Protein arginine and lysine methylation in *Trypanosoma cruzi*

Rafael Fogaça de Almeida[1], Aline Castro Rodrigues Lucena[2], Michel Batista[2,3], Fabricio Klerynton Marchini[2,3], Lyris Martins Franco de Godoy[1,*]

[1] Laboratório de Biologia Molecular e Sistêmica de Tripanossomatídeos, Instituto Carlos Chagas, Fundação Oswaldo Cruz, Curitiba, 81350010, Parana, Brazil; [2] Laboratório de Ciências e Tecnologias Aplicadas em Saúde, Instituto Carlos Chagas, Fundação Oswaldo Cruz, Curitiba, 81350010, Parana, Brazil; [3] Plataforma de Espectrometria de Massas, Instituto Carlos Chagas, Fundação Oswaldo Cruz, Curitiba, 81350010, Parana, Brazil.

* Corresponding author:

Lyris Martins Franco de Godoy; Rua Prof. Algacyr Munhoz Mader, 3775, CIC 81350-010 Curitiba/PR, Brasil; Tel.:+55(41)3316-3230; Fax: +55(41)3316-3267

e-mail: lyris.godoy@fiocruz.br

Abbreviations: ABC, ammonium bicarbonate; ADMA, asymmetric dimethyl-arginine; ATP, Adenosine triphosphate; BP, Biological Process; CC, Cellular component; COG, Clusters of Orthologous Groups; DTT, 1,4-Dithiothreitol; ESI, Eletrospray ionization; GO, Gene ontology; HPLC, High performance liquid chromatography; IAA, Iodoacetamide; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC-MS/MS, Liquid chromatography with tandem mass spectrometry; MF, Molecular function; MMA, monomethyl-arginine; MS, Mass spectrometry; MeCN, acetonitrile; PBS, Phosphate-buffered saline; PKMT, protein lysine methyltransferase; PRMT, protein arginine methyltransferase; PTM, Post-translational modification; Rme1, monomethyl-arginine; Rme2a, asymmetric dimethyl-arginine; Rme2s, symmetric dimethyl-arginine; RBP, RNA binding protein; SAM, S-adenosylmethionine; SDMA, symmetric dimethyl-arginine; SDS, Sodium dodecyl sulfate.

Abstract

Post-translational methylation of proteins, which occurs in arginines and lysines, modulates several biological processes at different levels of cell signaling. Recently, methylation has been demonstrated in the regulation beyond histones, for example, in the dynamics of protein-protein interaction and protein-nucleic acid. However, the presence and role of protein methylation in *Trypanosoma cruzi*, the etiologic agent of Chagas disease, has not yet been elucidated. In this work, we applied mass spectrometry based-proteomics (LC-MS/MS) to describe the arginine and lysine methylproteome of *T. cruzi*. In epimastigotes, 1336 methylsites in 878 methylated proteins were identified by LC-MS/MS. The analyzes show that protein methylation impacts different biological processes, with emphasis on translation. Separately, protein arginine methylation is related to oxireduction and carbohydrate metabolism, while lysine methylation impacts the protein synthesis. In addition, 50 methylated proteins have been previously described with phosphorylation sites in *T. cruzi*, represented by RNA binding proteins, sterol methyltransferase activity and calpain peptidases. This work represents the first proteomic analysis of *T. cruzi* methylproteome and is the first to characterize lysine methylation in trypanosomatids. Collectively, these data inform about new fundamental biological aspects of this organism impacted by protein methylation.

Keywords: LC-MS/MS; Protein methylation. Proteomics; Trypanosoma cruzi.

Statement of the significance

Trypanosoma cruzi is a protozoan parasite that causes Chagas' disease in humans and throughout its life cycle faces different environment changes and protein methylation is an important post-translational modification by which cells respond and adapt to the environment. We applied a mass spectrometry-based proteomics and report the first proteomic analysis of both arginine and lysine methylproteome in *T. cruzi* being the first large-scale characterization of lysine methylation in trypanosomes. Our data demonstrate that the methylation of proteins in *T. cruzi* is broad and impacts different cellular processes, mainly the protein synthesis. This study represents a significant advance on the importance of protein methylation in trypanosomes.

1. Introduction

Post-translational modifications (PTMs) impact different biological functions and significantly contribute to cellular homeostasis and environment adaptation [1]. Among them, an important PTM is protein methylation [2,3], which occurs in lysine and arginine residues and impacts fundamental cellular events, from gene transcription and RNA processing to protein translation and cell signaling [4].

Arginine and lysine methylation are catalyzed by S-adenosylmethionine (SAM)-dependent protein arginine methyltransferases (PRMTs) and protein lysine methyltransferases (PKMTs), respectively [5]. In humans, the most common PRMT is the type I, which catalyzes the formation of monomethyl-arginine (MMA/Rme1) transferring a methyl group from SAM to guanidino nitrogen, and can add a second methyl group at the same nitrogen, resulting an asymmetric dimethyl-arginine (ADMA/Rme2a) [6]. Type II PRMTs also catalyze the synthesis of MMA and can add one second methyl group at the terminal nitrogen adjacent, resulting in symmetric dimethyl-arginine (SDMA/Rme2s) [6]. Protein lysine methyltransferases (PKMTs) catalyze the addition of up to three methyl groups at the ε-amino group of lysine residue (Kme1, Kme2 and Kme3). Most known human PRMTs and PKMTs can methylate both histone and non-histone proteins [7] and are potential drug targets due to the fundamental roles in the cell biology [8].

The kinetoplastid protozoan parasite *Trypanosoma cruzi* is the etiologic agent of Chagas disease [9], an illness that is estimated to affect about 6 to 7 million people worldwide [10] and represents a significant burden to the health-care system [11]. No vaccine is available and treatment is carried out with inefficient and highly toxic drugs [12]. In addition, *T. cruzi* is a parasite that has a complex life cycle (such another trypanosomes like *T. brucei* and *Leishmania* sp.) with different stages of development and different hosts therefore is necessary to adapt to different conditions and for that is an interesting organism to study the biological impact of PTMs. Recently, the presence of a variety of PTMs in *T. cruzi* histones has been demonstrated, with acetylation and methylation being the most abundant types in these proteins [13,14].

Large-scale of non-histone methylation analysis although relatively well known in higher eukaryotes [15–19], in protozoan parasites are scarce and the function of arginine and lysine methylation in these organisms is still poorly understood, especially for the kinetoplastids. In *Plasmodium falciparum*, arginine [20] and lysine [21] methylated proteins are involved in diverse biological pathways, for example, RNA metabolism, protein synthesis, transport, proteolysis, protein folding and chromatin organization. In the deep-branching *Giardia duodenalis* [22], interestingly, methyl-arginines (and arginine-methyltransferases) are absent, but their biological functions appear to be compensated by methyl-lysines, which are regulated between its different life-cycle stages. For the TriTryps, studies about protein methylation are only available for *Leishmania* and *T. brucei*. For *Leishmania*, only 19 methylated proteins were identified in *L. donovani*, among them are RNA binding proteins, ribosomal proteins, Elongation factor 1-alpha [23] and, more recently, 40 putative PRMT7 targets were described in *L. major* [24], including 17 RNA binding proteins (RBPs), introducing the importance of arginine methylation in the RNA metabolism of this parasite. In *T. brucei*, approximately 10% of the proteome contains methyl-arginines, which potentially impact diverse cellular pathways in this parasite [25,26], but no global study about lysine methylation is available. Here, we applied a bottom-up proteomic approach (GeLC-MS/MS) to comprehensively characterize the *T. cruzi* methylproteome of arginine and lysine residues. Using mass spectrometry (MS) and bioinformatics analyses we identified more than a thousand of methylsites present in proteins involved in many biological functions in *T. cruzi* Furthermore, we compare the functions of both arginine and lysine methylproteomes. It sheds light to important processes that are potentially regulated by protein methylation and provides another promising landscape to understanding the biology of this pathogen.

2. Materials and Methods

2.1 Cell culture

T. cruzi Dm28c epimastigotes in exponential growth phase were cultured in Liver Infusion Tryptose (LIT) medium [27] supplemented with 10% fetal bovine serum and incubated without agitation at 28 °C. Exponential epimastigotes were obtained in the order of 4.3 x10^9 cells.

2.2 Protein extraction, separation and digestion

Cells were washed in PBS, resuspended in lysis buffer (4% SDS, 10 mM Tris-HCl pH 7.5, 100 mM DTT) in a proportion of 240 µL/3 x 10^8 cells, vortexed for 30 seconds, heated for 3 min at 95 °C and sonicated for 1 hour at room temperature. To remove debris, samples were then centrifuged at 20,000 xg at 20 °C for 5 min and the supernatant (protein extract) was transferred to a clean tube. *T. cruzi* protein extract (25 ug) was separated by SDS-PAGE (5 - 20% acrylamide) in five lanes (5 µg/lane) and stained with Coomassie Blue. Each gel lane was sliced horizontally into 10 fractions covering the different molecular weight ranges and each fraction was cut into 1 mm^3 cubes, which were transferred to a clean microfuge tube and submitted to in-gel digestion as previously described [28]. Briefly, the gel pieces were destained twice with 25 mM ammonium bicarbonate (ABC) in 50% ethanol, dehydrated with 100% ethanol, reduced with 10 mM DTT in 50 mM ABC, alkylated with 55 mM iodoacetamide in 50 mM ABC and digested with 12.5 ng/µL trypsin in 50 mM ABC and incubated for 16h at 37ºC. Digestion was stopped by adding trifluoroacetic acid (TFA) to a final concentration of 0.5%.Peptides were extracted twice with 30% acetonitrile (MeCN) in 3% TFA and twice with 100% MeCN, them dried in a Speed Vac (Thermo Scientific) and desalted using C18 StageTips [29] prior to nanoLC- ESI-MS/MS.

2.3 NanoLC-ESI-MS/MS analysis

Tryptic peptides were separated by online reverse phase (RP) nanoscale capillary liquid chromatography (nanoLC) and analyzed by electrospray mass spectrometry in tandem (ESI-MS/MS). The samples were analyzed at the mass spectrometry facility P02-004/Carlos Chagas Institute - Fiocruz Parana with an Easy-nLC 1000 system (Thermo Fisher Scientific) connected to an LTQ-Orbitrap XL ETD mass spectrometer (Thermo Fisher Scientific) equipped with a nanoelectrospray ion source (Phoenix S&T). Peptide mixtures were injected in a 30 cm analytical column (75 µm inner diameter) in-house packed with C18 resin (ReproSil-Pur C18-AQ 2.4 µm), eluted from 5 to 40% MeCN in 5% DMSO and 0.1% formic acid in a 120 min gradient. The nanoLC column was heated at 60 °C to improve both chromatographic resolution and reproducibility. Peptides were ionized by nanoelectrospray (voltage, 2.7 kV) and injected into the mass spectrometer. Mass spectra were obtained by Data-Dependent Acquisition (DDA), with an initial Orbitrap scan with resolution R = 15,000 followed by MS/MS of the 10 most intense ions in LTQ (Iontrap). These precursor ions are fragmented by Collision-Induced Dissociation (CID) with normalized collision energy = 35%, activation time of 30 ms and activation q = 0.25. Singly charged precursor ions were rejected. Parallel to the MS2 was conducted a full scan in Orbitrap with a resolution R = 60,000 (mass range m/z 300-2000) and selected ions were dynamically excluded for 90 seconds. The lock mass option [30], in presence of DMSO peaks [31], was used in all full scans to improve mass accuracy of precursor ions.

2.4 Search and identification of peptide/protein

The analysis of the raw data from the LC-MS/MS was performed by MaxQuant platform [32], version 1.5.2.8. The proteins were identified using a database with a total of 19,242 protein sequences of *T. cruzi* CL Brener, downloaded on January 04, 2016 from UniProt

(www.uniprot.org). Contaminants (human keratins, BSA and porcine trypsin) were added to the database, as well as their reverse sequences. The search for methylated sites follows the criteria: MS tolerance of 20 ppm (Orbitrap), MS/MS tolerance of 0.5 Da (Iontrap), allowing for 2 missed cleavages. The peptides length searched with at least 7 amino acids. Trypsin with specific cleavage. Carbamidomethylation of cysteines was determined as fixed modification. Monomethylation and dimethylation of lysines and arginines and trimethylation of lysines were searched as variable modifications, as well as oxidation of methionines and N-terminal acetylation. A false discovery rate (FDR) of 1% was applied to the protein, peptide and methylsite level.

2.5 Bioinformatics analyses of *T. cruzi* methylproteome

The functional classification and enrichment analyzes of Gene Ontology (GO)[33], Clusters of Orthologous Groups (COG) [34] and Kyoto Encyclopedia of Genes and Genomes (KEGG)[35] were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) tool [36,37], version 6.8 (https://david.ncifcrf.gov). For the enrichment analysis, FDR was applied for multiple tests correction , as proposed by Benjamini and Hochberg [38] with cutoff ≤ 0.05. Protein consensus sequences analyses were performed using the iceLogo [39] with the *T. cruzi* Swiss-Prot proteome as reference set. The protein-protein interactions were visualized using STRING[40] in Cytoscape (v.3.8) [41].

3. Results

3.1 GeLC-MS/MS analysis reveals abundant arginine and lysine methylation in *T. cruzi* epimastigotes

We applied a typical bottom-up proteomic approach to characterize the methylproteome of arginine and lysine residues of *T. cruzi* epimastigotes (Fig. 1A). Our analysis of the total *T. cruzi* protein extract resulted in the acquisition of 116405 MS/MS spectra, 51323 (approximately 44%) of which were identified when confronted with the database. After exclusion of the identified contaminants and reverse sequences, a total of 20577 peptides and 4133 proteins were identified. Among them, 878 methylated proteins (Table S1) and total of 1336 methylsites (Table S2) were detected, 1263 (94.5%) with a high localization probability score (≥0.75) according with MaxQuant analysis [42] . Among the methylated proteins identified (878), the majority was identified bearing monomethylation (Fig. 1B), being monomethyl sites (58.3%) the most abundant, followed by dimethyl (24.8%) and trimethyl sites (5.6%). Moreover, some methylated proteins were identified with more than one type of methylation: mono and dimethylation sites (5.2%), mono and trimethyl sites (0.3%), di and trimethyl sites (5.1%) and mono, di and trimethylation sites (0.5%). Among the methylation sites (Fig. 1C), 52% represent monomethylation (363 Kme, 337 Rme), 38% dimethylation (187 Kme2, 320 Rme2) and 10% trimethylation (129 Kme3). Overall, methylation sites were almost equally distributed between arginine (657; 49%) and lysine (679; 51%) residues (Fig. 1D).

3.2 *T. cruzi* arginine and lysine methylsites are surrounded by different amino acid pattern

In order to identify amino acids patterns surrounding methylsites in *T. cruzi*, we investigate the amino acids frequencies surrounding sites and different amino acids are more frequent in the vicinity of arginine (Fig. 2A) and lysine (Fig. 2B). Methyl-arginines are enriched upstream for methionine (M/Met), cysteine (C/Cys) and arginine (Arg) and downstream for arginines (R/Arg). Methyl-lysine sites are enriched for glutamic acid (E/Glu at 192 positions -1, +2, +3 and +4) and cysteine (C/Cys at positions -1, -2, -3 and -4 positions).

3.3 Methylated proteins impact numerous processes in *T. cruzi*

In order to identify the pathways potentially regulated by methylated proteins in *T. cruzi* epimastigotes, we performed functional classification and enrichment analysis of proteins based on GO, COG and KEGG terms. For cellular component (CC) analysis the GO classification terms were available for a subset of 285 (32.4%) of the methylated proteins, distributed in 59 (Table S3) and enriched in 5 (Table S4) different classes. The methylated proteins are localized across the entire cell, membranes, cytoplasm, nucleus and other organelles (Fig. S1A), while the enrichment analysis revealed that they are mainly involved with ribosomes, nucleosomes and cytoplasm (Fig. 2C). For the biological process (BP) analysis, the GO classification terms were available for 295 (33.6%) of the methylated proteins, distributed in 143 (Table S3) and enriched in 11 (Table S4) different classes. Methylation proteins of *T. cruzi* was mainly related to translation, protein folding, cellular movement, pathogenesis, membrane and intracellular transport, redox metabolism and a wide variety of other processes (Fig. S1B). For the enriched biological processes are spotlighted the response to stress, metabolism of amino acids and carboxylic acid, among others, such translation and oxidation-reduction process (Fig. 2D). For the molecular function (MF) analysis, the GO classification terms were available for a subset of 415 (47,2%) of the methylated proteins, distributed in 210 (Table S3) and enriched in 8 (Table S4) different classes. In agreement with the cellular component and biological processes results, the most represented (Fig. S1C) and enriched (Fig. 2C) molecular functions affected by

methylation included nucleic acid, metabolite and ion binding, among others, which are ultimately related to protein synthesis, cellular redox homeostasis and other metabolic and catabolic processes.

The analysis of the *T. cruzi* methylproteome against the other two systematic gene functional databases (COG and KEGG) support the multiple functions associated to methylated proteins described above. The COG classification showed clusters of proteins with functions commonly associated to methylproteomes analysis in other species [26], such RNA metabolism and intracellular trafficking and cytoskeleton (Fig. S2A). The KEGG classification identified multiple pathways associated to the methylated proteins (Fig. S2B) and, interestingly, biosynthesis of secondary metabolites and antibiotics are also well represented.

In order to ensure the comparison of our data with previously reported, we reanalyzed the *T. brucei* data [26] with the same tools and parameters used for *T. cruzi* in the present work. As expected, the methylated proteins in arginine are associated with the cytoskeleton and involved with locomotion, including the presence of kinesins. In comparison with *T. brucei*, the processes impacted by the arginine methylproteome in *T. cruzi* presents marked differences (Table 1).

3.4 Lysine methylation impact different processes compared to arginine methylation in *T. cruzi*

To evaluate whether different cellular functions are being regulated by the arginine and lysine methylproteomes differently, we performed enrichment analysis with the both subproteomes separately. As can be seen, a wider variety of functional categories have been enriched among the lysine-methylated proteins in relation to the arginine-methylated proteins. In addition, the terms enriched are indeed different between the two subproteomes. Arginine methylation mainly impacts oxidation processes and metabolism of carbohydrates, as illustrated by the enrichment of classes such as *oxidation-reduction* and *carbohydrate metabolic process* (BP), *Proton-transporting two-sector ATPase complex* and *catalytic domain* (CC) and *Pyridoxal phosphate*

binding (MF), among others (Table 1). On the other hand, lysine methylation appears to be more strongly associated with the translation and other protein-related processes, as indicated by the enrichment of several related classes, such as amino acid transport and metabolism (COG), translation (BP), ribosome (CC), structural constituent of ribosome (MF) and ribosome (KEGG) (Table 2).

3.5 Protein-protein interaction analysis present a global view of the methylproteome network

In order to evaluate the protein network impacted by methylated proteins in *T. cruzi*, we performed protein-protein interaction (PPIs) analysis. The total proteins identified (878) only a subset of 187 proteins presents interaction annotation in the STRING database. Proteins bearing K or R methylation (or both) vastly interact with each other (Fig. 3). The highest degree of interaction occurs for methylated proteins related to translation, such as ribosomal proteins and translation initiation factors. Interestingly, clusters of protein interaction bearing only one type of methylation were observed. For example, the network related to aminoacyl-tRNA synthesis, which is mainly R-Me modified, similarly to that shown for othereukaryotes [15] and the small ones related to splicing and fatty acid metabolism, which have only K-methylated proteins. In addition, the interaction network of the K-Me and R-Me subproteomes separately (Fig. S4), to which 95 and 93 proteins could be mapped, respectively. Despite the similar number of proteins mapped on both maps, the clusters of the R-Me map are more connected than the clusters of the K-Me map, which has smaller and isolated clusters. Detailed data about protein interaction analysis are available in Table S5.

3.6 Methylation presents co-occurrence with phosphorylation in *T. cruzi* proteins

We investigate the co-occurrence of methylation and other PTMs, we compared our methylproteome with the phosphoproteome of *T. cruzi* available in the literature. Of the 878 methylated proteins identified in our study, 50 were also found to be phosphorylated in epimastigotes [43] among them there are cytoskeleton proteins, kinases, RNA-binding proteins, calpain cysteine peptidases and hypothetical conserved proteins (Table 3).

To evaluate which pathways and processes are related to the proteins that present co-occurrence of methylation and phosphorylation, we performed a functional clustering analysis based on the annotations in the INTERPRO, SMART, GO and KEGG databases, using the DAVID [37] tools. The analysis identified three enriched functional clusters (Table S6). In the first cluster there are proteins involved in RNA binding and belong to family proteins with RNA recognition motifs (RRM), the second cluster groups proteins related to sterol methyltransferase activity and the third one represents proteins with cysteine peptidase activity.

4. Discussion

PTMs can be found across the entire cell and are involved in various processes in eukaryotes [44]. Nonetheless, in trypanosomes only a few large-scale studies have been conducted to elucidate the global impact of specific PTMs at the cellular level, mainly for phosphorylation [43,45–48] and acetylation [49]. Methylproteome characterization studies have generally analyzed methylation in arginine (R) and lysine (K) separately and different processes have been identified impacted by the modification in these different residues. To date, while R-methylated proteins are involved in RNA processing, transcriptional regulation and repair of DNA damage, K-methylation has been extensively described in many histone residues, playing a role in the regulation of chromatin compaction and gene transcription. In our work we characterize methylation in both residues and reported the first large-scale analysis of lysine and arginine methylation of *T. cruzi* proteins.

Our data shows that different types of protein methylation are widely found in *T. cruzi*, were identified monomethylated, dimethylated and trimethylated proteins. These methylated proteins are located in different cell compartments and present a wide range of functions, suggesting an influence of the methylation in several biological processes of the parasite, including metabolism, oxireduction, protein folding, and RNA metabolism, were identified here, for the first time, in large-scale.

Under the conditions applied in this study, arginine and lysine methylation occurred almost in equal abundance in the substrates, however the biological processes that impact on *T. cruzi* are different. Arginine-methylated proteins are involved in several processes, in particular, oxireduction and carbohydrate metabolism, whereas lysine-methylated proteins are mainly involved in the protein synthesis. This last one is fundamental to differential gene regulation processes in trypanosomes. Due the post transcriptional events regulating the gene expression in *T. cruzi*, the presence of RNA binding proteins is even more crucial. RNA binding proteins are known substrates of PRMT in different organisms [18,26,50–52]. We identified proteins involved in RNA processing with methylsites and, analyzing data previously reported, these proteins also have phosphosites (Table 3). Lott and colleagues [26] reported RNA processing impacted by arginine methylated proteins in *T. brucei* procyclic form, while the main represented function was proteins associated with cytoskeleton and locomotion. Here, the main function represented by the majority of methylated proteins identified was protein synthesis, and, interestingly, mainly in lysine methylated proteins. Indeed, the same PTM can affect different processes in the cell, thereby this dissimilarity can be explained by the differences in parasite's biology in a PTM level. For example, recently, Schenkman et al. [49] analyzing lysine acetylation in trypanosomes, revealed that protein acetylation is involved in very distinct set of acetylated proteins when comparing *T. cruzi* and *T. brucei.* Therefore, these functional differences between methylproteomes (and acetylomes) are compatible to the broad role of PTMs in cell signaling already demonstrated in other eukaryotes and that seems to also exist between trypanosomes [44].

Methylated proteins also impact important process like RNA processing, that is responsible for direct assembly of multiprotein complexes on primary transcripts, mature mRNAs, and stable ribonucleoprotein components of the RNA processing machinery [53]. Recently, Amorim and colleagues [47] demonstrated that EF-1-α and EF-2 are phosphorylated in *T. cruzi* during the final phase of metacyclogenesis. This demonstrates that different post-translational modifications are present and may be regulating different forms of *T. cruzi* throughout their life cycle as a response to environment changes. Here, we identified 9 proteins methylated in *translation elongation factor activity* GO-MF, among them elongation factor 1-alpha (EF-1-α) (Tc00.1047053511367.360), elongation factor 1-gamma (EF-1-gamma) (Tc00.1047053510163.20), elongation factor 2 (EF-2) (Tc00.1047053508169.20) and transcription elongation factor-like protein (Tc00.1047053507715.30). These elongation factors are key pieces of translation process and, especially EF-1- α, may act beyond the canonical process in eukaryotes, including in *T. cruzi* [54] and therefore can be potential drug targets, such as EF-2 in *P. falciparum* [55].

Other important class of protein is kinases, that are key mediators of signal transduction and we detected methylsites in 18 proteins annotated with kinase function, among them *serine/threonine* and *protein kinase activity* classes with 11 and 7 proteins, respectively (Table S3). Interestingly, 3 of proteins related to kinases identified here (Q4E2T4, Q4D9W8 and Q4CXF6) present GO related to transmembrane and integral component of membrane. An *in silico* approach also found kinases with annotation of a transmembrane domain in *T. brucei*, suggesting that kinases, if located on the surface of trypanosomatids, can phosphorylate host molecules or parasites to modify its environment [56].

Our analysis of amino acids patterns surrounding methylation sites reveal different residues in K-Me and R-Me. The vicinity of lysine methylated sites are enriched for glutamic acids (E/Glu), something also seen in *P. falciparum* [21] and *Giardia duodenalis* [22] and arginine methylated sites are enriched upstream for arginine (R/Arg) and downstream for methionine (M/Met). Well-defined patterns of amino acids, such “RGG”, are not always surrounding targets of PRMTs [57]

and, likewise, are not always defined for PKMT [58]. Our data for *T. cruzi* reinforces that the writers target motifs seem to be broader in methyl-arginine and for methyl-lysine the motifs seem to carry a certain ancestry of older eukaryotes.

Another important aspect of PTMs in cell signaling is the crosstalk [44]. It is known, for example, that arginine methylation affects other methylated residues, such as lysines themselves, and other modifications such as acetylation and phosphorylation [50]. In addition, there is an interrelationship between the methylation of lysines 4 and 79 of histone H3 (H3K4 and H3K79) with the ubiquitination of lysine 123 of histone H2B (H2BK123) in *Saccharomyces cerevisiae*, where the non-ubiquitination of K123 prevents methylation of K4 and K79 [59]. There is also antagonism between arginine methylation and serine phosphorylation in the C-terminal domain of RNA polymerase II, which impacts the transcription of specific genes in mammals [60]. Among the effects of lysine methylation and interaction with phosphorylation, crosstalk between lysine 810 (K810) methylation in retinoblastoma tumor suppressor protein, that preventing recognition by a cyclin-dependent kinase (Cdk) in serine residue, does not allow the phosphorylation of this substrate, controlling the progression of the cell cycle [61].

We found sites of methylation and phosphorylation in 50 proteins when compared with data from the *T. cruzi* phosphoproteome, suggesting the occurrence of different types of PTMs even in non-histone proteins, this leads to believe that there is a potential crosstalk between these modifications working to adjust the function of the protein. Although the methylation and phosphorylation sites are not so close in the primary sequence (Table 3) they may be close within the tertiary structure of the protein and together influence the interaction with other binding molecules or they may be further away to prevent interference, so that modifying enzymes can independently bind to the respective sites [62]. Our functional clusterization analysis of proteins with methylation and phosphorylation sites grouped with highest stringency three clusters related to RNA binding proteins, sterol methyltransferase activity and calpain peptidases (Table S6). Related to first cluster the presence of proteins methylated involved with RNA metabolism is commonly reported and in *T. brucei* where was identified 10 protein with RRM [26], the

proteins identified in *T. cruzi* are all arginine-methylated proteins. These proteins interact with RNA typically through aromatic residues within the RNP1 and RNP2 domains located on conserved β-sheets and are directly involved in RNA processing [53,63]. In this regard, the methylation typically impacts mainly the protein-protein interaction and subcellular localization, but does not directly impact protein-RNA interactions. On the other hand, with phosphorylation site occurring in the same protein, the protein function or recruitment could be, perhaps, modified. The second cluster have proteins with sterol methyltransferase activity, these enzymes are involved in the ergosterol and related 24-alkyl sterols metabolism. The steroid biosynthesis is a target of inhibitors for the treatment of both Chagas's disease and leishmaniasis [64,65]. The third cluster have proteins from the family of calpains, which is an intracellular protease involved in many cellular functions that are regulated by calcium. Our results of methylproteome and the analysis of previously phosphoproteome data of *T. cruzi* provide candidates to study of potential crosstalk between methylation and phosphorylation in *T. cruzi*. Further characterization is necessary to identify such crosstalk, for example through protein interaction assays, to identify partners linking the specific sites of methylation and phosphorylation [66].

Collectively, our data gave another status for protein methylation in the biology of *T. cruzi* and indicate that it has great potential to be at the level of other PTMs, such acetylation and phosphorylation, acting in the cell regulation. Finally, our data show that fundamental processes are being regulated by protein methylation in *T. cruzi* and that this modification should be further investigated, in order to reveal key parts in the biology of this parasite and potential chemotherapeutic candidates.

5. Associated Data

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [67] partner repository with the dataset identifier PXD034140.

References

1. Karve, T. M., Cheema, A. K. (2011). Small Changes Huge Impact: The Role of Protein Posttranslational Modifications in Cellular Homeostasis and Disease. *Journal of Amino Acids*, *2011*(207691), 1–13. https://doi.org/10.4061/2011/207691

2. Paik, W. K., Paik, D. C., Kim, S. (2007). Historical review: the field of protein methylation. *Trends in Biochemical Sciences*, *32*(3), 146–152. https://doi.org/10.1016/j.tibs.2007.01.006

3. Paik, W. K., Kim, S. (1971). Protein methylation. *Science.*, *174*(6), 114–119. https://doi.org/10.1016/0955-0674(93)90080-A

4. Murn, J., Shi, Y. (2017). The winding path of protein methylation research: milestones and new frontiers. *Nature Reviews Molecular Cell Biology*, *18*(8), 517–527. https://doi.org/10.1038/nrm.2017.35

5. Boriack-Sjodin, P. A., Swinger, K. K. (2016). Protein methyltransferases: A distinct, diverse and dynamic family of enzymes. *Biochemistry*, *55*(11), 1557–1569. https://doi.org/10.1021/acs.biochem.5b01129

6. Morales, Y., Cáceres, T., May, K., Hevel, J. M. (2016). Biochemistry and regulation of the protein arginine methyltransferases (PRMTs). *Archives of Biochemistry and Biophysics*, *590*, 138–152. https://doi.org/10.1016/j.abb.2015.11.030

7. Biggar, K. K., Li, S. S.-C. (2014). Non-histone protein methylation as a regulator of cellular signalling and function. *Nature Reviews Molecular Cell Biology*, *16*(1), 5–17. https://doi.org/10.1038/nrm3915

8. Li, X., Wang, C., Jiang, H., Luo, C. (2019). A patent review of arginine methyltransferase inhibitors (2010–2018). *Expert Opinion on Therapeutic Patents*, *29*(2), 97–114. https://doi.org/10.1080/13543776.2019.1567711

9. Chagas, C. (1909). Nova tripanozomiaze humana: estudos sobre a morfolojia e o ciclo evolutivo do Schizotrypanum cruzi n. gen., n. sp., ajente etiolojico de nova entidade morbida do homem. *Memórias Do Instituto Oswaldo Cruz*, *1*(2), 159–218. https://doi.org/10.1590/S0074-02761909000200008

10. WHO. (2021). *American Trypanosomiasis (Chagas Disease)*. WHO (World Health Organization). http://www.who.int/en/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis)

11. Bartsch, S. M., Avelis, C. M., Asti, L., Hertenstein, D. L., Ndeffo-Mbah, M., Galvani, A., Lee, B. Y. (2018). The economic value of identifying and treating Chagas disease patients earlier and the impact on Trypanosoma cruzi transmission. *PLOS Neglected Tropical Diseases*, *12*(11), e0006809. https://doi.org/10.1371/journal.pntd.0006809

12. Pérez-Molina, J. A., Molina, I. (2018). Chagas disease. *The Lancet*, *391*(10115), 82–94. https://doi.org/10.1016/S0140-6736(17)31612-4

13. Picchi, G. F. A., Zulkievicz, V., Krieger, M. A., Zanchin, N. T., Goldenberg, S., de Godoy, L. M. F. (2017). Post-translational Modifications of Trypanosoma cruzi Canonical and Variant Histones. *Journal of Proteome Research*, *16*(3), 1167–1179. https://doi.org/10.1021/acs.jproteome.6b00655

14. de Almeida, R. F., Fernandes, M., de Godoy, L. M. F. (2021). An updated map of

Trypanosoma cruzi histone post-translational modifications. *Scientific Data*, *8*(1), 93. https://doi.org/10.1038/s41597-021-00818-w

15. Larsen, S. C., Sylvestersen, K. B., Mund, A., Lyon, D., Mullari, M., Madsen, M. V, … Nielsen, M. L. (2016). Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells. *Science Signaling*, *9*(443), 1–15. https://doi.org/10.1126/scisignal.aaf7329

16. Lanouette, S., Mongeon, V., Figeys, D., Couture, J. (2014). The functional diversity of protein lysine methylation. *Molecular Systems Biology*, *10*(4), 724. https://doi.org/10.1002/msb.134974

17. Wang, K., Zhou, Y. J., Liu, H., Cheng, K., Mao, J., Wang, F., … Zou, H. (2015). Proteomic analysis of protein methylation in the yeast Saccharomyces cerevisiae. *Journal of Proteomics*, *114*, 226–233. https://doi.org/10.1016/j.jprot.2014.07.032

18. Uhlmann, T., Geoghegan, V. L., Thomas, B., Ridlova, G., Trudgian, D. C., Acuto, O. (2012). A Method for Large-scale Identification of Protein Arginine Methylation. *Molecular & Cellular Proteomics*, *11*(11), 1489–1499. https://doi.org/10.1074/mcp.M112.020743

19. Cao, X., Arnaudo, A. M., Garcia, B. A. (2013). Large-scale global identification of protein lysine methylation in vivo. *Epigenetics*, *8*(5), 477–485. https://doi.org/10.4161/epi.24547

20. Zeeshan, M., Kaur, I., Joy, J., Saini, E., Paul, G., Kaushik, A., … Malhotra, P. (2017). Proteomic Identification and Analysis of Arginine-Methylated Proteins of Plasmodium falciparum at Asexual Blood Stages. *Journal of Proteome Research*, *16*(2), 368–383. https://doi.org/10.1021/acs.jproteome.5b01052

21. Kaur, I., Zeeshan, M., Saini, E., Kaushik, A., Mohmmed, A., Gupta, D., Malhotra, P. (2016). Widespread occurrence of lysine methylation in Plasmodium falciparum proteins at asexual blood stages. *Scientific Reports*, *6*(35432), 1–24. https://doi.org/10.1038/srep35432

22. Emery-Corbin, S. J., Hamey, J. J., Ansell, B. R. E., Balan, B., Tichkule, S., Stroehlein, A. J., … Jex, A. R. (2020). Eukaryote-Conserved Methylarginine Is Absent in Diplomonads and Functionally Compensated in Giardia. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msaa186

23. Rosenzweig, D., Smith, D., Myler, P. J., Olafson, R. W., Zilberstein, D. (2008). Post-translational modification of cellular proteins during Leishmania donovani differentiation. *Proteomics*, *8*, 1843–1850. https://doi.org/10.1002/pmic.200701043

24. Ferreira, T. R., Dowle, A. A., Parry, E., Alves-ferreira, E. V. C., Hogg, K., Kolokousi, F., … Walrad, P. B. (2020). PRMT7 regulates RNA-binding capacity and protein stability in Leishmania parasites. *Nucleic Acids Research*, *48*(10), 5511–5526. https://doi.org/10.1093/nar/gkaa211

25. Fisk, J. C., Li, J., Wang, H., Aletta, J. M., Qu, J., Read, L. K. (2013). Proteomic Analysis Reveals Diverse Classes of Arginine Methylproteins in Mitochondria of trypanosomes. *Molecular & Cellular Proteomics*, 302–311. https://doi.org/10.1074/mcp.M112.022533

26. Lott, K., Li, J., Fisk, J. C., Wang, H., Aletta, J. M., Qu, J., Read, L. K. (2013). Global proteomic analysis in trypanosomes reveals unique proteins and conserved cellular processes impacted by arginine methylation. *Journal of Proteomics*, *91*, 210–225. https://doi.org/10.1016/j.jprot.2013.07.010

27. Camargo, E. P. (1964). Growth and differentiation in Trypanosoma cruzi. I. Origin of metacyclic trypanosomes in liquid media. In *Rev Inst Med trop São Paulo* (Vol. 6, pp. 93–100). http://www.imtsp.fm.usp.br/portal/portal/images/pdf/revista/vol06/93-100.pdf

28. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V, Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, *1*(6), 2856–2860. https://doi.org/10.1038/nprot.2006.468

29. Rappsilber, J., Mann, M., Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols*, *2*(8), 1896–1906. https://doi.org/10.1038/nprot.2007.261

30. Olsen, J. V, de Godoy, L. M. F., Li, G., Macek, B., Mortensen, P., Pesch, R., … Mann, M. (2005). Parts per Million Mass Accuracy on an Orbitrap Mass Spectrometer via Lock Mass Injection into a C-trap. *Molecular & Cellular Proteomics*, *4*(12), 2010–2021. https://doi.org/10.1074/mcp.T500030-MCP200

31. Hahne, H., Pachl, F., Ruprecht, B., Maier, S. K., Klaeger, S., Helm, D., … Kuster, B. (2013). DMSO enhances electrospray response, boosting sensitivity of proteomic experiments. *Nature Methods*, *10*(10), 989–991. https://doi.org/10.1038/nmeth.2610

32. Cox, J., Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*(12), 1367–1372. https://doi.org/10.1038/nbt.1511

33. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., … Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, *25*, 25–29.

34. Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V, … Natale, D. A. (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*, *14*, 1–14. https://doi.org/10.1186/1471-2105-4-41

35. Kanehisa, M., Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *28*(1), 27–30. https://doi.org/10.1093/nar/28.1.27

36. Huang, D. W., Sherman, B. T., Lempicki, R. A. (2009). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, *37*(1), 1–13. https://doi.org/10.1093/nar/gkn923

37. Huang, D. W., Sherman, B. T., Lempicki, R. A. (2009). Systematic and integrative analysis of

large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211

38. Benjamini, Y., Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*, *57*(1), 289–300. https://doi.org/10.2307/2346101

39. Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., Gevaert, K. (2009). Improved visualization of protein consensus sequences by iceLogo. *Nature Methods*, *6*(11), 786–787. https://doi.org/10.1038/nmeth1109-786

40. Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., … Mering, C. von. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, *47*(D1), D607–D613. https://doi.org/10.1093/nar/gky1131

41. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., … Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303

42. Olsen, J. V, Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., Mann, M. (2006). Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell*, *127*, 635–648. https://doi.org/10.1016/j.cell.2006.09.026

43. Marchini, F. K., de Godoy, L. M. F., Rampazzo, R. C. P., Pavoni, D. P., Probst, C. M., Gnad, … Krieger, M. A. (2011). Profiling the Trypanosoma cruzi Phosphoproteome. *PLoS ONE*, *6*(9), 1–10. https://doi.org/10.1371/journal.pone.0025381

44. Minguez, P., Parca, L., Diella, F., Mende, D. R., Kumar, R., Helmer-Citterich, M., … Bork, P. (2012). Deciphering a global network of functionally associated post-translational

modifications. *Molecular Systems Biology*, *8*(599), 1–14. https://doi.org/10.1038/msb.2012.31

45. Urbaniak, M. D., Martin, D. M. a, Ferguson, M. a J. (2013). Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of Trypanosoma brucei. *Journal of Proteome Research*, *12*(5), 2233–2244. https://doi.org/10.1021/pr400086y

46. Nakayasu, E. S., Gaynor, M. R., Sobreira, T. J. P., Ross, J. A., Almeida, I. C. (2009). Phosphoproteomic analysis of the human pathogen Trypanosoma cruzi at the epimastigote stage. *Proteomics*, *9*(13), 3489–3506. https://doi.org/10.1002/pmic.200800874

47. Amorim, J. C., Batista, M., Cunha, E. S., Lucena, A. C. R., De Paula Lima, C. V, Sousa, K., … Marchini, F. K. (2017). Quantitative proteome and phosphoproteome analyses highlight the adherent population during Trypanosoma cruzi metacyclogenesis. *Scientific Reports*, *7*(9899), 1–12. https://doi.org/10.1038/s41598-017-10292-3

48. Lucena, A. C. R., Amorim, J. C., de Paula Lima, C. V., Batista, M., Krieger, M. A., de Godoy, L. M. F., Marchini, F. K. (2019). Quantitative phosphoproteome and proteome analyses emphasize the influence of phosphorylation events during the nutritional stress of Trypanosoma cruzi: the initial moments of in vitro metacyclogenesis. *Cell Stress and Chaperones*, *24*(5), 927–936. https://doi.org/10.1007/s12192-019-01018-7

49. Moretti, N. S., Cestari, I., Anupama, A., Stuart, K., Schenkman, S. (2018). Comparative Proteomic Analysis of Lysine Acetylation in Trypanosomes. *Journal of Proteome Research*, *17*(1), 374–385. https://doi.org/10.1021/acs.jproteome.7b00603

50. Bedford, M. T., Clarke, S. G. (2009). Protein arginine methylation in mammals: who, what, and why. *Molecular Cell*, *33*(1), 1–13. https://doi.org/10.1016/j.molcel.2008.12.013

51. Boisvert. (2003). A Proteomic Analysis of Arginine-methylated Protein Complexes. *Molecular*

& Cellular Proteomics, 1319–1330. https://doi.org/10.1074/mcp.M300088-MCP200

52. Blackwell, E., Ceman, S. (2012). Arginine methylation of RNA-binding proteins regulates cell function and differentiation. *Molecular Reproduction and Development*, *79*(3), 163–175. https://doi.org/10.1002/mrd.22024

53. Varani, G., Nagai, K. (1998). RNA recognition by RNP proteins during RNA processing. *Annu. Rev. Biophys. Biomol. Struct.*, *27*(1), 407–445. https://doi.org/10.1146/annurev.biophys.27.1.407

54. Alves, L. R., Oliveira, C., Goldenberg, S. (2015). Eukaryotic translation elongation factor-1 alpha is associated with a specific subset of mRNAs in Trypanosoma cruzi. *BMC Microbiology*, *15*(104), 1–9. https://doi.org/10.1186/s12866-015-0436-2

55. Cowell, A. N., Winzeler, E. A. (2019). Advances in omics-based methods to identify novel targets for malaria and other parasitic protozoan infections. *Genome Medicine*, *11*(63), 1–17. https://doi.org/https://doi.org/10.1186/s13073-019-0673-3

56. Parsons, M., Worthey, E. A., Ward, P. N., Mottram, J. C. (2005). Comparative analysis of the kinomes of three pathogenic trypanosomatids: Leishmania major, Trypanosoma brucei and Trypanosoma cruzi. *BMC Genomics*, *6*(127), 1–19. https://doi.org/10.1186/1471-2164-6-127

57. Wooderchak, W. L., Zang, T., Zhou, Z. S., Acuña, M., Tahara, S. M., Hevel, J. M. (2008). Substrate Profiling of PRMT1 Reveals Amino Acid Sequences That Extend Beyond the “RGG” Paradigm. *Biochemistry*, *47*(36), 9456–9466. https://doi.org/10.1021/bi800984s

58. Bremang, M., Cuomo, A., Agresta, A. M., Stugiewicz, M., Spadotto, V., Bonaldi, T. (2013). Mass spectrometry-based identification and characterisation of lysine and arginine methylation in the human proteome. *Molecular BioSystems*, *9*(9), 2231. https://doi.org/10.1039/c3mb00009e

59. Farooq, Z., Banday, S., Pandita, T. K., Altaf, M. (2016). The many faces of histone H3K79 methylation. *Mutation Research. Reviews in Mutation Research*, *6*(2), 356–372. https://doi.org/10.1016/j.mrrev.2016.03.005

60. Sims, R. J., Rojas, L. A., Beck, D., Bonasio, R., Schüller, R., Drury, W. J., … Reinberg, D. (2011). The C-terminal domain of RNA polymerase II is modified by site-specific methylation. *Science (New York, N.Y.)*, *332*(6025), 99–103. https://doi.org/10.1126/science.1202663

61. Carr, S. M., Munro, S., Kessler, B., Oppermann, U., La Thangue, N. B. (2011). Interplay between lysine methylation and Cdk phosphorylation in growth control by the retinoblastoma protein. *The EMBO Journal*, *30*(2), 317–327. https://doi.org/10.1038/emboj.2010.311

62. Korkuć, P., Walther, D. (2017). Towards understanding the crosstalk between protein post-translational modifications: homo- and heterotypic PTM pair distances on protein surfaces are not random. *Proteins*, *85*(1), 78–92. https://doi.org/https://doi.org/10.1002/prot.25200

63. Maris, C., Dominguez, C., Allain, F. H. T. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS Journal*, *272*(9), 2118–2131. https://doi.org/10.1111/j.1742-4658.2005.04653.x

64. Braga, M. V, Magaraci, F., Lorente, S. O., Gilbert, I., de Souza, W. (2005). Effects of Inhibitors of Delta24(25)-Sterol Methyl Transferase on the Ultrastructure of Epimastigotes of Trypanosoma cruzi. *Microscopy and Microanalysis*, *11*(6), 506–515. https://doi.org/10.1017/S143192760505035X

65. Magaraci, F., Jimenez, C. J., Rodrigues, C., Rodrigues, J. C. F., Braga, M. V., Yardley, V., … Gilbert, I. H. (2003). Azasterols as Inhibitors of Sterol 24-Methyltransferase in Leishmania Species and Trypanosoma cruzi. *Journal of Medical Chemistry*, *24*(25), 4714–

4727. https://doi.org/10.1021/jm021114j

66. Li, W., Wang, H.-Y., Zhao, X., Duan, H., Cheng, B., Liu, Y., … , Du, H.-N. (2019). A methylation-phosphorylation switch determines Plk1 kinase activity and function in DNA damage repair. *Science Advances*, *5*(3), eaau7566. https://doi.org/10.1126/sciadv.aau7566

67. Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., … Vizcaíno, J. A. (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Research*, *50*(D1), D543–D552. https://doi.org/10.1093/nar/gkab1038

Acknowledgments

The authors thank Fiocruz for using the Technological Platforms Network. Also thank CNPQ, CAPES and Fiocruz for financial support.

Conflict of Interest

The authors declare no conflict of interest.

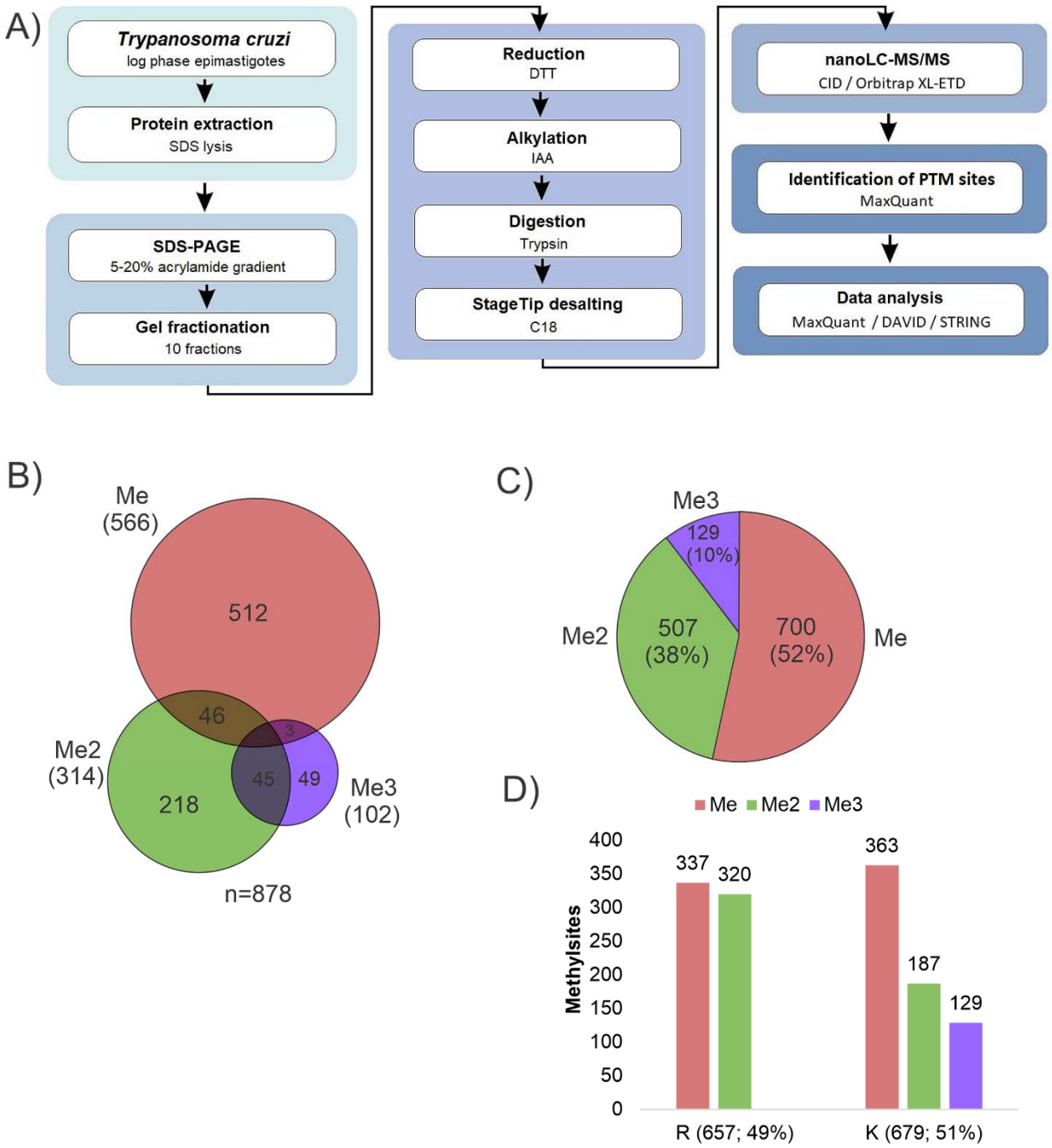


Figure 1. Workflow and global numbers of *T. cruzi* methylproteome A) Proteomic workflow applied to characterize *T. cruzi* arginine and lysine methylproteomes. B) Venn diagram showing the overlap of methylation types (Me, Me2 and Me3) in the identified methylated proteins; C) Distribution of the identified

methylation sites for the methylation type (Me, Me2 and Me3); D) Distribution of the methylation sites (Me, Me2 and Me3) in the amino acids arginine (R) and lysine (K).

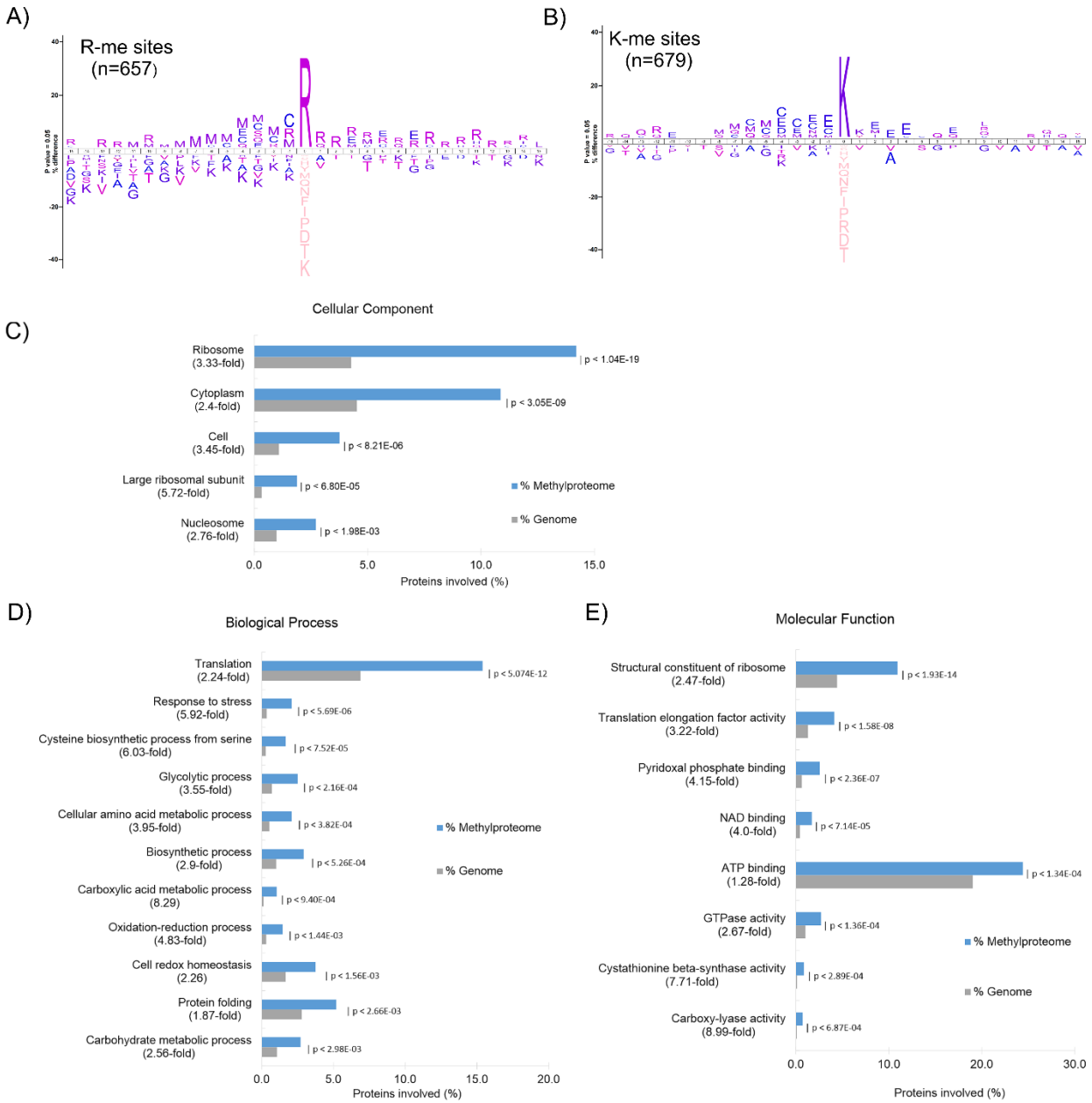


Figure 2. Characterization of *T. cruzi* methylproteome. Frequency of residues in modification window ±15 residues from the (A) arginine and lysine (B) modification site, only significant amino acids are shown (p-value 0.05). The central aminoacid (R or K at position 0) indicates the site modified in peptides. All of the flanking amino acid residues are indicated with the letters differing in height, with the differences

corresponding to their frequencies in the respective position. GO-based enrichment analysis of methylated proteins according to cellular localization (C), biological processes (D) and molecular functon (E). P<0.05 indicates a significant difference.

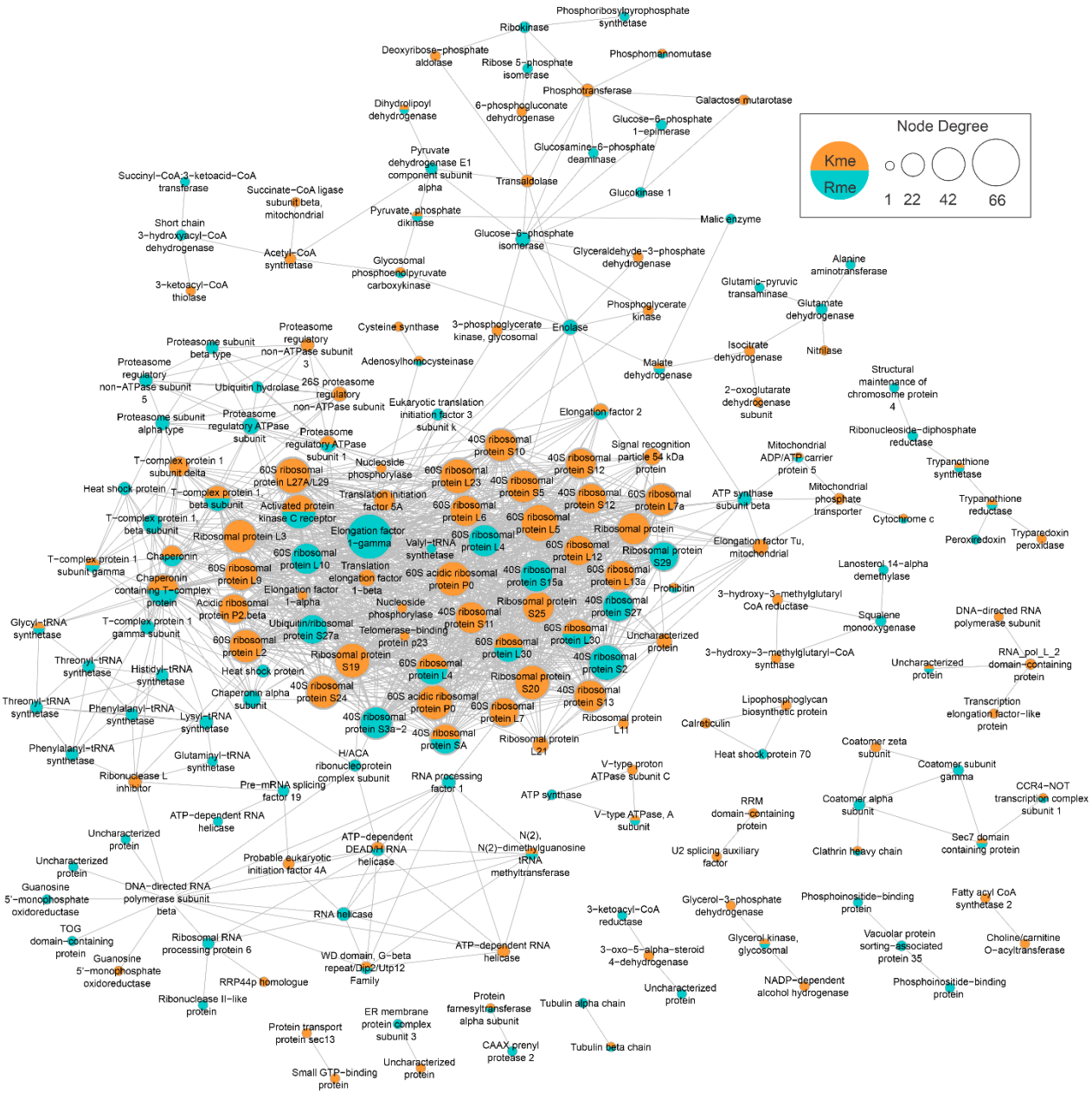


Figure 3. Interaction network of methylated proteins in *T. cruzi.* Visualization of functional connection between methylated proteins bearing K-Me sites (orange) or (and) R-Me sites (turquoise). Node degree correspond to number of interactions of the protein. Detailed data are listed in Table S5.

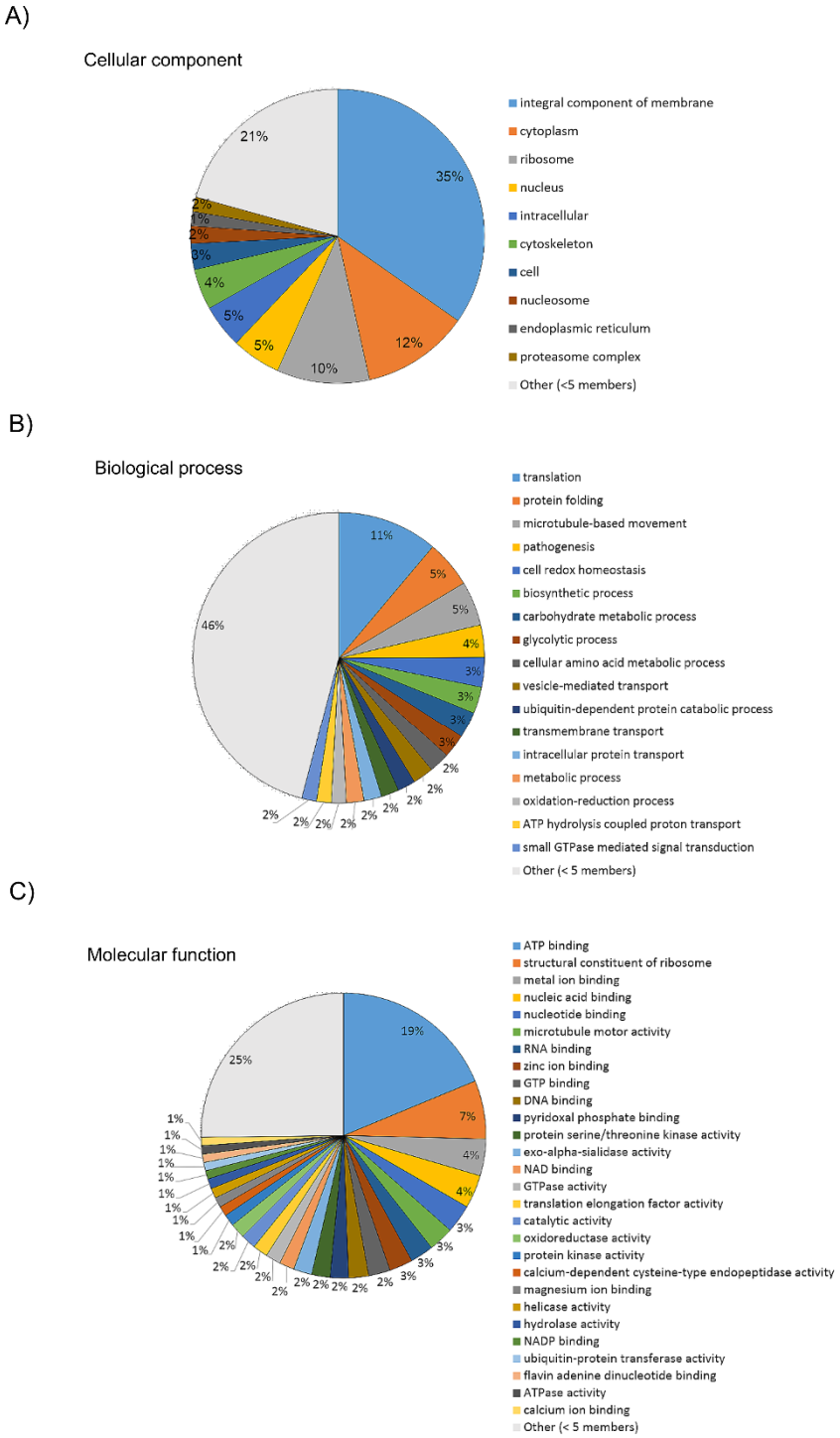


Figure S1. Functional classification of *T. cruzi* methylated proteins. Methylproteome classification accordingly with cellular component (A), biological process (B) and molecular function (C). Classes containing ≥5 proteins were illustrated individually, while classes containing <5 proteins were grouped in the "other" category. The complete classification list is available in Table S3.

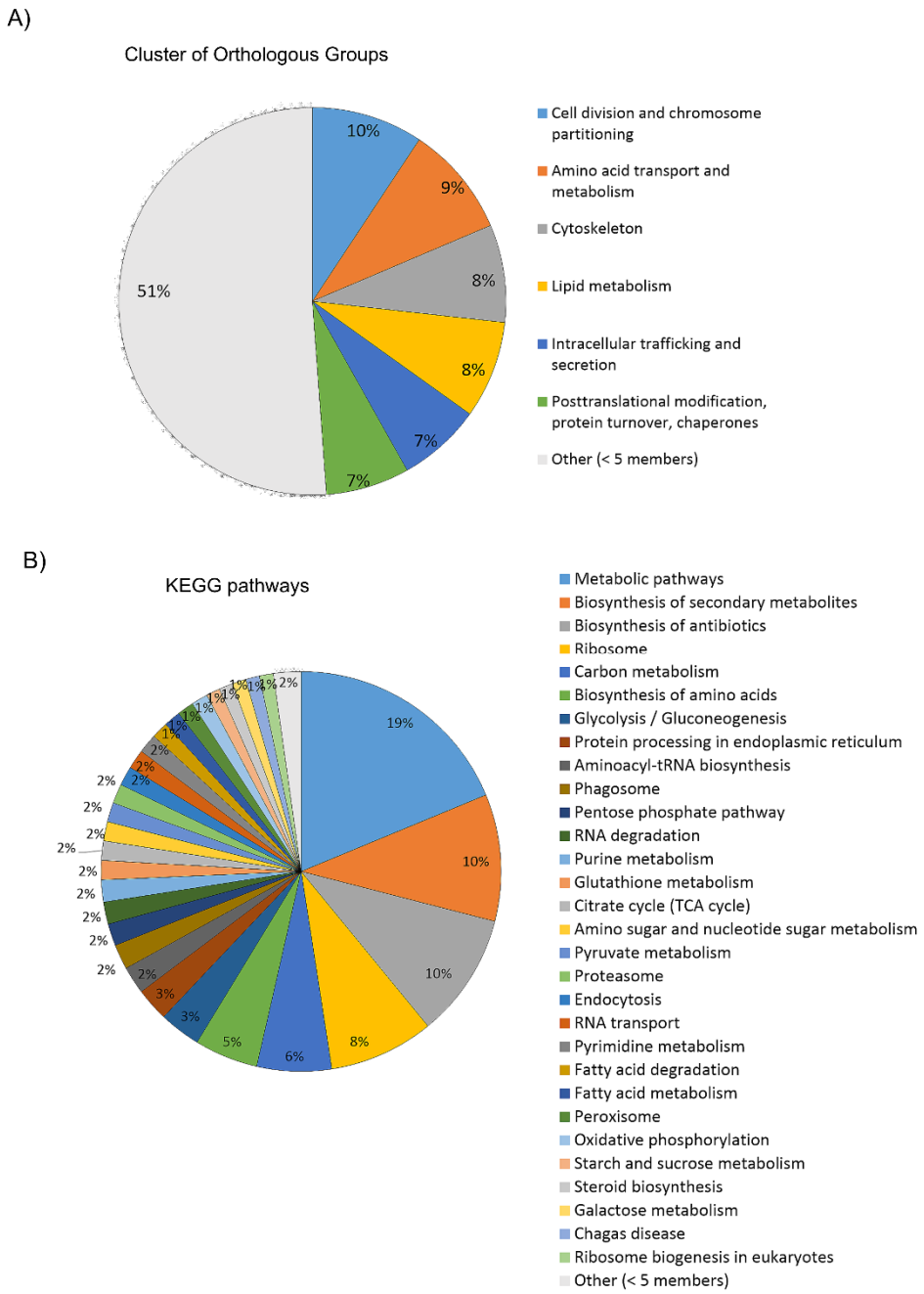


Figure S2. Functional pathways of *T. cruzi* methylated proteins. Classification of methylated proteins identified in this study with their terms predicted for the Clusters of Orthologous Groups (A) and for the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (B). Classes containing ≥5 proteins were illustrated individually, while classes containing <5 proteins were grouped in the "other" category. The complete classification list is available in Table S3.

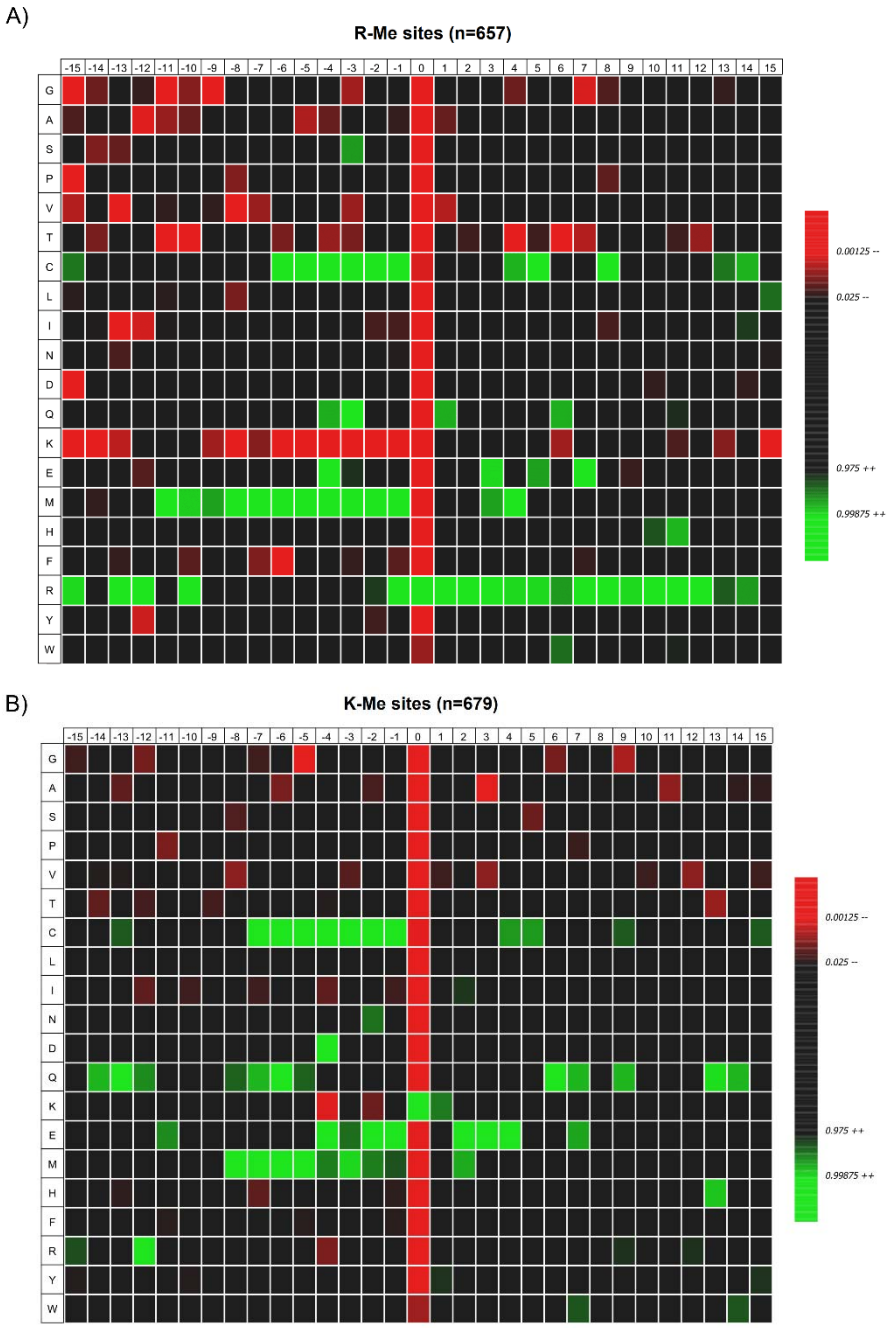
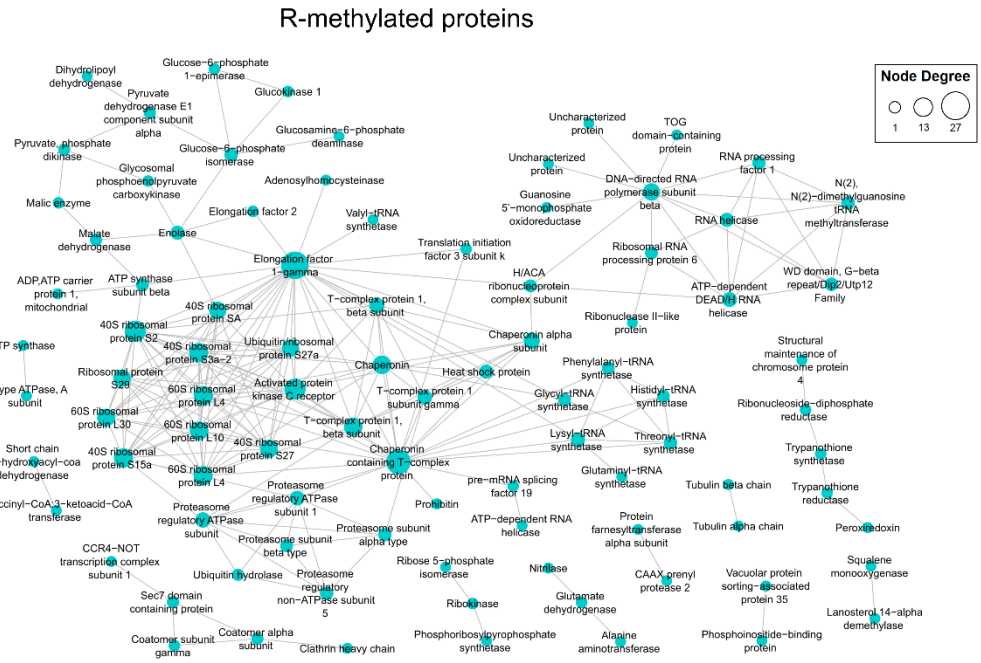


Figure S3. Heatmap results of iceLogo analysis. Increased or decreased amino acid frequencies are shown in a gradient of respectively green or red shades for K-Me (A) and R-Me (B) sites. Related to Fig. 2A and 2B).

A)



B)

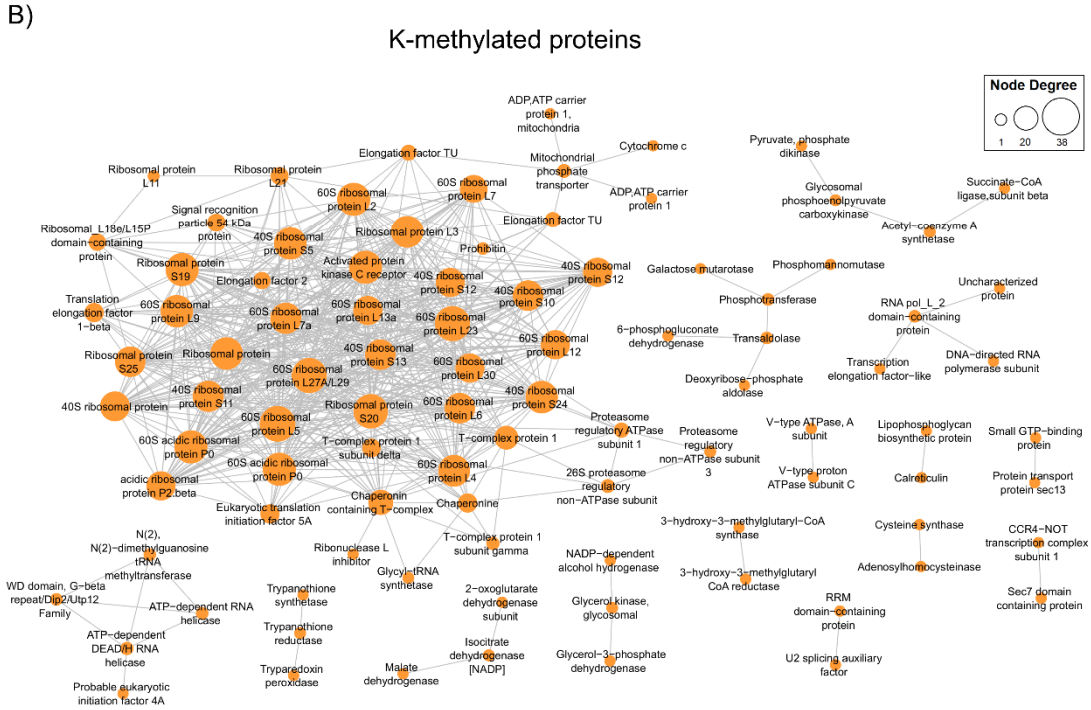


Figure S4. Interaction network maps of arginine (A) and lysine (B) methylated proteins in *T. cruzi*. Node degree correspond to number of interactions of the protein. Detailed data are listed in Table S5. Detailed data are listed in Table S5.

Table 1 - Functional categories enriched among *T. cruzi* arginine-methylated proteins

Category	Term	Count	Fold Enrichment	PValue	BH*
Biological Process	Oxidation-reduction process	6	8.37	4.00E-04	4.11E-02
	Biosynthetic process	10	4.19	4.29E-04	2.23E-02
	Carbohydrate metabolic process	10	3.99	6.31E-04	2.18E-02
Cellular Component	Cytoplasm	34	3.25	2.00E-09	8.80E-08
	Cytosolic small ribosomal subunit	4	14.32	1.92E-03	4.15E-02
	Proton-transporting two-sector ATPase complex, catalytic domain	3	25.06	4.59E-03	4.94E-02
Molecular Function	Pyridoxal phosphate binding	12	6.21	1.72E-06	2.55E-04
	L-alanine:2-oxoglutarate aminotransferase	4	20.19	4.59E-04	3.34E-02

*Benjamini-Hochberg ≤0.05.

Table 2 - Functional categories enriched among *T. cruzi* lysine-methylated proteins

Category	Term	Count	Fold Enrichment	PValue	BH*
Clusters of Orthologous Groups	Amino acid transport and metabolism	11	2.89	2.94E-03	4.89E-02
Biological Process	Translation	50	2.66	8.93E-11	8.84E-09
	Cysteine biosynthetic process from serine	8	10.62	1.66E-06	8.23E-05
	Response to stress	7	7.3	1.80E-04	5.92E-03
	Glycolytic process	9	4.69	3.96E-04	9.76E-03
Cellular Component	Ribosome	45	4.22	1.04E-16	4.77E-15
	Large ribosomal subunit	9	10.97	5.25E-07	1.13E-05
	Cytoplasm	30	2.65	1.97E-06	2.82E-05
	Cell	12	4.41	6.31E-05	6.78E-04
	Endoplasmic reticulum	5	7.24	4.04E-03	3.42E-02
	Proteasome activator complex	3	23.16	5.37E-03	3.79E-02
Molecular Function	Structural constituent of ribosome	52	2.97	1.19E-12	1.84E-10
	GTPase activity	16	4	6.02E-06	4.67E-04
	Cystathionine beta-synthase activity	6	13.73	1.73E-05	8.94E-04
	Translation elongation factor activity	16	3.16	1.16E-04	4.47E-03
	NADP binding	7	7.01	2.64E-04	8.14E-03
KEGG Pathways	Ribosome	56	1.87	6.17E-07	4.01E-05
	Biosynthesis of amino acids	26	2.06	3.41E-04	1.10E-02
	Biosynthesis of antibiotics	47	1.61	3.99E-04	8.61E-03
	Cysteine and methionine metabolism	16	2.64	4.80E-04	7.76E-03
	Glycolysis / Gluconeogenesis	19	2.34	5.92E-04	7.66E-03
	Glycine, serine and threonine metabolism	11	2.75	3.98E-03	4.23E-02

*Benjamini-Hochberg ≤0.05.

Table 3. Proteins with co-occurrence of methylation and phosphorylation in *T. cruzi*

Description	Uniprot Ids	Gene Ids	Me site (K, R) position	Ph site (S, T, Y) position
calpain cysteine peptidase, putative	Q4CPQ6	Tc00.1047053505985.9	R1356	S346
calpain-like cysteine peptidase, clan CA, family C2, putative	Q4CW64	Tc00.1047053509003.30	K72	S20
cysteine peptidase, Clan CA, family C2, putative	Q4D6J0	Tc00.1047053506721.30	R3348; R3357	S970; S1889; T4555
cysteine peptidase, Clan CA, family C2, putative	Q4E0D8	Tc00.1047053508999.190	K732	S822
	Q4CL00	Tc00.1047053510957.9		
cAMP-specific phosphodiesterase, putative	Q4DUD4	Tc00.1047053506625.80	R635; R636	S485
cytoskeleton-associated protein CAP5.5, putative	Q4DJS7	Tc00.1047053509237.130	R214	S759; S769
cytoskeleton associated protein, putative	Q4CSI2	Tc00.1047053506441.20	R231; K364	S91; T114; S615; Y618; S619; S653; S674; S892; S927; T1053; S1055; S1075
cytoskeleton associated protein, putative	Q4DJ92	Tc00.1047053511815.170	R322; K374	S427; S428
	Q4DVF9	Tc00.1047053506859.170		
cytoskeleton associated protein, putative	Q4DVF8	Tc00.1047053506859.180	R157; K209	S14; T21
cytoskeleton associated protein, putative	Q4D770	Tc00.1047053508265.100	R231; K364	S91; T114; S653; S695; S987; S1005; S1132; T1137; S1139; S1150; S1153; S1154; S1159
Dpy-30 motif containing protein, putative	Q4CYE7	Tc00.1047053507629.30	R938; R946	S767
dynein heavy chain,	Q4DX29	Tc00.1047053508831.4	K124; R126;	S483

putative	Q4CPP9	Tc00.1047053508275.9	K131	
eukaryotic initiation factor 5a, putative	Q4E4N4	Tc00.1047053506925.120	K42	S2
		Tc00.1047053506925.130		
flagellum targeting protein kharon1, putative (fragment)	Q4DNH8	Tc00.1047053509791.189	R146	S77; S79
flagellum targeting protein kharon1, putative	Q4D3F7	Tc00.1047053508719.70	K146	S77; S79; S276; S332; S344; S381
Gem-associated protein 2, putative	Q4D0P8	Tc00.1047053510301.40	K305; K313	S97; S99
	Q4D4R7	Tc00.1047053504163.70		
heterogeneous nuclear ribonucleoprotein H/F, putative	Q4DGN8	Tc00.1047053504157.10	K147	S123; S125; T126; T128; S339
	Q4DFM4	Tc00.1047053511109.130		
leucine rich repeat protein, putative	Q4E246	Tc00.1047053508707.310	R512; K548; K632	S898; S906; S910
Mitochondrial import receptor subunit ATOM69, putative	Q4D015	Tc00.1047053511803.40	K89	S502; T505
p25-alpha, putative	Q4CVJ1	Tc00.1047053504199.20	R96	S128; S130; S132
	Q4DIP8	Tc00.1047053506635.130		
phosphoprotein phosphatase, putative	Q4DDM9	Tc00.1047053507601.10	K625	S5; S8; S9; S270; T272
prostaglandin F2alpha synthase	Q4E4V7	Tc00.1047053508461.80	R9	S215
	Q4CNI1	Tc00.1047053507617.9		
protein kinase A catalytic subunit isoform 2, putative	Q4E4T3	Tc00.1047053508461.280	K18; K160	T174; T230
protein kinase, putative	Q4D5A8	Tc00.1047053511633.70	R387; R389	S318; S319; S322
protein associated with differentiation 8, putative	Q4DZJ0	Tc00.1047053508799.270	R412	T588
	Q4CUG0	Tc00.1047053509713.10		
	Q4DZI9	Tc00.1047053508799.280	R401	
Putative intraflagellar transport protein A1	Q4CP52	Tc00.1047053503539.20	R4; R7	S362; S363; S365; S436; S441; S442; S445; Y447
pyruvate phosphate dikinase, putative	Q4E0Q0	Tc00.1047053506297.190	R365; R587; K871	T481; S482; S680
	Q4E3P5	Tc00.1047053510101.140		

retrotransposon hot spot (RHS) protein, putative	Q4E3H4	Tc00.1047053507611.10	K553	S32; S34; T36
RNA guanylyltransferase, putative	Q4D3H7	Tc00.1047053507511.30	K445	S30; S34
RNA-binding protein, putative	Q4E5D6	Tc00.1047053511277.580	R8; R251	S302; S304; S307; S308; S412; S414
RNA-binding protein, putative	Q4D7Y1	Tc00.1047053510755.120	R213	S7; S9; S188
	Q4D3A9	Tc00.1047053508413.50		
RNA-binding protein, putative	Q4D488	Tc00.1047053509317.60	R36	S375; S377; S379
	Q4DT61	Tc00.1047053511621.50		
RNA-editing-associated protein 1, putative	Q4DZM3	Tc00.1047053504147.224	K362; K375	S188; S189; S193; S194
sterol 24-c-methyltransferase, putative	Q4CLW8	Tc00.1047053505683.10	R91	S235
	Q4CMB7	Tc00.1047053504191.10		
	Q4CM63	Tc00.1047053510185.10		
tyrosine aminotransferase	Q4CVI0	Tc00.1047053510795.10	K60; R196	S77
	Q4E4E7	Tc00.1047053510187.30		
		Tc00.1047053510187.40		
ubiquitin hydrolase, putative	Q4DG39	Tc00.1047053510761.70	R540; K544	S947; S949; S952; T955; S956
hypothetical protein	Q4DWC8	Tc00.1047053509733.60	R8	S10; S13; S144; Y146; T183; S185; S248; S402; S411
hypothetical protein, conserved	Q4CYL3	Tc00.1047053504423.30	K308; R408	S57; S618; T622; S813
hypothetical protein, conserved	Q4D0U3	Tc00.1047053508051.20	R1419	S256; S260; T263; S265; S266; S267; S1403; S1406
hypothetical protein, conserved	Q4DRF1	Tc00.1047053508547.160	K756; K760	S15; S81; T83
hypothetical protein, conserved	Q4DIU6	Tc00.1047053511389.50	R237; R239	T221
hypothetical protein, conserved	Q4DPA5	Tc00.1047053509647.130	K7; R9	S91; S95

hypothetical protein, conserved	Q4CTC8	Tc00.1047053507569.10	R266; K308; R408	S57; S241; T246; S618; T622; S810
hypothetical protein, conserved	Q4DG58	Tc00.1047053511693.20	R200	S90; S92
hypothetical protein, conserved	Q4E2Q5	Tc00.1047053504153.280	R171	S366; S370
hypothetical protein, conserved	Q4D2L5	Tc00.1047053511435.40	R115; R118; R126	S88; T90; S91; S305; S311; S352; S355; S359
hypothetical protein, conserved	Q4D442	Tc00.1047053510733.50	K70; K77	S484
hypothetical protein, conserved	Q4CWG6	Tc00.1047053508145.49	R4; R8	S1788; S1792; S1793; S1796
	Q4D106	Tc00.1047053504243.30		
hypothetical protein, conserved	Q4CY87	Tc00.1047053510001.20	R507; K508	S17; S21; S48; S51; S56; S497; S502
	Q4D238	Tc00.1047053507529.20		
hypothetical protein, conserved	Q4E5E4	Tc00.1047053511277.490	R7; R10	S186; S189
	Q4E1X3	Tc00.1047053507993.40		

Me: methylation; Ph: Phosphorylation.