⌘ *Author's Choice*

The Phosphoproteome of Bloodstream Form *Trypanosoma brucei*, Causative Agent of African Sleeping Sickness[S]

Isabelle R. E. Nett‡§, David M. A. Martin‡, Diego Miranda-Saavedra¶, Douglas Lamont‡, Jonathan D. Barber‡, Angela Mehlert‡‖, and Michael A. J. Ferguson‡**

The protozoan parasite *Trypanosoma brucei* is the causative agent of human African sleeping sickness and related animal diseases, and it has over 170 predicted protein kinases. Protein phosphorylation is a key regulatory mechanism for cellular function that, thus far, has been studied in *T. brucei* principally through putative kinase mRNA knockdown and observation of the resulting phenotype. However, despite the relatively large kinome of this organism and the demonstrated essentiality of several *T. brucei* kinases, very few specific phosphorylation sites have been determined in this organism. Using a gel-free, phosphopeptide enrichment-based proteomics approach we performed the first large scale phosphorylation site analyses for *T. brucei*. Serine, threonine, and tyrosine phosphorylation sites were determined for a cytosolic protein fraction of the bloodstream form of the parasite, resulting in the identification of 491 phosphoproteins based on the identification of 852 unique phosphopeptides and 1204 phosphorylation sites. The phosphoproteins detected in this study are predicted from their genome annotations to participate in a wide variety of biological processes, including signal transduction, processing of DNA and RNA, protein synthesis, and degradation and to a minor extent in metabolic pathways. The analysis of phosphopeptides and phosphorylation sites was facilitated by in-house developed software, and this automated approach was validated by manual annotation of spectra of the kinase subset of proteins. Analysis of the cytosolic bloodstream form *T. brucei* kinome revealed the presence of 44 phosphorylated protein kinases in our data set that could be classified into the major eukaryotic protein kinase groups by applying a multilevel hidden Markov model library of the kinase catalytic domain. Identification of the kinase phosphorylation sites showed conserved phosphorylation sequence motifs in several kinase activation segments, supporting the view that phosphorylation-based signaling is a general and fundamental regulatory process that extends to this highly divergent lower eukaryote. *Molecular & Cellular Proteomics 8: 1527–1538, 2009.*

Phosphorylation by eukaryotic protein kinases (PKs)[1] is the most intensively studied posttranslational modification in eukaryotic cells mainly because reversible phosphorylation regulates almost every cellular process (1). PKs make up one of the largest protein superfamilies, comprising the protein-serine/threonine kinases and protein-tyrosine kinases (2). Knowledge of the fundamental role of protein kinases for cell survival in multicellular organisms has led to functional studies of this enzyme class in parasitic protozoa to search for alternative drug targets to treat tropical diseases (3). *Trypanosoma brucei* causes human African trypanosomiasis (also known as African sleeping sickness) and is responsible for ~30,000 deaths per annum (4). Initial bioinformatics analysis of the *T. brucei* genome revealed the presence of 156 conventional protein kinases (ePks) and 20 atypical protein kinases (aPKs) (5). Our characterization of the *T. brucei* genome using a multilevel hidden Markov model (HMM) library of the kinase catalytic domain (6) followed by manual curation identified 170 ePKs and 12 aPKs.

From the ‡Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee DD1 5EH, Scotland, United Kingdom and the ¶Department of Haematology, Cambridge Institute for Medical Research, University of Cambridge, Hills Road, Cambridge CB2 0XY, United Kingdom

⌘ *Author's Choice*—Final version full access.

Received, December 4, 2008, and in revised form, March 31, 2009

Published, MCP Papers in Press, April 4, 2009, DOI 10.1074/mcp.M800556-MCP200

[1] The abbreviations used are: PK, protein kinase; ePK, conventional protein kinase; aPK, atypical protein kinase; HMM, hidden Markov model; SCX, strong cation exchange; TiO_2, titanium dioxide; FDR, false discovery rate; *Try*PP-DB, *T. brucei* PhosphoProteomics database; RACK1, receptor for activated protein kinase C-1; LTQ, linear trap quadrupole; ELM, Eukaryotic Linear Motif; DYRK, dual specificity tyrosine phosphorylation-regulated kinase; CDK, cyclin-dependent kinase; MAP, mitogen-activated protein; CAMK, Ca^{2+}/calmodulin-dependent protein kinase; GSK3, glycogen synthase 3 kinase; ERK, extracellular signal-regulated kinase; MAPK, mitogen-activated protein kinase; CRK, cell division cycle 2-related protein kinase; MEK, mitogen-activated protein kinase/extracellular signal-regulated kinase kinase; MEKK, MEK kinase; NEK, NIMA-related kinase; STE, homologues of yeast sterile 7, 11, and 20 kinases; *Tb*TRACK, *T. brucei* receptor for activated protein kinase C-1.

Over the past few years, research on *T. brucei* protein kinases has largely focused on cell cycle-regulating enzymes because of their potential as targets for African sleeping sickness (3, 7). Indeed several protein kinases have been shown to be essential for the organism (8–12). Most of the above mentioned protein kinases were characterized using techniques such as RNA interference, overexpression of the kinase and/or a kinase-dead version, *in vitro* and *in vivo* kinase assays, and immunofluorescence microscopy. However, despite these encouraging investigations, the literature for *T. brucei* proteins with a predicted role in signal transduction lacks information about their *in vivo* phosphorylation sites. Knowledge of protein phosphorylation sites, however, is crucial for a complete functional characterization of protein kinases and kinase-substrate relationships (1, 13).

At the start of this study, only seven phosphorylation sites had been mapped in any of the trypanosomatids: Mehlert *et al.* (14) had found that six of seven threonine residues of the Gly-Pro-Glu-Glu-Thr repeat-procyclin are phosphorylated using MALDI mass spectrometry, and da Cunha *et al.* (15) had determined one phosphorylation site at Ser^{12} of *Trypanosoma cruzi* histone H1 protein by electrospray MS. Over the past few years phosphoproteome studies in other organisms using phosphopeptide enrichment in combination with high accuracy mass spectrometry have revealed thousands of novel protein phosphorylation sites and given new insights into the extent and function of this posttranslational modification (16–20).

In this study, we used a gel-free peptide-based phosphoproteomics approach to identify 491 trypanosomal phosphoproteins and, where possible, their precise phosphorylation sites. In addition to phosphorylation on serine and threonine residues, we identified several phosphorylated tyrosine residues in *T. brucei* proteins despite the lack of conventional tyrosine kinases in the *T. brucei* genome (5). Furthermore phosphorylation in the activation segment of protein kinases was found to be conserved in *T. brucei*, suggesting that at least some signaling pathways described in other eukaryotes are conserved in this ancient and highly divergent eukaryote.

EXPERIMENTAL PROCEDURES

Cell Preparation and Lysis—Bloodstream form *T. brucei* glycosylphosphatidylinositol-specific phospholipase C-null mutant strain RUMP 528 cells (21) were isolated from infected rats and purified over DEAE-cellulose (22). Cells were washed in *T. brucei* dilution buffer (5 mM KCl, 80 mM NaCl, 1 mM $MgSO_4$, 20 mM Na_2HPO_4, 2 mM NaH_2PO_4, 20 mM glucose, pH 7.4) and osmotically lysed using water containing protease and phosphatase inhibitors at a ratio of 1×10^9 cells:1 ml of lysis buffer (5 mM EDTA, 1 μg/ml E-64 protease inhibitor, 0.1 mM N^{α}-tosyl-L-lysinyl-chloromethyl ketone, 1 mM benzamidine, 0.1 mM freshly prepared PMSF, Phosphatase Inhibitor Mixture II (Calbiochem)). The lysate was left at room temperature for 5 min before cooling on ice water. Insoluble material was removed by centrifugation at 100,000 $\times$ *g* for 1 h at 4 °C. The supernatant was transferred to a fresh tube and the protein concentration was determined using the Coomassie Plus Protein Assay Reagent (Pierce).

Protein Digestion—About 5 mg of cytosolic protein was precipitated by adding a 50% aqueous solution of ice-cold TCA to the sample to a final concentration of 10% TCA (w/v). Precipitation was performed at 4 °C overnight, and the protein pellet was washed twice with an equal volume of ice-cold water. The dried pellet was redissolved and denatured in 20 μl 8 M urea, 200 mM ammonium bicarbonate, 20 mM dithiothreitol, and reduction was performed for 2 h at room temperature. Cysteine residues were alkylated by adding an equal volume of 100 mM iodoacetamide (50 mM final concentration) for 1 h in the dark. The reduced and alkylated protein solution was diluted 10-fold with water (final concentrations of 0.4 M urea and 10 mM ammonium bicarbonate) and digested overnight at 30 °C with trypsin (Worthington) at a trypsin:protein ratio of 1:100 (w/w).

Phosphopeptide Enrichment—Strong cation exchange (SCX) fractionation of tryptic peptides was performed on a 3.0-mm $\times$ 20-cm column (Poly LC, Columbia, MD) containing 5-μm polysulfoethyl aspartamide beads with a 200-Å pore size as described previously (23). The first 15 SCX fractions were freeze dried and subjected to titanium dioxide (TiO_2) chromatography as follows. Each SCX fraction was redissolved in 50 μl of 200 mg/ml 2,5-dihydroxybenzoic acid dissolved in 80% acetonitrile, 2% trifluoroacetic and incubated for 5 min with ~2 mg of TiO_2 beads that had been prewashed in the same solution. The beads were then transferred into a $ZipTipC_{18}$ (Millipore) and washed twice with 45 μl of 80% acetonitrile, 2% trifluoroacetic. Bound peptides were eluted with 30 μl of 0.6% ammonia solution, pH 10.5, and the eluted phosphopeptides that absorbed to the underlying C_{18} material of the ZipTip were eluted with 10 μl of 30% acetonitrile followed by 10 μl of 50% acetonitrile, 0.05% trifluoroacetic acid. The combined eluates were dried, reconstituted in 20 μl of 5% formic acid, vortexed vigorously, and diluted to a final concentration of 1% formic acid with water for LC-MS/MS analysis.

Liquid Chromatography-Mass Spectrometry—Liquid chromatography was performed on a fully automated UltiMate Nano LC System (Dionex) fitted with a 1 $\times$ 5-mm PepMap C_{18} trap column and a 75-μm $\times$ 15-cm reverse phase PepMap C_{18} analytical nanocolumn (LC Packings, Dionex). Samples were loaded in 0.1% formic acid (buffer A) and separated using a binary gradient consisting of buffer A and buffer B (90% acetonitrile, 0.08% formic acid). Peptides were eluted with a linear gradient from 5 to 40% buffer B over 130 min. The HPLC system was either coupled to a hybrid quadrupole time-of-flight mass spectrometer (QSTAR XL, Applied Biosystems) or an LTQ-Orbitrap mass spectrometer (Thermo Electron), both equipped with a nanospray ionization source. The QSTAR XL mass spectrometer was operated in a data-dependent mode, which consisted of a MS survey scan (*m/z* 400–2000) for 1 s followed by four 2-s MS/MS scans (*m/z* 60–1800) of the four most intense doubly or triply charged ions exceeding 10 counts. Former target ions were excluded for 180 s. The mass spectrometer was calibrated using 1 pmol/μl Glu-fibrinopeptide B. In the LTQ-Orbitrap mass spectrometer a survey scan was performed over a split mass range (*m/z* 300–800 and *m/z* 800–2000) in the Orbitrap analyzer (R = 60,000), each scan triggering five MS^2 LTQ acquisitions of the five most intense ions using multistage activation on the neutral loss of 98 and 49 thomsons. The Orbitrap mass analyzer was internally calibrated on the fly using the lock mass of polydimethylcyclosiloxane at *m/z* 445.120025.

Data Processing—Peak list files of the SCX-TiO_2 experiments produced in the Mascot generic format were concatenated using the Mascot daemon engine (Matrix Science, London, UK) for QSTAR XL spectra. Raw files obtained from the LTQ-Orbitrap were converted to Mascot generic files using Raw2msm software (a gift from Prof. Matthias Mann) before merging into a single file. Data files were searched against concatenated forward and reversed sequence databases consisting of amino acid sequences from *T. brucei* strain 927 and translated open reading frames from *T. brucei* strain 427 using

the Mascot searching engine (Mascot V2.1, Matrix Science). This database was then filtered to allow a non-redundant protein set to be determined by clustering at the 95% level of a self-match bit score. A detailed description of the database construction is in supplemental Material S1. The search criteria for QSTAR XL data were as follows: one missed cleavage was allowed; carbamidomethylation (Cys) was set as a fixed modification; oxidation (Met), *N*-acetylation (protein), phosphorylation (Ser/Thr/Tyr), and pyro-Glu (N-terminal Glu) were set as variable modifications; precursor ion mass tolerance was 1.1 Da; and fragment ion mass tolerance was set to 0.5 Da. For LTQ-Orbitrap data the precursor ion mass tolerance was set to 20 ppm and 0.8 Da for all MS^2 spectra acquired in the LTQ mass analyzer.

Data Analysis and Validation—The Mascot results output file was parsed with custom Perl scripts using the Mascot parser (Matrix Science) and stored in a relational database, the Mascot Large Results Viewer, to be able to view and extract information from large proteomics data sets in a timely manner. Peptide match data for the forward sequences that met quality control criteria were imported into a separate database system, the *T. brucei* PhosphoProteomics database (*Try*PP-DB), to allow further curation and analysis of the spectra. Imported data were filtered on peptide accurate mass (delta mass within 0.1 Da, allowing a mass shift of 1 Da for the ^{13}C isotope) and on Mascot hit rank. If the Mascot score difference between the highest ranking peptide sequence and a lower ranked isoform was 20 or more, the lower ranking peptide entry was excluded from further analysis. A Mascot score threshold was determined as the 1% false discovery rate (FDR1) by calculating the false discovery rate for each experiment in the following way. All peptide hits from forward and reverse databases corresponding to proteins in the non-redundant set and meeting the quality criteria described above were ranked by Mascot score. FDR1 was determined as the lowest score *n* where the proportion of reverse hits with score greater than *n* was less than 1% of the number of forward hits with score greater than *n*. Peptides with scores below this threshold were excluded from the analysis.

The assignment of phosphopeptides and annotation of individual phosphorylation sites were performed automatically using in-house developed software on all phosphopeptide matches with a Mascot score >25. For phosphopeptides belonging to the kinase subset of proteins, confirmation of assignment and site annotation were also performed manually. The automated assignment and annotation method is described in detail elsewhere,[2] but in summary, the assignment of a phosphopeptide requires the following criteria to be met. (i) A series of either four or five of six sequential b- or y-ions ions must be present. (ii) For proline-containing peptides, at least one enhanced fragmentation N-terminal to the Pro residue (excluding peptides containing Pro-Pro and Gly-Pro sequences) must be present. (iii) At least three singly charged peptide fragment ions that have lost phosphoric acid (−98 Da) must be present. Only those peptides with MS/MS spectra that met these criteria were taken to the second level of phosphorylation site annotation.

For the annotation of a phosphorylation site within a peptide sequence, the modified site had to be determined unambiguously by observing distinct site-determining ions, *i.e.* those derived from a transition between two fragment ions as a result of a neutral loss of phosphoric acid from threonine and/or serine residues that could not be assigned to an alternative phosphorylation site. Phosphotyrosine residues were assigned according to the presence of a unique transition between two fragment ions of 243 Da.

Additionally for the manual inspection of the spectra of the kinase subset of proteins, phosphopeptides that did not show at least three fragment ions with loss of phosphoric acid were accepted if the peptide length was less than 12 amino acids and/or in the absence of a good b-ion series if the modification site was within the first four amino acids of the N terminus. The observation of a strong precursor ion and/or the detection of a precursor ion without the phosphate group(s) was also used as an advisory criterion if there were less than three −98 Da fragment ions in the spectrum.

Bioinformatics Analysis—All *T. brucei* sequences were analyzed with the freely available phosphorylation site prediction programs "NetPhos" and "Scansite" using the default settings. Predicted phosphorylation sites were recorded in *Try*PP-DB. In addition, regular expression patterns matching known phosphorylation sites were obtained from the ELM database (25) and from the list described by Macek *et al.* (20) and matches to the *T. brucei* protein set identified using the European Molecular Biology Open Software Suite program fuzzpro (26). All *T. brucei* 927 sequences were aligned against the phospho.ELM database (27) of phosphoproteins with blastp. Regions showing similarity were parsed, and sites were recorded where known phosphorylation sites in the phospho.ELM database aligned with serine, threonine, or tyrosine residues. Storage of these data in a relational database facilitated subsequent analysis of the prevalence of known kinase substrate motifs and an evaluation of the efficacy of substrate site prediction programs.

T. brucei Kinome Characterization—Version 4 of the annotated proteome of *T. brucei* (release, May 2006) was downloaded from the Sanger Institute. The peptide sequences were scanned through a highly sensitive and specific multilevel HMM library (6) of the protein kinase superfamily followed by expert curation. Assignment of putative protein kinases to the main ePK and aPK groups was done by using the E-value cutoffs specific for each group as described previously (6).

T. brucei Phosphoprotein Categories—Functional assignment of the phosphoproteins was performed manually from primary scientific publications because the majority of the phosphoproteins detected in this study have not been mapped to Gene Ontology categories.

RESULTS

Characterization of the T. brucei Kinome—Using a highly sensitive and specific multilevel HMM (see "Experimental Procedures" for details) we identified 170 ePKs and 12 aPKs (see Table I). No protein kinases were found to contain more than one kinase catalytic domain. Our *T. brucei* kinome analysis represents an enhancement of the previously published kinome (5), which had listed 156 ePKs and 20 aPKs. We identified kinases that had not been included in the previous version of the *T. brucei* proteome (*e.g.* Tb05.5K5.80) and thus could not have been identified by Parsons *et al.* (5). Moreover two ePKs included in their study (5) (Tb927.7.3190 and Tb927.3.1300) are not included in the present study as they do not seem to possess kinase catalytic domains. Our list of aPKs is shorter because we only include those aPK families that have been shown to possess kinase catalytic activity (Alpha, PDHK, PIKK, and RIO). To assign homology relationships to the *T. brucei* kinases within each group, we carried out phylogenetic analyses to compare the kinome of the parasite with the previously characterized kinomes of *Homo sapiens* and *Dictyostelium discoideum* (6), *Encephalitozoon cuniculi*, *Saccharomyces cerevisiae*, and *Schizosaccharomy-*

[2] D. M. A. Martin, I. R. E. Nett, J. D. Barber, and M. A. J. Ferguson, manuscript submitted for publication.

TABLE I

The kinome of T. brucei compared with other unicellular parasitic and free living organisms and with the human kinome

The proteomes of the organisms below were extracted from the following resources: *T. brucei*, GeneDB (44), version 4, 9192 peptides; *H. sapiens*, Ensembl (45), version NCBI36.53, 47,509 peptides; *S. cerevisiae*, *Saccharomyces* Genome Database (46), version February 20, 2009, 6717 peptides; *S. pombe*, GeneDB (44), version February 19, 2009, 5016 peptides; *E. cuniculi*, Ref. 47, 1996 peptides; *Plasmodium falciparum*, PlasmoDB (24), version 5.5, 5460 peptides; *T. gondii*, ToxoDB (34), version 5.0, 23,941 peptides. AGC group, including cyclic nucleotide and calcium-phospholipid-dependent kinases, ribosomal S6-phosphorylating kinases, G protein-coupled kinases, and all close relatives of these groups; CAMK, calmodulin-regulated kinases; CK1 group, casein kinase 1 and close relatives; CMGC group, including cyclin-dependent kinases, mitogen-activated protein kinases, glycogen synthase kinases (GSK3), and CDK-like kinases; RGC group, receptor guanylate cyclase; STE, including many kinases functioning in MAP kinase cascades; TK, tyrosine kinases; TKL group, tyrosine kinase-like kinases; PIKK group, phosphatidylinositol 3-kinase-related kinases; PDHK, pyruvate dehydrogenase kinases; RIO, "right open reading frame"; Alpha, exemplified by myosin heavy chain kinase of *D. discoideum*.

	ePKs											aPKs					Kinome as percentage of proteome
	AGC	CAMK	CK1	CMGC	RGC	STE	TK	TKL	Other	NEK	Total	PIKK	RIO	PDHK	Alpha	Total	
																	%
T. brucei	22	28	5	47	0	28	0	0	20	20	170	6	2	2	2	12	2.0
H. sapiens	82	95	12	68	5	61	91	48	6	10	478	6	3	5	6	20	1.0
S. cerevisiae	20	37	4	25	0	14	0	0	14	1	115	5	2	2	0	9	1.8
S. pombe	20	28	5	26	0	13	0	0	16	1	109	5	2	1	0	8	2.3
E. cuniculi	4	5	2	12	0	0	0	1	4	1	29	2	1	0	0	3	1.6
P. falciparum	14	20	1	19	0	1	0	5	23	4	87	1	2	0	0	3	1.6
T. gondii	82	84	9	66	0	14	0	17	50	8	330	15	5	6	0	26	1.5

ces pombe (28) and the kinomes of nine Apicomplexan parasites[3] (see Table I for comparison).

As previously found by Parsons *et al.* (5), no protein kinases of the groups RGC (receptor guanylate cyclase), TK (tyrosine kinase), or TKL (tyrosine kinase-like kinase) were found. Receptor guanylate cyclases are likely to be a late metazoan innovation, whereas tyrosine kinases and tyrosine kinase-like kinases might have predated the radiation of eukaryotes and have been subsequently lost in a number of lineages (6). However, tyrosine phosphorylation is likely to be carried out by dual specificity protein kinases (such as Wee and dual specificity tyrosine phosphorylation-regulated kinase (DYRK) family kinases) in *T. brucei.* The largest group was found to be the CMGC group, which includes all the cyclin-dependent kinases (CDKs) and mitogen-activated protein (MAP) kinase cascades. It must be noted that all Apicomplexa and the intracellular parasite *E. cuniculi* lack canonical MAP kinase signaling modules, which have been presumably lost in these parasites (28). Therefore, the presence of MAP kinase cascade homologues (supplemental Table S1) suggests that this conserved signaling route is important in *T. brucei.* Because trypanosomatids transcribe most of their genes in large polycistronic units, signaling cascades in these parasites may function in posttranscriptional regulation.

Acquisition of Mass Spectrometric Data—Cytosolic proteins of bloodstream form *T. brucei* were digested with trypsin in solution, and phosphopeptides were first separated by low pH SCX HPLC. The first 15 SCX fractions were further enriched for phosphopeptides by TiO_2 chromatography. Each enriched fraction was subjected to LC-MS/MS three times, twice on a QSTAR XL mass spectrometer and once on an LTQ-Orbitrap instrument. Concatenated data files were used to search an in-house *T. brucei* protein database consisting of forward and reverse sequences with Mascot software. The output data were stored in a relational database (*Try*PP-DB) after quality filtering on precursor ion mass, the relative Mascot score of phosphopeptide isoforms, and a 1% FDR based on decoy sequences (see "Experimental Procedures" for details).

Phosphoproteome Data Analysis—The MS/MS spectra of putative phosphopeptides were subjected to an automated assessment using a set of stringent criteria (see Fig. 1) that produced a list of "autovalidated" phosphopeptides. Spectra of putative phosphopeptides belonging to proteins annotated in GeneDB and/or UniRef with the word "kinase" were also manually inspected using the same criteria, modulated with expert knowledge, to produce a list of "manually validated" phosphopeptides. The MS/MS spectra of autovalidated phosphopeptides were further filtered automatically to assess whether a precise phosphorylation site(s) could be annotated (Fig. 1). As a test set of biological interest, all of the MS/MS spectra of the manually validated phosphopeptides were also manually inspected to determine whether unambiguous phosphorylation sites could be annotated. From these analyses, the automated method identified 91 unambiguous precise kinase phosphorylation sites belonging to 42 kinase proteins, whereas full manual inspection identified 99 sites on 47 kinase proteins (excluding *Tb*TRACK; see below) (see Fig. 1). Examples of MS/MS spectra for an auto- and manually validated phosphopeptide and one that only survived manual validation are shown in Fig. 2. Of the 91 autovalidated phosphorylation sites, eight sites on threonine residues were absent from the manually validated list, suggesting that the automated method, although stringent, has a false positive rate of

[3] D. Miranda-Saavedra, D. M. Martin, G. J. Barton, and C. Doerig, manuscript submitted for publication.

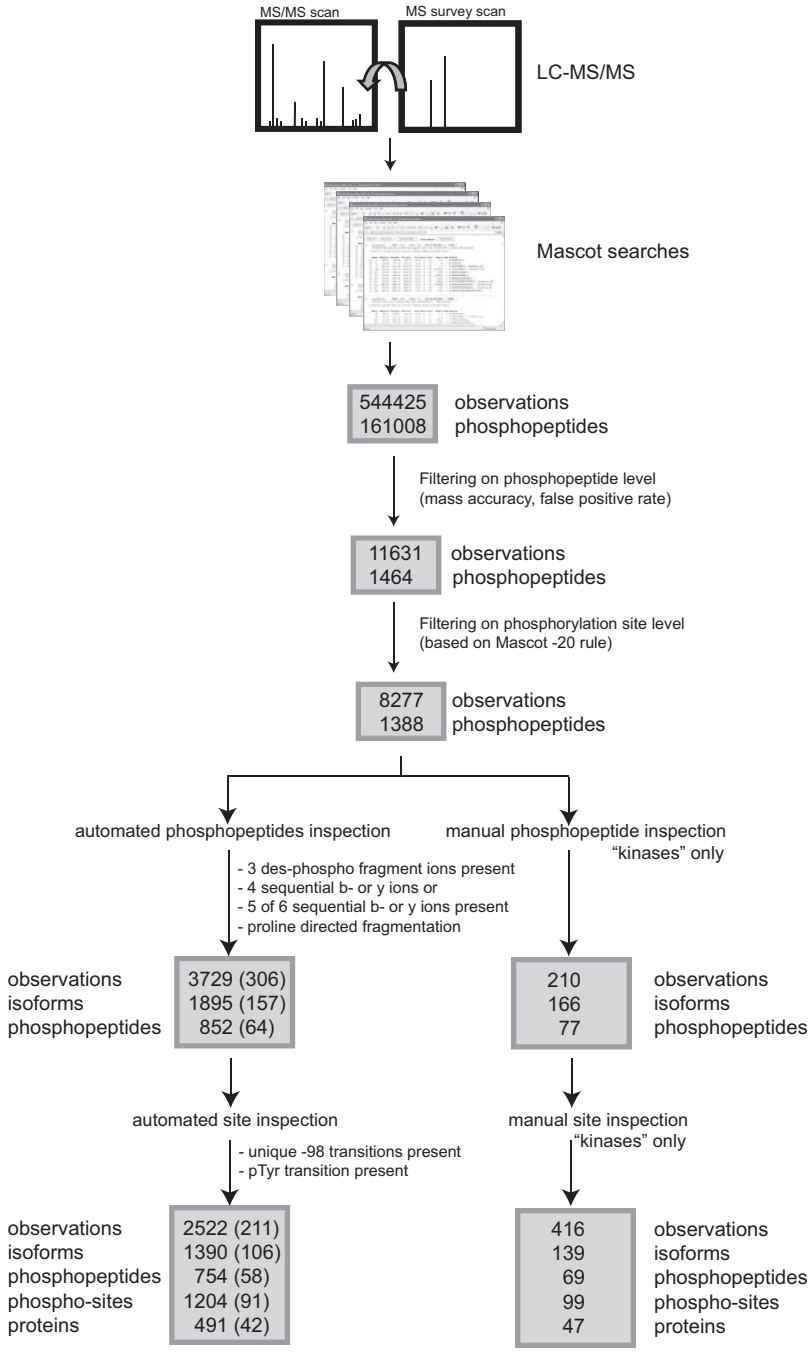
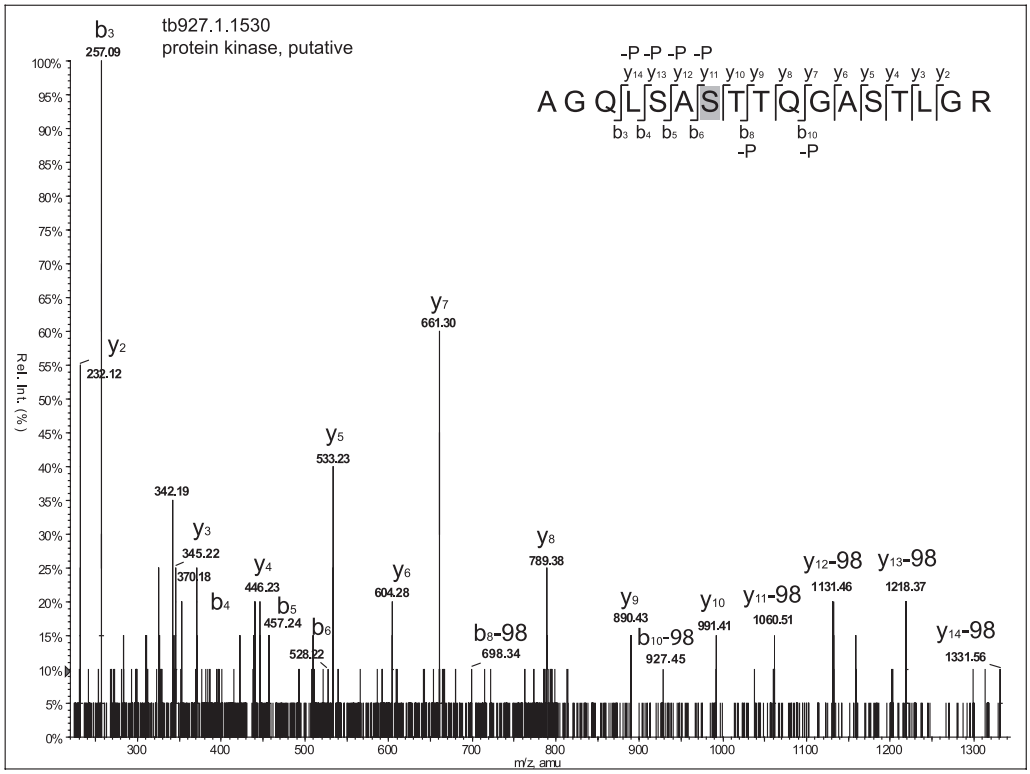


FIG. 1. **Phosphoproteome work flow.** The diagram illustrates the effects of the various data filtering and criteria regimes used to collapse the raw data into a list of automatically and manually assigned phosphopeptides and automatically and manually annotated phosphorylation sites. The numbers in *parentheses* on the *left side* of the diagram refer to automatically processed data for the kinase subset of proteins. *Numbers* on the *right side* of the diagram refer to manually processed data for the same kinase subset of proteins. In the diagram, the number of "*observations*" refers to the number of Mascot hits to potentially mono- or multiply phosphorylated peptides; the number of "*isoforms*" refers to the total number of unique phosphopeptides and those with potentially different phosphorylation site modification patterns; the number of "*phosphopeptides*" refers to the number of phosphopeptides regardless of multiple phosphorylation and potential isoforms; the number of "*phospho-sites*" refers to the number of distinct annotated phosphorylation sites; the number of "p*hosphoproteins*" refers to the number of proteins containing one or more annotated phosphorylation sites or assigned phosphopeptides. The automated phosphopeptide assignment and phosphorylation site annotation method is described in detail elsewhere.[2] *pTyr*, phosphotyrosine.

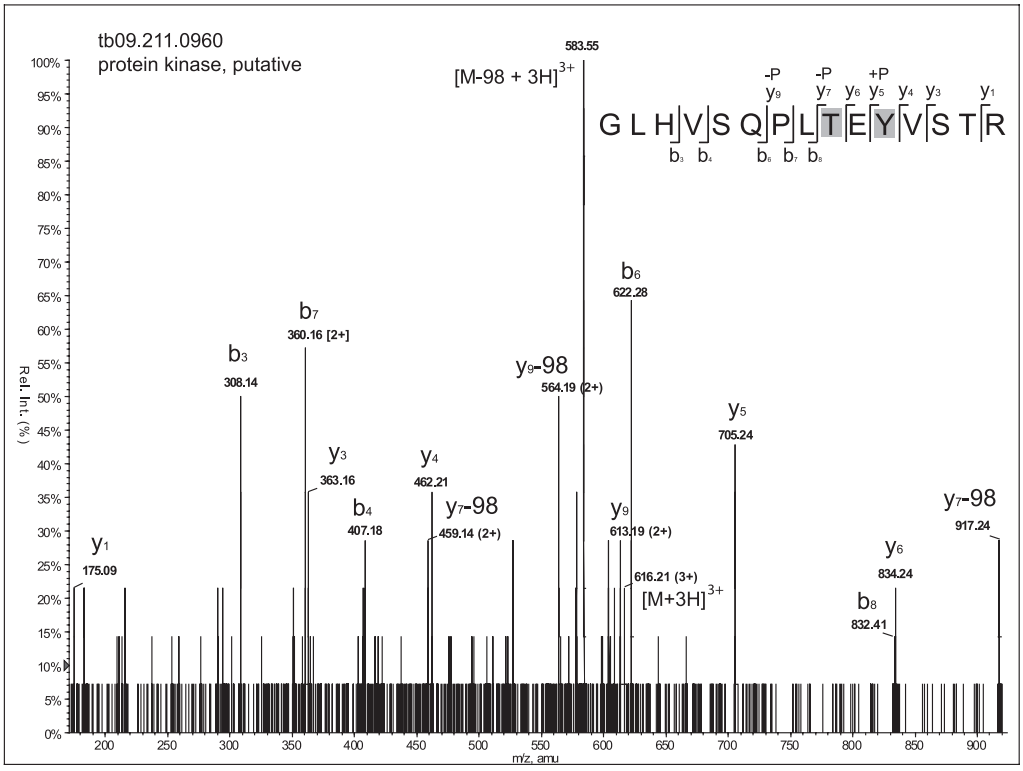
about 10% and needs further development particularly in the assignment of appropriate threshold settings of MS/MS spectra.[2] Of the remaining 449 non-kinase *T. brucei* phosphoproteins detected in our study, the location of phosphate groups on specific serine/threonine/tyrosine residues was predicted automatically for 1113 phosphorylation sites. This is a useful resource for the field of trypanosome research with the caveat that automated phosphopeptide identifications should be confirmed by manual inspection of the supporting MS/MS spectra in the database before being used experimentally.

The phosphoproteins detected in our phosphoproteomics screen of the cytosolic fraction of bloodstream form *T. brucei* are predicted from their genome annotations to participate in a wide variety of biological processes, including signal transduction, processing of DNA and RNA, protein synthesis, and degradation and to a lesser extent in metabolic pathways. Significantly overrepresented were phosphoproteins with hitherto unknown functions (MS hits to "hypothetical proteins") (see Fig. 3). A complete list of identified phosphoproteins, phosphopeptides, and phosphorylation sites (supple-

A


tb927.1.1530
protein kinase, putative
A G Q L S A S T T Q G A S T L G R
-P -P -P -P
Rel. Int. (%)
m/z, amu

B


tb09.211.0960
protein kinase, putative
G L H V S Q P L T E Y V S T R
[M-98 + 3H]3+
[M+3H]3+
Rel. Int. (%)
m/z, amu

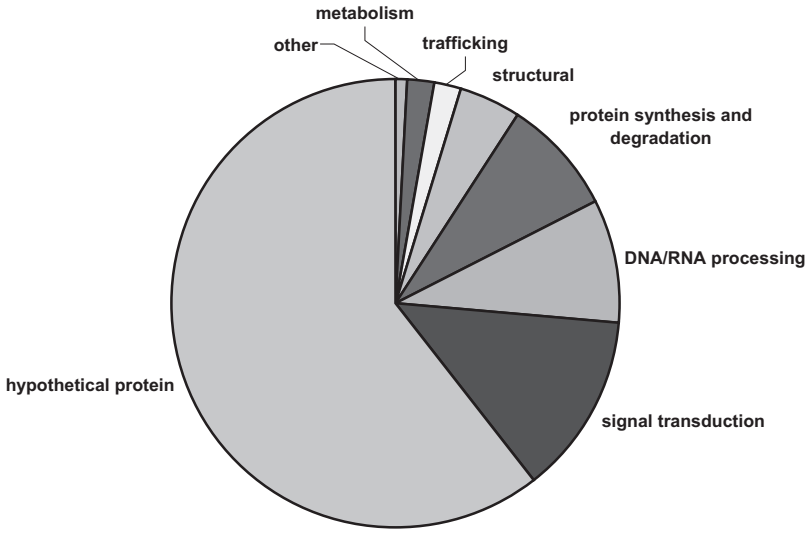


FIG. 3. **Biological processes of *T. brucei* cytosolic phosphoproteins as predicted from their genome annotations.** The majority of phosphoproteins were found to be hypothetical proteins of unknown function with the remainder divided among the indicated cellular functions.

mental Table S2) as well as annotated mass spectra (supplemental Material S2) can be found in the supplemental material and is also accessible via an on-line extract from the *Try*PP-DB.

Analysis of the T. brucei Cytosolic Kinase Phosphoproteome—Manual annotation of phosphopeptides belonging to the kinase subset of phosphoproteins led to the identification of 43 phosphorylated ePKs, one phosphorylated aPK (see Table II), and three phosphorylated non-protein kinases involved in nucleoside/nucleotide metabolism (Table III). The identified protein kinases could be categorized into seven major ePK groups using a multilevel HMM library of the kinase catalytic domain and represented about 25% of all predicted ePKs in *T. brucei.* The majority of kinases identified in our phosphoproteomics screen were members of the CMGC group (18 of 47 predicted members). Five protein kinases were classified as CAMK members (28 predicted), seven protein kinases were found to belong to the AGC group (22 predicted), and another six were designated as STE group members (28 predicted). In addition, we identified three members of the NEK kinase group (20 predicted) and another three protein kinases of the "other" kinase group (40 predicted), and one protein kinase could be assigned to the CK1 group (five predicted members). We also found the *T. brucei* homologue of RACK1 (receptor for activated protein kinase C-1) to be phosphorylated (see Table II). RACK1 is involved in recruiting signaling proteins to specific membrane sites (29) and has been shown to be required for cytokinesis in *T. brucei* (30).

The distribution of all unambiguously assigned phosphorylation sites from 491 proteins was found to be 906 phosphoserine sites (75%), 259 phosphothreonine sites (21.5%), and 39 phosphotyrosine sites (3.5%) as determined by our automated assignment method (see Table IV). Manual spectral validation of the kinase subset of proteins resulted in the identification of 63 phosphoserine, 24 phosphothreonine, and 13 phosphotyrosine residues yielding a serine/threonine/tyrosine distribution of 63/24/13%. In this case, most of the phosphotyrosine sites (12 of 13) were found on members of the CMGC kinase group. Approximately half of the ePKs were found to be multiply phosphorylated, ranging from two distinct phosphorylation sites to a maximum of nine sites found on two proteins (Tb09.211.2410 and Tb10.61.1180) (see Table II).

In this study we also examined phosphorylation site predictions, which have been described in the literature (20, 25, 31), using our *T. brucei* phosphoproteome data set. Several sets of kinase recognition motifs, described as either a regular expression or as a probabilistic model, were applied to our data set, and 95 patterns and motifs specific to individual kinases showed at least one match to *T. brucei* phosphoproteins. The 95 patterns form a redundant set with the same pattern being represented in multiple resources. Likewise there may be many patterns for a single kinase class or

FIG. 2. **Example of an MS^2 spectrum for an auto- and manually validated phosphopeptide (*A*) and an MS^2 spectrum that was only accepted by manual inspection but failed in the automated validation process (*B*).** Shown in *A* is the MS^2 spectrum of the serine-phosphorylated (pS) peptide AGQLSApSTTQGASTLGR of protein kinase Tb927.1.1530 acquired on a QSTAR XL mass spectrometer. Peptide ions derived from b- and y-ion fragmentation are indicated. Neutral loss of phosphoric acid was observed at Ser^{321} (indicated with a *gray box*). Shown in *B* is the fragmentation spectrum of the diphosphorylated peptide GLHVSQPLpTEpYVSTR of protein kinase Tb09.211.0960 measured on a QSTAR XL mass spectrometer. Phosphorylation at the threonine (pT) and tyrosine (pY) residue of the TEY motif could be deduced because of the neutral loss of phosphoric acid starting from the y_7 ion ($-P$) and the observed mass increment of 243 Da ($+P$), respectively. This peptide was rejected by the automated validation method because of the presence of only two singly charged desphospho fragment ions (see Footnote 2). *Rel. Int.*, relative intensity.

TABLE II

Protein and lipid kinases and phosphorylation sites detected in our global phosphoproteome study

Numbers in parentheses indicate the number of observations. The PIKK family kinase Tb927.1.1930 belongs to the aPK superfamily. See Table I legend for group descriptions.

Tb accession number	Protein kinase group	Manually annotated phosphorylation sites
Tb09.211.2410	AGC	Ser^{302} (1),[a,b] Thr^{306} (2), Ser^{307} (1),[b] Ser^{311} (1),[a,b] Tyr^{312} (1),[b] Ser^{315} (2),[a] Ser^{320} (2),[a,b] Thr^{324} (1),[a,b] Ser^{326} (1)[a]
Tb10.70.2260	AGC	Ser^{467} (2)[a,b]
Tb11.01.1030	AGC	Ser^{520} (1)[a,b]
Tb11.02.2210 (PKA-R)	AGC	Ser^{181} (1), Thr^{183} (1), Ser^{218} (1)[a,b]
Tb927.3.2440/Tb03.48o8.470	AGC	Ser^{229} (1)[a,b]
Tb927.7.5770/Tb07.10c21.210	AGC	Ser^{298} (1)[a]
Tb11.03.0340	AGC	Ser^{137} (1),[a,b] Ser^{138} (1)[a,b]
Tb927.4.5310/Tb04.24m18.60	NEK	Ser^{195} (1)[a,b]
Tb10.70.0970	NEK	Thr^{195} (1)
Tb10.70.7860	NEK	Ser^{602} (1)[a]
Tb10.05.0200	CAMK	Ser^{17} (1)[b]
Tb10.70.3410	CAMK	Ser^{659} (1)[a,b]
Tb927.2.1820/27h14.30	CAMK	Ser^{310} (1),[b] Ser^{312} (2),[b] Ser^{314} (1)[a,b]
Tb927.3.4560/Tb03.26j7.970	CAMK	Ser^{482} (1),[a,b] Ser^{486} (1)
Tb09.160.4770	CAMK	Ser^{345} (1),[b] Thr^{385} (1), Ser^{418} (1)
Tb11.02.4860	Other	Ser^{61} (1), Ser^{602} (1),[b] Ser^{604} (1),[b] Thr^{608} (1)[a,b]
Tb927.7.6310/Tb07.2f2.640	Other	Ser^{462} (2),[a,b] Thr^{469} (4)[a,b]
Tb10.61.1180	Other	Ser^{277} (1),[b] Ser^{301} (1),[a,b] Ser^{304} (1),[b] Ser^{951} (1),[a,b] Thr^{953} (1),[a] Thr^{955} (2),[a] Thr^{958} (1), Ser^{960} (1),[b] Ser^{963} (1)[a,b]
Tb927.5.800/Tb05.28f8.750	CK1	Ser^{19} (1)[b]
Tb10.61.3140	CMGC	Ser^{2} (1), Tyr^{187} (2)[b]
Tb11.01.4230	CMGC	Thr^{41} (1)[a,b]
Tb11.02.0640	CMGC	Tyr^{269} (1)
Tb10.70.2210	CMGC	Tyr^{34} (1)[a,b]
Tb10.70.4250	CMGC	Ser^{369} (1)[b]
Tb10.100.0230	CMGC	Ser^{65} (2),[b] Ser^{66} (6),[a,b] Ser^{69} (5),[a] Thr^{70} (3),[a,b] Ser^{72} (5),[b] Tyr^{345} (1)[b]
Tb10.61.0250	CMGC	Tyr^{192} (1)[b]
Tb09.211.4890	CMGC	Ser^{2} (3), Ser^{4} (3), Ser^{5} (2),[b] Thr^{9} (1)
Tb10.61.1850	CMGC	Thr^{163} (1),[b] Tyr^{165} (1),[b] Ser^{329} (1)[a,b]
Tb10.61.1520	CMGC	Tyr^{868} (1)
Tb09.211.0960	CMGC	Thr^{160} (1), Tyr^{162} (1)[b]
Tb11.01.8550	CMGC	Thr^{158} (1),[b] Tyr^{160} (1)[b]
Tb927.7.7360/Tb07.30d13.430	CMGC	Tyr^{57} (1)[a,b]
Tb927.7.3880/Tb07.6c8.130	CMGC	Ser^{13} (1)[a]
Tb927.6.4220/Tb06.26g9.1110	CMGC	Thr^{207} (1), Tyr^{209} (1)[b]
Tb927.6.1780/Tb06.28p18.710	CMGC	Ser^{7} (1),[b] Thr^{186} (2), Ser^{187} (2),[b] Tyr^{188} (2)[b]
Tb927.2.4510/30m24.325	CMGC	Ser^{750} (1)[a,b]
Tb927.3.1850/Tb03.30p12.590	CMGC	Ser^{117} (1)
Tb10.61.2490	STE	Ser^{644} (1)
Tb09.211.2260	STE	Ser^{372} (1),[b] Thr^{374} (1), Ser^{466} (2),[a,b] Thr^{468} (1),[a] Thr^{470} (2)[a,b]
Tb927.1.1530	STE	Ser^{321} (1)[b]
Tb11.01.5880	STE	Ser^{145} (1),[a,b] Ser^{147} (1),[a,b] Ser^{307} (1)[a,b]
Tb927.8.5950/Tb08.11j15.920	STE	Thr^{14} (1)[a,b]
Tb927.3.2060/Tb03.30p12.1060	STE	Ser^{344} (1)[a,b]
Tb927.1.1930	PIKK family	Ser^{1710} (1)[b]
Tb11.01.3170	TRACK	Thr^{10} (1), Ser^{18} (1),[a] Thr^{24} (1)[a,b]

[a] Sites predicted by Scansite.

[b] Sites predicted by NetPhos.

phosphopeptide-binding domain. A detailed list of all pattern match statistics is given in supplemental Table S3. The motifs matching sites that are least frequently occupied (nil or very low phosphorylation) tend to be those matching kinases and systems not identified in the parasite (see supplemental Table S3). In contrast, almost every motif corresponding to kinases known to be present in *T. brucei* shows strong positive hits. A more comprehensive analysis taking into account sequence and interaction context for each protein is beyond the scope of this analysis.

Two broad spectrum phosphosite predictors, NetPhos and Scansite, were also applied to our phosphoproteome data

TABLE III
Metabolic kinases and phosphorylation sites detected in our global phosphoproteome study

Numbers in parentheses indicate the number of observations.

Tb accession number	Protein description	Manually annotated phosphorylation sites
Tb927.6.2360/Tb06.4M18.780	Adenosine kinase, putative	Ser^{344} (1)[a]
Tb09.211.0350	Adenylate kinase, putative	Ser^{134} (1),[a,b] Ser^{136} (1)[a]
Tb11.01.7800	Nucleoside-diphosphate kinase	Thr^{93} (1)

[a] Sites predicted by NetPhos.
[b] Sites predicted by Scansite.

TABLE IV
Phosphorylation site distributions

Ratios of phosphorylated serine/threonine/tyrosine residues occurring on total proteins (all proteins) and the kinases (kinase set) are displayed. Phosphorylation sites were assigned manually for the kinase set and automatically (auto) for all proteins, applying stringent criteria for validation.

	Phosphoserine	Phosphothreonine	Phosphotyrosine
All proteins (auto)	906 (75%)	259 (21.5%)	39 (3.5%)
Kinase set (manual)	63 (63%)	24 (24%)	13 (13%)
Kinase set (auto)	48 (52.7%)	32 (35.2%)	11 (12.1%)

set. Taking the results of the Scansite and NetPhos prediction algorithms together, 75% of all phosphorylation sites for the manually validated kinase subset of proteins were found by the prediction programs using the default settings (see Tables II and III). The fact that 25% of the manually validated sites were not predicted by the programs may indicate that *T. brucei* has some kinases with novel substrate specificity, perhaps providing scope for the identification of new motifs for hitherto uncharacterized kinases.

DISCUSSION

A growing number of phosphoproteome studies have shed light on the extent and function of protein phosphorylation in various organisms and contributed to an increased understanding of complex cell signaling pathways. This study presents the first large scale analyses of a *T. brucei* phosphoproteome and, as expected, confirms the generality of serine/threonine/tyrosine phosphorylation in eukaryotes. We used mass spectrometry-based technologies to analyze 491 *T. brucei* phosphoproteins based on the detection of 852 unique phosphopeptides. Most *T. brucei* phosphorylation sites were detected on serine residues (75%) and threonine residues (21.5%) and to a lesser extent on tyrosine residues (3.5%). In comparison, recent studies using mass spectrometry-based approaches found this ratio to be 86/12/2% in HeLa cells (32) and 70/20/10% in prokaryotic cells (19, 20). Thus, *T. brucei* displays phosphoserine and phosphotyrosine contents more similar to prokaryotes than higher eukaryotes.

Of the 491 *T. brucei* phosphoproteins, 44 kinases were detected. Interestingly about half of the kinases were found to be phosphorylated at more than one site, suggesting that the well established paradigm that protein kinases act as molecular "switchboards" integrating signaling pathways (1) extends to this divergent protozoan parasite. At 182 protein kinases, the kinome of *T. brucei* is larger than those of *S. cerevisiae* (~1.5 times (28)), *S. pombe* (~1.5 times (28)), the malaria parasite (~2 times (33)), and the intracellular parasite *E. cuniculi* (~1.7 times (28)). This suggests that cellular signaling by reversible protein phosphorylation has evolved a sophisticated and central role in *T. brucei*. The protozoan parasite *Toxoplasma gondii* has a larger kinome (356 PKs),[2] but in relation to its proteome, the *T. gondii* complement is smaller (1.5 *versus* 2.0% for *T. brucei*) (see Table I).

Typical protein-protein-interacting domains mediating phosphopeptide recognition and thereby integrating signals within the cell have been shown to be present in *T. brucei*. For example, the phosphothreonine-binding forkhead-associated domain (FHA) is present in one of the *T. brucei* CAMKs (5) and a phosphoserine/threonine-binding polo box domain is present in the parasite polo-like kinase (*Plk*) homologue *Tb*PLK (11, 35). Furthermore isoforms of 14-3-3 proteins that bind to specific phosphoserine and phosphothreonine motifs have been described in *T. brucei* (36). However, common phosphotyrosine-binding domains, such as the Src homology 2 (SH2) domain and phosphotyrosine-binding (PTB) domain, are notably absent in the *T. brucei* genome consistent with the lack of identifiable receptor and non-receptor tyrosine kinases in this organism. On the other hand, we identified 13 phosphorylated tyrosine residues on bloodstream form *T. brucei* protein kinases by our mass spectrometry-based approach. The vast majority of these sites were found on proteins belonging to the CMGC protein kinase group, including cyclin-dependent kinases, MAP kinases, glycogen synthase 3 kinase (GSK3), and DYRKs.

Cell division cycle 2-related protein kinases (CRK1, CRK2, and three CRK3) were found to be phosphorylated on Tyr^{16}, Tyr^{57}, and Tyr^{34}, respectively. Interestingly these phosphorylated tyrosine sites are flanked by Ser^{15} in CRK1, Ser^{56} in CRK2, and Thr^{33} in CRK3, which could correspond to the conserved human CDK1 Thr^{14} and Tyr^{15} residues. Phosphorylation of the conserved Tyr^{15} of CDK1 by the Wee1 dual specificity tyrosine kinase inactivates the CDK complex (37), and as Wee1 kinase homologues are also encoded by the *T. brucei* genome, it is quite possible that CRKs in *T. brucei* are regulated by a similar mechanism.

We also identified phosphorylated tyrosine residues in the activation loops of putative *T. brucei* DYRKs. This family of protein kinases activate themselves by autophosphorylation of a critical tyrosine residue in their activation segment and phosphorylate their substrates on serine and threonine residues (38, 39). DYRKs generally contain an Y*X*Y motif in their activation loop as part of the conserved sequence Y*X*YIQSRFY. In three *T. brucei* DYRKs we found a phosphorylated tyrosine (pY) residue within the sequences FTpYIQSRFY (for Tb11.02.0640), SSpYVQSRCY (for Tb10.100.0230), and HTpYVQSRYY (for Tb10.61.1520). The presence of a phosphorylated tyrosine residue in the putative activation loops of the *T. brucei* DYRK homologues strongly suggests that these are active dual specificity protein kinases in the parasite.

Human GSK3β protein kinase also contains a tyrosine phosphorylation site (Tyr^{216}) in the activation loop generated by an intramolecular autophosphorylation event, which is important for activity (40). The tyrosine phosphorylation site is part of the sequence NVSYICS, and the *T. brucei* GSK3β homologue was found to be phosphorylated at Tyr^{187} within the sequence NVApYICS. Moreover phosphorylation of human GSK3β by p90rsk1 (90-kDa ribosomal S6 kinase-1) *in vivo* on Ser^{9} inactivates the kinase (41). Interestingly the *T. brucei* GSK3β homologue was found to be phosphorylated at the N-terminal Ser^{2} residue, which could represent a comparable GSK3β inhibitory site. Both the detection of a phosphorylated tyrosine residue in the activation loop and a phosphorylated serine residue at the N terminus suggests that the mechanism of kinase activation and inactivation may be similar in *T. brucei* to that in mammals. In addition, we identified the *T. brucei* GSK3α phosphotyrosine residue Tyr^{277} within the sequence NVPpYIFS that corresponds to the human GSK3α Tyr^{279} in NVSpYICS.

We also obtained information on the phosphorylation status of members of the *T. brucei* MAP kinase family. The signature TEY motif specific for members of the ERK and MAPK/MAK/MRK overlapping kinase (MOK) families was found to be phosphorylated in two *T. brucei* kinases (Tb10.61.1850 and Tb11.01.8550), including the previously characterized ERK-like, CRK-like kinase-1 (ECK1) (10). Members of the male germ cell-associated kinase (MAK) and MAK-related kinase (MRK) families as well as KKIALRE and KKIANRE kinases show modest homology to both MAP kinases and CDKs, which all share a TDY motif in the activation loop (42). Our phosphoproteomics studies revealed the dual phosphorylation of the TDY motif of the functionally characterized *T. brucei* MAPK5 (Tb927.6.4220) (43). Interestingly we also detected phosphorylated T*X*Y motifs in *T. brucei* MAP kinases in which the threonine and tyrosine residues were separated by a serine as in Tb927.6.1780 (TSY). The *T. brucei* genome contains several STE family members, which are thought to represent MEKs (STE7), MEK activator kinases MEKKs (STE11), and upstream kinases of MEKKs (STE20) (5). Our study identified six STE kinases, suggesting that this enzyme family is regulated by phosphorylation, and together with the observation of MAP kinase expression indicates that phosphorylation of STE kinases might influence their ability to interact with their MAP kinase substrates.

In summary, we show, for the first time, that known kinase activating and deactivating phosphorylation sites are conserved in *T. brucei* suggesting that at least some segments of known signal transduction pathways are conserved in this organism. This is despite the fact that we currently have no idea what the upstream receptors or downstream effectors of these pathways might be in this parasite. For example, our study shows the expression and phosphorylation of the *D. discoideum* ERK1 homologue kinase (Tb10.61.0250) and also the presence of phosphorylated STE kinases, which points to the existence of an ERK1/2 signal transduction pathway in *T. brucei*. However, canonical signal transduction via ERK1/2 is mediated by receptor tyrosine kinases and G protein-coupled receptors via Raf1, which is a tyrosine kinase-like kinase. The tyrosine kinase-like kinase group and G protein-coupled receptors as well as proteins with phosphotyrosine-binding domains are absent from the *T. brucei* genome, indicating that any trypanosome equivalent of an ERK1/2 pathway must be activated by a completely different mechanism. The study by Parsons *et al.* (5) identified 10 putative receptor-linked protein kinases among which were proteins belonging to the STE kinase family that might represent the upstream activator proteins of *T. brucei* ERKs. However, the extracellular stimuli triggering signal transduction cascades in *T. brucei* are still unknown, making the mapping of stimulus/response networks extremely difficult. Hopefully the data set provided from this work together with the *T. brucei* kinome will assist researchers in designing experiments to map signal transduction pathways in this medically important organism.

Acknowledgments—We thank Dr. Nick Morrice and Dr. Rune Linding for helpful discussions and Dr. Tom Walsh for expert systems support. We also thank Prof. Carol MacKintosh for reading the manuscript and valuable feedback and Prof. Geoff Barton for support.

S The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

§ Supported by a Wellcome Trust Prize Ph.D. studentship.

‖ Supported by a Medical Research Council program grant.

** Supported by a Wellcome Trust Program Grant 085622 and Strategic Award 083481. To whom correspondence should be addressed. Tel.: 44-1382-344219; Fax: 44-1382-348896; E-mail: m.a.j.ferguson@dundee.ac.uk.

REFERENCES

1. Cohen, P. (2000) The regulation of protein function by multisite phosphorylation: a 25 year update. *Trends Biochem. Sci.* **25,** 596–601
2. Hanks, S. K., and Hunter, T. (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9,** 576–596
3. Naula, C., Parsons, M., and Mottram, J. C. (2005) Protein kinases as drug targets in trypanosomes and *Leishmania. Biochim. Biophys. Acta* **1754,** 151–159
4. Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gürtler, R. E., McKerrow, J., Reed, S., and Tarleton, R. (2008) Kinetoplastids: related protozoan

pathogens, different diseases. *J. Clin. Investig.* **118,** 1301–1310
5. Parsons, M., Worthey, E. A., Ward, P. N., and Mottram, J. C. (2005) Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major, Trypanosoma brucei* and *Trypanosoma cruzi. BMC Genomics* **6,** 127
6. Miranda-Saavedra, D., and Barton, G. J. (2007) Classification and functional annotation of eukaryotic protein kinases. *Proteins* **68,** 893–914
7. Hammarton, T. C. (2007) Cell cycle regulation in *Trypanosoma brucei. Mol. Biochem. Parasitol.* **153,** 1–8
8. Tu, X., and Wang, C. C. (2004) The involvement of two cdc2-related kinases (CRKs) in *Trypanosoma brucei* cell cycle regulation and the distinctive stage-specific phenotypes caused by CRK3 depletion. *J. Biol. Chem.* **279,** 20519–20528
9. Tu, X., and Wang, C. C. (2005) Pairwise knockdowns of cdc2-related kinases (CRKs) in *Trypanosoma brucei* identified the CRKs for G1/S and G2/M transitions and demonstrated distinctive cytokinetic regulations between two developmental stages of the organism. *Eukaryot. Cell* **4,** 755–764
10. Ellis, J., Sarkar, M., Hendriks, E., and Matthews, K. (2004) A novel ERK-like, CRK-like protein kinase that modulates growth in Trypanosoma brucei via an autoregulatory Cterminal extension. *Mol. Microbiol.* **53,** 1487–1499
11. Hammarton, T. C., Kramer, S., Tetley, L., Boshart, M., and Mottram, J. C. (2007) Trypanosoma brucei Polo-like kinase is essential for basal body duplication, kDNA segregation and cytokinesis. *Mol. Microbiol.* **65,** 1229–1248
12. Hall, B. S., Gabernet-Castello, C., Voak, A., Goulding, D., Natesan, S. K., and Field, M. C. (2006) TbVps34, the trypanosome orthologue of Vps34, is required for Golgi complex segregation. *J. Biol. Chem.* **281,** 27600–27612
13. Pawson, T., and Scott, J. D. (2005) Protein phosphorylation in signaling: 50 years and counting. *Trends Biochem. Sci.* **30,** 286–290
14. Mehlert, A., Treumann, A., and Ferguson, M. A. (1999) *Trypanosoma brucei* GPEETPARP is phosphorylated on six out of seven threonine residues. *Mol. Biochem. Parasitol.* **98,** 291–296
15. da Cunha, J. P., Nakayasu, E. S., Elias, M. C., Pimenta, D. C., Téllez-Iñón, M. T., Rojas, F., Muñoz, M. J., Manuel, M., Almeida, I. C., and Schenkman, S. (2005) Trypanosoma cruzi histone H1 is phosphorylated in a typical cyclin dependent kinase site accordingly to the cell cycle. *Mol. Biochem. Parasitol.* **140,** 75–86
16. Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., and White, F. M. (2002) Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae. Nat. Biotechnol.* **20,** 301–305
17. Nühse, T. S., Stensballe, A., Jensen, O. N., and Peck, S. C. (2003) Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. *Mol. Cell. Proteomics* **2,** 1234–1243
18. Moser, K., and White, F. M. (2006) Phosphoproteomic analysis of rat liver by high capacity IMAC and LC-MS/MS. *J. Proteome Res.* **5,** 98–104
19. Macek, B., Mijakovic, I., Olsen, J. V., Gnad, F., Kumar, C., Jensen, P. R., and Mann, M. (2007) The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis. Mol. Cell. Proteomics* **6,** 697–707
20. Macek, B., Gnad, F., Soufi, B., Kumar, C., Olsen, J. V., Mijakovic, I., and Mann, M. (2008) Phosphoproteome analysis of *E. coli* reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. *Mol. Cell. Proteomics* **7,** 299–307
21. Leal, S., Acosta-Serrano, A., Morita, Y. S., Englund, P. T., Böhme, U., and Cross, G. A. (2001) Virulence of *Trypanosoma brucei* strain 427 is not affected by the absence of glycosylphosphatidyl phospholipase C. *Mol. Biochem. Parasitol.* **114,** 245–247
22. Cross, G. A. (1984) *J. Cell. Biochem.* **24,** 79–90
23. Beausoleil, S. A., Jedrychowski, M., Schwartz, D., Elias, J. E., Villén, J., Li, J., Cohn, M. A., Cantley, L. C., and Gygi, S. P. (2004) Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 12130–12135
24. Aurrecoechea, C., Brestelli, J., Brunk, B. P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O. S., Heiges, M., Innamorato, F., Iodice, J., Kissinger, J. C., Kraemer, E., Li, W., Miller, J. A., Nayak, V., Pennington, C., Pinney, D. F., Roos, D. S., Ross, C., Stoeckert, C. J., Jr., Treatman, C., and Wang, H. (2009) PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Res.* **37,** D539–543
25. Puntervoll, P., Linding, R., Gemünd, C., Chabanis-Davidson, S., Mattingsdal, M., Cameron, S., Martin, D. M., Ausiello, G., Brannetti, B., Costantini, A., Ferrè, F., Maselli, V., Via, A., Cesareni, G., Diella, F., Superti-Furga, G., Wyrwicz, L., Ramu, C., McGuigan, C., Gudavalli, R., Letunic, I., Bork, P., Rychlewski, L., Küster, B., Helmer-Citterich, M., Hunter, W. N., Aasland, R., and Gibson, T. J. (2003) ELM server: a new resource for investigating short functional sites in modular eukaryotic proteins. *Nucleic Acids Res.* **31,** 3625–3630
26. Rice, P., Longden, I., and Bleasby, A. (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* **16,** 276–277
27. Diella, F., Cameron, S., Gemünd, C., Linding, R., Via, A., Kuster, B., Sicheritz-Pontén, T., Blom, N., and Gibson, T. J. (2004) Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins. *BMC Bioinformatics* **5,** 79
28. Miranda-Saavedra, D., Stark, M. J., Packer, J. C., Vivares, C. P., Doerig, C., and Barton, G. J. (2007) The complement of protein kinases of the microsporidium *Encephalitozoon cuniculi* in relation to those of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe. BMC Genomics* **8,** 309
29. Usacheva, A., Smith, R., Minshall, R., Baida, G., Seng, S., Croze, E., and Colamonici, O. (2001) The WD motif-containing protein receptor for activated protein kinase C (RACK1) is required for recruitment and activation of signal transducer and activator of transcription 1 through the type I interferon receptor. *J. Biol. Chem.* **276,** 22948–22953
30. Rothberg, K. G., Burdette, D. L., Pfannstiel, J., Jetton, N., Singh, R., and Ruben, L. (2006) The RACK1 homologue from *Trypanosoma brucei* is required for the onset and progression of cytokinesis. *J. Biol. Chem.* **281,** 9781–9790
31. Obenauer, J. C., Cantley, L. C., and Yaffe, M. B. (2003) Scansite 2.0: proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res.* **31,** 3635–3641
32. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127,** 635–648
33. Ward, P., Equinet, L., Packer, J., and Doerig, C. (2004) Protein kinases of the human malaria parasite *Plasmodium falciparum*: the kinome of a divergent eukaryote. *BMC Genomics* **5,** 79
34. Gajria, B., Bahl, A., Brestelli, J., Dommer, J., Fischer, S., Gao, X., Heiges, M., Iodice, J., Kissinger, J. C., Mackey, A. J., Pinney, D. F., Roos, D. S., Stoeckert, C. J., Jr., Wang, H., and Brunk, B. P. (2008) ToxoDB: an integrated Toxoplasma gondii database resource. *Nucleic Acids Res.* **36,** D553–556
35. Kumar, P., and Wang, C. C. (2006) Dissociation of cytokinesis initiation from mitotic control in a eukaryote. *Eukaryot. Cell* **5,** 92–102
36. Inoue, M., Nakamura, Y., Yasuda, K., Yasaka, N., Hara, T., Schnaufer, A., Stuart, K., and Fukuma, T. (2005) The 14-3-3 proteins of Trypanosoma brucei function in motility, cytokinesis, and cell cycle. *J. Biol. Chem.* **280,** 14085–14096
37. Krupa, A., Preethi, G., and Srinivasan, N. (2004) Structural modes of stabilization of permissive phosphorylation sites in protein kinases: distinct strategies in Ser/Thr and Tyr kinases. *J. Mol. Biol.* **339,** 1025–1039
38. Kentrup, H., Becker, W., Heukelbach, J., Wilmes, A., Schürmann, A., Huppertz, C., Kainulainen, H., and Joost, H. G. (1996) Dyrk, a dual specificity protein kinase with unique structural features whose activity is dependent on tyrosine residues between subdomains VII and VIII. *J. Biol. Chem.* **271,** 3488–3495
39. Lochhead, P. A., Sibbet, G., Kinstrie, R., Cleghon, T., Rylatt, M., Morrison, D. K., and Cleghon, V. (2003) dDYRK2: a novel dual-specificity tyrosine phosphorylation regulated kinase in Drosophila. *Biochem. J.* **374,** 381–391
40. Cole, A., Frame, S., and Cohen, P. (2004) Further evidence that the tyrosine phosphorylation of glycogen synthase kinase-3 (GSK3) in mammalian cells is an autophosphorylation event. *Biochem. J.* **377,** 249–255
41. Stambolic, V., and Woodgett, J. R. (1994) Mitogen inactivation of glycogen synthase kinase-3 beta in intact cells via serine 9 phosphorylation. *Biochem. J.* **303,** 701–704
42. Miyata, Y., Akashi, M., and Nishida, E. (1999) Molecular cloning and characterization of a novel member of the MAP kinase superfamily. *Genes Cells* **4,** 299–309

43. Domenicali Pfister, D., Burkard, G., Morand, S., Renggli, C. K., Roditi, I., and Vassella, E. (2006) A mitogen-activated protein kinase controls differentiation of bloodstream forms of *Trypanosoma brucei. Eukaryot. Cell* **5,** 1126–1135
44. Hertz-Fowler, C., Peacock, C. S., Wood, V., Aslett, M., Kerhornou, A., Mooney, P., Tivey, A., Berriman, M., Hall, N., Rutherford, K., Parkhill, J., Ivens, A. C., Rajandream, M. A., and Barrell, B. (2004) GeneDB: a resource for prokaryotic and eukaryotic organisms. *Nucleic Acids Res.* **32,** D339–343
45. Flicek, P., Aken, B. L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T., Down, T., Dyer, S. C., Eyre, T., Fitzgerald, S., Fernandez-Banet, J., Gräf, S., Haider, S., Hammond, M., Holland, R., Howe, K. L., Howe, K., Johnson, N., Jenkinson, A., Kähäri, A., Keefe, D., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Megy, K., Meidl, P., Overduin, B., Parker, A., Pritchard, B., Prlic, A., Rice, S., Rios, D., Schuster, M., Sealy, I., Slater, G., Smedley, D., Spudich, G., Trevanion, S., Vilella, A. J., Vogel, J., White, S., Wood, M., Birney, E., Cox, T., Curwen, V., Durbin, R., Fernandez-Suarez, X. M., Herrero, J., Hubbard, T. J., Kasprzyk, A., Proctor, G., Smith, J., Ureta-Vidal, A., and Searle, S. (2008) Ensembl 2008. *Nucleic Acids Res.* **36,** D707–714
46. Hirschman, J. E., Balakrishnan, R., Christie, K. R., Costanzo, M. C., Dwight, S. S., Engel, S. R., Fisk, D. G., Hong, E. L., Livstone, M. S., Nash, R., Park, J., Oughtred, R., Skrzypek, M., Starr, B., Theesfeld, C. L., Williams, J., Andrada, R., Binkley, G., Dong, Q., Lane, C., Miyasato, S., Sethuraman, A., Schroeder, M., Thanawala, M. K., Weng, S., Dolinski, K., Botstein, D., and Cherry, J. M. (2006) Genome Snapshot: a new resource at the Saccharomyces Genome Database (SGD) presenting an overview of the Saccharomyces cerevisiae genome. *Nucleic Acids Res.* **34,** D442–445
47. Katinka, M. D., Duprat, S., Cornillot, E., Méténier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretaillade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J., and Vivarès, C. P. (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi. Nature* **414,** 450–453