Published in final edited form as:

Trends Parasitol. 2018 November ; 34(11): 945–960. doi:10.1016/j.pt.2018.08.003.

***Plasmodium* Parasites Viewed Through Proteomics**

Kristian E. Swearingen^{1,2} and Scott E. Lindner^{3,*}

¹. Institute for Systems Biology, Seattle, WA 98109

². Center for Infectious Disease Research, Seattle, WA 98109

³. Department of Biochemistry and Molecular Biology, Center for Malaria Research, Pennsylvania State University, University Park, PA 16802

Abstract

Early sequencing efforts that produced the genomes of several species of malaria parasites (*Plasmodium* genus) propelled transcriptomic and proteomic efforts. In this review, we focus upon some of the exciting proteomic advances from studies of *Plasmodium* parasites over approximately the last decade. With improvements to both instrumentation and data processing capabilities, long standing questions about the forms and functions of these important pathogens are rapidly being answered. In particular, global and subcellular proteomics, quantitative proteomics, and the detection of post-translational modifications have all revealed important features of the parasite's regulatory mechanisms. Finally, we provide our perspectives on future applications of proteomics to *Plasmodium* research, as well as suggestions for further improvement through standardization of data deposition, analysis, and accessibility.

Keywords

Plasmodium; malaria parasites; proteomics; mass spectrometry; post-translational modifications; quantitative proteomics

A Wave of Proteomic Discovery

The last decade has seen a surge in proteomic efforts that coincided with incredible advances in mass spectrometry instrumentation and data analysis software. At the cusp of this technological wave, the malaria research community foresaw the power in harnessing these advances to improve and utilize quantitative proteomic approaches, to assess the composition of subcellular compartments, and to develop the means to robustly detect **post-translational modifications** (PTMs, see Glossary) on proteins [1–3]. These goals have been met, and in some ways exceeded. However, as rightly noted by Sinden at that time, having these catalogues of data without an understanding of the parasite and knowledge of how to

*Correspondence: Scott.Lindner@psu.edu (S.E. Lindner).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

integrate the two will limit the power of proteomic studies [2]. Now, with increasingly large and complex catalogues of proteomic data available, it is even more important to ground these data in a fundamental knowledge of what the parasite must do to infect and transmit, and how it might accomplish that. In this way, proteomics can be a powerful way to fill in mechanistic details, and at best, provide testable models for the mechanisms and processes involved. Here we cover many of the recent global, subcellular, quantitative, and PTM-focused proteomic efforts with *Plasmodium* parasites (summarized in Table 1), but regrettably cannot include all relevant publications that resulted from this incredible wave of proteomic discovery.

A Primer on Proteomics

While the term “proteomics” broadly refers to any number of bioanalytical techniques for identifying, characterizing, or quantifying proteins, in the context of systems biology and the other “omics”, e.g. genomics, transcriptomics, metabolomics, etc., proteomics almost exclusively refers to liquid chromatography (LC) coupled to tandem mass spectrometry (MS/MS). The most widely used approach for identification of proteins is so-called “**shotgun**” or “bottom-up” proteomics. In this method, proteins are enzymatically cleaved into peptides, most commonly with trypsin, and separated chromatographically prior to MS. In most **LC-MS/MS** experiments, peptides are loaded onto a reversed-phase column, typically C18, and separated according to hydrophobicity via a mobile phase consisting of water and a linearly increasing percentage of acetonitrile. Eluate is evaporated and peptides are ionized via electrospray as they exit the column, causing them to be introduced into the mass spectrometer as gas-phase ions. The standard shotgun proteomics method employs data-dependent analysis (DDA): first, peptide ions entering the MS are separated according to their mass-to-charge ratio (m/z), producing a precursor mass or MS^1 spectrum; then, the most abundant ion is isolated and fragmented, typically by collision-induced dissociation (CID), in which collisions with neutral gas molecules cause the peptide to fragment along the peptide backbone; finally, the m/z of the fragment ions is analyzed, producing a fragment or MS^2 mass spectrum. This process is repeated on the next most abundant ion, and so on, serially isolating and fragmenting peptide ions one-at-a-time. The masses of the peptide fragments observed in an MS^2 spectrum combined with the mass of the intact peptide precursor ion observed in an MS^1 spectrum may be used to determine the sequence of the peptide, which in turn may be used to infer the identity of the protein from which it was derived. The thousands of MS^2 spectra produced by DDA experiments may be analyzed through any number of open-source or commercial MS/MS database matching search engines (for a current listing of available software, see [4]) and analysis suites for assigning peptide spectrum matches (PSMs), inferring the identity of the parent protein, and assessing the false discovery rate (FDR) among the PSMs and protein inferences. While DDA experiments are primarily qualitative, it is possible to extract information about the relative abundance of individual proteins within and between samples using any of a number of so-called “label-free” quantification methods that infer peptide and protein abundance from various combinations of chromatographic peak height, spectral intensity, and the number of PSMs identifying a protein [5, 6]. Fold-change quantification by DDA can be aided by labeling peptides with isobaric but isotopically distinct tags, e.g. **iTRAQ** or **TMT** [7]. Using

this method, multiple samples (eight-channel iTRAQ is common) may be combined and analyzed by a single LC-MS/MS run, making it useful for experiments quantifying differential protein expression across multiple time points, treatment conditions, or cell types.

Other specialized and emerging MS-based proteomics techniques are also employed for protein identification, e.g. top-down proteomics, in which proteins are analyzed intact with no enzymatic cleavage, and data-independent acquisition (DIA) methods, in which all peptide precursors in narrow m/z bands are simultaneously fragmented rather than isolating individual precursors serially. However, at present shotgun proteomics using a DDA approach is by far the most widely used method for protein discovery, and is the sole method employed to date for *Plasmodium* proteomics. Accordingly, this review will focus on shotgun proteomics methods. Similarly, targeted proteomics methods for protein quantification such as selected reaction monitoring (SRM) and parallel reaction monitoring (PRM) [8] have yet to be widely employed for analysis of *Plasmodium*, and as such will not be discussed here.

Because there is no protein corollary to PCR by which the amount of protein in a sample may be multiplied, the number of proteins that may be reliably identified and quantified by LC-MS/MS is directly affected by several technical factors, including but not limited to: the amount, complexity, and concentration dynamic range of the sample; the duty cycle and sensitivity of the mass spectrometer; and the efficiency (i.e. peak width) of LC separation. Until recently, LC-MS instrumentation has been the predominant limiting factor to the efficiency of protein discovery; the number of proteins identified from a given organism by LC-MS/MS has historically always been far below the number of genes predicted to be expressed. However, the number of proteins detectable from a given amount of material within a given amount of analysis time has steadily increased over time with the advent of increasingly faster and more sensitive mass spectrometers, as well as advances in chromatography instruments and techniques capable of producing peak capacities to match the faster mass spectrometer duty cycles. Indeed, the most recent version of the Orbitrap mass spectrometer is reported to have identified over 10,000 proteins from mouse brain tissue in a single 100 min LC-MS/MS run [9].

Although LC-MS/MS instrumentation is becoming less and less of a limiting factor, *Plasmodium* proteomics efforts have and will continue to face major technical hurdles to producing sufficient amounts of purified sample for analysis. Even for laboratory strains that can be cultured *in vitro* or in animal models, *Plasmodium* samples are time-consuming and expensive to produce in large quantities, especially in the mosquito and **liver stages**. Additionally, as an obligate parasite, measures must always be taken to minimize contamination from host material. Although the relatively small size of the *Plasmodium* genome and streamlined number of protein-coding genes (5,000 – 6,000 ORFs) makes it very tractable for LC-MS/MS analysis, contamination with mammalian or arthropod material drastically increases the complexity of the sample and may even mask the signal from parasite proteins. All of the above reasons contribute to the relatively small number of *Plasmodium* proteomics studies published to date as well as the paucity of proteomics data for exoerythrocytic stages and for species other than *P. falciparum*.

Proteomics Can Inform Drug Discovery and Vaccine Candidate Development Efforts

Proteomics ultimately can indicate what proteins are present in a sample, and by inference, perhaps also inform us about the biology of specific life cycle stages and the parasite as a whole. These data are especially meaningful when the identity and/or functions of these proteins is known, or can be bioinformatically predicted. Ultimately, however, it is important to stress that *proteomics simply provides a parts list*, and that extrapolations to biological functions and mechanisms must be supported by additional experimentation. However, obtaining such a “catalog of parts” can be a powerful step toward advancing crucial malaria control efforts, such as the characterization of novel drugs and vaccine candidates.

Ideally, new anti-malarial drugs will target a protein present in all life cycle stages of the parasite to have the widest reaching effect. However, as drug treatment of mosquitoes is impractical, targeting proteins that are present during all stages of a human infection (**asexual blood stage**, sexual stage, liver stage) would also have a profound and perhaps nearly equal impact. Because of this, extensive proteomic efforts have been made to understand the blood stages in *P. falciparum* and *P. vivax*, as well as species that infect rodents as model systems (Figure 1, Table 1). However, very little is known about the proteome of the clinically silent liver stage that initiates a human infection, with only a single study published from work on *P. yoelii* late liver stage infections of mice (discussed below).

Currently new drugs are often discovered through large-scale robotic screening of commercial libraries for anti-plasmodial effects (e.g. a reduction of *in vitro* asexual blood stage growth of *P. falciparum*), or by development of compounds targeted to a specific protein that is essential to the parasite. Lead candidates that are downselected from large-scale screens (such as those included in the Malaria Box and Pathogen Box) are further studied by genomic and metabolomic assays in order to identify the mode-of-action and the protein(s) that may be affected by the compound [10–14]. Proteomics may further validate whether these protein targets may be the true targets in a given life cycle stage, or indicate that these compounds might be effective during an unexpected time of the development and transmission of the parasite.

Substantial efforts are also being made to identify an effective, long lasting, and strain transcending vaccine for malaria. Strong progress has been made with both subunit vaccines (those consisting of recombinant protein(s)) and live parasites attenuated through irradiation and/or gene deletions [15]. A leading approach uses genetically attenuated parasites that are lacking genes essential to liver stage development, which has provided 100% sterile protection in rodent models, as well as strongly encouraging results in controlled human infections [16, 17]. Selecting the best genes to delete has proven challenging, as the gene(s) must be essential for liver stage development, but yet dispensable for all other life cycle stages so that the attenuated **sporozoite** form of the parasite can be produced at scale. Thus, having a proteomic view of all life cycle stages would be extremely beneficial for prioritizing candidate genes for deletion in future rounds of vaccine development. As we describe in this review, many life cycle stages in most *Plasmodium* species remain

uncharacterized or under-characterized given the capabilities of current instrumentation and computational approaches (Figure 1). Chief among these gaps is a near vacuum of proteomic data for the liver stage itself, which is exceedingly challenging but now experimentally tractable.

Global Proteomics of *Plasmodium*'s Life Cycle Stages

Shortly after the release of nearly complete *Plasmodium* genomes, two major efforts to characterize the proteomes across multiple life cycle stages were published [18, 19]. In both, the total number of proteins detected across life cycle stages was large (e.g. 1800–2400 proteins), but were largely derived from blood stage parasites that could be provided in extremely large masses and with high purity. These findings were groundbreaking, and despite the technical limitations of instrumentation and data analysis, they provided a first window into the genomic clustering of protein expression, expression of antigenically variable proteins in both blood stages and sporozoites, and the translational repression mechanism that that sexual stage female **gametocytes** use with a number of their most abundant mRNAs. It is unlikely that a single, comprehensive publication like these will occur again, but instead, we will achieve a comprehensive view across life cycle stages from the assembly of data of specific stages across publications. However, in order to compile these data robustly, both the raw and processed data must be made fully available, and preferably processed in a consistent, robust manner. Until this is achieved, we urge the community to carefully consider the thresholds used for data reporting in each paper, and to download and independently process the raw data using the desired levels of stringency. Here we briefly describe the impact of recent proteomic efforts of human-infectious *P. falciparum* and rodent-infectious *P. yoelii* and *P. berghei* model species (distilled in Table 1), as much of the literature of *P. vivax* proteomics has been recently reviewed [20].

Asexual Blood Stage Parasites

By far, the easiest stages on which to conduct proteomic analyses are the **asexual blood stages**, where large amounts of sample are readily produced. All clinical symptoms of malaria present themselves during the asexual blood stage, which follows after the initiating liver stage infection [21]. Uncomplicated malaria (fever, chills, sweating and nausea) is caused by the synchronous infection and lysis of red blood cells, whereas complicated malaria (coma, low birth weight or death of unborn babies, death of infected individual) is due to parasite sequestration in vital tissues and the vasculature. Thus, proteomic analyses of these aspects of parasite biology are important to potentially ameliorating the pathologies associated with this disease. Production of large amounts of sample mass for proteomics is especially straightforward for *P. falciparum*, where robust methods to synchronize and isolate parasites in tight windows of development are routine. This is also reflected in the wealth of proteomic studies of this stage, and detailed catalogues of proteins that largely reflect the transcriptional cascade program (listed in Figure 1, Table 1). While analogous synchronization and purification methods exist for rodent-infectious *P. berghei* and *P. yoelii*, most studies to date have utilized mixed asexual blood stage parasites to limit vertebrate animal use, and because efficient methods to separate schizonts from gametocytes are lacking [18, 22–24]. While studies of asexual blood stage parasite are important in their own

right (all clinical symptoms of malaria are attributed to this stage), many laboratories also use asexual blood stage samples first as they develop new proteomic approaches for *Plasmodium*, such as quantitative proteomics, surface proteomics, protein/ligand identification, and detection of PTMs (discussed below).

Gametocytes

Advances made in asexual blood stages are often next extended to study the male and female gametocyte stages, as it is also relatively straightforward to produce these parasites at scale. The study of the differentiation of parasites into male and female stages is a rapidly expanding and important area, as the development of interventions that inactivate one or both sexes can have a substantial effect in reducing host-to-vector parasite transmission. Recent work has identified several commitment factors and other molecular mechanism that promote the differentiation of parasites from an asexual blood stage cycle into a linear sexual development pathway (reviewed in [25]). However, several fundamental questions of gametocyte biology to which proteomic approaches could be applied remain unanswered or are incompletely addressed. For instance: *What defines the male and female developmental paths? What prompts a single parasite to commit all of its progeny to become just one sex in the next cycle? What developmental cue prompts female gametocytes to induce specific translational repression of transcripts important to establishing the infection within the mosquito?* Following the first comparative studies of gametocytes by Florens *et al.* and Hall *et al.*, a recent flurry of studies have more clearly defined enriched male and female gametocyte markers, which then provided a means to select and separate the sexes late in gametocytogenesis to permit studies of what may make them transmissible to mosquitoes [18, 19, 26–30]. Additional technical advances are allowing these observations to be made earlier and earlier in gametocytogenesis, especially in *P. falciparum* where the developmental window is much larger (10 days vs 30 hours in rodent-infectious species). In addition to the presence/absence/tuning of protein abundance, recent work exploring possible epigenetic effects critical to this process has possibly provided another layer of regulation for this process [31].

Gametes, Zygotes, Ookinetes, and Oocysts

Proteomic studies of the early mosquito stages were prioritized because these parasite stages (**gametes, zygote, ookinete**) are not enclosed by a host membrane, and thus can still be targeted by human antibodies present in the blood meal. Compared with the blood stages, the mosquito stages of *Plasmodium* are even more difficult to produce in the numbers and purity necessary for analysis by LC-MS/MS. Accordingly, fewer proteomics studies have included mosquito stages, and the proteome coverage has been low compared to that achieved for blood stages. Nonetheless, proteomic analysis of the surface-exposed proteins of these stages has revealed potential vaccine antigens, as well as likely protein-protein interactions between male and female gametes [32, 33]. Despite the impact that these global and surface proteomic data could have upon therapeutics, studies of these stages have been limited to rodent-infectious *P. berghei* and bird-infectious *P. gallinaceum* [32, 34–36]. Although challenging, there is no apparent technical reason why these studies could not be done in *P. falciparum*, and thus they should be done instead of relying upon comparisons of **syntenic** orthologues between species. In the same vein, the only proteomic characterization

of the **oocyst** stage to date was produced by Hall and colleagues in *P. berghei* as part of their comparative study across the life cycle in 2005 [18]. As with other stages (e.g. sporozoites, liver stage parasites), the presence of substantial host/vector material hampered identification of oocyst proteins (277 *Plasmodium* proteins identified from 1000 mosquito midguts). Because there is exciting and under-explored biology that occurs within the oocyst, such as the formation of cytoplasmic islands/sporoblasts and the budding of sporozoites from them, revisiting the oocyst proteome with current instrumentation and software is certainly warranted.

Sporozoites

Transmission of the parasite from infected mosquitoes to a new mammalian host is via the sporozoite stage of the parasite. Hundreds or thousands of sporozoites can form and develop within a single oocyst found upon the mosquito midgut. In turn, a single mosquito can harbor one or a few oocysts as is seen in the field, or up to hundreds of oocysts in optimized laboratory conditions [37]. After maturation in the oocyst, sporozoites will egress from the midgut *en masse* in a semi-synchronous manner, and will enter the mosquito hemocoel where they become highly infectious [38]. Upon identification of the mosquito salivary gland, parasites invade this tissue and there remain poised for transmission. We and others have prioritized proteomic studies of the development of sporozoites in these environments as a means to understand how the sporozoite becomes highly infectious, and also to define how the sporozoite becomes and remains prepared for transmission.

In proteomic analyses of the multiple life stages of *P. yoelii* and *P. berghei*, the number of proteins detected for oocysts, oocyst sporozoites, and salivary gland sporozoites was considerably lower (fewer than 500) than for blood stages analyzed in the same work [18, 19, 33]. This poor coverage was attributed to heavy contamination with mosquito material. Salivary gland sporozoite protein detection was markedly improved in a multi-life stage analysis of *P. falciparum* by first purifying sporozoites on a Renograffin 60 discontinuous gradient [39, 40]. In our proteomic analyses of salivary gland sporozoites, we employed a discontinuous Accudenz gradient to enrich sporozoites, significantly reducing the extent of contamination from the mosquito and its microbiome [41]. Combined with further advances in mass spectrometry instrumentation, this purification method enabled detection of as many as 2,000 proteins from salivary gland sporozoites obtained from *P. falciparum*, *P. yoelii*, and *P. vivax* [42, 43]. Additional technical advances in sample preparation and LC-MS/MS will be required to achieve anything close to complete proteome coverage. Among the notable findings revealed by the above efforts was the presence of a sizeable complement of stage-specific proteins, including those specific to the mosquito stages of the parasite life cycle as well as those specific to each of the unique forms the parasite takes during this stage, i.e. ookinete, oocyst, oocyst sporozoite, and salivary gland sporozoite. Many of the important sporozoite-specific proteins characterized to date are related to invasion and motility, e.g. CSP, TRAP, P36/P52, CelTOS, SIAP, SPECT, and SPELD. Furthermore, cross-species comparisons have shown that these essential sporozoite proteins are conserved, and even show similar transcript and protein expression levels across species [42, 43]. Notably, however, there are many highly expressed, sporozoite-specific proteins that are not conserved, even among species that target the same vertebrate host (e.g. *P. vivax* and *P.*

falciparum [42–44]). In addition to there being evidence of divergent adaptation and specialization, such species-specific proteins also represent the need for species-specific interventions and highlight the challenge of developing pan-species vaccines and antimalarial drugs.

Liver Stage Parasites

The start of a new infection of humans (and other vertebrates) begins when a sporozoite is introduced into the skin, after which it locates and enters the blood stream to passively travel to the liver (reviewed in [45]). The sporozoite can then specifically and rapidly bind to hepatocytes in the short time they are within the vasculature of the liver. Following non-productive traversal of one or more hepatocytes, the parasite can productively infect a single hepatocyte. Ultimately, a single parasite that completes liver stage development is sufficient to cause a full blown asexual blood stage infection. Thus understanding how to intervene and completely arrest liver stage parasites is crucial to global malaria control strategies. Many of the leading malaria vaccine candidate rely upon arresting the parasite completely within the liver stage, either through live attenuated parasite vaccines or subunit vaccines that can induce both antibody and T-cell responses (reviewed in [16, 46]). However, conducting proteomics of liver stage parasites is exceedingly challenging due to the difficulty in separating parasites from host material. Despite this, Tarun and colleagues identified a proteome of late (40 hrs) and very late (50 hrs) *P. yoelii* liver stage parasites, which was possible due to the very large size of these parasites (stretching the hepatocyte to nearly four times its normal volume) and the presence of an integrated GFP reporter that permitted FACS selection [47]. This work revealed proteins that are present in these late liver stage parasites, and suggested that if *P. falciparum* or *P. vivax* orthologues were expressed similarly, these could be prioritized candidates for genetic attenuation. Studies of very early, early, and mid liver stage parasites have not been published, and this key portion of the life cycle is under-described. Moreover, liver stage proteomics of *P. falciparum* and *P. vivax* parasites, especially of the dormant **hypnozoite** stage of *P. vivax*, *P. ovale*, and *P. cynomolgi*, would be exceptional contributions to the global eradication agenda [48]. These should be technically feasible now with recent advances in humanized mouse models and *in vitro* micropatterned cell culture systems [49–51].

Proteomics of Membrane-Bound Compartments

Gaining subcellular resolution of where these proteins reside has largely relied upon using density gradients to purify the target organelles, coupled with targeted validation experiments that are important to confirm a subset of these findings. Using *P. falciparum* asexual blood stage parasites, the nucleus, apicoplast, food vacuole, Maurer's clefts, and other compartments have been characterized [52–57]. Additionally, work to characterize *P. berghei* ookinete micronemes and asexual blood stage secreted proteins has been reported [23, 33, 58].

More recently, **BioID** and similar approaches have gained favor (reviewed in [59]). This system targets a promiscuous biotin ligase to a protein/compartments, which generates a cloud of reactive biotin that is largely contained to a labeling radius (10nm) and membrane-

bound compartments. In the last few years, this has been used to define the gametocyte egressome, osmiophilic bodies, the parasitophorous vacuole and its membrane (PVM) [60–62]. As these systems can provide spatial information about intracellular complexes and structures, we anticipate these will gain in popularity, especially as new variants of the biotin ligases with better labeling properties (e.g. TurboID) are described [63].

Finally, a “compartment” that warrants additional consideration is the surface of the invasive forms of the parasite: the merozoite, ookinete, and sporozoite. These are the forms of the parasite that are not enclosed by a host cell, and thus their surface-exposed proteins are accessible to antibody-based interventions.

Our understanding of asexual blood stage merozoite surface proteins is considerable, as blood stage parasites can be made in bulk and these proteins are critical for parasite survival in the extracellular environment and infection of the new RBC [64]. Perhaps due to this wealth of data, few efforts have been published describing these surface-exposed proteins by mass spectrometric approaches. However, a particularly fruitful approach focused upon GPI-anchored proteins found in detergent-resistant membrane domains of extracellular merozoites, or reciprocally by using radiolabeled GPI precursors [65, 66]. In these studies, several of the merozoite surface protein (MSP) proteins, along with 6-Cys family proteins and rhoptry proteins were identified, thus validating (and roughly quantifying) old targets and providing new ones. As the purpose of identifying surface-exposed proteins is to ultimately produce them as recombinant antigens, work by the Wright laboratory that has identified correlates of expression in common mammalian cell-based systems is exciting and is an important extension of this effort [67, 68].

In assessing the ookinete surface, Wass *et al.* recognized that the use of fractionation approaches for surface proteomics can also include proteins found in the apical organelles. Instead, a membrane-impermeable, biotinylated, amine-reactive tag was used to label surface exposed proteins, which has gained traction in the field as a preferred approach [33]. Through comparisons between the global and surface ookinete proteomes, many appreciated surface/secreted proteins were clearly detected (e.g. CTRP, p28, p25, LAP1, LAP2, SOAP, chitinase), as well as several uncharacterized, conserved proteins that warrant additional study. The authors noted a large number of typically cytosolic proteins also being labelled, due to issues with membrane integrity during the labeling process. While many of these hits can reasonably be discounted for this reason, others have been validated to be *bona fide* surface proteins (described below) and caution is urged.

Among the first sporozoite proteins to be characterized were CSP and TRAP. Given that these proteins are highly abundant, essential for invasion, immunogenic, and localized at the sporozoite surface, it was reasoned that other sporozoite surface proteins would also be viable targets to understand and potentially disrupt invasion. Our proteomic analysis of *P. falciparum* and *P. yoelii* salivary gland sporozoites [43] included a pilot study to identify surface-exposed proteins by the biotin-labeling approach Wass and colleagues employed for the ookinete surface proteome [33]. We followed this work with an exhaustive survey of the *P. falciparum* salivary gland sporozoite surface proteome employing improved methodology and more rigorous quantitative analysis of enrichment, and have since employed these

methods to identify surface proteins on *P. vivax* salivary gland sporozoites [42, 69]. The Honavar and Lindner laboratories employed the same methods to more comprehensively survey the *P. yoelii* salivary gland sporozoite surface proteome, and used the combined data of the above studies to develop a semi-supervised model for predicting sporozoite surface proteins based on protein primary sequence [70]. In all studies that use this surface labeling approach, many spurious protein identifications have been made, including known cytoplasmic proteins (e.g. histone and ribosomal proteins). While some of these may be contaminants due to parasites lacking full membrane integrity, many of the seemingly dubious identifications have been determined by molecular biology methods to be truly surface-exposed on salivary gland sporozoites, including the chaperone HSP20, the enzyme GAPDH, members of the inner membrane complex, and poly(A)-binding protein 1 (PABP1) [42, 71–73]. While most of the putative surface proteins identified in the above studies have yet to be experimentally validated, the evidence to date suggests that the complement of sporozoite surface proteins includes unexpected “moonlighting” cytosolic proteins, potentially with alternate, stage-specific functions.

Post-Translational Modifications

One of the unique capabilities of mass spectrometry-based proteomics is high-throughput detection of protein PTMs. Importantly, in most cases the exact residue modified can be determined from the mass spectra. In theory, any chemical moiety that is covalently attached to a protein may be detected by mass spectrometry, though in practice, automated MS/MS sequence database searching of shotgun data provides the most robust results for PTMs that consistently modify a small number of residues or motifs, that withstand CID without fragmenting or falling off, and whose mass is not isobaric with other PTMs or amino acid substitutions. Additionally, detection of protein PTMs is greatly enhanced when modified peptides can be enriched prior to LC-MS/MS. Accordingly, the most widely studied PTM in *Plasmodium* proteomics (and proteomics generally) is phosphorylation. Multiple studies of *P. falciparum* intraerythrocytic stages have used standard phosphopeptide enrichment techniques followed by LC-MS/MS to demonstrate dynamic regulation of development and invasion during this phase of asexual reproduction [74–78]. Recently, a sub-minute time course of *P. berghei* gametogenesis showed that this metamorphosis is preceded by rapid changes in protein phosphorylation, implicating several kinases as regulators of the process [35].

Other PTMs in *Plasmodium* have also been revealed by enrichment approaches. Proteome-wide assessment of lysine acetylation in *P. falciparum* trophozoites was performed after chromatographic and immunoprecipitation (IP) enrichment of modified peptides, revealing that the PTM is widespread throughout the proteome and likely plays a role in transcriptional regulation. IP enrichment also aided in detection of arginine methylation in *P. falciparum* trophozoites, which also appears to be widespread in the proteome and may have a variety of functions [79]. Subcellular enrichment also aids in PTM detection: LC-MS/MS analysis of histones isolated from *P. falciparum* sexual and asexual stages identified extensive acetylation, methylation, and ubiquitination of lysine, revealing dynamic epigenetic regulation of these stages [31, 80]. Similar strategies have also enabled detection of acetylated proteins [81, 82].

While detection of PTMs is greatly enhanced by pre-enrichment, PTMs may also be detected from *post hoc* analysis of shotgun proteomics data from un-enriched samples. At present, the only published phosphoproteomic data for salivary gland sporozoites was obtained from researching the raw proteomics data from *P. falciparum* and *P. vivax* salivary gland sporozoites [42]. Relatively few phosphoproteins were identified compared to what would be possible with enrichment, but the data nonetheless revealed phosphorylation of transcriptional regulators and gliding machinery, suggesting a role for phospho-signaling in regulation of these processes. These same datasets were also searched for N-terminal protein acetylation, providing evidence of protein processing for regulation of protein function and secretion [83]. Further, the evidence for O-fucosylation and C-mannosylation of CSP and TRAP in *P. falciparum* and *P. vivax* was obtained from manual interpretation of the raw mass spectra from these same proteomics datasets [42, 69]. Future studies employing enrichment strategies will reveal many more PTMs in other life stages and species, but in the interim, there is much information that has yet to be extracted from re-analysis of existing datasets.

PTMs that are not amenable to direct detection by MS may be detected by replacing the PTM with a chemical tag, especially one that enables IP enrichment. For example, nitric oxide-based signaling by cysteine S-nitrosylation is difficult to detect by proteomics because the S-NO bond is highly labile. Detection of S-nitrosylation in *P. falciparum* trophozoites was achieved by chemically replacing nitrous groups with biotin and enriching modified proteins by IP [84]. LC-MS/MS of these proteins showed that glycolysis pathways were a major target of the PTM, and led to the discovery that GAPDH activity was inhibited by S-nitrosylation of its active site cysteine. O-GlcNAc is similarly difficult to detect due to its high lability in the gas phase, but replacing it with biotin via click chemistry enabled enrichment and identification of O-GlcNAcylated proteins in *P. falciparum* rings and trophozoites [85]. The proteins modified with this PTM were involved in glycolysis, protein folding, and cell shape, similar to functions seen in other organisms. Replacing problematic PTMs with biotin has also been used to deal with modification of proteins by hydrophobic lipids that make the modified peptides difficult to detect by standard methods. The modification of *P. falciparum* trophozoite proteins with farnesol (via prenylation), myristate, and palmitate has been detected by metabolically labeling parasites with a lipid analogs that could be replaced with biotin via click chemistry and enriching these proteins by IP [86–88]. These lipid modifications have been shown to be critical for parasite growth, and as such represent important targets for antimalarials.

Concluding Remarks: A Call for Standardized Data Deposition, Analysis, and Accessibility

Despite the depth and range of the studies summarized here, *Plasmodium* proteomics is still a nascent field compared to the advances achieved with more tractable organisms (see Outstanding Questions). The volume and quality of *Plasmodium* proteomics data is likely to grow at increasing rates as LC-MS instrumentation continues to overcome technical hurdles inherent to study of the parasite. As the body of data grows, it will be incumbent on the generators of the data to present the information to the broader research community in a

clear and useful manner. At a minimum, the raw mass spectrometry data from proteomics experiments should be made publically available in the repositories of the ProteomeXchange and experimental metadata should be provided in accordance with the minimum information about a proteomics experiment (MIAPE) principles developed by the Proteomics Standards Initiative (PSI) of the Human Proteome Organization (HUPO) [89–91]. However, to be of the greatest use to the *Plasmodium* research community, the aggregate information obtained from these experiments must be presented in manner that allows researchers to easily peruse the results without having to download and analyze the raw data themselves. At present, the proteins identified from many (but not all) *Plasmodium* proteomics datasets are compiled at PlasmoDB.org [92, 93]. The results are searchable and the identified peptides from every contributed dataset, including PTMs, are mapped to the predicted protein sequence in a visual interface. Importantly, this resource combines proteomics data with multiple other data types, e.g. genomics, transcriptomics, gene ontology, and phenotypes, an invaluable resource that reveals important information such as evidence for stage-specific expression and translational repression of proteins. The primary shortcoming of the proteomics data assembled to date at PlasmoDB is that it simply reports the number of peptides and spectra supplied by the researcher providing the dataset. Given the wide variation in quality of proteomics data and the lack of universally accepted standards for MS data processing, there is likely a high false discovery rate among the reported identifications, and at present there is no way for the user to easily assess the quality of the underlying data. A beneficial data curation model we would suggest adopting is that of PeptideAtlas [94], a multi-species proteomics data compendium. User-submitted raw MS data for a given organism is searched against the most up-to-date genome available, combined with all other submitted data acquired to date, then analyzed with open-source MS data analysis tools to produce a final list of high-confidence protein identifications that comprises the aggregate knowledge of all available (user-submitted) data. Having the repository perform the analyses rather than the submitters provides several advantages: stringent criteria can be used to ensure low false discovery rates; combining data from multiple experiments and labs will boost the confidence of true positives that are repeatedly identified while eliminating spurious identifications; the data can be searched for PTMs that were not included in the submitter's original analysis; and the data can be re-analyzed whenever necessary, such as when new data is submitted or when there are major updates to the reference genome. Positioned as we are at what is likely the cusp of a new era of *Plasmodium* proteomics, the standards developed by the *Plasmodium* community now for acquisition, analysis, and dissemination of proteomics data will impart immeasurable benefits to future work to eradicate malaria.

Glossary

Asexual Blood Stage:

The stage where *Plasmodium* parasites infect and develop within a vertebrate red blood cell. Morphologically these parasites transition from a ring stage to a trophozoite to a schizont, the latter of which can rupture the host membrane and release infectious merozoites to infect a new red blood cell.

BioID:

The use of a promiscuous biotin ligase to label proteins with biotin within some distance and/or compartment. Labeled proteins are readily captured on streptavidin-coated beads.

Gametocytes/Gametes:

The sexual stage of the parasite develops within the vertebrate host as a gametocyte precursor cell, which activates into a gamete upon transmission to a mosquito.

Hypnozoite:

A dormant form of the liver stage parasite found in some *Plasmodium* species.

Isotopic Labeling (iTRAQ/TMT):

Peptides are labeled with primary amine-reactive, stable isotope tags that are isobaric but that produce reporter ions of distinct masses upon fragmentation of the peptide by collision-induced dissociation. Multiple samples (e.g. different biological replicates, time points, or treatment conditions) can be labeled separately then pooled and analyzed by a single LC-MS/MS experiment

LC-MS/MS:

Liquid chromatography (LC) coupled to mass spectrometry (MS). Tandem mass spectrometry (MS/MS) refers to the technique of separating ions according to their mass-to-charge ratio (m/z), then fragmenting isolated ions within the mass spectrometer (typically by collision-induced dissociation (CID)) and separating the fragments according to their m/z .

Liver Stage Parasite:

A trophic form of the parasite found within a single hepatocyte that expands parasite numbers a thousand fold and ultimately releases merozoites that can infect red blood cells.

Oocyst:

A parasite stage associated with the mosquito midgut, inside of which sporogony is used to expand parasite numbers a thousand fold or more.

Ookinete:

An invasive form of the parasite found within the mosquito midgut, which can burrow through the midgut wall and re-establish an association with the midgut under the basal lamina.

Post-Translational Modifications (PTMs):

Changes to specific amino acids that affect their chemical properties. Common PTMs include phosphorylation, acetylation, methylation, and glycosylation.

Shotgun Proteomics:

An LC-MS/MS technique for identifying proteins from analysis of component peptides produced by enzymatic cleavage, typically with trypsin. The mass spectra of the fragmented peptides are searched against a database comprising the predicted gene products of the originating organism. Peptide spectrum matches (PSM) identify peptide sequences, which in turn are used to infer the proteins present in the sample. This is currently one of the most common proteomic approaches.

Sporozoite:

An invasive form of the parasite found first within oocysts, which egress, traverse through the mosquito hemocoel, and invade the mosquito salivary gland to await passive transmission in the saliva during a blood meal.

Syntenic Genes:

Genes that encode protein orthologues and are found in the same location and context in the genome across species.

Zygote:

Fusion of the male gamete (microgamete) and female gamete (macrogamete) produces a parasite with 2N chromosomal content.

References

1. Sims PF and Hyde JE (2006) Proteomics of the human malaria parasite *Plasmodium falciparum*. Expert review of proteomics 3, 87–95 [PubMed: 16445353]
2. Sinden RE (2004) A proteomic analysis of malaria biology: integration of old literature and new technologies. International journal for parasitology 34, 1441–1450 [PubMed: 15582521]
3. Johnson JR, et al. (2004) Proteomics in malaria. Journal of proteome research 3, 296–306 [PubMed: 15113107]
4. https://en.wikipedia.org/wiki/List_of_mass_spectrometry_software
5. Al Shweiki MR, et al. (2017) Assessment of Label-Free Quantification in Discovery Proteomics and Impact of Technological Factors and Natural Variability of Protein Abundance. Journal of proteome research 16, 1410–1424 [PubMed: 28217993]
6. Cox J, et al. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Molecular & cellular proteomics : MCP 13, 2513–2526 [PubMed: 24942700]
7. Zhang L and Elias JE (2017) Relative Protein Quantification Using Tandem Mass Tag Mass Spectrometry. Methods in molecular biology 1550, 185–198 [PubMed: 28188531]
8. Shi T, et al. (2016) Advances in targeted proteomics and applications to biomedical research. Proteomics 16, 2160–2182 [PubMed: 27302376]
9. Meier F, et al. (2018) BoxCar acquisition method enables single-shot proteomics at a depth of 10,000 proteins in 100 minutes. Nature methods
10. Cowell AN, et al. (2018) Mapping the malaria parasite druggable genome by using in vitro evolution and chemogenomics. Science 359, 191–199 [PubMed: 29326268]
11. Creek DJ, et al. (2016) Metabolomics-Based Screening of the Malaria Box Reveals both Novel and Established Mechanisms of Action. Antimicrobial agents and chemotherapy 60, 6650–6663 [PubMed: 27572396]
12. Allman EL, et al. (2016) Metabolomic Profiling of the Malaria Box Reveals Antimalarial Target Pathways. Antimicrobial agents and chemotherapy 60, 6635–6649 [PubMed: 27572391]
13. Van Voorhis WC, et al. (2016) Open Source Drug Discovery with the Malaria Box Compound Collection for Neglected Diseases and Beyond. PLoS pathogens 12, e1005763 [PubMed: 27467575]
14. Spangenberg T, et al. (2013) The open access malaria box: a drug discovery catalyst for neglected diseases. PloS one 8, e62906 [PubMed: 23798988]
15. Hoffman SL, et al. (2015) The march toward malaria vaccines. Vaccine 33 Suppl 4, D13–23 [PubMed: 26324116]
16. Kreutzfeld O, et al. (2017) Engineering of Genetically Arrested Parasites (GAPs) For a Precision Malaria Vaccine. Frontiers in cellular and infection microbiology 7, 198 [PubMed: 28620583]
17. Kublin JG, et al. (2017) Complete attenuation of genetically engineered *Plasmodium falciparum* sporozoites in human subjects. Science translational medicine 9

18. Hall N, et al. (2005) A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses. *Science* 307, 82–86 [PubMed: 15637271]
19. Florens L, et al. (2002) A proteomic view of the Plasmodium falciparum life cycle. *Nature* 419, 520–526 [PubMed: 12368866]
20. Bourgard C, et al. (2018) Plasmodium vivax Biology: Insights Provided by Genomics, Transcriptomics and Proteomics. *Frontiers in cellular and infection microbiology* 8, 34 [PubMed: 29473024]
21. Phillips MA, et al. (2017) Malaria. *Nature reviews. Disease primers* 3, 17050 [PubMed: 28770814]
22. Niikura M, et al. (2018) Comparative genomics and proteomic analyses between lethal and nonlethal strains of Plasmodium berghei. *Experimental parasitology* 185, 1–9 [PubMed: 29307564]
23. Pasini EM, et al. (2013) Proteomic and genetic analyses demonstrate that Plasmodium berghei blood stages export a large and diverse repertoire of proteins. *Molecular & cellular proteomics : MCP* 12, 426–448 [PubMed: 23197789]
24. Hart KJ, et al. (2018) Plasmodium male gametocyte development and transmission are critically regulated by general and transmission-specific members of the CAF1/CCR4/NOT complex. *bioRxiv*
25. Josling GA, et al. (2018) Regulation of Sexual Commitment and Gametocytogenesis in Malaria Parasites. *Annual review of microbiology*
26. Miao J, et al. (2017) Sex-Specific Biology of the Human Malaria Parasite Revealed from the Proteomes of Mature Male and Female Gametocytes. *Molecular & cellular proteomics : MCP* 16, 537–551 [PubMed: 28126901]
27. Suarez-Cortes P, et al. (2016) Comparative Proteomics and Functional Analysis Reveal a Role of Plasmodium falciparum Osmiophilic Bodies in Malaria Parasite Transmission. *Molecular & cellular proteomics : MCP* 15, 3243–3255 [PubMed: 27432909]
28. Lasonder E, et al. (2016) Integrated transcriptomic and proteomic analyses of P. falciparum gametocytes: molecular insight into sex-specific processes and translational repression. *Nucleic acids research* 44, 6087–6101 [PubMed: 27298255]
29. Tao D, et al. (2014) Sex-partitioning of the Plasmodium falciparum stage V gametocyte proteome provides insight into falciparum-specific cell biology. *Molecular & cellular proteomics : MCP* 13, 2705–2724 [PubMed: 25056935]
30. Khan SM, et al. (2005) Proteome analysis of separated male and female gametocytes reveals novel sex-specific Plasmodium biology. *Cell* 121, 675–687 [PubMed: 15935755]
31. Coetzee N, et al. (2017) Quantitative chromatin proteomics reveals a dynamic histone post-translational modification landscape that defines asexual and sexual Plasmodium falciparum parasites. *Scientific reports* 7, 607 [PubMed: 28377601]
32. Talman AM, et al. (2014) Proteomic analysis of the Plasmodium male gamete reveals the key role for glycolysis in flagellar motility. *Malaria journal* 13, 315 [PubMed: 25124718]
33. Wass MN, et al. (2012) Proteomic analysis of Plasmodium in the mosquito: progress and pitfalls. *Parasitology* 139, 1131–1145 [PubMed: 22336136]
34. Garcia CHS, et al. (2018) Dynamic molecular events associated to Plasmodium berghei gametogenesis through proteomic approach. *Journal of proteomics* 180, 88–98 [PubMed: 29155091]
35. Invergo BM, et al. (2017) Sub-minute Phosphoregulation of Cell Cycle Systems during Plasmodium Gamete Formation. *Cell reports* 21, 2017–2029 [PubMed: 29141230]
36. Patra KP, et al. (2008) Proteomic analysis of zygote and ookinete stages of the avian malaria parasite Plasmodium gallinaceum delineates the homologous proteomes of the lethal human malaria parasite Plasmodium falciparum. *Proteomics* 8, 2492–2499 [PubMed: 18563747]
37. Billingsley PF, et al. (1994) Relationship between prevalence and intensity of Plasmodium falciparum infection in natural populations of Anopheles mosquitoes. *The American journal of tropical medicine and hygiene* 51, 260–270 [PubMed: 7943543]
38. Sato Y, et al. (2014) Plasmodium berghei sporozoites acquire virulence and immunogenicity during mosquito hemocoel transit. *Infection and immunity* 82, 1164–1172 [PubMed: 24379288]

39. Lasonder E, et al. (2008) Proteomic profiling of Plasmodium sporozoite maturation identifies new proteins essential for parasite development and infectivity. *PLoS pathogens* 4, e1000195 [PubMed: 18974882]
40. Pacheco ND, et al. (1979) Rapid, large-scale isolation of Plasmodium berghei sporozoites from infected mosquitoes. *The Journal of parasitology* 65, 414–417 [PubMed: 39127]
41. Kennedy M, et al. (2012) A rapid and scalable density gradient purification method for Plasmodium sporozoites. *Malaria journal* 11, 421 [PubMed: 23244590]
42. Swearingen KE, et al. (2017) Proteogenomic analysis of the total and surface-exposed proteomes of Plasmodium vivax salivary gland sporozoites. *PLoS neglected tropical diseases* 11, e0005791 [PubMed: 28759593]
43. Lindner SE, et al. (2013) Total and putative surface proteomics of malaria parasite salivary gland sporozoites. *Molecular & cellular proteomics : MCP* 12, 1127–1143 [PubMed: 23325771]
44. Jex A, et al. (2017) Integrated transcriptomic, proteomic and epigenomic analysis of Plasmodium vivax salivary-gland sporozoites. *bioRxiv*
45. Lindner SE, et al. (2012) Malaria parasite pre-erythrocytic infection: preparation meets opportunity. *Cellular microbiology* 14, 316–324 [PubMed: 22151703]
46. Ballou WR (2009) The development of the RTS,S malaria vaccine candidate: challenges and lessons. *Parasite immunology* 31, 492–500 [PubMed: 19691554]
47. Tarun AS, et al. (2008) A combined transcriptome and proteome survey of malaria parasite liver stages. *Proceedings of the National Academy of Sciences of the United States of America* 105, 305–310 [PubMed: 18172196]
48. Markus MB (2015) Do hypnozoites cause relapse in malaria? *Trends in parasitology* 31, 239–245 [PubMed: 25816801]
49. Gural N, et al. (2018) In Vitro Culture, Drug Sensitivity, and Transcriptome of Plasmodium Vivax Hypnozoites. *Cell host & microbe* 23, 395–406 e394 [PubMed: 29478773]
50. Mikolajczak SA, et al. (2015) Plasmodium vivax liver stage development and hypnozoite persistence in human liver-chimeric mice. *Cell host & microbe* 17, 526–535 [PubMed: 25800544]
51. Vaughan AM, et al. (2012) Complete Plasmodium falciparum liver-stage development in liver-chimeric mice. *The Journal of clinical investigation* 122, 3618–3628 [PubMed: 22996664]
52. Oehring SC, et al. (2012) Organellar proteomics reveals hundreds of novel nuclear proteins in the malaria parasite Plasmodium falciparum. *Genome biology* 13, R108 [PubMed: 23181666]
53. Lamarque M, et al. (2008) Food vacuole proteome of the malarial parasite Plasmodium falciparum. *Proteomics. Clinical applications* 2, 1361–1374 [PubMed: 21136929]
54. Nyalwidhe J and Lingelbach K (2006) Proteases and chaperones are the most abundant proteins in the parasitophorous vacuole of Plasmodium falciparum-infected erythrocytes. *Proteomics* 6, 1563–1573 [PubMed: 16470785]
55. Vincensini L, et al. (2005) Proteomic analysis identifies novel proteins of the Maurer's clefts, a secretory compartment delivering Plasmodium falciparum proteins to the surface of its host cell. *Molecular & cellular proteomics : MCP* 4, 582–593 [PubMed: 15671043]
56. Sam-Yellowe TY, et al. (2004) Proteome analysis of rhoptry-enriched fractions isolated from Plasmodium merozoites. *Journal of proteome research* 3, 995–1001 [PubMed: 15473688]
57. Boucher MJ, et al. (2018) The proteome of the malaria plastid organelle, a key anti-parasitic target. *bioRxiv*
58. Lal K, et al. (2009) Characterisation of Plasmodium invasive organelles; an ookinete microneme proteome. *Proteomics* 9, 1142–1151 [PubMed: 19206106]
59. Kim DI and Roux KJ (2016) Filling the Void: Proximity-Based Labeling of Proteins in Living Cells. *Trends in cell biology* 26, 804–817 [PubMed: 27667171]
60. Schnider CB, et al. (2018) BioID Reveals Novel Proteins of the Plasmodium Parasitophorous Vacuole Membrane. *mSphere* 3
61. Khosh-Naucke M, et al. (2017) Identification of novel parasitophorous vacuole proteins in P. falciparum parasites using BioID. *International journal of medical microbiology : IJMM*
62. Kehr J, et al. (2016) Proteomic Analysis of the Plasmodium berghei Gametocyte Egressome and Vesicular bioID of Osmiophilic Body Proteins Identifies Merozoite TRAP-like Protein (MTRAP)

- as an Essential Factor for Parasite Transmission. *Molecular & cellular proteomics* : MCP 15, 2852–2862 [PubMed: 27371728]
63. Branon TC, et al. (2017) Directed evolution of TurboID for efficient proximity labeling in living cells and organisms. *bioRxiv*
 64. Beeson JG, et al. (2016) Merozoite surface proteins in red blood cell invasion, immunity and vaccines against malaria. *FEMS microbiology reviews* 40, 343–372 [PubMed: 26833236]
 65. Gilson PR, et al. (2006) Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite *Plasmodium falciparum*. *Molecular & cellular proteomics* : MCP 5, 1286–1299 [PubMed: 16603573]
 66. Sanders PR, et al. (2005) Distinct protein classes including novel merozoite surface antigens in Raft-like membranes of *Plasmodium falciparum*. *The Journal of biological chemistry* 280, 40169–40176 [PubMed: 16203726]
 67. Zenonos ZA, et al. (2014) Towards a comprehensive *Plasmodium falciparum* merozoite cell surface and secreted recombinant protein library. *Malaria journal* 13, 93 [PubMed: 24620899]
 68. Crosnier C, et al. (2013) A library of functional recombinant cell-surface and secreted *P. falciparum* merozoite proteins. *Molecular & cellular proteomics* : MCP 12, 3976–3986 [PubMed: 24043421]
 69. Swearingen KE, et al. (2016) Interrogating the *Plasmodium* Sporozoite Surface: Identification of Surface-Exposed Proteins and Demonstration of Glycosylation on CSP and TRAP by Mass Spectrometry-Based Proteomics. *PLoS pathogens* 12, e1005606 [PubMed: 27128092]
 70. El-Manzalawy Y, et al. (2016) PlasmoSEP: Predicting surface-exposed proteins on the malaria parasite using semisupervised self-training and expert-annotated data. *Proteomics* 16, 2967–2976 [PubMed: 27714937]
 71. Minns AM, et al. (2018) Nuclear, Cytosolic, and Surface-Localized Poly(A)-Binding Proteins of *Plasmodium yoelii*. *mSphere* 3
 72. Montagna GN, et al. (2012) Critical role for heat shock protein 20 (HSP20) in migration of malarial sporozoites. *The Journal of biological chemistry* 287, 2410–2422 [PubMed: 22139844]
 73. Cha SJ, et al. (2016) Identification of GAPDH on the surface of *Plasmodium* sporozoites as a new candidate for targeting malaria liver invasion. *The Journal of experimental medicine* 213, 2099–2112 [PubMed: 27551151]
 74. Lasonder E, et al. (2015) Extensive differential protein phosphorylation as intraerythrocytic *Plasmodium falciparum* schizonts develop into extracellular invasive merozoites. *Proteomics* 15, 2716–2729 [PubMed: 25886026]
 75. Pease BN, et al. (2013) Global analysis of protein expression and phosphorylation of three stages of *Plasmodium falciparum* intraerythrocytic development. *Journal of proteome research* 12, 4028–4045 [PubMed: 23914800]
 76. Lasonder E, et al. (2012) The *Plasmodium falciparum* schizont phosphoproteome reveals extensive phosphatidylinositol and cAMP-protein kinase A signaling. *Journal of proteome research* 11, 5323–5337 [PubMed: 23025827]
 77. Solyakov L, et al. (2011) Global kinomic and phospho-proteomic analyses of the human malaria parasite *Plasmodium falciparum*. *Nature communications* 2, 565
 78. Treeck M, et al. (2011) The phosphoproteomes of *Plasmodium falciparum* and *Toxoplasma gondii* reveal unusual adaptations within and beyond the parasites' boundaries. *Cell host & microbe* 10, 410–419 [PubMed: 22018241]
 79. Zeeshan M, et al. (2017) Proteomic Identification and Analysis of Arginine-Methylated Proteins of *Plasmodium falciparum* at Asexual Blood Stages. *Journal of proteome research* 16, 368–383 [PubMed: 27933903]
 80. Trelle MB, et al. (2009) Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite *Plasmodium falciparum*. *Journal of proteome research* 8, 3439–3450 [PubMed: 19351122]
 81. Cobbold SA, et al. (2016) Proteome-wide analysis reveals widespread lysine acetylation of major protein complexes in the malaria parasite. *Scientific reports* 6, 19722 [PubMed: 26813983]

82. Miao J, et al. (2013) Extensive lysine acetylation occurs in evolutionarily conserved metabolic pathways and parasite-specific functions during *Plasmodium falciparum* intraerythrocytic development. *Molecular microbiology* 89, 660–675 [PubMed: 23796209]
83. Arnesen T (2011) Towards a functional understanding of protein N-terminal acetylation. *PLoS biology* 9, e1001074 [PubMed: 21655309]
84. Wang L, et al. (2014) Protein S-nitrosylation in *Plasmodium falciparum*. *Antioxidants & redox signaling* 20, 2923–2935 [PubMed: 24256207]
85. Kupferschmid M, et al. (2017) Identification of O-GlcNAcylated proteins in *Plasmodium falciparum*. *Malaria journal* 16, 485 [PubMed: 29187233]
86. Gisselberg JE, et al. (2017) The Prenylated Proteome of *Plasmodium falciparum* Reveals Pathogen-specific Prenylation Activity and Drug Mechanism-of-action. *Molecular & cellular proteomics : MCP* 16, S54–S64 [PubMed: 28040698]
87. Wright MH, et al. (2014) Validation of N-myristoyltransferase as an antimalarial drug target using an integrated chemical biology approach. *Nature chemistry* 6, 112–121
88. Jones ML, et al. (2012) Analysis of protein palmitoylation reveals a pervasive role in *Plasmodium* development and pathogenesis. *Cell host & microbe* 12, 246–258 [PubMed: 22901544]
89. Deutsch EW, et al. (2017) The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. *Nucleic acids research* 45, D1100–D1106 [PubMed: 27924013]
90. Vizcaino JA, et al. (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nature biotechnology* 32, 223–226
91. Taylor CF, et al. (2007) The minimum information about a proteomics experiment (MIAPE). *Nature biotechnology* 25, 887–893
92. Warrenfeltz S, et al. (2018) EuPathDB: The Eukaryotic Pathogen Genomics Database Resource. *Methods in molecular biology* 1757, 69–113 [PubMed: 29761457]
93. Aurrecochea C, et al. (2009) PlasmoDB: a functional genomic database for malaria parasites. *Nucleic acids research* 37, D539–543 [PubMed: 18957442]
94. Kusebauch U, et al. (2014) Using PeptideAtlas, SRMAtlas, and PASSEL: Comprehensive Resources for Discovery and Targeted Proteomics. *Current protocols in bioinformatics* 46, 13 25 11–28
95. Zhang Y, et al. (2013) Proteomic analysis of *Plasmodium falciparum* schizonts reveals heparin-binding merozoite proteins. *Journal of proteome research* 12, 2185–2193 [PubMed: 23566259]
96. Silvestrini F, et al. (2010) Protein export marks the early phase of gametocytogenesis of the human malaria parasite *Plasmodium falciparum*. *Molecular & cellular proteomics : MCP* 9, 1437–1448 [PubMed: 20332084]
97. Bowyer PW, et al. (2011) Global profiling of proteolysis during rupture of *Plasmodium falciparum* from the host erythrocyte. *Molecular & cellular proteomics : MCP* 10, M110 001636
98. Florens L, et al. (2004) Proteomics approach reveals novel proteins on the surface of malaria-infected erythrocytes. *Molecular and biochemical parasitology* 135, 1–11 [PubMed: 15287581]
99. Nilsson Bark SK, et al. (2018) Quantitative Proteomic Profiling Reveals Novel *Plasmodium falciparum* Surface Antigens and Possible Vaccine Candidates. *Molecular & cellular proteomics : MCP* 17, 43–60 [PubMed: 29162636]
100. Acharya P, et al. (2009) A glimpse into the clinical proteome of human malaria parasites *Plasmodium falciparum* and *Plasmodium vivax*. *Proteomics. Clinical applications* 3, 1314–1325 [PubMed: 21136953]

Highlights

- The number and quality of mass spectrometry-based proteomics studies of *Plasmodium* has increased significantly due to substantial improvements in instrumentation and computational approaches. The high mass accuracy and duty cycle of these instruments now enables confident detection and identification of thousands of proteins and post-translational modifications.
- Proteomic characterization of sub-compartments of the parasite and its surface has provided opportunities to explore the mechanisms controlling parasite functions, and to identify new targets for therapeutic interventions.
- Studies have largely centered upon life cycle stages where sample mass is not limiting (e.g. *P. falciparum* blood stage). There is a need to expand our proteomic understanding of mosquito and liver stage parasites in essentially all *Plasmodium* species.
- A comprehensive data curation model should be applied to all raw datasets to standardize Plasmodium proteomics data and reduce false positives.

Outstanding Questions

- How much and in what ways does protein expression change over the course of the *Plasmodium* life cycle, and how similar are the proteomes of a specific life cycle stage among *Plasmodium* species?
- How are post-translational modifications used by the parasite to regulate its essential functions, and to evade host immune responses?
- In what ways does *Plasmodium* control translation during key life cycle transitions, such as transmission?
- What is the best way to distill and disseminate proteomics data to the larger *Plasmodium* research community?

Figure 1: Proteomic studies across the *Plasmodium* life cycle.

Illustrations of the major life cycle stages of *Plasmodium* parasites are provided, with a designation of whether total proteomics has been reported (filled circles) or not (empty circles) in four major species (Pf = *P. falciparum*, Pv = *P. vivax*, Pb = *P. berghei*, Py = *P. yoelii*) or others (key provided in the lower right). Hypnozoites are known to be present in *P. vivax*, *P. ovale*, and *P. cynomolgi* infections (empty circles), but not in *P. falciparum*, *P. berghei*, or *P. yoelii* infections (gray filled circles). It is noteworthy that while blood stage parasites have been extensively studied, mosquito stage and liver stage parasites have not. In some cases, all proteomic data for a life cycle stage is derived from a single study with a single species.

Table 1:

A description of proteomics efforts covered in this review.

Life Cycle Stage(s)	Description	Authors	Year	Ref#
<i>P. falciparum</i>				
T, Mz, G, SG-SPZ	Comparative Study	Florens <i>et al.</i>	2002	[19]
Sch	Heparin-binding proteins	Zhang <i>et al.</i>	2013	[95]
R,T,S	Phosphoproteomics, Isobaric tags	Pease <i>et al.</i>	2013	[75]
T,G	Gametocyte protein export	Silvestrini <i>et al.</i>	2010	[96]
Mz	Phosphoproteomics	Lasonder <i>et al.</i>	2015	[74]
Mz	Surface proteomics, GPI-anchored proteins	Sanders <i>et al.</i>	2005	[66]
Mz	Surface proteomics, GPI-anchored proteins	Gilson <i>et al.</i>	2006	[65]
Sch	Study of RBC membrane rupture, Proteolysis	Bowyer <i>et al.</i>	2011	[97]
M-ABS	Surface proteomics	Florens <i>et al.</i>	2004	[98]
M-ABS	Quantitative proteomics, Surface proteomics	Nilsson Bark <i>et al.</i>	2018	[99]
M-ABS	Field isolates	Acharya <i>et al.</i>	2009	[100]
G, FG	Late (Stage V) Gametocytes	Tao <i>et al.</i>	2014	[29]
MG, FG	Separated gametocytes	Khan <i>et al.</i>	2005	[30]
MG, FG	FACS separated sexes, Translational Repression in FG	Lasonder <i>et al.</i>	2016	[28]
MG, FG	FACS separated sexes	Miao <i>et al.</i>	2017	[26]
G	Osmiophilic bodies	Suarez-Cortes <i>et al.</i>	2016	[27]
M-ABS	Apicoplast proteome (BioID)	Boucher <i>et al.</i>	2018	[57]
R,T,S	Nuclear proteome	Oehring <i>et al.</i>	2012	[52]
Mz	Rhoptry	Sam-Yellowe <i>et al.</i>	2004	[56]
M-ABS (RBC Ghosts)	Maurer's cleft proteome	Vincensini <i>et al.</i>	2005	[55]
T,Sch	Food vacuole proteome	Lamarque <i>et al.</i>	2008	[53]
T	Parasitophorous vacuole proteome	Nyalwidhe and Lingelbach	2006	[54]
M-ABS	Parasitophorous vacuole proteome (BioID)	Khosh-Naucke <i>et al.</i>	2017	[61]
M-ABS	Parasitophorous vacuole membrane proteome (BioID)	Schnider <i>et al.</i>	2018	[60]
M-ABS	Histone post-translational modifications	Trelle <i>et al.</i>	2009	[80]
R, T, S, G	Histone post-translational modifications	Coetzee <i>et al.</i>	2017	[31]
Sch	Phosphoproteomics	Solyakov <i>et al.</i>	2011	[77]
Sch	Phosphoproteomics	Lasonder <i>et al.</i>	2012	[76]
Mz	Phosphoproteomics	Lasonder <i>et al.</i>	2015	[74]
Sch	Total proteomics, Phosphoproteomics	Treeck <i>et al.</i>	2011	[78]
T	Acetylation	Miao <i>et al.</i>	2013	[82]
T	Acetylation	Cobbold <i>et al.</i>	2016	[81]
R,T,S	Arginine methylation	Zeeshan <i>et al.</i>	2017	[79]
T	Prenylation	Gisselberg <i>et al.</i>	2017	[86]
M-ABS	O-GlcNAcylation	Kupferschmid <i>et al.</i>	2017	[85]

Life Cycle Stage(s)	Description	Authors	Year	Ref#
<i>P. berghei</i>				
M-ABS, G	Comparative Study, Translational Repression in FG	Hall <i>et al.</i>	2005	[18]
M-ABS	Secreted proteome	Pasini <i>et al.</i>	2013	[23]
G	Egressome, Osmiophilic body proteomes (BioID)	Kehrer <i>et al.</i>	2016	[62]
Sch, G	Total proteomics	Niikura <i>et al.</i>	2018	[22]
M-ABS	Total proteomics	Hart <i>et al.</i>	2018	[24]
<i>P. falciparum</i>				
SG-SPZ	Comparative study	Florens <i>et al.</i>	2002	[19]
OO-SPZ, SG-SPZ	Total proteomics	Lasonder <i>et al.</i>	2008	[39]
SG-SPZ	Total and surface proteomics, Phosphoproteomics, N-acetylation	Lindner and Swearingen <i>et al.</i>	2012	[43]
SG-SPZ	Surface proteomics, Glycosylation	Swearingen <i>et al.</i>	2016	[69]
<i>P. berghei</i>				
Ook, OO, SG-SPZ	Comparative study	Hall <i>et al.</i>	2005	[18]
microG	Total proteomics	Talman <i>et al.</i>	2014	[32]
micro/macroG	Phosphoproteomics	Invergo <i>et al.</i>	2017	[35]
micro/macroG	iTRAQ proteomics	Garcia <i>et al.</i>	2018	[34]
Ook	Ookinete microneme	Lai <i>et al.</i>	2009	[58]
Ook	Ookinete surface proteomics	Wass <i>et al.</i>	2012	[33]
<i>P. yoelii</i>				
SG-SPZ	Total and surface proteomics, Phosphoproteomics, N-acetylation	Lindner and Swearingen <i>et al.</i>	2012	[43]
SG-SPZ	Surface proteomics, Predictive model for surface exposure	El-Manzalawy <i>et al.</i>	2016	[70]
<i>P. vivax</i>				
SG-SPZ	Thai field isolate, Total proteomics, Surface proteomics	Swearingen <i>et al.</i>	2017	[42]
SG-SPZ	Multi-omics comparative study	Jex <i>et al.</i>	2017	[44]
<i>P. gallinaceum</i>				
Z, Ook	Total proteomics, Ookinete secreted proteins	Patra <i>et al.</i>	2008	[36]
<i>P. yoelii</i>				
LS 40 hour, LS 50 hour	Late liver stage proteomics	Tarun <i>et al.</i>	2008	[47]

Footnote: Abbreviations used - M-ABS: Mixed Asexual Blood Stage; R: Ring Stage; T: Trophozoite; Sch: Schizont; Mz: Merozoite; G: Gametocyte; MG: Male Gametocyte; FG: Female Gametocyte; microG: Male Gamete (Microgamete); macroG: Female Gamete (Macrogamete); Z: Zygote; Ook: Ookinete; OO: Oocyst; OO-SPZ: Oocyst Sporozoite; SG-SPZ: Salivary Gland Sporozoite; LS: Liver Stage Parasite.