth characterization of specific organelles during an infection time course, while eliminating the need to perform organellar enrichment and fractionation.

Clinical and Therapeutic Applications—Proteomic approaches have strongly supported basic research into the molecular mechanisms of infectious disease, which ultimately can aid in the identification of new therapeutic targets. However, MS-based proteomics also provides means for analyzing clinical samples, and for profiling the immune system antigenic responses to help vaccine development. Improvement in the speed and accuracy of pathogen infection diagnoses will help curb the rise in multidrug resistance bacteria and serve to assess the severity of outbreaks because of emerging viruses. In the last decade, several MS-based approaches have been under development to support these efforts. Building upon the early work of microbe mass fingerprinting mentioned above, MALDI-Tandem MS systems with high resolution and accurate mass can identify species-specific candidate biomarker proteins, even when limited genome or protein sequence information is available (90). Another technique, PCR-electrospray ionization mass spectrometry (PCR-ESI/MS), is being developed for improved diagnosis of infections as a complementary approach to routine cell culture and microscopic-based assays. PCR-ESI/MS has broad-range detection capability directly from patient samples, and allows increased throughput through multiplexing of PCR primers designed against multiple pathogens (91). However, several drawbacks still exist, including the possibility of sample-to-sample contamination and high initial cost investment. From a therapeutic perspective, proper coordination of innate and adaptive immune systems and orchestration of antigen peptide presentation is critical for mounting a successful host defense against infection or priming the immune system following vaccination. Given that antigenic targets are peptides, when combined with other omics techniques, mass spectrometry can provide a systems level perspective of the body’s immune response (92).

It is clear that whether the application of proteomics is for understanding protein dynamics in cellular model systems, or for assessing clinical outcomes in human individuals, in the future, proteomics will increasingly be paired with other omics techniques (genomics, transcriptomics, metabolomics, lipidomics, etc) to assemble a systems biology perspective of host-pathogen responses (93–95). The integration of genomics and proteomics with disease phenotypes has the potential to establish disease or individual-specific proteotypes, which may aid in making therapies more specific to the disease or individual (82). Overall, as the multiplicity of omics studies increases, the challenge will not be in which system to apply these methods, but how to properly integrate, visualize, and interpret the complex outputs. Future work that establishes flexible yet comprehensive platforms for integrating large-scale data sets will be a vital component in establishing the proteotype signatures of host-pathogen relationships.

* This work was supported by the NIH, NIGMS grant GM111411 and NHBLI grant HL127640, and the Edward Mallinckrodt, Jr. Foundation Scholar Award to IMC. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

§ To whom correspondence should be addressed: 210 Lewis Thomas Laboratory Department of Molecular Biology Princeton University Princeton, NJ 08544. Tel.: +1 609 2589417; Fax: +1 609 2584575; E-mail: icristea@princeton.edu.

REFERENCES

1. WHO. (2015) World Health Statistics 2015
2. Frieri, M., Kumar, K., and Boutin, A. (2016) Antibiotic resistance. J. Infect. Public Health. http://dx.doi.org/10.1016/j.jiph.2016.08.007
3. Greco, T. M., Diner, B. A., and Cristea, I. M. (2014) The impact of mass spectrometry-based proteomics on fundamental discoveries in virology. Annu. Rev. Virol. 1, 581–604
44. Mirrashidi, K. M., Elwell, C. A., Verschueren, E., Johnson, J. R., Frando, A., Yu, X., Decker, K. B., Barker, K., Neunuebel, M. R., Saul, J., Graves, M., Kuhner, S., van Noort, V., Betts, M. J., Leo-Macias, A., Batisse, C., Rode, Molecular & Cellular Proteomics 16 Supplement 4
49. Cristea, I. M., Moorman, N. J., Terhune, S. S., Cuevas, C. D., O’Keefe, 50. Vashist, S., Urena, L., Chaudhry, Y., and Goodfellow, I. (2012) Identification 51. Han, Y., and Garcia, B. A. (2013) Combining genomic and proteomic 52. Phillips, S. L., Soderblom, E. J., Bradrick, S. S., and Garcia-Blanco, M. A. (2016) Identification of proteins bound to dsDNA and RNA in vivo reveals 53. Viktorovskaya, O. V., Greco, T. M., Cristea, I. M., and Thompson, S. R. (2013) 54. Upadhyay, A., Dixit, U., Manvar, D., Chaturvedi, N., and Pandey, V. N. (2013) Affinity capture and identification of host cell factors associated with hepatitis C virus (+) strand subgenomic RNA. Mol. Cell. Proteomics 12, 1539–1552 55. Han, Y., and Garcia, B. A. (2013) Combining genomic and proteomic approaches for epigenetics research. Epigenetics 8, 439–452 56. de Jesus, T. C., Calderano, S. G., Vitorino, F. N., Llanos, R. P., Lopes, M. C., de Araujo, C. B., Thiemann, O. H., Reis, M. D., Elias, M. C., and da Cunha, J. P. (2016) Quantitative proteomic analysis of replicative and non-replicative forms reveals important insights into chromatin biology of Trypanosoma cruzi. Mol. Cell. Proteomics 15, 15–27 57. Richards, A. L., Merrill, A. E., and Coon, J. J. (2015) Proteome sequencing 58. Heaton, N. S., Moshkina, N., Fenouil, R., Gardner, T. J., O’Shea, S., Kelleher, D., 59. Dave, K. A., Norris, E. L., Bukreyev, A. A., Headlam, M. J., Buchholz, U. J., Singh, T., Collins, P. L., and Gorman, J. J. (2014) A comprehensive proteomic view of responses of A549 type II alveolar epithelial cells to human respiratory syncytial virus infection. Mol. Cell. Proteomics 13, 3250–3269 60. Norris, E. L., Headlam, M. J., Dave, K. A., Smith, D. D., Bukreyev, A., Singh, T., Jayakody, B. A., Chappell, K. J., Collins, P. L., and Gorman, J. J. (2016) Proteome-specific insights into cellular proteome regulation. Mol. Cell. Proteomics 15, 3297–3320 61. Ribet, D., and Cossart, P. (2016) Pathogen-mediated posttranslational modifications: a re-emerging field, Cell 164, 694–702 62. Hutchinson, E. C., Denham, E. M., Thomas, B., Trudgian, D. C., Hester, S. S., Ridlova, G., York, A., Turrel, L., and Fodor, E. (2012) Mapping the phosphoproteome of influenza A and B viruses by mass spectrometry. PLoS Pathogens 8, e1002993 63. Davido, D. J., Leib, D. A., and Schaffer, P. A. (2002) The cyclin-dependent kinase inhibitor roscuvastatin inhibits the transactivating activity and alters the posttranslational modification of herpes simplex virus type 1 ICP0. J. Virol. 76, 1077–1088 64. Kulej, K., Avgousti, D. C., Weltzman, M. D., and Garcia, B. A. (2015) Characterization of histone post-translational modifications during viral infection using mass spectrometry-based proteomics. Methods 90, 8–20 65. Lin, A. E., Greco, T. M., Dohner, K., Sodeik, B., and Cristea, I. M. (2013) A proteomic perspective of innate viral protein regulation: pUL46 tegument protein is targeted for degradation by ICP0 during herpes simplex virus type 1 infection. Mol. Cell. Proteomics 12, 3237–3252 66. Rowles, D. L., Tsai, Y. C., Greco, T. M., Lin, A. E., Li, M., Yeh, J., and Cristea, I. M. (2015) DNA methyltransferase DNMT3A associates with viral proteins and impacts HSV-1. Proteomics 15, 1968–1982 67. Bell, C., Desjardins, M., Thibault, P., and Radtke, K. (2013) Proteomics analysis of Herpes Simplex virus type 1-infected cells reveals dynamic changes of viral DNA sensors. J. Biol. Chem. 290, 26412–26421 68. Bagdonate, I., Norden, R., Joshi, H. J., King, S. L., Vakhruchev, S. Y., Olofsson, S., and Wandall, H. H. (2016) Global mapping of O-glycosylation of varicella zoster virus, human cytomegalovirus, and Epstein–Barr virus. J. Biol. Chem. 291, 12014–12028 69. Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) How many drug targets are there? Nat. Rev. Drug Discov. 5, 993–996 70. Lund, B., Leth-Larsen, R., Jensen, O. N., and Ditzel, H. J. (2009) Efficient isolation and quantitative proteomic analysis of cancer cell plasma membrane proteins for identification of metastasis-associated cell surface markers. J. Proteome Res. 8, 3078–3090 71. Larsson, T., Bergstrom, J., Nilsson, C., and Karlsson, K. A. (2000) Use of an affinity proteomics approach for the identification of low-abundant bacterial adhesins as applied on the Lewis(a/b) -binding adhesin of Helicobacter pylori. FEBS Lett. 464, 15–19 72. Gudelski-O’Regan, N., Greco, T. M., Cristea, I. M., and Shenk, T. (2012) Increased expression of LDL receptor-related protein 1 during cytomegalovirus infection reduces virion cholesterol and infectivity. Cell Host Microbe 12, 86–96 73. Shimogawa, M. M., Saeda, E. A., Vaishast, A. A., Barshop, W. D., Wohlschlegel, J. A., and Hill, K. L. (2015) Cell surface proteomics provides insight into stage-specific remodeling of the host-parasite interface in Trypanosoma brucei. Mol. Cell. Proteomics 14, 1977–1988 74. Wollscheid, B., Bauchsl-Fluck, D., Henderson, C., O’Brien, R., Bibel, M., Schiess, R., Aebersold, R., and Watts, J. D. (2009) Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. Nat. Biotechnol. 27, 378–386 75. Jean Beltran, P. M., and Cristea, I. M. (2014) The life cycle and pathogenesis of human cytomegalovirus infection: lessons from proteomics. Expert Rev. Proteomics 11, 697–711 76. Vincent, H. A., Ziehr, B., and Moorman, N. J. (2016) Human cytomegalovirus strategies to maintain and promote mRNA translation. Viruses 8, 97 77. Crow, M. S., Lunn, K. K., Cheng, X., Song, B., and Cristea, I. M. (2016) Diverse mechanisms evolved by DNA viruses to inhibit early host defenses. Crit. Rev. Biochem. Mol. Biol. 51, 452–481 78. Weeke, M. P., Tomas, G., Polunin, D., Poon, E., Zako, J., and Gyorke, S. (2015) Quantitative temporal viromics: an approach to investigate host-pathogen interaction. Cell 157, 1460–1472 79. McAlister, G. C., Nusinow, D. P., Jedrychowski, M. P., and Gorman, J. J. (2014) Multi-Notch MS3 enables accurate, sensitive, and multiplexed detection of
differential expression across cancer cell line proteomes. Anal. Chem. 86, 7150–7158
80. Thompson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A. K., and Hamon, C. (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal. Chem. 75, 1895–1904
81. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol. Cell. Proteomics 3, 1154–1169
82. Aebersold, R., and Mann, M. (2016) Mass-spectrometric exploration of proteome structure and function. Nature 537, 347–355
83. Weisbrod, C. R., Chavez, J. D., Eng, J. K., Yang, L., Zheng, C., and Bruce, J. E. (2013) In vivo protein interaction network identified with a novel real-time cross-linked peptide identification strategy. J. Proteome Res. 12, 1569–1579
84. DeBlasio, S. L., Chavez, J. D., Alexander, M. M., Ramsey, J., Eng, J. K., Mahoney, J., Gray, S. M., Bruce, J. E., and Cilia, M. (2015) Visualization of host-poliovirus interaction topologies using protein interaction reporter technology. J. Virol. 90, 1973–1987
85. Wu, X., AP, Chavez, J. D., Schweppe, D. K., Zheng, C., Weisbrod, C. R., Eng, J. K., Murail, A., Lee, S. A., Ramage, E., Gallagher, L. A., Kulasekara, H. D., Edrozo, M. E., Kamischke, C. N., Brittnacher, M. J., Miller, S. I., Singh, P. K., Manoil, C., and Bruce, J. E. (2016) In vivo protein interaction network analysis reveals porin-localized antibiotic inactivation in Acinetobacter baumannii strain AB5075. Nat. Commun. 7, 13414
86. Jean Beltran, P. M., Mathias, R. A., and Cristea, I. M. (2016) A portrait of the human organelle proteome in space and time during cytomegalovirus infection. Cell Syst. 3, 361–373
87. Dunkley, T. P., Watson, R., Griffin, J. L., Dupree, P., and Lilley, K. S. (2004) Localization of organelle proteins by isotope tagging (LOPIT). Mol. Cell. Proteomics 3, 1128–1134
88. Rhee, H. W., Zou, P., Udeshi, N. D., Martell, J. D., Mootha, V. K., Carr, S. A., and Ting, A. Y. (2013) Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science 339, 1328–1331
89. Chen, C. L., Hu, Y., Udeshi, N. D., Lau, T. Y., Wirtz-Peitz, F., He, L., Ting, A. Y., Carr, S. A., and Perrimon, N. (2015) Proteomic mapping in live Drosophila tissues using an engineered ascorbate peroxidase. Proc. Natl. Acad. Sci. U. S. A. 112, 12093–12098
90. Wynne, C., Fenselau, C., Demirev, P. A., and Edwards, N. (2009) Top-down identification of protein biomarkers in bacteria with unsequenced genomes. Anal. Chem. 81, 9633–9642
91. Wolk, D. M., Kaleta, E. J., and Wysocki, V. H. (2012) PCR-electrospray ionization mass spectrometry: the potential to change infectious disease diagnostics in clinical and public health laboratories. J. Mol. Diagnostics 14, 296–304
92. Purcell, A. W., and Gorman, J. J. (2004) Immunoproteomics: Mass spectrometry-based methods to study the targets of the immune response. Mol. Cell. Proteomics 3, 193–208
93. Villar, M., Ayllon, N., Alberdi, P., Moreno, A., Moreno, M., Tober, R., Mateos-Hernandez, L., Weisheit, S., Bell-Sakyi, L., and de la Fuente, J. (2015) Integrated Metabolomics, Transcriptomics and Proteomics Identifies Metabolic Pathways Affected by Anaplasma phagocytophilum Infection in Tick Cells. Mol. Cell. Proteomics 14, 3154–3172
94. Schubert, O. T., and Aebersold, R. (2015) Microbial proteome profiling and systems biology: Applications to Mycobacterium. Adv. Exp. Med. Biol. 883, 235–254
95. Kim, M., Rai, N., Zorraquino, V., and Tagkopoulos, I. (2016) Multi-omics integration accurately predicts cellular state in unexplored conditions for Escherichia coli. Nat. Commun. 7, 13090