Northumbria Research Link

1   **Extensive differential protein phosphorylation as intraerythrocytic _Plasmodium_**

2   **_falciparum_ schizonts develop into extracellular invasive merozoites**

3

4

5   Edwin Lasonder[1,*], Judith L. Green[2], Munira Grainger[2], Gordon Langsley[3,4,*], Anthony A.

6   Holder[2,*]

7

8

9

10   [1]School of Biomedical and Healthcare Sciences, A405 Portland Square, Plymouth

11   University, Drake Circus, Plymouth, Devon PL4 8AA, United Kingdom.

12   [2]Division of Parasitology, MRC National Institute for Medical Research, The Ridgeway,

13   Mill Hill, London, NW7 1AA, United Kingdom.

14   [3]Laboratoire de Biologie Cellulaire Comparative des Apicomplexes, Faculté de

15   Médicine, Université Paris Descartes - Sorbonne Paris Cité, France.

16   [4]Inserm U1016, CNRS UMR8104, Institut Cochin, Paris, 75014 France.

17

18   *Corresponding authors

19

Abstract

Pathology of the most lethal form of malaria is caused by *Plasmodium falciparum* asexual blood stages and initiated by merozoite invasion of erythrocytes. We present a phosphoproteome analysis of extracellular merozoites revealing 1765 unique phosphorylation sites including 785 sites not previously detected in schizonts. The observed differential phosphorylation between extra and intraerythrocytic life cycle stages was confirmed using both phospho-site and phospho-motif specific antibodies and is consistent with the core motif [K/R]xx[pS/pT] being highly represented in merozoite phosphoproteins. Comparative bioinformatic analyses highlighted protein sets and pathways with established roles in invasion. Within the merozoite phosphoprotein interaction network a sub-network of 119 proteins with potential roles in cellular movement and invasion was identified and suggested that it is co-regulated by a further small sub-network of protein kinase A (PKA), two calcium-dependent protein kinases (CDPKs), a phosphatidyl inositol kinase (PI3K) and a GCN2-like eIF2-kinase with a predicted role in translational arrest and associated changes in the ubiquitinome. To test this notion experimentally, we examined the overall ubiquitination level in intracellular schizonts versus extracellular merozoites and found it highly upregulated in merozoites. We propose that alterations in the phosphoproteome and ubiquitinome reflect a starvation-induced translational arrest as intracellular schizonts transform into extracellular merozoites.

1. Introduction

The most lethal form of human malaria is caused by *Plasmodium falciparum*. This parasite has a complex life cycle in both mosquito and human host, where asexual multiplication and development within red blood cells (RBC) is responsible for disease pathology. Intra-RBC parasite development is initiated by merozoite invasion of erythrocytes, as merozoites are first released into the blood stream from merosomes in the liver [1] and subsequently released at the end of each cycle of multiplication in RBC. Preventing blood stage infection by targeting merozoites is an attractive intervention strategy to develop treatments to alleviate disease. Therefore, understanding merozoite development and the identification of new anti-malarial therapeutic targets are of the utmost importance [2].

The merozoite is well adapted for RBC invasion [3]. It has the pellicular structure typical of members of the Apicomplexa phylum including an apical end with secretory organelles: micronemes, rhoptries and dense granules [4]. RBC invasion is a multi-step process that is initiated by merozoite attachment, followed by reorientation and sequential discharge of the contents of the apical organelles [5, 6]. During erythrocyte entry, a process that is controlled in part by calcium and cAMP fluxes [7, 8], a moving junction is formed between the parasite and host cell surfaces, and parasite proteins are transferred to the newly invaded erythrocyte [8-10]. After invasion the parasite resides within a parasitophorous vacuole and develops from a 'ring' to a trophozoite that degrades host cell haemoglobin and performs DNA synthesis. The onset of mitosis produces the multinucleated schizont, and following cell division individual merozoites

65   are formed, which are then released at the end of the cycle to infect new RBC [11].

66   Within erythrocytes the primary amino acid nutrient source for the developing parasite is

67   haemoglobin, which is degraded in the food vacuole. This structure is elaborated anew

68   following invasion [12] and haemoglobin digestion starts only several hours post-

69   invasion [13]. In late stage multinucleated "segmenter" schizonts that have undergone

70   cytokinesis, the food vacuole is 'pinched off' from the developing merozoites. Therefore,

71   merozoites and post-invasion early ring forms are deprived of haemoglobin-derived

72   amino acids for several hours. However, organelle degradation by autophagy likely

73   provides the developing merozoite and early ring stages with some nutrient [14].

74

75   Protein kinases, phosphatases and signal transduction pathways are integral to

76   regulation of the parasite life cycle; for example, both merozoite egress from the host

77   cell and erythrocyte invasion are governed by protein phosphorylation [15, 16].

78   Unravelling protein phosphorylation and signalling pathways in merozoites requires

79   large-scale phosphoproteome studies that are now feasible using mass spectrometry.

80   Significant efforts have been made in recent years to generate phosphoproteome data

81   principally for schizonts, leading to the identification of more than 12,000 unique protein

82   phosphorylation sites [17-22]. Our previous large scale phosphoproteome analysis of *P.*

83   *falciparum* schizonts by liquid chromatography tandem mass spectrometry (LC-MS/MS)

84   revealed extensive phosphatidylinositol and cAMP-dependent protein kinase (PKA)

85   signalling and identified three novel PKA substrates associated with the glideosome

86   motor complex that is implicated in driving the parasite into the RBC. These data

87   support a role for cAMP as an important regulator of host cell invasion [7, 18].

88   In this study we present a large-scale analysis of protein phosphorylation in extracellular

89   merozoites and its comparison with established schizont phosphoproteomes to deepen

90   our understanding of merozoite biology. We use bioinformatics to analyse

91   phosphorylation sites and establish a merozoite phosphoprotein interaction network.

92   Reassuringly this revealed a sub-network of merozoite proteins with predicted roles in

93   cellular movement and invasion. Interestingly, it also revealed a smaller sub-network of

94   five kinases known in other cellular systems to regulate starvation-induced translational

95   arrest and autophagy, both of which involve protein ubiquitination [23-26]. We confirm

96   global changes in phosphorylation patterns between schizonts and merozoites using

97   both commercial phospho-site specific antibodies and a specific phosphoprotein

98   antibody. Importantly, we show that differential phosphorylation is linked to highly

99   elevated ubiquitination in merozoites and we propose that alterations in the

100   phosphoproteome and ubiquitinome reflect starvation-induced translational arrest in

101   merozoites.

102

103

2. Materials and Methods

2.1 Parasites.

P falciparum 3D7 line parasites were cultured *in vitro* as described previously [27]. Parasite populations were synchronised by treatment with sorbitol, Percoll gradient centrifugation and passage through a magnetic column, allowing the purification of merozoites as described [27, 28].

2.2 In solution protease digestion and peptide purification using the Filter Assisted Sample Preparation (FASP) method.

Approximately 2-3 x10^9 merozoites were lysed in 1 ml 2 % SDS in 50 mM Tris-HCl pH 7.3 in the presence of phosphatase and protease inhibitors. Halt™ phosphatase inhibitor cocktail (Thermo Fisher) and complete protease inhibitor cocktail (Roche) were used to suppress phosphatase and protease enzymatic activities. The Halt™ phosphatase cocktail is a concentrated solution of sodium fluoride, sodium orthovanadate, sodium pyrophosphate and β-glycerophosphate, and was diluted 100-fold for the final concentration. The complete protease inhibitor cocktail tablet contains AEBSF, aprotinin, bestatin, E-64, leupeptin, pepstatin A and EDTA, and was diluted according to the specifications of the manufacturer. The merozoite lysate was heated to 95ºC for 5 min and further processed according to the FASP protocol [29]. The lysate was divided and transferred to 12 0.5 ml Amicon Ultra 30K spin filter units (30 kDa cut-off, Millipore), each of which was washed with 0.4 ml 8 M urea in 100 mM Tris-HCl pH 8.0 resulting in a final composition of 0.34% SDS and 5.8 M urea, then reduced with 10

127 mM DTT at room temperature (RT) for 25 min, concentrated by centrifugation at 14,000

128 *g* and alkylated in the dark with 50 mM iodoacetamide in 100 mM Tris-HCl pH 8.0

129 containing 8 M urea at RT for 25 min. The samples were concentrated by centrifugation

130 and diluted 10-fold with 100 mM Tris-HCl, 8 M urea, pH 8.0. This step was repeated 6

131 times to fully remove SDS from the lysate solution. Then, 2.5 µg lysC protease (Wako)

132 per sample was added in 100 mM Tris-HCl, 4 M urea, pH 8.0, and incubated overnight

133 at room temperature. Samples were further diluted with 50 mM ammonium bicarbonate

134 to a final concentration of 2 M urea and then incubated with 5 µg trypsin (Promega) for

135 18 h. The tryptic digest was acidified in 0.1 % (vol/vol) trifluoroacetic acid (TFA) and the

136 peptides purified and desalted with C18 Hypersil 500 mg SPE columns (Thermo

137 Fisher). Resultant peptides were separated by strong anion exchange (SAX) membrane

138 filters into six fractions on the basis of their isoelectric points. Peptides were eluted

139 stepwise with Britton and Robinson buffer solutions of decreasing pH: 8.0, 6.0, 5.0, and

140 4.0, respectively, followed by a final elution with 1 % TFA. All peptide fractions were

141 acidified immediately with 0.1 % TFA and then the peptides were purified again using

142 C18 Hypersil 500 mg SPE columns.

143

144 **2.3 Phosphopeptide enrichment**.

145 Peptides obtained from 2-3 x 10^9 purified merozoites, corresponding to 3 mg protein as

146 determined by a micro BCA protein assay kit (Thermo Fisher) were incubated with

147 Titansphere 10 µm TiO_2 beads (GL Sciences, Inc., Japan) to selectively purify

148 phosphorylated peptides, as described for the schizont phosphoproteome analysis [18].

149 Prior to incubation the TiO_2 beads were washed with 80% (vol/vol) acetonitrile, 0.1%

150 TFA, re-suspended in 30 mg/ml dihydroxybenzoic acid in 80% acetonitrile, 0.1% TFA

151 and diluted 1:4 with 0.1% TFA. Twenty microlitres of the slurry containing about 1 mg

152 beads were added to peptide solutions in 1.5 ml tubes and incubated under continuous

153 shaking for 1-2 h. Afterwards, the bead slurry was transferred to a micro spin column

154 made from a 200 µl precision pipette tip with an inserted 2mm fused silica frit (50 µm

155 internal diameter). Unbound fractions were collected following centrifugation of the spin

156 column and reincubated with freshly prepared TiO_2 bead slurry for 1-2 h. This step was

157 performed three times leading to four micro spin columns with TiO_2 beads per peptide

158 SAX fraction. The columns were washed three times by centrifugation with 100 µl 30%

159 acetonitrile, 3% TFA, followed by three washes with 100 µl 80% acetonitrile, 0.3% TFA.

160 Bound phosphopeptides were eluted from the beads with 100 µl 5 % NH_4OH, 5 %

161 piperidine, and 5 % pyrrolidine [30]. Peptide solutions were then acidified with TFA and

162 the peptides purified with STAGE tips [31] and dissolved in 0.1% TFA and 10 mM

163 EDTA.

164

165 **2.4 Liquid chromatography tandem mass spectrometry (LC-MS/MS).**

166 Tandem mass spectrometry experiments were performed with the linear ion trap

167 cyclotron resonance Fourier transform (LTQ-Ultra FT) mass spectrometer (Thermo

168 Fisher, Bremen, Germany) coupled to the nano EASY LC system (Proxeon, Denmark)

169 with 15 cm 100 µm internal diameter PicoTip columns (New Objective, Woburn, USA)

170 packed with 3 µm Reprosil C18 beads (Dr. Maisch GmbH, Ammerbuch, Germany).

171 Peptides were separated by liquid chromatography using a gradient from 92% buffer A

172 /8% buffer B to 73% buffer A /27% buffer B (where buffer A is 0.5% acetic acid in water

173 and buffer B is 0.5% acetic acid in acetonitrile) with a flow-rate of 300 nl/min over 90

174 min. A voltage of 2.2 kV was applied for electrospray ionisation. Data–dependent

175 acquisition was performed for switching automatically between MS, MS2 and

176 phosphoric acid neutral loss triggered MS3 scan modes. Full-scan MS spectra of intact

177 peptides (m/z 350–1500) with an automated gain control accumulation target value of

178 1,000,0000 ions were acquired in the Fourier transform ion cyclotron resonance cell

179 with a resolution of 50,000. The four most abundant ions were sequentially isolated and

180 fragmented in the linear ion trap by applying collision induced dissociation using an

181 accumulation target value of 10,000, a capillary temperature of 100°C, and a normalized

182 collision energy of 27%. Multi-stage activation was switched on for neutral loss

183 dependent MS3 fragmentation on the masses of phosphoric acid at charge states 2+,

184 3+ and 4+. The four most abundant ions were sequentially fragmented under identical

185 settings as for MSMS mode with a normalized collision energy of 40 %. A dynamic

186 exclusion of ions previously sequenced within 180 s was applied. All unassigned charge

187 states and singly charged ions were excluded from fragmentation. Sequencing

188 thresholds were set at 500 counts for MS2 and 5 counts for MS3. The mass

189 spectrometry data have been submitted to the ProteomeXchange Consortium

190 (http://proteomecentral.proteomexchange.org).

191

192

2.5 Mascot Peptide identification and MaxQuant validation.

Tandem mass spectrometry data were processed using the same procedure we applied previously to the schizont phosphoproteome [18]. Briefly, Mascot generic peak lists were generated by MaxQuant version 1.013.13 (http://maxquant.org//) [32] and submitted to Mascot version 2.2 (Matrix Science) to search the *P. falciparum* database downloaded from http://plasmodb.org/plasmo/ and supplemented with the human International Protein Index (IPI) database (http://www.ebi.ac.uk/IPI) and frequently observed contaminants and concatenated with reversed copies of all entries. The following search parameters were applied: peptide mass 10 ppm, MSMS mass accuracy of 0.5 Da; enzyme cleavage: trypsin allowing 2 miscleavage sites; fixed modification for cysteines by carboxyamidomethylation and variable modifications to enable detection of phosphorylation at serine, threonine and tyrosine; oxidation of methionine; deamidation of glutamine and asparagine and protein N-terminal acetylation. Mascot search results were processed further by MaxQuant [32], where peptides were filtered; requiring a minimal Mascot peptide score of 20 in combination with the probability that the identification is wrong, the posterior error probability (PEP) of 0.025, a minimal charge state of 2 and no more than two variable modifications per peptide. Peptides and proteins with a false discovery rate (FDR) better than l % were accepted. The phosphorylation site localization probability >0.75 was applied to a obtain phosphopeptide FDR of 1 % [33].

2.6 Gene ontology analysis.

All *P. falciparum* gene ontology analyses were performed with the software package Ontologizer [34] (http://compbio.charite.de/index.php/ontologizer2.html), with the following Open Biological Ontology and Gene association components from http://www.geneontology.org: gene ontology v1.2.obo, goslim_generic.obo and Gene_association.GeneDB–Pfalciparum_2011-5-31. The OPI GO terms were taken from and rearranged to a gene association compatible file. Ontologizer was used to identify overrepresented GO terms for the phosphoproteome relative to the background of the *P. falciparum* proteome (~5500 proteins). GO term enrichment was computed by the parent-child union approach and corrected for multiple testing by the Benjamini and Hochberg method, and was considered significant for adjusted p-values lower than 0.05 [34].

2.7 Motif analysis.

Phosphorylation sites were categorized by their chemical properties as acidic, basic, proline-directed, tyrosine or other by a decision tree method from [18, 35] as follows: 1/ Get the 6 neighbouring amino acids before and after the phosphorylation site. 2/ pY at position 0 then classify as "Tyrosine." 3/ P at +1 then classify as "Proline-directed." 4/ positions +1 to +6 contain more than one D and E residues then classify as "Acidic." 5/ K or R at position -3 then classify as "Basic." 6/ D or E at +1, +2, or +3 then classify as "Acidic." 7/ between -6 and -1 more than 2 K or R residues then classify as "Basic." 8/ remaining peptides classify as "Other".

237   Merozoite phosphorylation motifs were identified using MotifX [36] that tested for motif

238   overrepresentation in phosphorylated peptides with the following parameters:

239   phosphorylation motif window = 13 amino acids, P-value threshold =$1*10^{-4}$ for S and T

240   residues, $1*10^{-3}$ for Y residues, motif fold increase ≥ 2, a motif frequency > 5, and a

241   background of all *P. falciparum* proteins. MotifX analysis was performed for a normal

242   and a degenerate amino acid set. The degenerate amino acid set was enabled for

243   conservative amino acid substitutions within the motif window according to: A=AG,

244   D=DE, F=FY, K=KR, I=ILVM, Q=QN, S=ST, C=C, H=H, P=P, W=W. When different

245   motifs were found for a peptide by the analyses with different amino acid residues,

246   priority was given to the motif with the highest MotifX score. Sequence logos were

247   generated with Weblogo 3 at http://weblogo.threeplusone.com/create.cgi. The motifs

248   were matched to known protein kinase target motifs using CompariMotif [37] at

249   http://bioware.ucd.ie/~compass/biowareweb/Server_pages/comparimotif.php, and

250   matches with the highest scores were considered as potential links between

251   phosphorylation motifs and protein kinases.

252

253   **2.8 Phosphorylated protein interaction network analysis**.

254   The merozoite phosphoproteome interaction network was constructed from all *P.*

255   *falciparum* protein-protein interaction data with a minimum confidence level of 0.15

256   downloaded from the STRING database version 9.0 [38] that were matched with the

257   phosphorylated *P. falciparum* proteins identified in this study and visualized in

258   Cytoscape version 2.8.3 [39]. The merozoite phospho-interactome was analyzed for

259 highly connected nodes with the molecular complex detection (MCODE) clustering

260 algorithm [40] that was available as a Cytoscape plug-in using default parameters.

261

262 **2.9 Phospho-site, Phospho-motif and Protein Ubiquitin Western blots**

263 Antibodies to the peptide PQRKPL*SIEESF based on the sequence of amino acid

264 residues 41 to 52 of myosin tail domain interacting protein (MTIP) and containing a

265 phosphoserine corresponding to Ser47 [41] were prepared in a sheep and purified by

266 affinity chromatography on the phosphopeptide coupled to a solid support (University of

267 Dundee). Late schizonts and merozoites were purified using a magnet and lysed in 30

268 cell-pellet volumes of NP40 lysis buffer (0.5% Nonidet P40, 150 mM NaCl, 50 mM Tris-

269 HCl pH 8.0) containing protease inhibitors (Roche). Following incubation on ice for 20

270 min, the soluble protein solution was collected after centrifugation at 15,000 g for 20 min

271 at 4ºC. The merozoite lysate was either untreated or treated with Lambda phosphatase

272 (New England Biolabs), using five units phosphatase/µg protein and incubation for 20

273 min at 30ºC, then the enzyme was heat inactivated at 65ºC for 10 min. Lysates

274 containing approximately 10 µg protein were resolved on 10% NuPAGE Bis-Tris gels in

275 MOPS buffer (Life Technologies) under reducing conditions and the proteins transferred

276 to nitrocellulose using an iBlot system (Life Technologies). The blot was blocked

277 overnight in phosphate buffered saline containing 1% (wt/vol) BSA and 0.2% (vol/vol)

278 Tween-20 (PBST). For western blotting the sheep anti-phosphopeptide antibodies (1

279 µg/ml) were pre-incubated with 10 µg/ml non-phosphorylated peptide for 1 h at room

280 temperature to deplete antibodies not specific for the phosphoserine and then incubated

281 with the blot. MTIP was also detected with a rabbit anti-MTIP specific antibody that has

282 been described previously [5]. All antibody dilutions were made in PBST; washes were

283 with the same buffer. After extensive washing, the appropriate horseradish peroxidase

284 (HRP)-conjugated anti-IgG secondary antibodies (Sigma) were used to detect bound

285 antibody using enhanced chemiluminescence (GE Healthcare) and fluorography.

286 Antibodies recognizing phospho-PKA (motif 3), phospho-PKB/Akt (motif 4), phospho-

287 PKD (motif 11) and phospho-tyrosine motifs were obtained from Cell Signaling

288 Technologies. Schizont and merozoite cell pellets were resuspended in isotonic 0.15%

289 saponin to lyse red blood cell membranes. Parasites were pelleted by centrifugation at

290 2400 g and then schizonts were resuspended in 10 volumes and merozoites in 50

291 volumes of lysis buffer (0.5% NP40, 150 mM NaCl, 0.5 mM EDTA, 10 mM Tris HCl pH

292 8.0). Protein concentrations were determined using a detergent-compatible protein

293 assay (DC-protein assay, Biorad), and 10 µg of each sample were resolved on 10%

294 polyacrylamide gels as described above. Following protein transfer to nitrocellulose and

295 blocking overnight, antibodies diluted 1:1000 were added and incubated with the blot for

296 1 h at room temperature. Bound antibodies were detected with HRP-linked species-

297 specific secondary antibodies (Biorad) at a dilution of 1:5000, as described above.

298 To analyse ubiquitination in parasites, a time course of late stage schizonts was

299 prepared. Briefly, parasites were synchronized to a 1 h window and harvested at

300 appropriate time points during the cycle. After initial lysis of the RBC membrane using

301 0.15% saponin followed by centrifugation of the sample at 2400 g, parasites were lysed

302 in a buffer containing 4% CHAPS, 1% DTT, 6 M urea, and 2 M thiourea. Samples were

303 subjected to three freeze-thaw cycles, and then centrifuged at 100,000 g. The

304 concentration of protein in the supernatant was determined using the Biorad Protein

305   Assay, and 2 µg samples of each protein fraction were resolved on a 12% NuPAGE

306   Bis-Tris gel in MOPS buffer. Following transfer to nitrocellulose and blocking as

307   described above, an anti mono- and polyubiquitinylated conjugate monoclonal antibody

308   directly conjugated to HRP (FK2H, Enzo Life Sciences) was used at 1000-fold dilution

309   and detected using ECL. CDPK1 was detected using a specific rabbit polyclonal

310   antibody as described previously [41].

311

312   **3. Results**

313

314   **3.1 The phosphoproteome of *P. falciparum* merozoites**

315   Free merozoites were purified after release from highly synchronous populations of

316   schizonts [27, 28], and processed for mass spectrometric phosphoproteome analysis by

317   the gel-free protein digestion procedure, Filter Assisted Sample Preparation (FASP).

318   Merozoites were solubilised in 0.34 % SDS/ 5.8M urea and the proteins digested with

319   the proteolytic enzymes lysC and trypsin. Phosphopeptides were affinity-purified using

320   TiO_2 beads, and then subjected to liquid chromatography-tandem mass spectrometry

321   (LC-MS/MS) measurements. Peptides were identified by Mascot searches [42] of

322   MS/MS spectra against protein databases comprising all *P. falciparum* and human

323   proteins. Preliminary peptide identifications were validated in reverse database

324   searches using MaxQuant [32], resulting in identification of 1765 *P. falciparum*

325   phosphopeptides with phosphorylation site localization [33] probabilities >0.75 and

326   representing 740 distinct proteins (Figure 1, Table S1). Peptide identification criteria

327   were: a Mascot score greater than 20, posterior error probability (PEP) less than 0.025,

328   and a phosphopeptide False Discovery Rate (FDR) less than 0.01. A small number of

329   phosphorylated human proteins were identified by these validation criteria: 24 proteins,

330   comprising 36 sites (data not shown). Almost all detected merozoite phosphopeptides

331   were mono-phosphorylated (99.6%), which is likely explained in part by the observed

332   bias in enrichment by TiO_2 beads for such peptides [43]. The phospho-amino acid

333   distribution of 80% pSer and 19.6% pThr resembled that of our previous analysis of

334   schizonts using similar methodology [18], with 84.4% and 13.2%, respectively. Tyrosine

335   phosphorylation was less frequent in merozoites (0.4 % of phospho-sites) than in

336   schizonts (2.4 %). The observed low frequency of tyrosine phosphorylation in

337   merozoites highlights our stringent validation criteria since *Plasmodium* parasites lack

338   tyrosine kinases. *Plasmodium* proteins are probably phosphorylated by human tyrosine

339   kinases, and we speculate that host tyrosine kinase activity on extracellular merozoites

340   is likely lower than during intracellular schizont development, where we found a

341   frequency of 2.4 %. A comparative analysis of our merozoite data with the combined

342   data representing 12,252 phosphorylation sites derived from schizonts [18-22] revealed

343   a similar distribution in this combined set, with frequencies of 81.8 % (pSer), 14.4 %

344   (pThr), and 3.8% (pTyr), respectively (Figure 1). The merozoite phosphoproteome

345   contained 785 phosphorylated sites (Figure 1) that were not found among the 12,252

346   phosphorylation sites reported for schizonts. Several sites in merozoite proteins

347   implicated in erythrocyte invasion were found to be phosphorylated in both schizonts

348   and merozoites such as S610 of AMA1, and others such as S1240 of Sub2 appear to

349   be specifically phosphorylated in schizonts. However phosphorylation at some sites in

350   such proteins, for example residues 1024, 1246 and 1261 in MSP1; 1299 in EBA181;

16

351   and 137, 139, and 296 in RAP1, appears to be specific to the merozoite fraction (Table

352   S1); the functional significance of this is unknown. The proteome data generated in this

353   study comes from a single biological experiment measuring purified merozoites and

354   contains phosphopeptides with not more than 1 % false identifications.

355

356   **3.2 Functional annotation of the merozoite phosphoproteome**

357   The potential relevance of the merozoite phosphoproteome to parasite invasion biology

358   was assessed by functional enrichment analyses of the 740 phosphorylated merozoite

359   proteins relative to the full *P. falciparum* proteome (~ 5500 proteins). Firstly, a

360   comparative GO enrichment analysis was performed on extra (merozoites)- versus

361   intra-erythrocytic (schizonts) life cycle stages to identify any GO terms that were

362   significantly altered for the merozoite phosphoproteome using an extended set of

363   putative Gene Ontology (GO) annotations obtained from Ontology-based Pattern

364   Identification (OPI) clustering [44]. GO terms enriched in the merozoite

365   phosphoproteome (this study) were compared with enriched terms in the ensemble of

366   schizont phosphoproteomes represented by 12,252 phosphorylation sites detected

367   previously [18-22]. Statistical relevance in GO term enrichment between merozoites and

368   schizonts was assessed by Fisher exact tests (Table S2). The GO terms more than two-

369   fold enriched in the merozoite- relative to schizont-phosphoproteomes are shown in

370   Figure 2. We note that the enriched GO terms 'protein secretion', 'secretion', and

371   'secretion by cell' were not found to be statistically relevant, and neither are GO terms

372   represented by less than five proteins in the merozoite phosphoproteome (*e.g.* 'ATPase

373 activity, coupled to transmembrane movement of ions' and 'rhoptry'), even though the

374 differences appear striking.

375 In the category 'Molecular Function' (red columns, Figure 2 and see Table S3) we found

376 differences between merozoites and schizonts for terms associated with

377 phosphorylation such as 'lipid kinase activity' ($p < 0.05$) and 'calmodulin-dependent

378 protein kinase activities' ($p < 0.1$), as well for 'protein degradation', 'threonine-type

379 peptidase' and 'endopeptidase activities' ($p < 0.05$). GO terms in the category 'cellular

380 components' (green columns) were enriched ($p < 0.1$) in invasion-related components

381 such as 'apical part of cell' and 'apical complex' and significantly enriched ($p < 0.05$) for

382 'ubiquitin-mediated degradation' and 'proteasome core complex'. Finally, 16 biological

383 process (blue columns) were identified as enhanced in the merozoite phosphoproteome

384 including 11 with statistical relevance ($p < 0.05$) and 1 with $p < 0.1$. These terms included

385 'cellular response to stress' and three terms related to 'movement and locomotion' that

386 were most enriched ($p < 0.05$) in the merozoite phosphoproteome.

387 A second functional enrichment analysis was carried out with manually curated

388 pathways from the Metabolic Malaria Pathway database (MMPD)

389 (http://priweb.cc.huji.ac.il/malaria/). Enrichment in the merozoite phosphoproteome was

390 compared with enrichment in the ensemble of schizont phosphoproteomes (Figure 3;

391 Table S3). More than two-fold difference in enrichment ($p < 0.05$) between merozoites

392 and schizonts was observed for 'histone acetylation', and 'phosphoinositides and

393 membrane traffic'. The first term suggests an epigenetic change in gene expression in

394 merozoites. We previously highlighted extensive phosphatidyl inositol signalling in

395 schizonts [18] and now with the identification of eleven phosphorylated proteins involved

18

396   in inositol phosphate metabolism the current analysis argues that phosphatidyl inositol

397   signalling is even more elevated in merozoites. This would likely strongly activate

398   phosphatidyl inositol 3-phosphate kinase (PI3K)/Vps34 [45]. Importantly, two terms

399   'Genes coding for components of the proteasome degradation machinery' and

400   'proteasome-mediated proteolysis of ubiquitinated proteins' were also significantly

401   enriched (p< 0.05).

402

403   **3.3 Protein interaction network of the merozoite phosphoproteome**

404   We demonstrated previously that specific phosphorylated protein complexes in *P.*

405   *falciparum* schizonts have the potential to regulate most of the biological activities at this

406   stage [18]. Therefore, the merozoite phosphoproteome was mapped onto predicted

407   protein-protein interactions from the STRING database [38] to construct a scale-free

408   protein-interaction network of phosphorylated merozoite proteins. STRING is a protein-

409   protein interaction database of known interactions (from high throughput experiments

410   and literature mining) and predicted interactions (from genomic context and co-

411   expression) currently covering approximately 1100 species including *P. falciparum*. The

412   constructed merozoite interactome is composed of 682 nodes/proteins covering 92% of

413   the phosphoproteome and 19,584 edges corresponding to protein-protein interactions.

414   The global interaction network was analysed for the presence of densely connected

415   sub-networks using the molecular complex detection (MCODE) clustering algorithm

416   [40]. This allowed the identification of 20 sub-networks (Table S4), with the majority

417   interconnected via the hub of 4 proteins present in MCODE 13 (Figure 4A). Putative

418   biological functions of the different MCODE clusters were assigned by MMPD

419 enrichment analyses, as described above. This resulted in potential functional

420 annotations for 12 MCODE clusters representing 50% (369 proteins) of the merozoite

421 phosphoproteome (Figure 4B). Their assignments are in good agreement with the

422 MMPD enrichment analysis (Figure S1), and in some cases reinforced the proposed

423 biological functions. For example, MCODE cluster 16 comprised of 4 phosphorylated

424 proteins involved in inositol phosphate metabolism is 34-fold more enriched in the

425 cluster than in the full merozoite phosphoproteome (Figure 4B). Similarly, the hub

426 cluster MCODE 13 is 30-fold more enriched and is composed of three kinases

427 (PF3D7_1428500, a putative GCN2; PF3D7_0515300 [PI3K]; and PF3D7_0610600,

428 calcium dependent protein kinase 2 [CDPK2]), and the regulatory subunit

429 (PF3D7_1223100) of cAMP-dependent protein kinase [PKA]).

430

431 MCODE cluster 1 was identified as the largest sub-network composed of 119 phospho-

432 proteins with potentially 5739 protein-protein interactions (Figure 4A). It is under-

433 represented in schizont phosphorylation sites ($p < 0.05$) (Table S4), indicating an

434 enhanced role for this cluster in merozoite-related processes. Consistently, 23 proteins

435 had MMPD terms such as 'components of the linear motor responsible for merozoite

436 motility in invasion', 'domains of merozoite surface proteins', 'functional annotation of

437 merozoite invasion-related proteins', 'subcellular localization of proteins involved in

438 invasion' and 'genes coding for GPI-anchored membrane proteins'. Interestingly, four

439 proteins were associated with the ATG autophagic pathway (Figure 4B). Involvement of

440 proteins in the MCODE 1 cluster with cell invasion was further supported by enrichment

441 in GO annotations for 'locomotion and movement' (9.2-fold) covering 49 proteins of

442 which 13 were annotated to 'the apical complex' (12.0-fold), pointing again to a role in

443 invasion (Table S5). The functional significance for 'locomotion and movement' of this

444 sub-network was further supported by large differences in fold enrichment between

445 MCODE cluster 1 (9.2-fold) and the merozoite phosphoproteome (2.0-fold) showing that

446 these putative functions are concentrated in the sub-network. The detection of highly

447 enriched protein kinase activities for PKA (11.9-fold) and CDPK (9.3-fold) suggests a

448 regulatory role for these kinases in controlling MCODE cluster 1 activity. MCODE1

449 activities are likely co-regulated (depicted with yellow circles) by MCODE13 with for

450 example, PKAr of MCODE13 regulating PKAc catalytic activity in MCODE1 (see Fig.

451 S1). Similarly, the autophagy pathway may be co-regulated by MCODE 13 via direct

452 interaction of CDPK5 and PKAr with ATG8.

453 The presence in MCODE1 of PKG (PF3D7_1436600), GCα (PF3D7_1138400) and

454 PKAc (PF3D7_0934800), together with CDPK1 (PF3D7_0217500) and CDPK5

455 (PF3D7_1337800) indicates that in merozoites there is significant crosstalk between the

456 secondary messenger cyclic nucleotides and calcium [7, 8]. We have noted previously

457 that in schizonts both CDPK1 and GAP45 (PF3D7_1222700) could be phosphorylated

458 *in vitro* by PKA [18]. The presence of the putative zinc finger transcription factor KROX1

459 (PF3D7_1209300) in MCODE1 now suggests that this crosstalk could have a

460 transcriptional outcome, as cAMP and Ca^{2+} surges do on activating the transcription

461 factor Creb in mammalian cells [46, 47].

462

463

3.4 Phosphorylation motif analysis.

The merozoite phosphoproteome data set described here together with an ensemble of the reported phosphoproteomes of schizonts [18-22] enabled us to compare schizont to merozoite phosphorylation motifs. Firstly, we compared phosphorylation motif classes based on their chemical properties using a decision tree previously applied to our schizont phosphoproteome [18] (Figure 5). The motif classes 'basic-directed' and 'other' are overrepresented in merozoites (fold-changes: 2.0 and 1.3, respectively) compared to schizonts, while acidic-, proline- and tyrosine–directed classes are underrepresented (fold-changes: -2.0, -1.5 and -7.3, respectively). Tyrosine phosphorylation is almost absent in merozoites with only 9 sites found, suggesting that tyrosine kinase activity may modify parasite proteins in schizonts, but not merozoites (see also Figure 6).

Secondly, phosphorylation sites identified in merozoites were classified using the phosphorylation motif finding algorithm MotifX [36]. We identified 33 phosphorylation motifs (Table S6) of which only five were reported in schizonts [18] [19, 21]. Within the top 10 most abundant motifs we find 5 variants of the core motif [K/R]xx[S/T] covering 32.4 % of all phosphopeptides detected in merozoites. The 33 motifs contain no acidic residues (Asp and Glu), in contrast to schizonts, where 14 motifs were rich in these residues [18]. When the 33 motifs were mapped to known kinase target sites using CompariMotif [37], several ambiguous matches were observed (Table S7). However, the core motif ([K/R]xx[S/T]) exactly matched to known calcium/calmodulin-dependent protein kinase II and PKA substrates, while imperfect matches were observed to other AGC kinases such as PKC, PKD, PKG and Akt/PKB.

486   A clear trend emerged among the top 8 most enriched phosphorylation sites implicating

487   upregulated calcium/calmodulin-dependent protein kinase and PKA activities in

488   merozoites compared to schizonts (Figure 5B). Furthermore, we found that MCODE

489   cluster 1 was 1.4-fold (p < 0.05) enriched for sites with the [K/R]xx[S/T] core motif

490   relative to the full merozoite phosphoproteome (Table S4). No other MCODE cluster

491   was found to be statistically significantly enriched for either the core motif or merozoite

492   specific sites (Table S4), strengthening the notion that cluster 1 is enriched for

493   merozoite-specific functions.

3.5 Differential phosphorylation between schizonts and merozoites at the site-specific level.

The above global phospho-motif view of differential phosphorylation between schizonts and merozoites was confirmed at a specific phospho-site level by examining the phosphorylation status of Ser47 in MTIP (PF3D7_1246400), which is phosphorylated by CDPK1 [41]. Serine 47 (KPL*S47) in MTIP occurs within the core motif (K/R]xx[*S/T]) that we have shown above is enriched in merozoites. The specific anti-phospho-S47 antibody recognised MTIP in merozoites, but not in schizonts and this recognition was phosphatase sensitive, clearly identifying S47 of MTIP as being specifically phosphorylated only in merozoites (Figure 6A).

3.6 Experimental testing of phospho-motif enrichment analysis confirms differential phosphorylation between schizonts and merozoites *in vivo*

The differential distribution of specific phosphorylation motifs between schizont (S) and merozoite (M) proteins was also confirmed using a panel of commercial phospho-motif specific antibodies (Figure 6B). Remarkable differences in phosphorylation patterns between schizonts and merozoites were observed with antibodies specific for a typical PKA substrate (motif 3: RRx[pS/pT]), a typical PKB/Akt substrate (motif 4: Rxx[pS/pT]) and a typical PKD substrate (motif 11: LxRxx[pS/pT]). Notably, the abundant core motif [K/R]xx[pS/pT] identified by MotifX analysis (section 3.4) would be recognised by antibodies to phospho-motifs 3, 4 and 11 (Figure 6B). So, three independent phospho-motif specific antibodies show that the core motif (K/R]xx[S/T]) is distributed differently between schizont and merozoites proteins *in vivo*. Proteins containing phospho-tyrosine

517 were difficult to detect using a specific anti-phosphotyrosine antibody, consistent with

518 detection of only 9 sites in the merozoite phosphoproteome.

519

520 **3.7 Experimental testing of GO and MMPD enrichment analyses shows that**

521 **proteins are highly ubiquitinated in merozoites compared to schizonts**

522 Our various comparative bioinformatics analyses indicated that merozoites are enriched

523 in phosphoproteins involved in 'cellular response to stress', 'proteasome core complex',

524 'proteasome-mediated proteolysis' (Figure 2), and the 'autophagic pathway' (Figure 4B);

525 a collection of terms that suggested heightened ubiquitin-mediated protein degradation

526 might be taking place in merozoites. To test this prediction we probed protein extracts

527 prepared from schizonts (42- and 45-hours post-invasion) and free merozoites with an

528 antibody reactive with both mono- and poly-ubiquitinated proteins (Figure 6C). In

529 merozoites there was a very marked increase in the level of ubiquitination compared to

530 schizonts.

531

532

4. Discussion

The *P. falciparum* life cycle is complex with dynamic protein phosphorylation at different stages. Following erythrocyte invasion the intracellular parasite develops through ring, trophozoite and schizont stages prior to the release of extracelluar merozoites that invade new erythrocytes. Signal transduction pathways regulating development of the intraerythrocytic parasite have been identified by large scale phosphorylation screens of particular stages with a focus on schizonts [18-22]. In this study we have generated a phosphoproteome of extracellular merozoites, and compared it with an ensemble of schizont phosphoproteomes [18-22]. We used a subtractive approach to identify differential protein phosphorylation in merozoites, and identified 785 sites phosphorylated in merozoites and 980 phosphorylation sites shared between merozoites and schizonts. The large number of phosphorylation sites detected in the ensemble of schizont phosphoproteomes (12,252) compared to the 1785 sites detected in merozoites indicates with a reasonable degree of certainty that the 785 sites described in this study are merozoite specific.

At the phosphopeptide level we observed large differences in the distribution of specific phosphorylation motifs, with 28 motifs not previously observed in schizonts. In addition, the core motif sequence [K/R]xx[pS/pT] was highly enriched in merozoites being contained in 35 % of the 28 phospho-motifs; a significant enrichment that directly implicates upregulated calcium-dependent protein kinase and cAMP-dependent PKA activities in the biology of merozoites. The overall differential distribution of phosphorylated motifs was confirmed *in vivo* using commercial specific phospho-motif

556   antibodies to probe schizont and merozoite protein extracts. Three independent anti-

557   phospho-motif antibodies confirmed that schizonts and merozoites are differentially

558   phosphorylated. Consistently, the core motif [K/R]xx[pS/pT] that's enriched in

559   merozoites is common to all 3 phospho-motifs likely explaining some of the common

560   phospho-bands identified by western blot. This global view of differential

561   phosphorylation does not identify given phosphoproteins, so we examined S47 in MTIP,

562   as this residue occurs within the core motif imbedded in KPL*SIE. A specific phospho-

563   S47 antibody demonstrated that S47 in MTIP is clearly differentially phosphorylated

564   between schizonts (off) and merozoites (on). Thus, *in vivo* differential phosphorylation

565   occurs between schizonts and merozoites both at the site-specific and global levels.

566

567   We argue that differential phosphorylation translates into functional differences in

568   parasite biology that are highlighted by comparative gene ontology and pathway

569   enrichment analyses. In both cases we observed an increase in terms involved in

570   invasion biology (Figure 3, Table S2 and Table S3). This finding was corroborated by

571   cluster analysis of the merozoite phosphoprotein interaction network, where the largest

572   MCODE cluster (119 proteins) is also enriched for invasion biology. Besides invasion

573   biology we find the ATG autophagy pathway also enriched in MCODE1, and together

574   these observations lead us to propose that developing merozoites undergo starvation-

575   induced translational arrest, since they no longer have access to amino acids derived

576   from haemoglobin degradation. The marked increased in ubiquitination is consistent

577   with autophagic digestion of organelles no longer required by merozoites to invade.

578   Organelle and protein degradation would provide the necessary nutrient (lipid and

579   amino acids) for merozoites to survive the hostile extracellular environment, before they

580   elaborate a new food vacuole to digest host cell proteins such as haemoglobin [12, 13].

581   This upregulation of ubiquitination in extracellular merozoites compared to schizonts

582   has not been reported previously; in an earlier study ubiquitination was examined in

583   intracellular ring, trophozoite and schizont stages but merozoites were not included in

584   this analysis [48].

585

586

587   **Acknowledgements**

597

598 **FIGURE LEGENDS**

599

600 <u>Figure 1:</u> The merozoite phosphoproteome of *Plasmodium falciparum.*

601 <u>A:</u> Schematic overview of the phosphoproteome workflow. Steps in the workflow to

602 identify phosphopeptides in protease digests derived from 2-3*10^9 purified *P. falciparum*

603 merozoites. After lysis proteins were digested in solution with lysC and trypsin and

604 partially purified using FASP. Peptides were fractionated by strong anion exchange

605 (SAX) disks, enriched for phosphopeptides by TiO_2 and measured by liquid

606 chromatography tandem mass spectrometry. Phosphopeptides were identified by

607 Mascot database searches and validated by MaxQuant at 1 % FDR. A detailed

608 description of the procedure is provided in Experimental Procedures.

609 <u>B:</u> Venn diagram depicting overlap in protein phosphorylation sites between merozoite

610 and schizonts ensemble from previous studies [18-22].

611

612 <u>Figure 2</u>. GO annotation of the phosphoproteomes of asexual blood stages.

613 Comparative gene ontology enrichment analysis between the phosphoproteomes of

614 merozoites and schizont stages. Enrichment is shown for OPI-GO terms relative to all

615 5500 *P. falciparum* proteins. Statistical significance was tested using a one–tailed

616 Fisher exact test.

617

618

29

619   Figure 3. Pathway annotation of the phosphoproteomes of asexual blood stages.

620   Comparative pathway enrichment analysis between the phosphoproteomes of

621   merozoites and schizont life cycle stages. Enrichment is shown for increased MMPD

622   pathway terms relative to all 5500 *P. falciparum* proteins. Statistical significance was

623   tested using a one–tailed Fisher exact test (FET). We accepted pathways with p (FET)

624   < 0.005 in this figure. Pathways with lower significance (0.005 < p < 0.05) are provided

625   in Table S3.

626

627   Figure 4. The protein interaction network of the merozoite phosphoproteome.

628   A: Sub-networks of the phosphoprotein interaction network in merozoites. Protein

629   interaction data were downloaded from STRING and superimposed on the merozoite

630   phosphoproteome. Twenty sub-networks were identified by the MCODE clustering

631   algorithm. Here MCODE clusters are shown that interact with the central kinase cluster

632   MCODE 13.

633   B: MMPD pathway annotation of all MCODE clusters. A pathway enrichment analysis

634   was performed against the background of all 5500 proteins, and a clustering analysis

635   was performed to display functional divergence between MCODE clusters. Enriched

636   pathways in MCODE 1 are shown in yellow.

637

638

639 Figure 5: Protein phosphorylation motifs detected in merozoites.

640 A: Distribution of phosphorylation classes (acidic, basic, proline-directed, tyrosine and

641 other) in blood stages with merozoites in blue bars and schizonts in red bars.

642 B: Core phosphorylation motifs from the top 10 most abundant motifs in merozoites

643 represented by 13-mer sequence logos. Amino acids of the core motif sequence [K/R] x

644 x [S/T] are highlighted by red boxes.

645

646 Figure 6. Differential phosphorylation and ubquitination of proteins in schizonts and

647 merozoites.

648 A: Western blot of schizont (S) and merozoite (M) protein extracts probed with an

649 antibody specific for phosphorylated Ser47 in MTIP. The antibody only reacted with

650 MTIP in merozoite extracts despite the presence of the protein in both schizonts and

651 merozoites. Treatment of the merozoite extract with phosphatase (+) abolished the

652 antibody reactivity confirming that phosphorylation of Ser47 was essential for antibody

653 recognition. The presence of MTIP in both schizont and merozoite extracts is displayed

654 with an antibody raised to recombinant MTIP.

655 B: Western blot of schizont (S) and merozoite (M) protein extracts probed with

656 phosphorylated phospho-motif-specific antibodies. The bands recognised by the

657 antibodies specific to motifs 3, 4 and 11 showed a differential pattern between schizonts

658 and merozoites, whereas the anti-phosphotyrosine antibody showed only weak

659 reactivity with both schizont and merozoite extracts. Anti-CDPK1 antibodies were used

660 as a loading control.

661   <u>C</u>: Western blot of schizont protein extracts made 42 and 45 h post-invasion together

662   with protein extracts made from free merozoites (M), probed with an antibody to mono-

663   and poly-ubiquitinated proteins. The strong reactivity with merozoite proteins indicates

664   an increased level of ubiquitination in merozoites compared to schizonts. Anti-CDPK1

665   antibodies were used as a loading control in the bottom panel.

666   The mobility of molecular mass markers (in kDa) is indicated on the left side of the

667   panels.

668

669 **Supplementary data**

670

671 **Table S1**: Phosphopeptide identification in the merozoite phosphoproteome.

672

673 **Table S2**: Gene Ontology enrichment results comparing the merozoite

674 phosphoproteome with the schizont phosphoproteome ensemble.

675

676 **Table S3**: MMPD pathway enrichment comparing the merozoite phosphoproteome with

677 the schizont phosphoproteome ensemble.

678

679 **Table S4**: List of MCODE proteins.

680

681 **Table S5**: GO enrichment results for MCODE cluster 1.

682

683 **Table S6**: Phosphorylation motifs detected in merozoites.

684

685 **Table S7**: Kinase substrates predicted by CompariMotif.

686

687 **Figure S1**. Protein- Protein Interactions between MCODE 1 and MCODE 13

688

References

[1] Sturm, A., Amino, R., van de Sand, C., Regen, T., *et al.*, Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science* 2006, *313*, 1287-1290.

[2] Miller, L. H., Ackerman, H. C., Su, X. Z., Wellems, T. E., Malaria biology and disease pathogenesis: insights for new treatments. *Nature medicine* 2013, *19*, 156-167.

[3] Bannister, L. H., Mitchell, G. H., Butcher, G. A., Dennis, E. D., Cohen, S., Structure and development of the surface coat of erythrocytic merozoites of Plasmodium knowlesi. *Cell and tissue research* 1986, *245*, 281-290.

[4] Cowman, A. F., Berry, D., Baum, J., The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *J Cell Biol* 2012, *198*, 961-971.

[5] Baum, J., Richard, D., Healer, J., Rug, M., *et al.*, A conserved molecular motor drives cell invasion and gliding motility across malaria life cycle stages and other apicomplexan parasites. *J Biol Chem* 2006, *281*, 5197-5208.

[6] Farrow, R. E., Green, J., Katsimitsoulia, Z., Taylor, W. R., *et al.*, The mechanism of erythrocyte invasion by the malarial parasite, Plasmodium falciparum. *Seminars in cell & developmental biology* 2011, *22*, 953-960.

[7] Dawn, A. S., S.; More, K.R.; Siddiqui, F.A.; Pachikara, N.; Ramdani, G.; Langsley, G.; Chitnis, C.E., The Central Role of cAMP in Regulating Plasmodium falciparum Merozoite Invasion of Human Erythrocytes. *PLoS Pathog* 2014, *10,* e1004520.

[8] Gao, X., Gunalan, K., Yap, S. S., Preiser, P. R., Triggers of key calcium signals during erythrocyte invasion by Plasmodium falciparum. *Nat Commun* 2013, *4*, 2862.

[9] Billker, O., Lourido, S., Sibley, L. D., Calcium-dependent signaling and kinases in apicomplexan parasites. *Cell Host Microbe* 2009, *5*, 612-622.

[10] Holder, A. A., Mohd Ridzuan, M. A., Green, J. L., Calcium dependent protein kinase 1 and calcium fluxes in the malaria parasite. *Microbes and infection / Institut Pasteur* 2012, *14*, 825-830.

[11] Gerald, N., Mahajan, B., Kumar, S., Mitosis in the human malaria parasite Plasmodium falciparum. *Eukaryot Cell* 2011, *10*, 474-482.

[12] Dluzewski, A. R., Ling, I. T., Hopkins, J. M., Grainger, M., *et al.*, Formation of the food vacuole in *Plasmodium falciparum*: a potential role for the 19 kDa fragment of merozoite surface protein 1 (MSP1$_{19}$). *PLoS One* 2008, *3*, e3085.

[13] Abu Bakar, N., Klonis, N., Hanssen, E., Chan, C., Tilley, L., Digestive-vacuole genesis and endocytic processes in the early intraerythrocytic stages of Plasmodium falciparum. *J Cell Sci* 2010, *123*, 441-450.

[14] Tomlins, A. M., Ben-Rached, F., Williams, R. A., Proto, W. R., *et al.*, Plasmodium falciparum ATG8 implicated in both autophagy and apicoplast formation. *Autophagy* 2013, *9*, 1540-1552.

[15] Gaur, D., Chitnis, C. E., Molecular interactions and signaling mechanisms during erythrocyte invasion by malaria parasites. *Current opinion in microbiology* 2011, *14*, 422-428.

[16] Jacot, D., Soldati-Favre, D., Does protein phosphorylation govern host cell entry and egress by the Apicomplexa? *International journal of medical microbiology : IJMM* 2012, *302*, 195-202.

[17] Lasonder, E., Treeck, M., Alam, M., Tobin, A. B., Insights into the Plasmodium falciparum schizont phospho-proteome. *Microbes and infection / Institut Pasteur* 2012, *14*, 811-819.

733 [18] Lasonder, E., Green, J. L., Camarda, G., Talabani, H.*, et al.*, The Plasmodium falciparum
734 schizont phosphoproteome reveals extensive phosphatidylinositol and cAMP-protein kinase A
735 signaling. *J Proteome Res* 2012, *11*, 5323-5337.
736 [19] Pease, B. N., Huttlin, E. L., Jedrychowski, M. P., Talevich, E.*, et al.*, Global analysis of
737 protein expression and phosphorylation of three stages of Plasmodium falciparum
738 intraerythrocytic development. *J Proteome Res* 2013, *12*, 4028-4045.
739 [20] Solyakov, L., Halbert, J., Alam, M. M., Semblat, J. P.*, et al.*, Global kinomic and phospho-
740 proteomic analyses of the human malaria parasite Plasmodium falciparum. *Nat Commun* 2011, *2*,
741 565.
742 [21] Treeck, M., Sanders, J. L., Elias, J. E., Boothroyd, J. C., The phosphoproteomes of
743 Plasmodium falciparum and Toxoplasma gondii reveal unusual adaptations within and beyond
744 the parasites' boundaries. *Cell Host Microbe* 2011, *10*, 410-419.
745 [22] Collins, M. O., Wright, J. C., Jones, M., Rayner, J. C., Choudhary, J. S., Confident and
746 sensitive phosphoproteomics using combinations of collision induced dissociation and electron
747 transfer dissociation. *Journal of proteomics* 2014, *103*, 1-14.
748 [23] Wek, R. C., Jiang, H. Y., Anthony, T. G., Coping with stress: eIF2 kinases and translational
749 control. *Biochemical Society transactions* 2006, *34*, 7-11.
750 [24] Funderburk, S. F., Wang, Q. J., Yue, Z., The Beclin 1-VPS34 complex--at the crossroads of
751 autophagy and beyond. *Trends in cell biology* 2010, *20*, 355-362.
752 [25] Harr, M. W., Distelhorst, C. W., Apoptosis and autophagy: decoding calcium signals that
753 mediate life or death. *Cold Spring Harb Perspect Biol* 2010, *2*, a005579.
754 [26] Tudisca, V., Simpson, C., Castelli, L., Lui, J.*, et al.*, PKA isoforms coordinate mRNA fate
755 during nutrient starvation. *J Cell Sci* 2012, *125*, 5221-5232.
756 [27] Taylor, H. M., Grainger, M., Holder, A. A., Variation in the expression of a Plasmodium
757 falciparum protein family implicated in erythrocyte invasion. *Infection and immunity* 2002, *70*,
758 5779-5789.
759 [28] Blackman, M. J., Purification of Plasmodium falciparum merozoites for analysis of the
760 processing of merozoite surface protein-1. *Methods in cell biology* 1994, *45*, 213-220.
761 [29] Wisniewski, J. R., Zougman, A., Nagaraj, N., Mann, M., Universal sample preparation
762 method for proteome analysis. *Nat Methods* 2009, *6*, 359-362.
763 [30] Kyono, Y., Sugiyama, N., Imami, K., Tomita, M., Ishihama, Y., Successive and selective
764 release of phosphorylated peptides captured by hydroxy acid-modified metal oxide
765 chromatography. *J Proteome Res* 2008, *7*, 4585-4593.
766 [31] Rappsilber, J., Ishihama, Y., Mann, M., Stop and go extraction tips for matrix-assisted laser
767 desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal*
768 *Chem* 2003, *75*, 663-670.
769 [32] Cox, J., Mann, M., MaxQuant enables high peptide identification rates, individualized
770 p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008, *26*,
771 1367-1372.
772 [33] Olsen, J. V., Mann, M., Improved peptide identification in proteomics by two consecutive
773 stages of mass spectrometric fragmentation. *Proc Natl Acad Sci U S A* 2004, *101*, 13417-13422.
774 [34] Bauer, S., Grossmann, S., Vingron, M., Robinson, P. N., Ontologizer 2.0--a multifunctional
775 tool for GO term enrichment analysis and data exploration. *Bioinformatics* 2008, *24*, 1650-1651.
776 [35] Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T.*, et al.*, A tissue-specific atlas
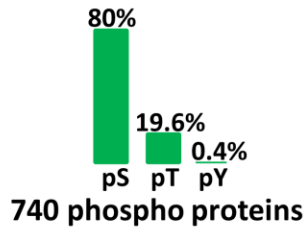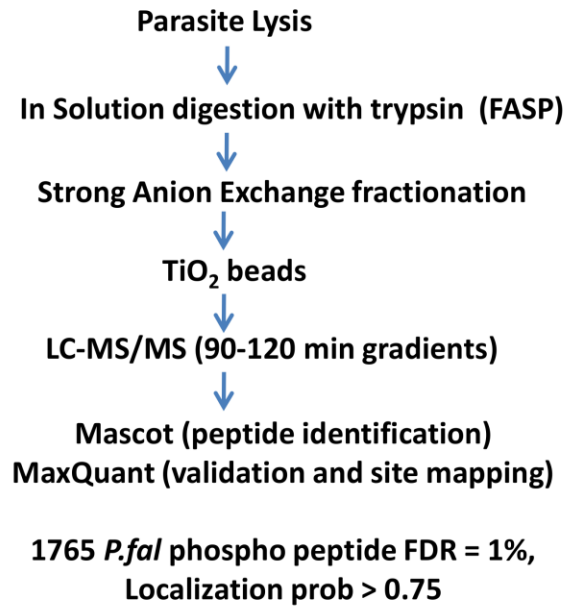777 of mouse protein phosphorylation and expression. *Cell* 2010, *143*, 1174-1189.

778   [36] Schwartz, D., Gygi, S. P., An iterative statistical approach to the identification of protein
779   phosphorylation motifs from large-scale data sets. *Nat Biotechnol* 2005, *23*, 1391-1398.
780   [37] Edwards, R. J., Davey, N. E., Shields, D. C., CompariMotif: quick and easy comparisons of
781   sequence motifs. *Bioinformatics* 2008, *24*, 1307-1309.
782   [38] Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M.*, et al.*, The STRING database in
783   2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids*
784   *Res* 2011, *39*, D561-568.
785   [39] Shannon, P., Markiel, A., Ozier, O., Baliga, N. S.*, et al.*, Cytoscape: a software environment
786   for integrated models of biomolecular interaction networks. *Genome Res* 2003, *13*, 2498-2504.
787   [40] Bader, G. D., Betel, D., Hogue, C. W., BIND: the Biomolecular Interaction Network
788   Database. *Nucleic Acids Res* 2003, *31*, 248-250.
789   [41] Green, J. L., Rees-Channer, R. R., Howell, S. A., Martin, S. R.*, et al.*, The motor complex of
790   Plasmodium falciparum: phosphorylation by a calcium-dependent protein kinase. *J Biol Chem*
791   2008, *283*, 30980-30989.
792   [42] Perkins, D. N., Pappin, D. J., Creasy, D. M., Cottrell, J. S., Probability-based protein
793   identification by searching sequence databases using mass spectrometry data. *Electrophoresis*
794   1999, *20*, 3551-3567.
795   [43] Thingholm, T. E., Jensen, O. N., Larsen, M. R., Enrichment and separation of mono- and
796   multiply phosphorylated peptides using sequential elution from IMAC prior to mass
797   spectrometric analysis. *Methods Mol Biol* 2009, *527*, 67-78, xi.
798   [44] Zhou, Y., Ramachandran, V., Kumar, K. A., Westenberger, S.*, et al.*, Evidence-based
799   annotation of the malaria parasite's genome using comparative expression profiling. *PLoS One*
800   2008, *3*, e1570.
801   [45] Vaid, A., Ranjan, R., Smythe, W. A., Hoppe, H. C., Sharma, P., PfPI3K, a
802   phosphatidylinositol-3 kinase from Plasmodium falciparum, is exported to the host erythrocyte
803   and is involved in hemoglobin trafficking. *Blood* 2010, *115*, 2500-2507.
804   [46] Poser, S., Storm, D. R., Role of Ca2+-stimulated adenylyl cyclases in LTP and memory
805   formation. *International journal of developmental neuroscience : the official journal of the*
806   *International Society for Developmental Neuroscience* 2001, *19*, 387-394.
807   [47] Shaywitz, A. J., Greenberg, M. E., CREB: a stimulus-induced transcription factor activated
808   by a diverse array of extracellular signals. *Annu Rev Biochem* 1999, *68*, 821-861.
809   [48] Ponts, N., Saraf, A., Chung, D. W., Harris, A.*, et al.*, Unraveling the ubiquitome of the
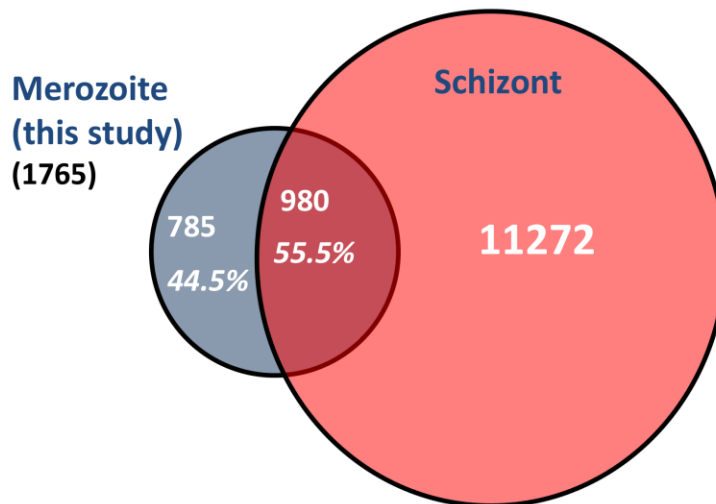810   human malaria parasite. *J Biol Chem* 2011, *286*, 40320-40330.

811

812

A

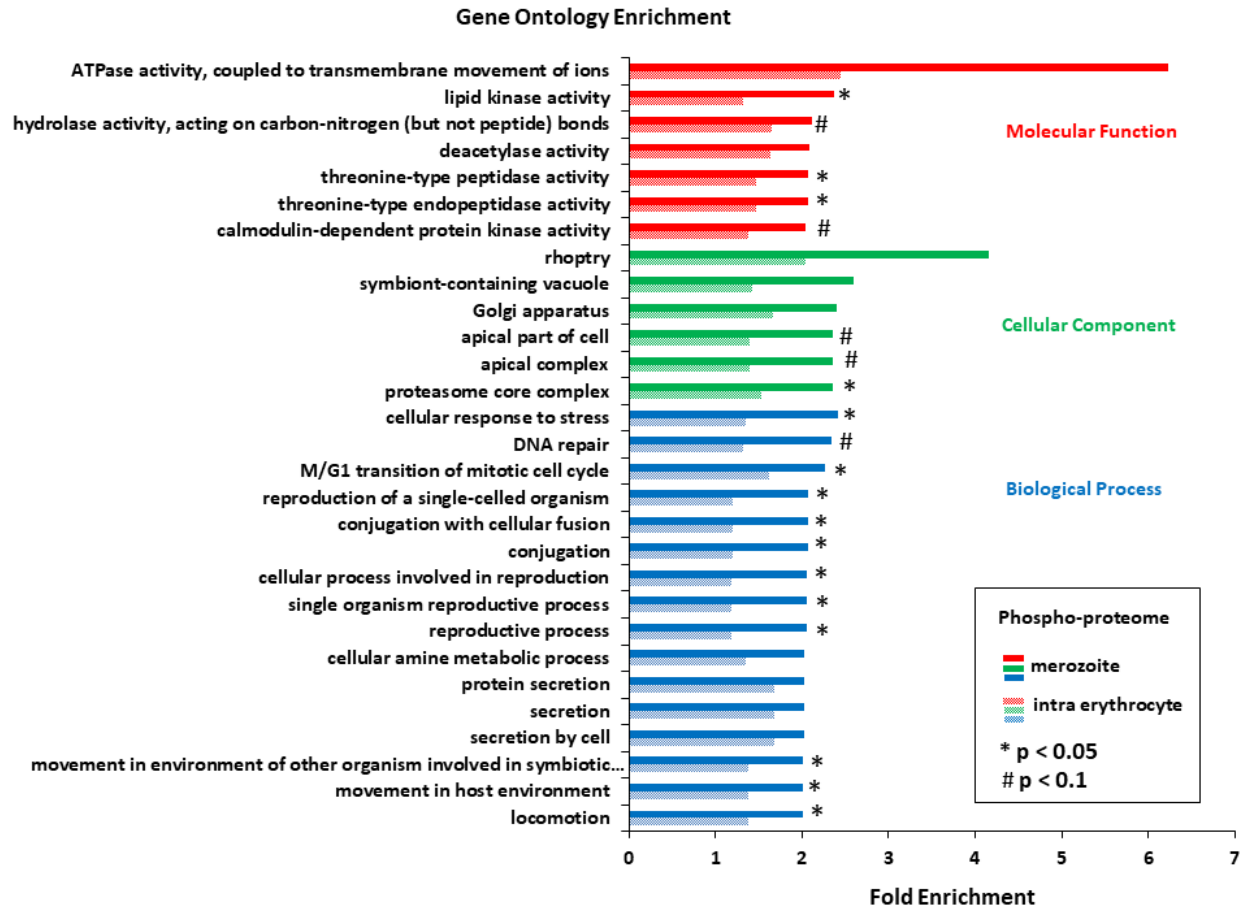**_P. fal_ merozoites phospho-proteome workflow**

Parasite Lysis
↓
In Solution digestion with trypsin (FASP)
↓
Strong Anion Exchange fractionation
↓
TiO$_2$ beads
↓
LC-MS/MS (90-120 min gradients)
↓
Mascot (peptide identification)
MaxQuant (validation and site mapping)

1765 _P.fal_ phospho peptide FDR = 1%,
Localization prob > 0.75

80%
19.6%
0.4%
pS   pT   pY
740 phospho proteins

B

PHOSPHO PEPTIDE SITES

Merozoite
(this study)
(1765)

Schizont

785
44.5%

980
55.5%

11272

Gene Ontology Enrichment

815

816

817   FIGURE 2

818

819

820

821

822

823

824

825

Malaria Metabolic Pathway Database enrichment
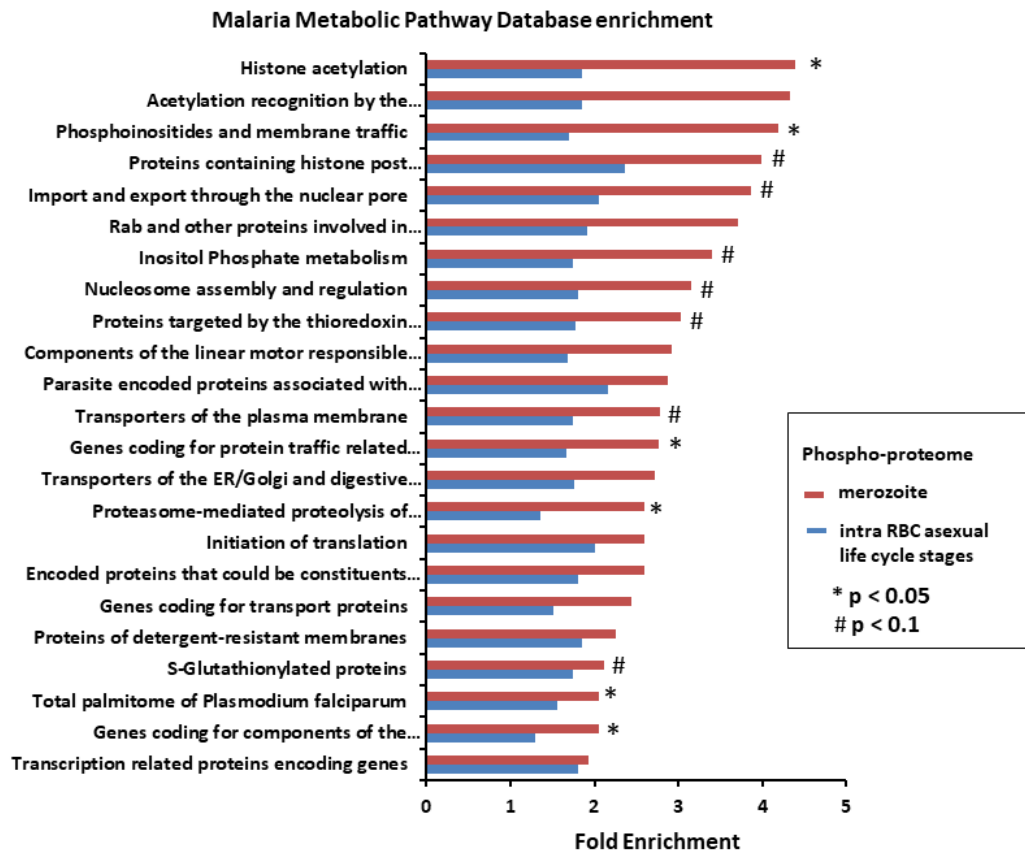
826

827   FIGURE 3

828

829

830

831

832

833

834

835

836
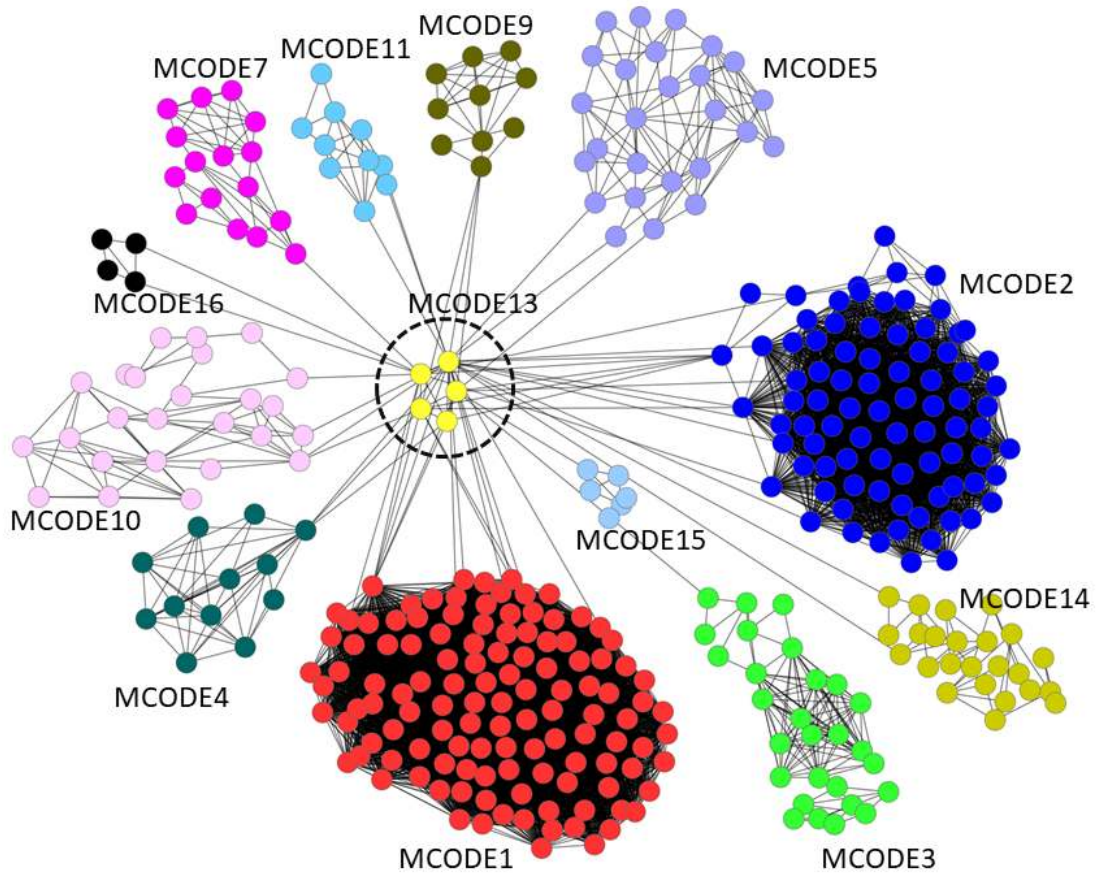
837

838

839   FIGURE 4A

840

841

842

843

844

845

846

847

848

849 FIGURE 4B

850

851   FIGURE 5

852

853

854

855

856

857
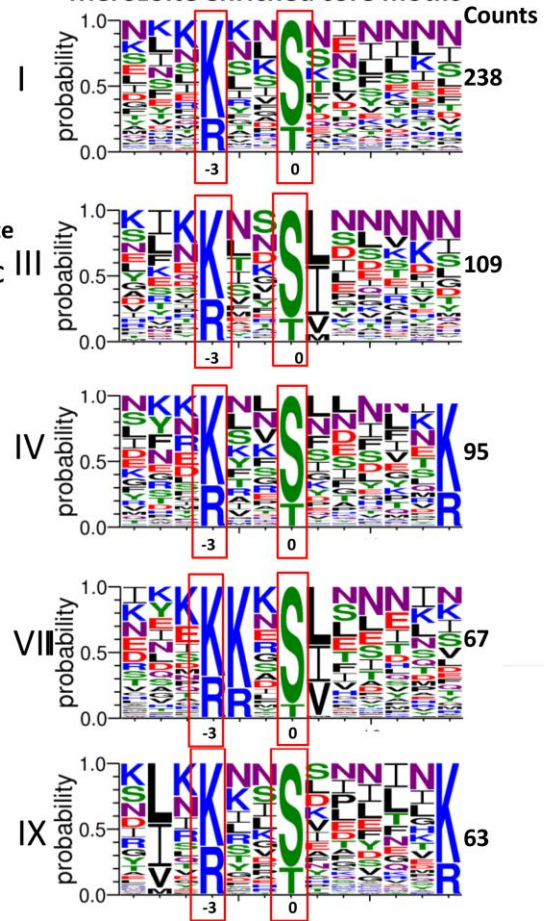
858

859

860

A Phosphorylation motif classes

B Merozoite enriched core motifs

861

862

863   FIGURE 6A

864

865

866

867

868

869

870

871

A.

S M   S M   M M   M M   phosphatase
                              - +   - +

37-   37-   37

25-   25   25
20-

Ser47P-MTIP   MTIP   Ser47P-MTIP   MTIP

B.

S M   S M   S M   S M   S M

97-

66-

45-

31-

22-

14-

motif 4   motif 3   motif 11   P-Tyr   CDPK1

C.

S   S   M
42 45

250-
150-
100-
75-   ubiquitin

50-

37-

75-   CDPK1
50-
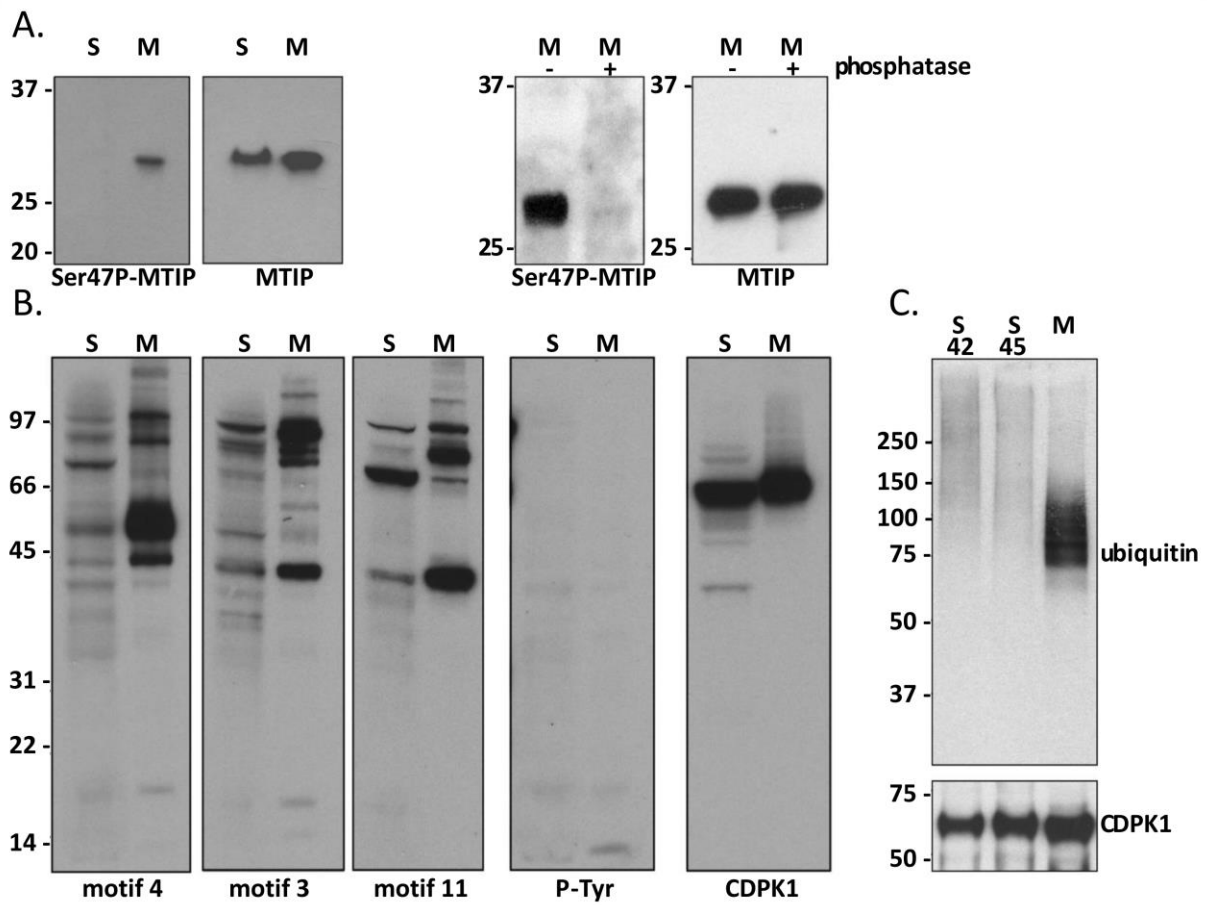
872
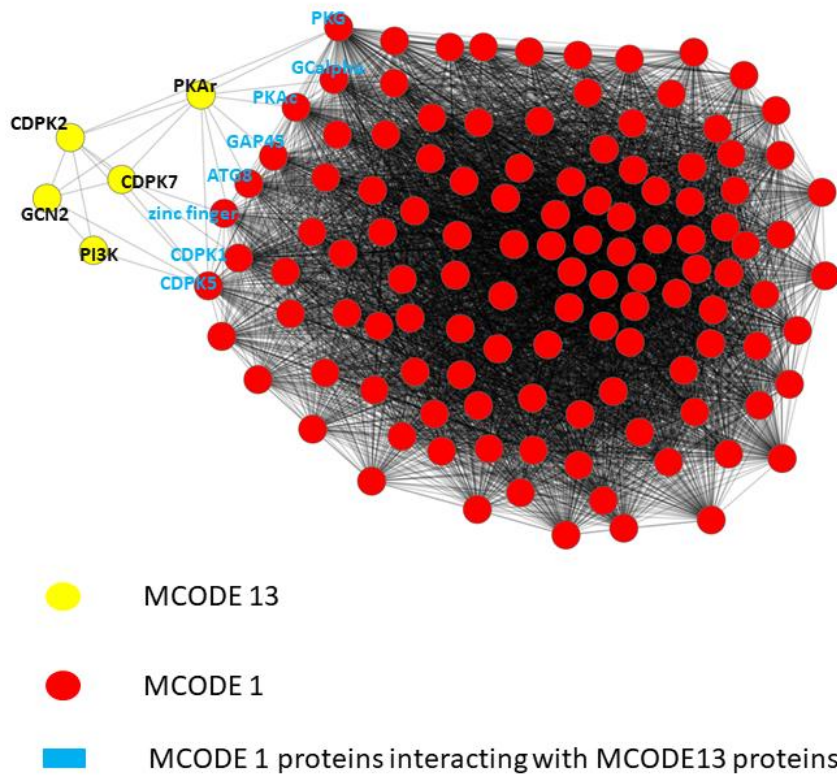
873

874   FIGURE 6B

875

876

877

878

879

880

881

882

44

FiGURE S1. Protein- Protein Interactions between MCODE 1 and MCODE 13

MCODE 13

MCODE 1

MCODE 1 proteins interacting with MCODE13 proteins

883