Originally published as:

Paape, D., Aebischer, T.
Contribution of proteomics of Leishmania spp. to the understanding of differentiation, drug resistance mechanisms, vaccine and drug development

DOI: 10.1016/j.jprot.2011.05.005

This is an author manuscript.
The definitive version is available at: http://www.sciencedirect.com/
Titel: Contribution of Proteomics of *Leishmania* spp. to the Understanding of Differentiation, Drug Resistance Mechanisms, Vaccine and Drug Development

Daniel Paape\(^1\), Toni Aebischer\(^2\)

\(^1\)Centre for Immunology and Infection, Department of Biology/Hull York Medical School, University of York, YO10 5DD, UK
\(^2\)Robert Koch-Institut, 13353, Berlin, Germany

Corresponding author: Daniel Paape

Daniel.paape@york.ac.uk
Abstract:

*Leishmania* spp., protozoan parasites with a digenetic life cycle, cause a spectrum of diseases in humans. Recently several *Leishmania* spp. have been sequenced which significantly boosted the number and quality of proteomic studies conducted. Here a historic review will summarize work of the pre-genomic era and then focus on studies after genome information became available. Firstly works comparing the different life cycle stages, in order to identify stage specific proteins, will be discussed. Identifying of post-translational modifications by proteomics especially phosphorylation events will be discussed. Further the contribution of proteomics to the understanding of the molecular mechanism of drug resistance and the investigation of immunogenic proteins for the identification of vaccine candidates will be summarized. Approaches of how potentially secreted proteins were identified are discussed. So far 30 – 35% of the total predicted proteome of *Leishmania* spp. has been identified. This comprises mainly the abundant proteins, therefore the last section will look into technological approaches how this coverage may be increased and what the gel-free and gel-based proteomics have to offer will be compared.
Introduction:

Genomics and transcriptomics were until recently unmatched in their power to enable highly parallel insight into the biology of cells and organisms. However, proteomics and more recently metabolomics have come of age and contribute and expand this knowledge.

The idea that mRNA levels are indicators of the respective protein expression/abundance, is often untrue and mRNA abundance can be misleading\(^1\), \(^2\) and do not inform about post-translational protein modifications that can dramatically change protein function. This is especially true for the trypanosomatids as these organisms regulate their protein expression and abundance post-transcriptionally\(^3\). The three major human disease causing trypanosomatids - conveniently referred to as TriTryps - are unicellular organisms of the genera *Leishmania* spp. and *Trypanosoma brucei* and *Trypanosoma cruzi* which cause morbidity and mortality in humans and life stock in the middle East, south and south-east Asia, middle and South-America as well as countries surrounding the Mediterranean sea and sub-Saharan Africa (data from WHO). *Leishmania* the causative agent of the leishmaniases has as the other trypanosomatids a digenetic life cycle and oscillates between the extracellular promastigote form transmitted by sand flies and the intracellular amastigote form occurring in vertebrate hosts. Currently there are 350 million people at risk of leishmaniases with a prevalence of 12 million people and an estimated 1.5 – 2 million new infections each year. There are 60,000 – 70,000 deaths each year mostly due to the most severe form of the leishmaniases the visceral leishmaniasis caused by *L. donovani* in India but also by *L. infantum* in northern Africa and southern Europe\(^4\),\(^5\). For a detailed review on *Leishmania* transmission and distribution see Alexander *et al*. and Reithinger *et al*. respectively\(^4\),\(^5\). Less severe disease outcomes are the mucocutaneous and the cutaneous leishmaniasis caused by *L. braziliensis* and *L. major* as well as *L. mexicana*, respectively.

Here, we like to review the information available on the proteome of *Leishmania* spp., a field that has greatly advanced since the genomic information became available in 2005\(^6\) and complement a recent review on proteomics of TriTryps\(^7\) by the latest data.

The focus will be a historical review of how proteomics in *Leishmania* spp. advanced over time and its contribution to the understanding of molecular changes in the different life cycle stages, drug resistance and to the identification of vaccine candidates and novel drug targets.

*Leishmania* spp. proteome profiling in pre-genomic era

*Leishmania* spp. have a core genome of 8178 genes shared between the three currently sequenced and annotated genomes of *L. major*, *L. infantum* and *L. braziliensis*, and only few genes are species specific, 5, 27 and 49, respectively\(^8\). Although it is known that host factors play also a part in disease
outcome, it is reasonable to assume that for the course of an infection differences in gene content or expression level in parasite species and isolates determine their respective virulence. Since the parasites regulate gene expression mainly at post-transcriptional stages, proteomics is thought to yield critical insight into the mechanisms of stage differentiation, species differences, virulence and drug resistance.

The first proteome maps of *Leishmania* were published in the early 1980s\(^9\)\(^{-11}\) long before the term proteomics was coined\(^{12}\). Handman *et al.* as well as Saravia and colleagues were investigating species differences and virulence factors\(^9\)\(^{-11}\). However, at the time no protein was identified as mass spectrometric technologies were not yet developed for large biomolecules, the genome/s of *Leishmania* were not yet sequenced and identifications were relying on for example Edman degradation. Identification of large biomolecules became only possible from the late 1980s onwards when two groups independently of each other developed matrix assisted laser desorption/ionization – time of flight (MALDI-TOF)\(^{13,14}\) analyses. The spectrum of analysis methods was then widened by the development of electrospray ionization (ESI)\(^{15}\). Several studies have led to stepwise improvement of methodology to investigate the parasites’ proteomes. Suitable lysis methods of *Leishmania* to resolve as many protein-species as possible were identified\(^{16}\). A number of detailed 2DE proteome maps were then produced. For example, Brobey and colleagues reported normalized maps with 1650 and 1530 protein spots for *L. amazonensis* and *L. major*, respectively and showed 4 of these as only present in *L. amazonensis*\(^ {17}\). Similarly, Gongora and colleagues produced highly resolved proteome 2DE maps of *L. (Viannia) guyanensis* and *L. (Viannia) panamensis* promastigotes\(^ {18}\) and identified ‘landmark protein spots’. These landmark protein spots were analysed by mass spectrometry and identified as 2 isoforms of heat shock protein 70 (hsp70), chaperonin (hsp60), ribosomal protein S12, kinetoplastid membrane protein 11 and a 13 kDa small myristoylated protein-3, putative (annotated as hypothetical at the time)\(^ {18}\). This enabled a comparison with clones of *L. guyanensis* that differed in metastatic capacities highlighting four protein species that may be involved in the different biologic behaviour\(^ {18}\).

However, it was the sequencing of the genomes of *L. major*\(^6\), *L. infantum* and *L. braziliensis*\(^8\), with data available since the beginning of this millennium that enabled the relatively straight forward identification of protein-species excised from 2DE gels by mass spectrometry or from peptides separated by liquid chromatography (LC) as discussed below.

**Identification of life cycle stage specific proteins**

To understand the differences during transformation and between the different life-cycle stages on the molecular level several groups resolved the proteome of *L. infantum*, *L. donovani* and *L.
mexicana promastigotes and amastigotes\textsuperscript{19-21} and promastigotes upon reception of the transformation signal\textsuperscript{22} or metacyclics\textsuperscript{23}.

A first study analysing protein spots excised from 2D gels by mass spectrometry investigated the different life cycle stages of L. infantum (promastigotes and axenic amastigotes) and 62 proteins species were reported to be differentially expressed \textsuperscript{19}. Genomes of L. major\textsuperscript{6}, L. infantum and L. braziliensis\textsuperscript{8} were not available at the time and obtained amino acid sequences were used for identification of homologous proteins by using the BLAST search algorithm that requires a very high degree of sequence conservation to unambiguously identify a protein, hence only two proteins (isocitrate dehydrogenase and triosephosphate isomerase) could be identified. Later the GeneDB database became publicly available which provided a platform for initially the L. major genome and later for all sequenced genomes\textsuperscript{24}. This presented a great ease for the field in general and led to a dramatic increase in studies conducted and proteins identified of Leishmania spp. (see also Figure. 1). For example, analysis of 101 protein spots of L. braziliensis promastigotes revealed that they originated from 75 genes\textsuperscript{25}. The transformation from promastigotes to axenic amastigotes of L. donovani was investigated by Bente \textit{et al.} and 115 regulated proteins species were reported. Analysis of the transformation process in L. mexicana revealed 147 regulated protein species\textsuperscript{23}. Stage specific differences in protein species abundance were detected in L. panamensis\textsuperscript{20}, where in axenic amastigotes 75 proteins species showed differential expression and in promastigotes 45 were solely detected or upregulated. A method to isolate truly intracellular amastigotes from infected bone marrow derived macrophages and lesions to perform proteomics was developed and reported by Paape and colleagues\textsuperscript{21}. The authors used DsRed expressing L. mexicana promastigotes to infect BALB/c mice or axenic amastigotes to infect macrophages. Amastigotes were purified and depleted of host cell material by fluorescent activated cell sorting by positive selection for the DsRed expressing parasites while at the same time selecting negatively for the main host cell contaminants, nucleic material using DAPI fluorescence as a negative selectable marker. The subsequent proteomic analysis revealed 34 proteins solely detected in amastigotes\textsuperscript{21}. Proteomes of L. major\textsuperscript{26} and L. braziliensis\textsuperscript{25} promastigotes were also analysed and methods to increase proteome coverage were investigated. Drummelsmith and colleagues extended the 2DE technique by performing IEF over narrower but overlapping pH ranges and resolved a total of 3696 distinct L. major protein spots. Several were mass spectrometrically analysed and 37 protein spots were identified as 29 proteins\textsuperscript{26}.

These studies highlight the fact that between the different life-cycle stages are “subtle” differences, protein species of similar MW or pI originating from the same gene, but also major differences in protein expression, up- or down regulation of expression and stage specific expression. Most of these studies produced 2DE maps to visualize and detect especially “subtle” differences that may be due to post-translational modifications (PTMs) affecting charged amino acids. Although more
proteins may be identified by gel free approaches (see below) these 2DE maps will continue to provide a significant resource that can help to identify changes at the protein species level. The current most comprehensive study investigating the molecular changes at the proteome level during differentiation of *L. donovani* promastigotes to axenic amastigotes were carried out by Rosenzweig and colleagues\textsuperscript{27} in the first gel free proteome study in *Leishmania* spp. This approach allowed to obtain higher proteome coverage and more precise quantitative information. Peptides of the digested proteome were labelled by the ‘isobaric tags for relative and absolute quantification (iTRAQ) method, this allowed the comparison of up to four proteomes at a time and determine changes in relative abundance of proteins over the course of differentiation. Parasites were harvested at 2.5, 5, 10, 15 and 24h post differentiation signal (pH 5.5 and 37 °C) and as axenic amastigotes (6 d post differentiation signal) and the proteome of these stages were compared by multidimensional LC-MS. In total, 1713 proteins could be identified, with 931 confirmed in biological repeats representing ~21% of the predicted *L. infantum* proteome. Of the identified proteins 934 had a known or predicted function and 779 had no known function. Almost 56% (969) of the identified proteins were expressed at all stages, with 289 gradually increased, 310 gradually decreased in abundance as parasites transformed from pro- into amastigotes. The other 370 were either not affected or expression increased and then decreased again. Many enzymes of the major metabolic pathways were identified and their expression pattern was analysed. Expression of glycosomal glycolytic enzymes increased in amastigotes whereas cytosolic glycolytic enzymes decreased suggesting a reduced glycolysis in amastigotes confirming early biochemical analyses using *ex vivo* prepared amastigotes\textsuperscript{28}. The glycosomal glycolytic enzymes are likely involved in gluconeogenesis, the glycosomal gluconeogenic enzymes pyruvate phosphate, dikinase phosphoenolpyruvate carboxykinase and fructose 1,6-bis-phosphahtase were also higher expressed in axenic amastigotes. Transamination of alanine could provide a source for pyruvate in the gluconeogenesis and the respective enzyme alanine aminotransfearase was found upregulated in axenic amastigotes. Enzymes involved in β-oxidation, the tricarboxylic acid cycle, mitochondrial respiration/ oxidative phosphorylation and amino acid catabolism were expressed at a higher level in axenic amastigotes. It was observed that the abundance of enzymes involved in pentose phosphate metabolism in axenic amastigotes was decreased. Proteins involved in translation (ribosomal proteins, translation elongation and initiation factors and tRNA-synthetases) remained unaltered in expression for the initial differentiation period (up to 10h) and then decreased in their abundance. Histones only increased in expression in the later stages of differentiation, post 10h. Of 10 detected calpain proteases, five decreased and three increased in abundance in axenic amastigotes, indicating a developmental switch in the predominant expression of calpains. Nineteen heat shock proteins and chaperones were identified which differed in their
expression levels, some increased in the initial differentiation period and returned then to promastigote levels or below, others did not change until later in differentiation when their expression returned to promastigote levels or below. A minor group also gradually decreased in abundance throughout differentiation and only hsp100 gradually increased in its expression. Also 20 proteins were reported to be up-regulated by more than 2-fold as early as 2.5h post differentiation signal, 13 of these have no known function, the others being D1 myo-inositol transporter, universal minicircle sequence binding protein, flavoprotein subunit-like protein, methionine synthetase, h1 histone-like, kinetoplast-associated protein-like protein and ascorbate-dependent peroxidase. This work clearly boosted insight into the molecular changes during the differentiation process as the parasite adapts to its new hostile environment within the lysosome-like parasitophorous vacuole. The most dramatic changes in protein abundance only occurred late in the transformation 24 – 144h post transformation signal.

The fluorescence based purification approach was developed further for gel free proteomic analysis. Thereby 1764 (~22% of theoretical *Leishmania* spp. proteome) proteins could be identified. Moreover, by adapting a label free algorithm to derive relative abundance values of the identified proteins, protein abundance data could be integrated e.g. into the *in silico* metabolic pathway map curated for *Leishmania* by McConville’s group. This approach confirmed predictions about metabolic adaptations to the intracellular habitat. The further integration of ‘omics’ data sets has just begun to yield interesting insight and in combination with metabolic network analyses will be very revealing in the future. Furthermore for the first time the experimental data confirmed the presence of a codon bias in these *Leishmania*, which is of importance when *Leishmania* spp. is used as an expression platform.

**Phosphoproteomics and parasite differentiation**

Post-translational modifications alter or extend the function and stability of the modified protein, it can further also influence its localization within the cell, hence identification and the implication of PTMs on a large scale became into recent focus. Phosphorylation/dephosphorylation of proteins by kinases/phosphatases at serine, threonine and/or tyrosine residue/s is a very important PTM. This regulatory element mediates signal transduction in the cell, activation/deactivation of proteins/enzymes, protein-protein interactions and also protein turn-over. This led to the hypothesis that phosphorylation is also an important mechanism involved in receiving/mediating the differentiation signal when the parasites is transmitted and has to adapt to a new environment. The importance of it was shown for a protein-tyrosine phosphatase from *L. donovani* were promastigotes and axenic amastigotes devoid of both alleles were mostly
unaffected in growth and differentiation but parasites were severely attenuated in their ability to establish themselves in the spleen or liver of BALB/c mice\textsuperscript{34}.

Morales and colleagues published the first analysis of the phosphoproteome of \textit{Leishmania}\textsuperscript{35}. Phosphoproteins of \textit{L. donovani} promastigotes and axenic amastigotes in logarithmic growth phase were enriched by immobilized metal affinity chromatography (IMAC) and proteomes separated by 2DE, the resulting gels were stained with ProQ diamond, a phosphoprotein specific stain, and for total proteins. In total 73 phosphoproteins were identified, corresponding to 55 proteins with an annotated function and 18 proteins annotated as hypothetical. It was also reported that in amastigotes relatively more proteins were phosphorylated compared to promastigotes 4% and 2.6% of total protein content, respectively\textsuperscript{35}. The identified proteins were classified in five functional groups: 1) stress and heat shock response; 2) RNA turnover, ribosomal biogenesis and protein degradation; 3) signalling; 4) metabolic processes and others and 5) hypothetical proteins. The authors also reported stage specific phosphorylation events with differential abundance of phosphoproteins, e.g. HSP-83, RNA helicase and glucose-6 phosphate dehydrogenase were upregulated in amastigotes whereas the protein phosphatase LmjF05.0100 was 3-fold upregulated in promastigotes; different isoforms in either life cycle stage and proteins exclusively phosphorylated in amastigotes, e.g. HSP100, Leishmania MAP kinase homolog LmaMPK10 and nucleotide diphosphate kinase b. Analysing of iTRAQ labelled \textit{L. donovani} promastigotes and axenic amastigote peptides\textsuperscript{27} was performed to detect PTMs of amino acid residues\textsuperscript{36}. This lead to the identification of 16 phosphorylated, 20 methylated, 26 acetylated and 13 glycosylated (fucosylated and hexosylated) peptides. This also implicated the involvement of phosphorylation events in the stress response, RNA and protein turnover\textsuperscript{36}.

Further differences in the phosphorylation were revealed by comparison of \textit{L. donovani} promastigotes and axenic amastigotes of IMAC enriched phosphoproteome by 2D-differential gel electrophoresis\textsuperscript{37}. In total 813 protein spots were detected with more than 700 present in three biological repeats. Of these 171 could be identified containing 55 not present in the previous study\textsuperscript{35}. Yeast orthologs of the obtained dataset were analysed with respect to their gene ontology (GO), this revealed six statistically significant GO categories to be overrepresented translation initiation, protein folding and catabolism, chromosome organization and biogenesis, organelar fusion and amino acid metabolism. Differential phosphorylation was observed for 318 protein spots, corresponding to 38% of the detected phosphoproteins. Amastigote proteins increasing in abundance/phosphorylation were almost exclusively protein chaperones e.g. HSP90/HSP83, HSP70 family members and stress induced protein STI1\textsuperscript{37}. Phosphorylation of HSPs might be involved in the regulation of HSP function in the parasite. For example, a highly conserved aspartic acid seen in the mouse or human HSP90 homologue is substituted by a serine that becomes phosphorylated in
amastigotes. The aspartic acid is proposed to correspond to a constitutive phosphorylation that locks the protein in a functionally distinct conformation.

In amastigotes the presence of a multimeric chaperone complex interacting with ribosomal proteins was reported with STI1 playing a crucial role in its formation. The complex might be involved in the control of protein translation\cite{37} or simply provide a platform of translation and formation of correctly folded proteins. Three phosphorylation sites of STI were identified (residues S15, T217 and S481) with residues S15 and S481 required for *L. donovani* viability in culture highlighting the significance of chaperone phosphorylation in the parasite\cite{37}.

The same group also analysed the phosphoproteome of *L. donovani* axenic amastigotes by gel free mass spectrometry. In two biological repeats 445 putative phosphoproteins were identified with 48\% annotated as hypothetical proteins. Based on *S. cerevisiae* orthologs (239) proteins were classified in functional GO categories with evidence for enrichment of the following biological processes: protein transport, translation, translation initiation, protein degradation and stress response. To verify for phosphorylation sites phosphoprotein fractions were enriched for phosphopeptides post trypsin digest on TiO$_2$-columns. Phosphopeptides were then analysed by 2D LC-MS/MS analysis and 126 proteins with 181 phosphorylation sites were identified\cite{38}. Serine residues were most frequently phosphorylated followed by threonine and tyrosine residues, 86\%, 12\% and 2\%, respectively, a figure becoming more accurate with a larger dataset. This is in an order that would be expected for higher eukaryotes\cite{39} but it is substantially different from *T. brucei* and *T. cruzi* where about one-third of the phosphorylated residues were found to be threonines\cite{40,41}. Whether this is of biological relevance or due to sample handling needs to be verified. Seventy eight percent of the proteins were phosphorylated at a single site, 19\% dually and 3\% triple phosphorylated. Several of the identified phosphorylation sites were not conserved or where the respective sequence was absent in other *Leishmania* spp.\cite{38} maybe reflecting species differences that may be linked to the different disease outcomes of tropisms of infection (cutaneous, visceral). The phosphorylation sites were even less conserved in comparison to other trypanosomatids (e.g. *T. cruzi, T. brucei, T. vivax* and *T. conglolense*). Interestingly *L. infantum* based on this analysis is closer related to *T. cruzi* than to *T. brucei* and *T. vivax*. The authors speculate that this might reflect the relevance of these sites for processes involved in intracellular survival/adaptation. On the other hand 40\% of the proteins are not present in the human genome and phosphorylation sites in 19 proteins were only present in the leishmanial homologue but not in the human protein, indicating that the responsible kinases may be potential drug targets. This idea is currently pursued by the Leishdrug consortium. Overall, the data suggest that PTMs are involved in the stress response in *Leishmania*\cite{37,38}.
Proteomics to study drug resistance in *Leishmania* spp.

Drug resistant parasites are a major problem in the field since leishmaniasis management relies on drugs\(^42\). However, the mechanistic reasons for drug resistance are only partially understood. Therefore several studies utilized comparative proteomics to look at the different protein expression pattern between drug susceptible and resistant parasites to elucidate the molecular mechanism of drug resistances.

Laboratory induced methotrexate drug resistance in *L. major* was investigated\(^26\). Methotrexate inhibits dihydrofolate reductase–thymidylate synthase (DHFR-TS) and to a much lesser extent pteridine reductase\(^43\). Comparison of 2DE maps of resistant and sensitive *L. major* parasite revealed a number of differentially expressed protein spots. The most striking was the presence of a 30 kDa protein spot with an approximate pi of 6.3 in the resistant line which was absent in the susceptible line. This protein spot was then identified as pteridine reductase\(^26\), over-expression of it is linked to methotrexate resistance\(^44, 45\). This was investigated in more depth in another comparative proteome study of *L. major* wt and two independent *in vitro* derived methotrexate resistant strains. Proteomes were resolved and 16 protein spots were analysed of which intensity differed by more than 2-fold in at least one of the two resistant strains when compared to the wild type parasites. Fifteen protein species, encoded by 14 genes, could be identified; interestingly some protein species were only present in one resistant strain but not the other, e.g. pteridine reductase 1 and argininosuccinate synthase, in one, and acyl-coa-dehydrogenase in the other strain\(^46\). Over-expression of argininosuccinate synthase and acyl-coa-dehydrogenase was the result of gene locus amplification events of genes in close structural proximity linked to methotrexate resistance pteridine reductase 1 and DHFR-TS, respectively. DHFR-TS is the primary target of methotrexate\(^47\) and gene amplification is one mechanism to respond to drug pressure\(^48, 49\).

It was also reported that enolase and S-adenosylmethionine synthetase were expressed at higher levels in resistant strains indicating that some responses to methotrexate selection pressure are conserved. The authors hypothesized that these differences constituted an early response to methotrexate exposure rather than resistance and compared the proteome of wild type parasites exposed for 24 h to an EC\(_{50}\) concentration of methotrexate with that of not exposed parasites. This comparison identified proteins belonging to the stress response (thiol-specific antioxidants) and again enolase and S-adenosylmethionine synthetase as previously detected in the resistant strains. Enolase as other glycolytic enzymes were reported to be over-expressed in response to several stresses like heat shock or osmotic stress\(^50, 51\). S-adenosylmethionine synthetase was linked to be involved in the acquisition of resistance\(^46\).

Vergnes and colleagues used 2DE resolved proteomes of antimony [Sb(V)] resistant and susceptible Indian field isolates\(^52\) to study resistance to this drug in *L. donovani*. In total 11 protein species were
identified seven up-regulated (heat shock protein 83; hsp70-related protein 1, mitochondrial precursor, chaperonin containing tcp1, subunit 8; 14-3-3 protein; ATP-dependent RNA helicase; hSP70, dipeptidyl-peptidase III) and five down-regulated (hsp70-related protein 1, mitochondrial precursor; hypothetical protein, unknown function; enolase, 20 S proteasome α5 subunit, SKCRP14.1 (small kinetoplastid calpain related protein)). Several of these proteins (hsp83, 14-3-3 protein and SKCRP14.1) have reported roles in programmed cell death in a number of organisms. This led to the hypothesis that programmed cell death is altered in the resistant strain. While over-expression of 14-3-3 did not confer resistance to Sb(V) increased hsp83 levels double the dose-dependent resistance of the parasites. Interestingly hsp83 over-expressing parasites were also cross-resistant to miltefosine. Over-expression of SKCRP14.1 increased the sensitivity to Sb(V) but decreased its susceptibility to miltefosine.\textsuperscript{52}

Sb(V) drug resistance was also investigated by comparative proteomics of Sb(V) resistant and susceptible \textit{L. donovani}. Promastigotes were lysed, separated into a ‘membrane enriched protein’ and a cytosolic protein fraction and the proteomes resolved by 2DE. Over expressed protein spots derived from the resistant parasites were analysed and of 12 protein spots in the membrane fraction 6 could be identified (cysteine-leucine rich protein, ABC and ABC1 transporter, hsp83, GPI protein transamidase and 60S ribosomal protein L23a(L25)). In the cytosolic fraction 14 of 18 protein spots originating from eight genes were identified (hsp70, carboxypeptidase, fructose-1,6-bisphosphate aldolase, enolase, proliferative cell nuclear antigen (PCNA), proteasome alpha 5 subunit and tubulin-beta chain).\textsuperscript{53}

Comparative proteomics was also used to study antimony resistance in \textit{L. infantum} axenic amastigote from resistant and susceptible strains to trivalent antimony [Sb(III)].\textsuperscript{54} 2DE maps were generated and differentially expressed proteins were analysed. Several proteins were down-regulated in the resistant parasites, LACK receptor, proteasome activator protein pa26 subunit, pyruvate kinase, kinetoplastid membrane protein 11 (KMP-11) and β-tubulin only one protein was over-expressed, argininosuccinate synthase. Argininosuccinate synthase is physically linked to an ABC transporter (P-glycoprotein A / MRPA) which was reported to be involved in antimony resistance.\textsuperscript{55, 56} The reason that the latter was not identified is likely to be related to the fact that it is a transmembrane protein which are generally under-represented in proteomic analyses. An increased protein turnover rate of KMP-11 was observed in the resistant parasites. It was suggested by the authors that the decreased expression/accelerated turnover rate of KMP-11 could lead to an increased export of Sb(III) due to an alteration of putative efflux systems\textsuperscript{57} or transporters, such as the aquaglyceroporin\textsuperscript{58}.
Identifying immunogenic proteins

Increasing resistance of parasites to available drugs highlights the need for developing a vaccine. While 1st generation whole cell vaccines have largely failed to show efficacy it is believed that subunit vaccines based on selected protein antigens can be developed. Therefore, it is important to identify new candidate antigens and the published proteomic data sets supposedly hold a valuable set of candidates. One good attribute for a vaccine is high abundance of the respective protein in the pathogenically important life cycle stage, the amastigote. It is also important to understand which leishmanial proteins are immunogenic and to what extent.

Gupta et al. analysed a fraction of soluble non-membranous *L. donovani* lysate of metacyclics which induced strong T cell responses from peripheral blood cells of individuals recovered from visceral leishmaniasis. The proteome of the soluble non-membranous lysate was resolved by 2DE and protein spots associated with the T-cell responses of 63 – 97.4 kDa and a pI range of 3-10 were excised and mass spectrometrically analysed. In total, 63 protein spots were excised and 52 protein species could be identified originating from 33 different genes among them were six encoding hypothetical proteins. Among the remaining ones there were for example heat shock proteins, reported to induce Th-1 cytokines in peripheral blood mononuclear cells of mucocutaneous and cutaneous leishmaniasis of infected patients, several glycolytic enzymes, protein disulfide-isomerase, calreticulin and kinesin-like protein (K39). In general proteins in this molecular weight range belong to the following functional classes: stress response, cytoskeleton assembly, energy metabolism, cell cycle control / proliferation and amino acid metabolism.

Immunogenic proteins of *L. infantum* promastigotes of mid-logarithmic growth phase were investigated by Dea-Ayuela and colleagues. To guide identification, 72 randomly chosen landmark protein spots were analysed by mass spectrometry of which 29 could be identified. These belonged to the following functional groups 1) carbohydrate metabolism, 2) proteolysis, 3) amino acid metabolism (i.e. succinyl-diaminopimelate desuccinylase), 4) fatty acid metabolism, 5) structural proteins, 6) stress response, 7) five hypothetical proteins and 8) other function (for example adrenodoxin reductase involved in the ergosterol biosynthesis). Immuno-blots were developed with anti-*L. infantum* rabbit hyperimmune sera and more than 50 protein spots between 90 and 30 kDa were detected. Comparison with silver stained gels revealed some of the immunogenic spots as propionyl-CoA-carboxylase, ATPase beta subunit, transketolase, a proteasome subunit, succinyl-diaminopimelate desuccinylase, a probable tubulin alpha chain, full length hsp70 and hsp70 related protein 1 as well as several proteins of unknown function.

Walden and colleagues identified *L. donovani* antigens recognised by serum antibodies of patients suffering from visceral leishmaniasis. Lysates of *in vitro* grown *L. donovani* promastigotes originally isolated from patients were separated by 2DE and then analyzed by protein immuno-blot.
stained 2D gels with a pH range of 4.5 – 7. 1067 protein spots were resolved and some 330 antigenic spots could be assigned in corresponding immuno-blots. Of these 330 antigenic spots 68 could be unequivocally matched to the 2D gels. Six protein spots could be identified by mass-spectrometry hsp70 (in two spots), gp63, eIF-4A, eEF2 and glucose regulated protein 78. Of note, the authors observed a high diversity in the antibody responses of the individual patients which prompted them to propose a recognition hierarchy for these antigens.

A protective immune response depends on the activation of leishmanicidal mechanisms in parasitized host cells. Reactive nitric oxide (NO) intermediates are thought to provide a main effector mechanism but its molecular targets remain to be identified. Thus, through proteomics NO effects, i.e. the effect of the NO donor 3-morpholinosydnonimine (SIN-1) on L. infantum promastigotes was investigated. Exposure of logarithmically growing promastigotes to 100 pM SIN-1 for 6 h killed 60-80% of the parasites. Comparative proteomics of the so treated parasites and untreated controls led to the identification of 22 proteins. Upon SIN-1 exposure enolase, LmjF36.4910 a hypothetical protein with unknown function, 20S proteasome beta 7 subunit and P28 protein became more abundant while hsp70, LmjF36.6360 hypothetical predicted protein, 20S proteasome alpha 5 unit and phosphomannomutase were decreased. The expression of 3-hydroxy-3-methyglutaryl-CoA synthase, peroxidoxin precursors and prostaglandine F2α synthase was induced. Interestingly the infectivity to J774 macrophages of promastigotes which survived the NO exposure was significantly increased.

**Secreted Proteins by Leishmania spp.**

*Leishmania* parasites have been shown to secrete factors that are relevant for nutrient acquisition, for interaction with their insect vectors and their vertebrate hosts. Thus, there is considerable interest in a comprehensive description of secreted proteins, an obvious application for proteomics. Silverman and co-workers studied the secretome of *L. donovani* promastigotes by using the stable isotope labelling of amino acids in culture (SILAC) methodology. Labelled stationary phase promastigotes were cultured in a serum free medium for 4 – 6 h and spent supernatant of the cultures was quantitatively compared to unlabelled proteins prepared from the pelleted cells. In total, 358 proteins were identified and based on the quantitative SILAC analysis 151 proteins were regarded as actively secreted. For most of these proteins the respective genes did not encode secretion signals and the authors proposed that bulk protein secretion in *Leishmania* spp. is via microvesicles, akin to exosomes, a suggestion that was further supported by electron microscopical analyses that identified such vesicles in the vicinity of the flagellar pocket. This concept is attractive given the implications that exosomes may have for the infection process but verification of it remains technically very challenging.
Cuervo and co-workers used 2DE to identify secreted proteins of *L. braziliensis* promastigotes\(^5\). Parasites were incubated for 3 h in serum free medium and the proteins in the supernatant of spent cultures were precipitated and subsequently subjected to 2DE. In Coomassie blue stained gels approximately 270 protein spots were detected of which 42 were identified originating from 35 genes. Bioinformatic analyses using SignalP3.0, SecretomeP2.0 and TargetP1.1, determined the presence or absence of signal peptides or predicted secretion via alternative pathways. Thereby 24 proteins appeared secreted via the classical (2 of 24) or non classical pathway (22 of 24). The absence of transmembrane domains was also analysed by predictive algorithms in these 42 identified protein species. The authors concluded that there is also an unconventional bacterial secretion pathway present in *Leishmania*, especially a Gram negative-like secretion pathway. The putative secreted proteins belonged to the following functional classes: stress response/chaperones, carbohydrate metabolism, protein synthesis, amino acid metabolism, proteolysis, antioxidants and others (for example dihydroorotate dehydrogenase, histidine secretory acid phosphatase, IgE-dependent histamine-releasing factor and LACK)\(^6\).

Paape *et al.* carried out a bioinformatic analysis assuming that protein secretion would require a classical N-terminal signal peptide and no transmembrane nor a GPI-anchor. In their proteome data set obtained from intracellular *L. mexicana* amastigotes, this approach predicted 143 secreted proteins\(^2\). It will undoubtedly require more work to investigate further the relative contributions of secretion signal mediated protein release versus exocytosis mechanisms but proteomics will help in the identification.

**Further developments**

Shortcomings of even the most comprehensive proteomic datasets produced to date for *Leishmania* spp. are a dramatic under representation of membrane proteins and, in comparison to transcriptomic data, a lack in sensitivity. Both aspects are recognised and several groups have investigated solutions.

**Prefractionation of parasite lysates to increase proteome coverage**

The disadvantage of separation highly complex mixtures by 2DE is that only abundant proteins will be detected as less abundant proteins may be masked or not detected and their identity remains elusive.

McNicoll and colleagues addressed this by analysing the proteomes of *L. infantum* promastigotes and amastigotes after reducing sample complexity by differential ammonium sulphate precipitations of the total lysate\(^6\). In total six fractions were obtained by collecting material precipitating at 0 – 40%, 40 – 50%, 50 – 60%, 60 – 70%, 70 – 80% and >80% ammonium sulphate concentration. The
desalted fractions were then separated by SDS-PAGE revealing distinct protein bands, with the bulk of proteins in the fraction from 40 – 80% ammonium sulphate that was then analysed be 2DE. The procedure proved highly reproducible and a total of 2261 and 2273 protein spots could be resolved for protein fractions generated from promastigotes and amastigotes respectively. Selected spots identified 168 protein species that originated from 71 different genes. Although fractionation clearly enriched for different proteins its full benefit remains to be shown.

To reduce complexity and enable identification of subproteomes, Foucher and colleagues used differential solubilization of membranes by digitonin for pre-fractionation. Digitonin is a non-ionic detergent which solubilizes sterols and at lower concentrations forms pores in the membrane. Depending on the concentration and accessibility as well as sterol composition of the membrane, different membrane types are solubilized. Further, digitonin even at higher concentration has no effect on the cytoskeleton. This characteristic was exploited by the authors. *Leishmania infantum* promastigotes and axenic amastigotes were treated with increasing concentrations (100 µM, 500 µM and 5 mM). The respective digitonin protein extracts as well as the whole cell extract and the non-digitonin solubilized material were separated by 2DE and resulted in 1921 protein spots for the whole cell extract. The serial digotonin extracts resolved 983, 883, 286 and 1086 protein spots for 100 µM, 500 µM, 5 mM and the non-solubilised proteins, respectively, and resulted in largely different protein spot pattern with little overlap. In total, 153 protein species were identified of which 47 were hypothetical. The subcellular localisation of identified protein species extracted with 100 µM digitonin was mainly cytoplasmic and proteins known to be found in the cytoplasm or within organelles. For the 500 µM digitonin extract the majority and for 5 mM all of the identified protein species were predicted to reside within organelles. The digitonin non-solubilized fraction was comprised mainly of protein species constituents of or associated with the cytoskeleton. The depletion of tubulins in the digitonin soluble material was also confirmed by Western analysis, it was virtually absent in the first two digitonin fractions but was detectable in the third fraction to a small degree compared to the non-soluble material. This differential lysis protocol of *Leishmania* spp. and its membranes thus may yield directly information about the likely subcellular localization of proteins.

Further developments included the use of pre-fractionation by liquid phase iso-electric focusing. *L. amazonensis* promastigotes lysates were prepared using this technique in order to increase the protein-spots of 2DE resolved proteomes and more protein species were resolved compared to only 2DE resolved proteomes.
**Improvement in electrophoresis techniques**

Brotherton and colleagues used free-flow electrophoresis (FFE) to increase proteome coverage of basic protein species of *L. infantum* promastigotes and axenic amastigotes. The lysates were subjected to FFE and collected fractions with a basic pH were combined into two pools pH7.0 – 9.0 and pH>9.0 and separated by 2DE. The pH 7.0-9.0 pool resulted in 1813 and 1823 resolved protein spots for promastigotes and amastigotes, respectively, with four unique and 14 upregulated in the extracellular stage and 15 unique and 33 upregulated in the intracellular stage. In the pooled fractions with a pH>9.0 623 and 646 protein spots were resolved for promastigotes and amastigotes, respectively, with seven unique and eight upregulated in the promastigote stage and five unique and three upregulated in amastigotes. Of these differentially expressed protein spots 18 from promastigotes and 25 from amastigotes were analysed by mass spectrometry and resulted in the identification of 87 and 63 protein species for pro- and amastigotes, respectively. It was also observed at proteome level that the glycolytic enzymes were upregulated in promastigotes whereas amastigotes express gluconeogenic enzymes and enzymes involved in β-oxidation at a higher level as previously reported. The authors reported also flagellar proteins to be over expressed in the promastigote stage which is probably not surprising.

**Gel free versus gel based proteomics**

Advantage of 2DE is the visualization of “break down” products as well as “multi spot trails” due to post translational modification. Lifecycle specific proteolytic products for α- and/or β-tubulin in amastigotes, eEF1α, hsp83 and Glucose 6 phosphate dehydrogenase, 6-1-pyruvate-5-carboxylate dehydrogenase, putative carboxypeptidase, elF4a and hypothetical protein LmjF08.1270 and for promastigotes putative proliferative cell nuclear antigen were reported. Stage specific break down products of various heat shock proteins were also observed. Proteasome α1 subunit was found in distinct protein spots for amastigotes and promastigotes at the same apparent pI but slightly different apparent molecular weights. A similar observation was reported for several different enolase isoforms some appeared to be promastigotes and some amastigote specific. Apparent differential post translational modifications were also reported for tryparedoxin peroxidase and triosephosphate isomerase.

Thus, while gel free methods are advantageous in the detection of proteins of lower and lowest abundance compared to gel based studies the latter still offer the advantage of being able to pre-select protein species for analysis and visualize differences at the protein species level due to post translational modifications that affect either size or charge of proteins.
In total 31-35% of the *Leishmania* spp. proteome predicted from the genome has been identified to date. Although this is a respectable number but the datasets still reflect mostly abundant proteins. In order to identify the other end of the spectrum, proteins of lower to lowest abundance, better prefractionation or selective depletion of (highly) abundant proteins are necessary as well as improved mass spectrometry techniques and methods, but will undoubtedly be available in the near future. Improvements seem also on the horizon to identify transmembrane proteins which are under-represented in the current data sets more readily.

Metabolic pathways such as $\beta$-oxidation of fatty acids were already known to be upregulated in the intracellular life cycle stage from biochemical analyses two decades ago but additional aspects of these pathways as revealed by proteomics will highlight less obvious facettes and will guide systems biology and help in the interpretation of future metabolomic data. Metabolomics on the other hand will support future proteomic studies and as an outcome novel drug targets are likely to be identified. Proteomic studies will also be useful to study further drug resistance in the field as it has already been shown to identify proteins.

As a field, it seems that a common standard to deposit primary data in public repositories or the umbrella of the EuPathDB initiative is wishful to enable direct interrogation of e.g. spectra by all interested researchers. In addition, defining an agreed standard to normalize protein levels e.g. between life cycle stages would also be helpful facilitating direct comparisons between studies.
References


Figure Legend:

Figure 1. Timeline summarizing publications of *Leishmania* spp. proteomics or of publications therefore relevant. *L.-* *Leishmania*, *ax.-* axenic, *promast. –* promastiogtes, *amast. –* amastigotes