

RESEARCH ARTICLE

Deep Insight into the Phosphatomes of Parasitic Protozoa and a Web Resource ProtozPhosDB

Tamanna Anwar, Samudrala Gourinath*

School of Life Sciences, Jawaharlal Nehru University, New Delhi, India

* sgourinath@mail.jnu.ac.in



OPEN ACCESS

Citation: Anwar T, Gourinath S (2016) Deep Insight into the Phosphatomes of Parasitic Protozoa and a Web Resource ProtozPhosDB. PLoS ONE 11(12): e0167594. doi:10.1371/journal.pone.0167594

Editor: Patricia Talamas-Rohana, Centro de Investigacion y de Estudios Avanzados del Instituto Politecnico Nacional, MEXICO

Received: August 8, 2016

Accepted: November 16, 2016

Published: December 8, 2016

Copyright: © 2016 Anwar, Gourinath. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: URL for uploaded data at Figshare is - <https://figshare.com/s/b3c143579758bfc8afeb>.

Funding: The authors are thankful to University Grants Commission (Government of India) for providing the funds for this project under Dr. D. S. Kothari postdoctoral fellowship scheme. Grant No. F.4-2/2006(BSR)/13-739/2012(BSR).

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Phosphorylation dynamically regulates the function of proteins by maintaining a balance between protein kinase and phosphatase activity. A comprehensive understanding of the role phosphatases in cellular signaling is lacking in case of protozoans of medical and veterinary importance worldwide. The drugs used to treat protozoal diseases have many undesired effects and the development of resistance, highlights the need for new effective and safer antiprotozoal agents. In the present study we have analyzed phosphatomes of 15 protozoans of medical significance. We identified ~2000 phosphatases, out of which 21% are uncharacterized proteins. A significant positive correlation between phosphatome and proteome size was observed except for *E. histolytica*, having highest density of phosphatases irrespective of its proteome size. A difference in the number of phosphatases among different genera shows the variation in the signaling pathways they are involved in. The phosphatome of parasites is dominated by ser/thr phosphatases contrary to the vertebrate host dominated by tyrosine phosphatases. Phosphatases were widely distributed throughout the cell suggesting physiological adaptation of the parasite to regulate its host. 20% to 45% phosphatome of different protozoa consists of ectophosphatases, i.e. crucial for the survival of parasites. A database and a webserver “ProtozPhosDB” can be used to explore the phosphatomes of protozoans of medical significance.

Introduction

Protozoan infection is a serious health problem in poor and developing countries with poor sanitary conditions. Protozoa are single-celled, microscopic organisms that can live freely or parasitic in nature. Parasitic protozoa are capable of multiplying in mammals, which contributes to their survival and also causes serious infections to develop from just a single organism. Protein phosphorylation and dephosphorylation are integral mechanism implicated in the regulation of cellular activities in any organism. Enzymes belonging to the class kinases and phosphatases that act antagonistically control the reversible phosphorylation. Protozoan parasites are capable of sustaining in diverse environmental conditions by exploiting the host’s surface proteins through phosphorylation/dephosphorylation of ser/thr and tyr residues, that helps in

cell growth inhibition and parasite-host interaction [1]. Kinases have been the focus of research forefront since long, but the emerging evidence suggests that phosphatases also play critical role in signaling network and cell-fate decisions. The versatility of phosphatase action emphasizes the need for understanding the network complexity of the signaling pathway in protozoan parasites as phosphatases are becoming extremely important both at research forefront and as targets for drug development. In this study, we have analyzed the phosphatomes of protozoan parasites belonging to the phyla Apicomplexa, Kinetoplastida and *Entamoeba histolytica*, *Giardia lamblia*, *Trichomonas vaginalis* and a microsporidium *Encephalitozoon cuniculi*, these are responsible for important diseases like Chagas' disease, giardiasis, leishmaniasis, amoebiasis, trichomoniasis, and sleeping sickness. The drugs used to treat these diseases are turning useless due to the identification of new drug-resistant strains. Moreover, prolonged treatments pose severe side effects [2, 3]. To reduce the effect of protozoan infections and dysentery, new drugs that preserve therapeutic efficiency and are far from side effects, are desired. Therefore, it is critical to search for more effective drug targets to treat infectious diseases caused by protozoan parasites.

Phosphorylation is critical for the development of the parasites including protozoan parasites. A bulk of data is available for kinases, on contrary, less has been known for phosphatases in unicellular parasites. The availability of full genomic sequences of several protozoan parasites has led this study to provide the first account of our current understanding of phosphatases in parasitic protozoans.

Materials and Methods

We performed preliminary analysis of the phosphatomes of several species of unicellular parasites that belong to Apicomplexa: *Plasmodium falciparum* (Pf); *Plasmodium vivax* (Pv); *Cryptosporidium parvum* (Cp); *Babesia bovis* (Bb), *Theileria parva* (Tp), *Toxoplasma gondii* (Tg), as well as Kinetoplastids: *Trypanosoma brucei* (Tb), *Trypanosoma cruzi* (Tc) and *Leishmania major* (Lm), *Entamoeba histolytica* (Eh), *Giardia lamblia* (Gl), *Trichomonas vaginalis* (Tv), a microsporidium *Encephalitozoon cuniculi* (Ec) and *Homo sapiens* (Hs). The complete set of predicted protein sequences from the ORFs of the selected protozoa genome has been obtained from UniProt release 2014_01 [4]. Phosphatases are classified based on catalytic signature motifs and structural domains, a deeper classification of phosphatases is mentioned in our previous work on *E. histolytica* and others [5–7].

The program CD-HIT was used to remove redundant sequences [8]. A sensitive sequence analysis method was used for analyzing the genomes of the above mentioned protozoans for PPs. Standalone version of InterProScan5 was used to make domain assignments for PP catalytic domain containing gene products. Each of the phosphatase domain containing proteins in all the protozoans were queried against the 18347 protein families and 7512 domains available in the InterPro database (InterPro 50.0) [9]. InterPro is a powerful tool for classification and function prediction of protein sequence. Protein sequences with phosphatase domain having e-value score of 10^{-4} were selected for the analysis. CELLO v.2.5 [10] was used for subcellular localization prediction. Structural domain analysis was carried out using Phyre2 (Protein Homology/AnalogY Recognition Engine) that is a web-based service for protein structure prediction [11]. To identify the molecular function, biological process and pathway of each protein, the PANTHER Classification System was used [12]. All the data obtained in the analysis of 15 parasitic protozoans was integrated into the database ProtozPhosDB. Venn diagram was made using InteractiVenn [13].

ProtozPhosDB is developed by using MySQL database software application along with PHP. For creating web interface, Bootstrap template was used, AJAX and JQuery were

deployed to make the database application more interactive. The website framework is quite portable and it can be easily transformed with any other list of proteins.

Results and Discussion

1.1 ProtozPhosDB

ProtozPhosDB is a public repository dedicated to the phosphatases encoded by completely sequenced parasitic protozoans. It is an on-line web portal developed to assist scientists in obtaining information about phosphatases in parasitic protozoans belonging to the phyla Apicomplexa, Kinetoplastida and few others with completed genome sequence. The web interface of ProtozPhosDB can be dynamically searched for phosphatases belonging to different families and sub-families or by their localization. There are three different types of searching from database;

1. Simple search: It can be done to search for any type phosphatase by entering keywords. The user should enter keyword (serine, threonine, tyrosine, etc.) in the search textbox then select search anywhere and hit on Go. This will output all the phosphatases of the searched family in all the fifteen protozoans. The option “match exactly” can be used to search for phosphatases from particular family (e.g. protein serine/threonine phosphatase, protein tyrosine phosphatase, etc.).
2. Advance search: In advanced search the user can retrieve the phosphatases in any of the fifteen parasitic protozoans by selecting the organism from the organism dropdown menu then from the family dropdown menu select the family of phosphatase and then sub-family from the sub-family dropdown menu and hit search button. The user can also search for all the phosphatases in a particular organism by selecting all from family and sub-family dropdown menu.
3. Search by localization: The database can also be searched according to localization of phosphatases. Here also the user has to select organism from the organism dropdown menu and then localization from the localization dropdown menu and hit search.

The results are displayed in the form of table indicating total number of entries. The results provide information regarding all the phosphatases of the selected family, the 3D structural domain, sub-cellular localization, molecular function/process and pathway annotation. The information regarding phosphatases in fifteen protozoans of medical importance is organized in single database to expedite retrieval of information regarding phosphatases.

1.2 Protozoan Parasite Proteome, Kinome and Phosphatome

Kinomes and phosphatomes of several of these unicellular parasites have been analyzed in earlier studies [14–20]. Comparison of phosphatomes (phosphatases identified in the present study) and kinomes of the selected unicellular parasites (Table 1) shows that phosphatomes of Apicomplexa is a bit larger than their kinome sizes except for *Pf* where kinome is slightly larger. In case of *Eh*, *Gl*, *Tv* and Kinetoplastids number of kinases is much larger than phosphatases, except for *Tc* (Table 1). We have identified larger number of phosphatases in selected unicellular parasites than reported (Fig 1). Number of kinases and phosphatases reported in earlier studies in these organisms is given in Table 1. Study of phosphatome in of *Pv*, *Py* and *Tbg* is reported for the first time in the present analysis, while in case of *Tg* and *Ec* very few phosphatases have been reported earlier.

Table 1. Comparison of proteome, kinome and phosphatome of protozoan parasites.

Organism	Uniprot Id	Proteome Size	Phosphatome Size Present Study	Phosphatome Size Reported	Kinome Size Reported
<i>Apicomplexa</i>					
<i>Cp</i>	353152	3805	73	24 [14, 15]	61 [19]
<i>Tg</i>	5811	8404	151	5 [14, 15]	135 [19]
<i>Bb</i>	5865	3960	58	14 [14, 15]	35 [19]
<i>Tp</i>	5875	4071	57	13 [14, 15]	38 [19]
<i>Py</i>	73239	7757	79	-	63 [19]
<i>Pv</i>	126793	5389	79	-	68 [19]
<i>Pf</i>	36329	5353	77	67 [16]	89 [19]
<i>Kinetoplastida</i>					
<i>Lm</i>	5664	8038	172	89 [14, 15]	199 [20]
<i>Tbg</i>	679716	9668	158	-	-
<i>Tbb</i>	1185	8587	161	79 [14, 15]	176 [20]
<i>Tcb</i>	353153	19242	230	88 [14, 15]	190 [20]
<i>Eh</i>	5759	7959	250	144 [14, 15]	331 [19]
<i>Tv</i>	5722	50191	482	210 [14, 15]	931 [19]
<i>Gl</i>	5741	11207	151	32 [14, 15]	278 [17]
<i>Ec</i>	284813	2008	35	8 [14, 15]	32 [19]

doi:10.1371/journal.pone.0167594.t001

1.3 Comparative Analysis of Protozoan Parasite Phosphatomes

Comparative analysis of the phosphatomes shows that Kinetoplastid parasites, *Eh*, *Gl* and *Tv* have larger phosphatomes than Apicomplexan parasites (Fig 1). The phosphatome of *Tv* being the largest followed by *Eh* and *Tc*. Apicomplexan parasites have relatively smaller phosphatomes with maximum number of phosphatases in *Tg* followed by *Plasmodium sp.*, the kinomes of these parasites are also small ranging from 35 kinases in *Bb* to 135 in *Tg* (Table 1) [21]. The

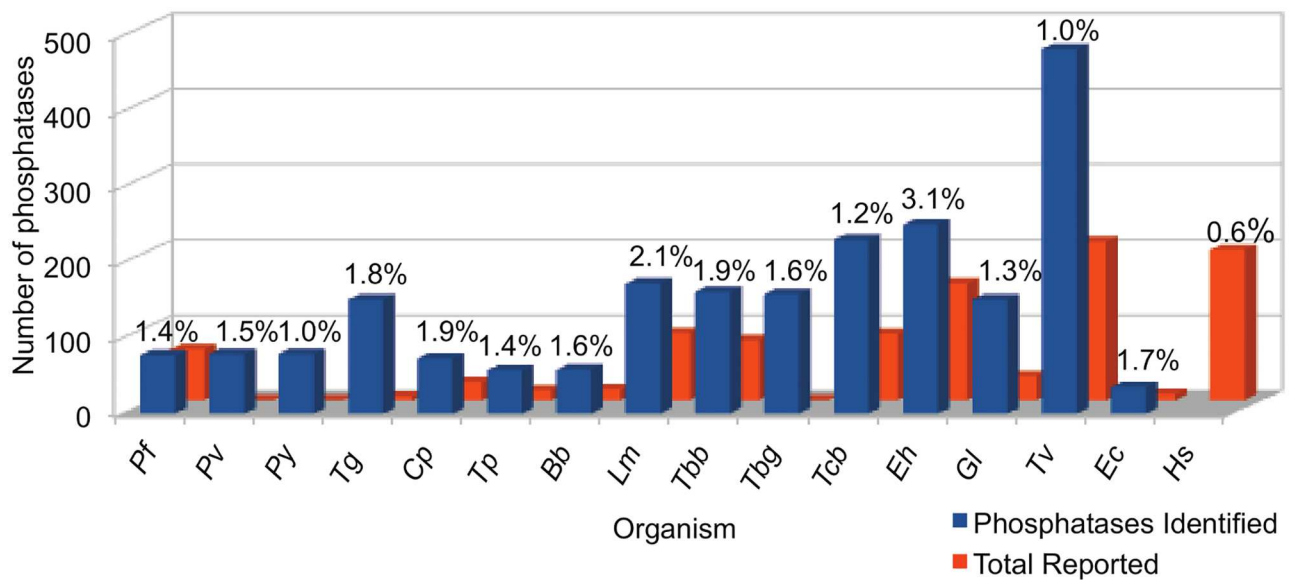


Fig 1. Bar diagram-representing comparison of phosphatases in the genomes of parasitic protozoa and human. Orange bars represents the number of phosphatases reported in previous studies. The percentage of phosphatase genes in the proteome complement is provided against every bar.

doi:10.1371/journal.pone.0167594.g001

phosphatomes vary in size between genera from 482 phosphatases in *Tv* to 35 in *Ec*. The phosphatome of *Eh* consists of 250 phosphatases that were reported in our earlier analysis [5, 6]. There was a huge variation in phosphatome of the genera *Trypanosoma*. The phosphatome of *Tb* strains lies in the range of 158 to 161 phosphatases, while the phosphatome of *Tc* consists of 230 phosphatases; their kinomes are composed of 176 kinases in *Tb* and 190 in *Tc* [22]. *Tc* is an intracellular parasite, it invades and reproduces in different type of cells like fibroblasts and macrophages, while, *Tb* that resides in the bloodstream of the mammalian host is a solely extracellular parasite. Since, these parasites have very distinct lifecycles, these organisms go through entirely different environments, leading to changes in expression of gene, protein levels and modification of proteins [23–25]. Protein phosphorylation actively regulates adaptive responses to intra cellular and extra cellular signals which shows marked changes during growth of these parasites [20, 26], indicating that *Tc* employs protein phosphorylation more aggressively as it needs to regulate the function of various different cell types while invading. Different species of the genus *Plasmodium* have comparatively smaller and similar phosphatomes comprising of 77 to 79 phosphatases (Table 1). In the earlier studies ratio of kinases to phosphatases in *Pf* is reported to be close to 2:1 [27], but in our analysis we have found ratio is ~1:1. Phosphatome of *Lm* has 172 phosphatases, while its kinome has 199 kinases [15], *Gl* has a phosphatome size of 151 phosphatases and its kinome consists of 278 kinases (Table 1). Human phosphatome has 199 proteins consisting of phosphatase domain [28]. There was a variation in apicomplexan phosphatomes size between different genera. Total number of phosphatases encoded in the genomes of the analyzed protozoan parasites along with the number of phosphatases reported in earlier studies [14, 15] and the percentage of phosphatome in the proteome of specific organisms is given in Fig 1. Comparison of the percentage of phosphatome in the respective proteome shows that, it varies from 1.0 to 3.1%. *Eh* has the highest percentage of phosphatases followed by *Lm* and *Cp*. The percentage of phosphatome in *Hs* is 0.6% of its proteome (Fig 1) [5, 28], which is much less than the selected parasites phosphatomes, indicating that vertebrates have various other mechanisms for the regulation of proteins but in protozoan parasites phosphorylation is an integral mechanism for protein regulation.

1.4 Correlation between Protozoan Parasite Proteome and Phosphatome

Analysis of the proteome size versus the phosphatome size reveals that the size of phosphatome increases with the increasing size of proteome, the only exception being *Eh*. The size of *Tc* proteome is more than double of *Eh* but still their phosphatomes are comparatively similar. A significant positive relation between phosphatome and proteome size was observed in the present study among the protozoan (Fig 2).

The size of the phosphatomes, to some extent reveals the significance of protein phosphorylation in signal transduction regulation and metabolic processes in these parasites. The difference in phosphatome size in case of *Eh* might be due to evolutionary pressure towards diversification of signal transduction pathways. The phosphatome of *Eh* is 1.3 times of human phosphatome [28]. It is striking to find that however the proteome size is highly dissimilar but their phosphatome size appear to be stable in the analyzed unicellular parasites.

1.5 Comparative Analysis and Distribution of Phosphatases in Protozoan Parasite

Parasitic protozoan have phosphatases belonging to all the major families of phosphatases including Serine/threonine phosphatases (STP), Protein Tyrosine Phosphatases (PTP), Endonuclease/Exonuclease/phosphatases (EEP), Acylphosphatase (ACP), Alkaline phosphatase

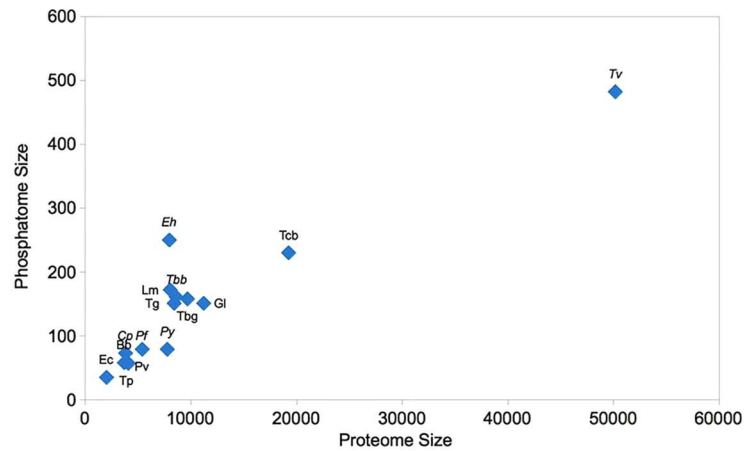


Fig 2. Correlation between proteome and phosphatome size of parasitic protozoa.

doi:10.1371/journal.pone.0167594.g002

(AlkP), Pyrophosphatases (PyroP) and Nucleoside phosphatase (NP), though their distribution vary in each protozoa (Fig 3).

The phosphatome of the analyzed parasitic protozoans is dominated by STP, revealing ser/thr phosphorylation is the most abundantly used in parasitic protozoans (Fig 3). In contrast the human phosphatome consists only 24.6% of phosphatases from STP family [28]. Ser/Thr phosphatases comprise of three unrelated sub-families Phosphoserine phosphatases (PPP) and Metal Dependent Phosphatases (PPM) and Aspartate-based phosphatases (FCP), the active sites of PPP and PPM sub-families are remarkably similar while the sub-families do not share any sequence homology, members of FCP are also ser/thr specific phosphatases [7, 29]. FCP consists aspartate based ser/thr phosphatases and haloacid dehalogenases (HAD) [29].

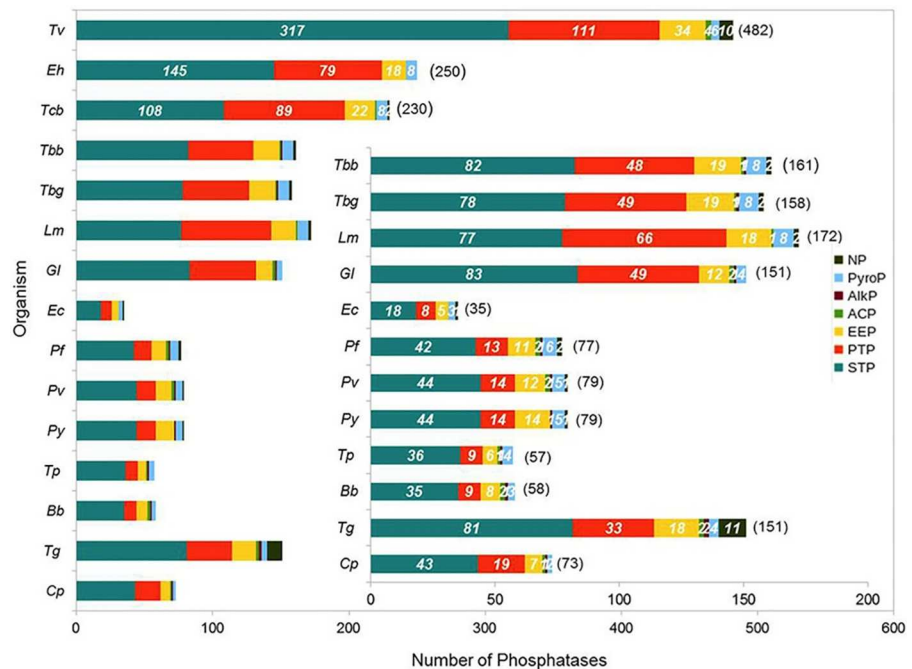


Fig 3. Distribution of protozoan parasite phosphatases into families of phosphatases.

doi:10.1371/journal.pone.0167594.g003

Among STP's, FCP sub-family was most represented in all the protozoans except for *Tv* and *Eh*. In Apicomplexa parasites members of the sub-family PPP and PPM were almost equivalent except for *Bb* and *Tg*. In *Eh*, *Gl* and Kinetoplastids phosphatases from the sub-family PPP were nearly double the number of PPM sub-family members. In case of *Tv* PPPs were more than double the number of FCPs, while PPMs were much less in comparison. In *Ec* also, FCPs were double that of PPPs, while PPMs were totally absent. FCP sub-family that dominated PPP phosphatome in all the parasitic protozoa was mostly composed of haloacid dehalogenases, suggesting that protozoans are generally rich in acid phosphatases.

PTP phosphatomes are relatively smaller than the STP phosphatome in all the unicellular parasites. *Tc* and *Lm* have highest percentage of PTP phosphatome, 39% and 38% respectively. Apicomplexa parasites generally have smaller PTP phosphatome (17% to 26%) in comparison to Kinetoplastids, except for *Tg* and *Cp* comprising 22% and 26% of their phosphatome, respectively. In contrary, the PTP phosphatome of humans is larger than STP phosphatome, containing 54.3% of PTPs [28]. Protein tyrosine phosphatases (PTP) consist of dual specificity phosphatase (DSPc), classical protein tyrosine phosphatase (PTPc), protein tyrosine phosphatase like (PTPLA), low molecular weight protein tyrosine phosphatase (LMPTP) and histidine phosphatase (HP) [30]. In general, PTPs from sub-families DSPc and HP were most abundant in all the parasitic protozoans, but in case of *C. parvum* and *T. parva* composition of HP were double that of DSPc. In general, PTPs from sub-families DSPc were most abundant in all the parasitic protozoans.

The investigated unicellular parasites have the large proportion of STPs and a low proportion of PTPs that is opposite to that of vertebrates having higher proportion of PTPs and lower proportion of STPs, suggesting different protein phosphorylation mechanisms in these unicellular parasites. Although PTPs comprises a smaller fraction compared to STPs, it plays an important role in signaling process regulating cell cycle control. Therefore, PTPs can possibly be potent targets for pharmacological approach.

If we look at the abundance of EEP, Kinetoplastids and *T. cruzi* has highest frequency of EEPs. Exonuclease-Endonuclease-Phosphatases (EEP) family consist of members from EEP sub-family and INPP sub-family. In *Trypanosoma sp.*, *Plasmodium sp.* and *Eh* number of EEPs is greater than INPPs. In case of *Tv* INPPs are around 4 times more in number than EEPs. In general, fewer numbers of INPPs were observed in *Ec* and Apicomplexa parasites (Fig 3).

Pyrophosphatase in parasitic protozoans consists of inorganic phosphatases in most of the organisms (Fig 3). Earlier reports also suggest that appropriate balance of intracellular inorganic phosphate is important for cellular metabolism [31]. Membrane bound inorganic pyrophosphatases were present in all the kinetoplastids, apicomplexa (*Plasmodium sp.*, *Bb*, *Tg*, *Tp*) and *Gl*. It has been reported that membrane proteins are critical targets for drugs and humans lack membrane-bound pyrophosphatases but they are vital for the survival of protozoan parasites [32], thus, the identified pyrophosphatases can be explored to understand their mechanism of action as well as these can be targeted against the parasitic organisms.

Phosphatases from several other families like acylphosphatase (ACP), alkaline phosphatase (AlkP) and nucleoside phosphatase (NP) were also observed in the analyzed unicellular parasites. Acylphosphatase is a ubiquitous small enzyme that acts on substrates like acetylphosphate, succinyl phosphate, 1,3-bisphosphoglycerate, carbamoylphosphate and β -aspartylphosphate membrane pump [33]. Acylphosphatases can be differentiated from acid phosphatases, alkaline phosphatases or protein phosphatases, as they cannot hydrolyze other phosphate esters or anhydrides. *T. vaginalis* has highest frequency of ACPs, others have either 1 or 2 members, while *Ec* and *Py* does not have ACPs. Alkaline phosphatase (AlkP) has the ability to remove phosphate groups from several kinds of molecules, such as proteins, nucleotides and alkaloids; at least one member was present in all the Apicomplexa parasites. In *Ec*, *Tv*, *Tc* and

Lm AlkPs were not seen, while in *Tb* strains and *Gl* a single member is found. Largest number of NPs was seen in *Tg* followed by *Tv*. *Trypanosoma* sp., *Plasmodium* sp., *Lm* and *Ec* have 1 or 2 members of NPs, while it is absent in *Gl*, *Tp*, *Cp* and *Bb* (Fig 3, Table A in S1 File).

Large differences in the phosphatomes between different genera of parasitic protozoans indicate considerable variability in the signaling pathways they are involved in. The phosphatome of humans is dominated by PTPs in contrary to the unicellular parasites having STPs in abundance [23].

1.6 Uncharacterized Phosphatome of Protozoan Parasites

In the present study we have analyzed 14 unicellular parasites and identified ~2000 phosphatases, out of which 21% are uncharacterized proteins. Highest percentage of uncharacterized proteins that were assigned to different families of phosphatases belongs to *Tp* (51.0%), followed by *Lm*, *Tc* and *Tbb* (Fig 4). For the sake of comparison we have taken the *Eh* phosphatome data that was analyzed in our previous work [5, 6]. A list of number of uncharacterized phosphatases classified as phosphatases in each genome is shown in Table B in S1 File.

1.7 Subcellular Localization of Phosphatases in Protozoan Parasites

Phosphatases can be characterized on the basis of their location in the cell, which can be intracellular i.e., secreted to the extracellular medium from the cell or membrane bound i.e., associated to plasma membrane [34, 35]. Phosphatases were found to be widely distributed throughout the cell (cytoplasm, mitochondria, nucleus, extracellular, plasma membrane, etc.). Highest distribution of phosphatases in the protozoans was seen in nuclear, plasma membrane and cytoplasmic regions (Table 2). Phosphatases were seen widely distributed of on the cell surface suggesting some physiological adaptation of the parasite to regulate its host. The details of the subcellular localization of each phosphatase can also be seen in ProtozPhosDB.

1.8 Molecular Function, Biological Process & Pathways Protozoan Parasite Phosphatases are Involved

Phosphatases are involved in plethora of biological functions. The functional analysis data of whole genomes was available for only 5 protozoans (*Pf*, *Lm*, *Tb*, *Eh* and *Tv*). We have seen that phosphatases were involved in 34 distinct functions. Kinase inhibitor activity, pyrophosphatase activity, kinase regulator activity, calcium ion binding activity, transferase activity, phosphoprotein phosphatase activity, cation transmembrane transporter activity, hydrolase activity, protein binding, nucleotide phosphatase activity, ion channel activity, nucleic acid binding activity were found to be conserved in all the analyzed protozoans. Several functions like RNA helicase activity, exodeoxyribonuclease activity, transmembrane activity, receptor protein kinase activity, endodeoxyribonuclease activity, phosphoric diester hydrolase activity, non-membrane spanning protein tyrosine kinase activity were specific to *Eh* only. Phosphatases having specific function in addition to the conserved functions, justifies the higher number of phosphatases in *Eh*. Some functions were specific to *Tv* only including cell communication, intracellular protein transport, monosaccharide metabolic process, phospholipid metabolic process, receptor-mediated endocytosis (Fig 5).

1.9 Ectophosphatases in Protozoan Parasites

Parasites come across various membranes and secreted molecules for attachment and invasion of the host cells. Ectophosphatases are potentially important phosphatase group, which is the phosphatase associated with the outer cell surface having their active sites oriented towards the

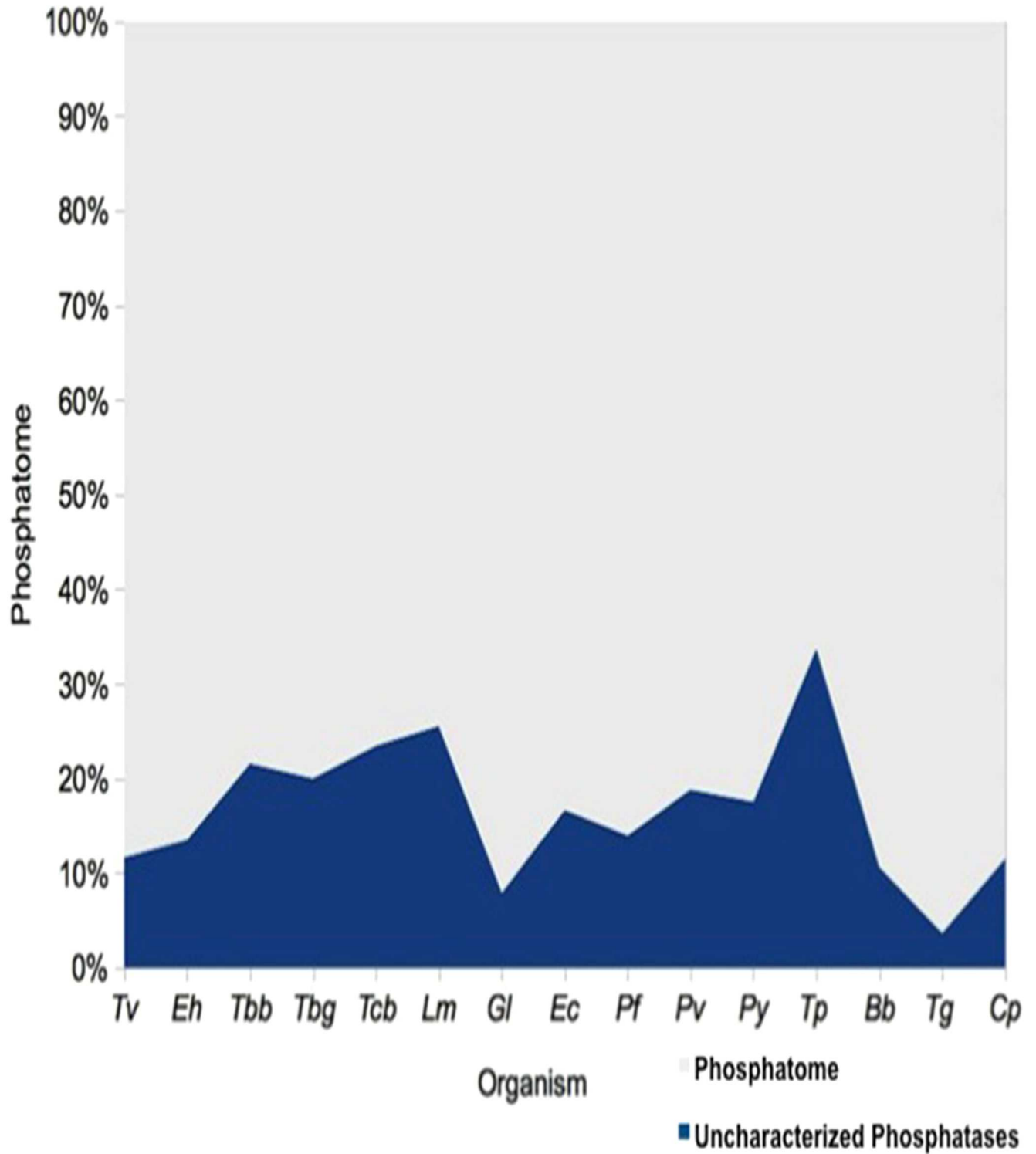


Fig 4. Percentage of uncharacterized proteins classified as phosphatases on the basis of domain identification.

doi:10.1371/journal.pone.0167594.g004

Table 2. Percentage of phosphatases distributed to different locations of the cell in protozoan parasites.

Sub-cellular Localization	<i>Lm</i>	<i>Tbb</i>	<i>Tbg</i>	<i>Tc</i>	<i>Pf</i>	<i>Pv</i>	<i>Py</i>	<i>Tg</i>	<i>Cp</i>	<i>Tp</i>	<i>Bb</i>	<i>Eh</i>	<i>Gl</i>	<i>Tv</i>	<i>Ec</i>
	Kinetoplastids				Apicomplexa							Others			
Cytoplasmic	23.3	29.8	27.8	25.2	14.3	16.5	7.6	21.2	20.5	33.3	31.0	30.7	28.5	38.6	48.6
Mitochondrial	11.6	11.2	11.4	13.0	5.2	5.1	8.9	6.0	2.7	–	3.4	0.4	3.3	3.7	–
Nuclear	32.0	26.1	26.6	19.1	54.5	54.4	48.1	47.7	46.6	26.3	25.9	33.9	21.9	23.9	20.0
Extracellular	5.8	6.2	5.7	5.7	–	3.8	–	2.6	–	1.8	–	1.6	1.3	4.6	–
Plasma Membrane	21.5	24.2	24.1	29.6	26.0	20.3	31.6	17.2	27.4	38.6	39.7	32.7	43.7	28.2	28.6
Glycosome	0.6	0.6	–	–	–	–	–	–	–	–	–	–	–	–	–
Vacuole	–	–	–	–	–	–	–	–	–	–	–	–	–	0.2	–
Uncharacterized	5.2	1.9	4.4	7.4	–	–	3.8	5.3	2.7	–	–	0.8	1.3	0.8	2.9

doi:10.1371/journal.pone.0167594.t002

external medium rather than the cytoplasm [36–38]. Ectophosphatases play an important role in cell proliferation, differentiation, parasite-host interaction, nutrition, ROS sensing and adhesion [39, 40]. Ectophosphatases have been reported in some protozoa parasites, including *Tb*, *Tc*, *Eh*, *Gl*, *Tv* and some *Leishmania* species [8, 39, 41–52]. A unique ectophosphatase cloned from *Tb* does not have any other similar protein beyond the genus *Trypanosoma* [1]. In the present analysis the phosphatases localized in the extracellular space and plasma membrane were considered as ectophosphatases and we found that 20 to 40 percent of the phosphatomes of different unicellular parasites consists of ectophosphatases (Table 2). In most of the organisms ectophosphatases are dominated by the members from HAD sub-family, though ectophosphatases consist of phosphatases from all the representative families. Earlier reports also suggest that the surface localization and low optimum pH of acidic phosphatases helps to function in acidic microenvironment [53–55], perhaps reflecting an adaptation to the extracellular environment by the parasite.

Conclusions

The genome sequence of parasitic protozoans provides a greater understanding of protein regulation and potential avenues for drug therapies and vaccine development. Lot of work has been done on the functional characterization of protein kinases; a better understanding of phosphatases and their involvement in parasite physiology and the identification of potential inhibitors to disrupt parasitic function are essential. This study provides the first account of our current understanding of protein phosphatomes of protozoan parasites having completed genomic sequences. The phosphatome size is positively correlated with the proteome size except for *Eh*. This may be due to extensive use of protein phosphorylation for regulation of proteins by *Eh*, while rest of the parasitic protozoans deploy other mechanism too. Thus, despite of having a small phosphatome, the parasite manages to maintain a large functional proficiency. The overall distribution of phosphatases within different protozoa genus seems to differ greatly. The distribution of phosphatases in different species of *Trypanosoma* was slightly variable, while *Plasmodium* sp. was having similar distribution of phosphatases. The differences in the phosphatomes of these parasites may be due to the variation in the living environments in which they exist. Interestingly, the phosphatome of these parasites have larger number of STPs unlike vertebrate host that is dominated by PTPs. Ectophosphatases apart from hydrolyzing extracellular phosphorylated substrates to change their function also have several alternative functions, such as adhesion, differentiation, participation in proliferation, virulence and infection. The identified ectophosphatases can be analyzed to understand the physiological roles of these enzymes in protozoan parasites. Further, ectophosphatases can be

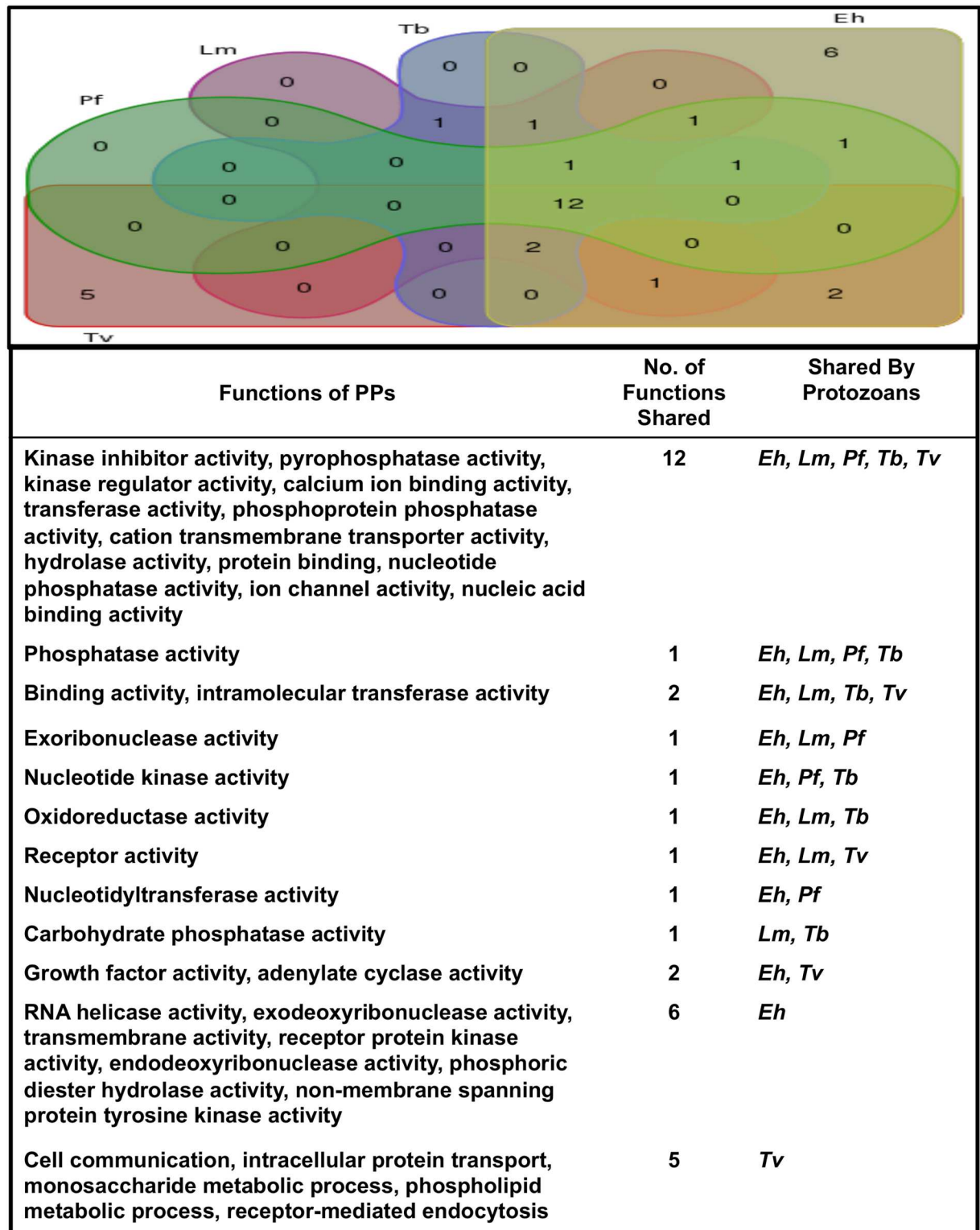


Fig 5. Venn diagram representing the analyses of conserved and specific functions of phosphatases in protozoan parasite.

doi:10.1371/journal.pone.0167594.g005

promising new drugs targets for treatment of these parasitic infections. Uncharacterized proteins are the loopholes in our understanding of biological processes. The dataset of uncharacterized proteins identified, as phosphatases can be further analyzed to understand physiological roles of these enzymes in the regulatory network of protozoan parasites. ProtozPhosDB can be used to access any information related to phosphatases in protozoans of medical and veterinary importance worldwide. The phosphatome data can be used to design experimental work to interpret the function of individual phosphatase in any particular parasitic protozoa and the signal transduction pathways they are involved in.

Supporting Information

S1 File. *Table A.* Distribution of protozoan parasite phosphatases into families and sub-families of phosphatases. *Table B.* Distribution of protozoan parasite uncharacterized proteins into phosphatase families.
(DOC)

Author Contributions

Conceptualization: TA SG.

Data curation: TA.

Formal analysis: TA.

Funding acquisition: SG.

Investigation: TA.

Methodology: TA.

Project administration: TA SG.

Resources: SG.

Software: TA.

Supervision: TA SG.

Validation: TA SG.

Visualization: TA.

Writing – original draft: TA.

Writing – review & editing: TA SG.

References

1. Bakalara N, Santarelli X, Davis C, Baltz T (2000) Purification, cloning, and characterization of an acidic ectoprotein phosphatase differentially expressed in the infectious bloodstream form of *Trypanosoma brucei*. *J Biol Chem* 275: 8863–8871. PMID: [10722732](#)
2. Andrews KT, Fisher G, Skinner-Adams TS (2014) Drug repurposing and human parasitic protozoan diseases. *Int J Parasitol Drugs Drug Resist* 4: 95–111. doi: [10.1016/j.ijpddr.2014.02.002](#) PMID: [25057459](#)
3. Herrera L (2014) *Trypanosoma cruzi*, the Causal Agent of Chagas Disease: Boundaries between Wild and Domestic Cycles in Venezuela. *Front Public Health* 2: 259. doi: [10.3389/fpubh.2014.00259](#) PMID: [25506587](#)
4. UniProt C (2015) UniProt: a hub for protein information. *Nucleic Acids Res* 43: D204–212. doi: [10.1093/nar/gku989](#) PMID: [25348405](#)

5. Anwar T, Gourinath S (2013) Analysis of the Protein phosphotome of *Entamoeba histolytica* reveals an intricate phosphorylation network. PLoS One 8: e78714. doi: [10.1371/journal.pone.0078714](https://doi.org/10.1371/journal.pone.0078714) PMID: [24236039](https://pubmed.ncbi.nlm.nih.gov/24236039/)
6. Anwar T, Gourinath S (2016) EhPPTome—Entamoeba histolytica Protein Phosphotome Database. IJBRA (IJBRA-138978) In Press.
7. Hotez PJ, Fenwick A, Savioli L, Molyneux DH (2009) Rescuing the bottom billion through control of neglected tropical diseases. Lancet 373: 1570–1575. doi: [10.1016/S0140-6736\(09\)60233-6](https://doi.org/10.1016/S0140-6736(09)60233-6) PMID: [19410718](https://pubmed.ncbi.nlm.nih.gov/19410718/)
8. Li Weizhong, Godzik Adam (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22: 1658–1659. doi: [10.1093/bioinformatics/btl158](https://doi.org/10.1093/bioinformatics/btl158) PMID: [16731699](https://pubmed.ncbi.nlm.nih.gov/16731699/)
9. Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, et al. (2011) InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Research 40: D306–312. doi: [10.1093/nar/gkr948](https://doi.org/10.1093/nar/gkr948) PMID: [22096229](https://pubmed.ncbi.nlm.nih.gov/22096229/)
10. Yu CS, Lin CJ, Hwang JK (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. Protein Science 13: 1402–1406. doi: [10.1110/ps.03479604](https://doi.org/10.1110/ps.03479604) PMID: [15096640](https://pubmed.ncbi.nlm.nih.gov/15096640/)
11. Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 4: 363–371. doi: [10.1038/nprot.2009.2](https://doi.org/10.1038/nprot.2009.2) PMID: [19247286](https://pubmed.ncbi.nlm.nih.gov/19247286/)
12. Mi Huaiyu, Poudel Sagar, Muruganujan Anushya Casagrande John T, Thomas Paul D. (2016). PANTHER version 10: expanded protein families and functions, and analysis tools. Nucl Acids Res 44(D1): D336–D342. doi: [10.1093/nar/gkv1194](https://doi.org/10.1093/nar/gkv1194) PMID: [26578592](https://pubmed.ncbi.nlm.nih.gov/26578592/)
13. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R (2015) InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics 16: 169. doi: [10.1186/s12859-015-0611-3](https://doi.org/10.1186/s12859-015-0611-3) PMID: [25994840](https://pubmed.ncbi.nlm.nih.gov/25994840/)
14. Andreeva AV, Kutuzov MA (2008) Protozoan protein tyrosine phosphatases. Int J Parasitol 38: 1279–1295. doi: [10.1016/j.ijpara.2008.04.003](https://doi.org/10.1016/j.ijpara.2008.04.003) PMID: [18547579](https://pubmed.ncbi.nlm.nih.gov/18547579/)
15. Kutuzov MA, Andreeva AV (2008) Protein Ser/Thr phosphatases of parasitic protozoa. Mol Biochem Parasitol 161: 81–90. doi: [10.1016/j.molbiopara.2008.06.008](https://doi.org/10.1016/j.molbiopara.2008.06.008) PMID: [18619495](https://pubmed.ncbi.nlm.nih.gov/18619495/)
16. Pandey R, Mohammed A, Pierrot C, Khalife J, Malhotra P, Gupta D (2014) Genome wide in silico analysis of *Plasmodium falciparum* phosphatome. BMC Genomics 15: 1024. doi: [10.1186/1471-2164-15-1024](https://doi.org/10.1186/1471-2164-15-1024) PMID: [25425018](https://pubmed.ncbi.nlm.nih.gov/25425018/)
17. Manning G, Reiner DS, Lauwaet T, Dacre M, Smith A, Zhai Y, et al. (2011) The minimal kinome of *Giardia lamblia* illuminates early kinase evolution and unique parasite biology. Genome Biol 12: R66. doi: [10.1186/gb-2011-12-7-r66](https://doi.org/10.1186/gb-2011-12-7-r66) PMID: [21787419](https://pubmed.ncbi.nlm.nih.gov/21787419/)
18. Miranda-Saavedra D, Barton GJ (2007) Classification and functional annotation of eukaryotic protein kinases. Proteins 68: 893–914. doi: [10.1002/prot.21444](https://doi.org/10.1002/prot.21444) PMID: [17557329](https://pubmed.ncbi.nlm.nih.gov/17557329/)
19. Martin DM, Miranda-Saavedra D, Barton GJ (2009) Kinomer v. 1.0: a database of systematically classified eukaryotic protein kinases. Nucleic Acids Res 37: D244–250. doi: [10.1093/nar/gkn834](https://doi.org/10.1093/nar/gkn834) PMID: [18974176](https://pubmed.ncbi.nlm.nih.gov/18974176/)
20. Parsons M, Valentine M, Deans J, Schieven GL, Ledbetter JA (1991) Distinct patterns of tyrosine phosphorylation during the life cycle of *Trypanosoma brucei*. Mol Biochem Parasitol 45: 241–248. PMID: [1710035](https://pubmed.ncbi.nlm.nih.gov/1710035/)
21. Doerig C, Meijer L, Mottram JC (2002) Protein kinases as drug targets in parasitic protozoa. Trends Parasitol 18: 366–371. PMID: [12377287](https://pubmed.ncbi.nlm.nih.gov/12377287/)
22. Parsons M, Worthey EA, Ward PN, Mottram JC (2005) Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*. BMC Genomics 6: 127. doi: [10.1186/1471-2164-6-127](https://doi.org/10.1186/1471-2164-6-127) PMID: [16164760](https://pubmed.ncbi.nlm.nih.gov/16164760/)
23. Andrade LO, Andrews NW (2005) The *Trypanosoma cruzi*-host-cell interplay: location, invasion, retention. Nat Rev Microbiol 3: 819–823. doi: [10.1038/nrmicro1249](https://doi.org/10.1038/nrmicro1249) PMID: [16175174](https://pubmed.ncbi.nlm.nih.gov/16175174/)
24. Matthews KR (2005) The developmental cell biology of *Trypanosoma brucei*. J Cell Sci 118: 283–290. doi: [10.1242/jcs.01649](https://doi.org/10.1242/jcs.01649) PMID: [15654017](https://pubmed.ncbi.nlm.nih.gov/15654017/)
25. Atwood JA 3rd, Weatherly DB, Minning TA, Bundy B, Cavola C, Opperdoes FR, et al. (2005) The *Trypanosoma cruzi* proteome. Science 309: 473–476. doi: [10.1126/science.1110289](https://doi.org/10.1126/science.1110289) PMID: [16020736](https://pubmed.ncbi.nlm.nih.gov/16020736/)
26. Dell KR, Engel JN (1994) Stage-specific regulation of protein phosphorylation in *Leishmania major*. Mol Biochem Parasitol 64: 283–292. PMID: [7935606](https://pubmed.ncbi.nlm.nih.gov/7935606/)
27. Wilkes JM, Doerig C (2008) The protein-phosphatome of the human malaria parasite *Plasmodium falciparum*. BMC Genomics 9: 412. doi: [10.1186/1471-2164-9-412](https://doi.org/10.1186/1471-2164-9-412) PMID: [18793411](https://pubmed.ncbi.nlm.nih.gov/18793411/)

28. Liberti S, Sacco F, Calderone A, Perfetto L, Iannuccelli M, Panni S, et al. (2013) HuPho: the human phosphatase portal. *FEBS J* 280: 379–387. doi: [10.1111/j.1742-4658.2012.08712.x](https://doi.org/10.1111/j.1742-4658.2012.08712.x) PMID: [22804825](https://pubmed.ncbi.nlm.nih.gov/22804825/)
29. Szoor B (2010) Trypanosomatid protein phosphatases. *Mol Biochem Parasitol* 173: 53–63. doi: [10.1016/j.molbiopara.2010.05.017](https://doi.org/10.1016/j.molbiopara.2010.05.017) PMID: [20594956](https://pubmed.ncbi.nlm.nih.gov/20594956/)
30. Veeramani S, Lee MS, Lin MF (2009) Revisiting histidine-dependent acid phosphatases: a distinct group of tyrosine phosphatases. *Trends Biochem Sci* 34: 273–278. doi: [10.1016/j.tibs.2009.03.002](https://doi.org/10.1016/j.tibs.2009.03.002) PMID: [19467874](https://pubmed.ncbi.nlm.nih.gov/19467874/)
31. Dick CF, Dos-Santos AL, Meyer-Fernandes JR (2011) Inorganic phosphate as an important regulator of phosphatases. *Enzyme Res* 2011: 103980. doi: [10.4061/2011/103980](https://doi.org/10.4061/2011/103980) PMID: [21755037](https://pubmed.ncbi.nlm.nih.gov/21755037/)
32. Kellosalo J, Kajander T, Kogan K, Pokharel K, Goldman A (2012) The structure and catalytic cycle of a sodium-pumping pyrophosphatase. *Science* 337: 473–476. doi: [10.1126/science.1222505](https://doi.org/10.1126/science.1222505) PMID: [22837527](https://pubmed.ncbi.nlm.nih.gov/22837527/)
33. Stefani M, Ramponi G (1995) Acylphosphate phosphohydrolases. *Life Chem Rep* 12: 271–301.
34. Andrews PD, Stark MJ (2000) Dynamic, Rho1p-dependent localization of Pkc1p to sites of polarized growth. *J Cell Sci* 113 (Pt 15): 2685–2693.
35. Fernandes AC, Soares DC, Saraiva EM, Meyer-Fernandes JR, Souto-Pradon T (2013) Different secreted phosphatase activities in *Leishmania amazonensis*. *FEMS Microbiol Lett* 340: 117–128. doi: [10.1111/1574-6968.12080](https://doi.org/10.1111/1574-6968.12080) PMID: [23305417](https://pubmed.ncbi.nlm.nih.gov/23305417/)
36. Meyer-Fernandes JR (2002) Ecto-ATPases in protozoa parasites: looking for a function. *Parasitol Int* 51: 299–303. PMID: [12243784](https://pubmed.ncbi.nlm.nih.gov/12243784/)
37. Gomes SAO, de Souza ALF, Silva BA, Kiffer-Moreira T, Santos-Mallet JR, Santos ALS, et al. (2006) *Trypanosoma rangeli*: Differential expression of cell surface polypeptides and ecto-phosphatase activity in short and long epimastigote forms. *Exp Parasitol* 112: 253–262. doi: [10.1016/j.exppara.2005.11.015](https://doi.org/10.1016/j.exppara.2005.11.015) PMID: [16442100](https://pubmed.ncbi.nlm.nih.gov/16442100/)
38. Meyer-Fernandes JR, da Silva-Neto MA, Soares Mdos S, Fernandes E, Vercesi AE, de Oliveira MM (1999) Ecto-phosphatase activities on the cell surface of the amastigote forms of *Trypanosoma cruzi*. *Z Naturforsch C* 54: 977–984. PMID: [10627997](https://pubmed.ncbi.nlm.nih.gov/10627997/)
39. Gomes MT, Lopes AH, Meyer-Fernandes JR (2011) Possible roles of ectophosphatases in host-parasite interactions. *J Parasitol Res* 2011: 479146. doi: [10.1155/2011/479146](https://doi.org/10.1155/2011/479146) PMID: [21603194](https://pubmed.ncbi.nlm.nih.gov/21603194/)
40. Cosentino-Gomes D, Meyer-Fernandes JR (2011) Ecto-phosphatases in protozoan parasites: possible roles in nutrition, growth and ROS sensing. *J Bioenerg Biomembr* 43: 89–92. doi: [10.1007/s10863-011-9334-y](https://doi.org/10.1007/s10863-011-9334-y) PMID: [21253843](https://pubmed.ncbi.nlm.nih.gov/21253843/)
41. Fernandes EC, Meyer-Fernandes JR, Silva-Neto MA, Vercesi AE (1997). *Trypanosoma brucei*: ecto-phosphatase activity present on the surface of intact procyclic forms. *Z Naturforsch C* 52: 351–358. PMID: [9232891](https://pubmed.ncbi.nlm.nih.gov/9232891/)
42. Fernandes EC, Granjeiro JM, Aoyama H, Fonseca FV, Meyer-Fernandes JR, Vercesi AE (2003). A metallo phosphatase activity present on the surface of *Trypanosoma brucei* procyclic forms. *Vet Parasitol* 118: 19–28. PMID: [14651871](https://pubmed.ncbi.nlm.nih.gov/14651871/)
43. Fernandes EC, Granjeiro JM, Taga EM, Meyer-Fernandes JR, Aoyama H (2003). Phosphatase activity characterization on the surface of intact bloodstream forms of *Trypanosoma brucei*. *FEMS Microbiol Lett* 220: 197–206. PMID: [12670681](https://pubmed.ncbi.nlm.nih.gov/12670681/)
44. Furuya T, Zhong L, Meyer-Fernandes JR, Lu HG, Moreno SN, Docampo R (1998). Ecto-protein tyrosine phosphatase activity in *Trypanosoma cruzi* infective stages. *Mol Biochem Parasitol* 92: 339–348. PMID: [9657337](https://pubmed.ncbi.nlm.nih.gov/9657337/)
45. Meyer-Fernandes JR, da Silva-Neto MA, Soares MD, Fernandes E, Vercesi AE, de Oliveira MM (1999). Ecto-phosphatase activities on the cell surface of the amastigote forms of *Trypanosoma cruzi*. *Z Naturforsch C* 54: 977–984. PMID: [10627997](https://pubmed.ncbi.nlm.nih.gov/10627997/)
46. Dutra PM, Couto LC, Lopes AH, Meyer-Fernandes JR (2006). Characterization of ecto-phosphatase activities of *Trypanosoma cruzi*: a comparative study between Colombian and Y strains. *Acta Trop* 100: 88–95. doi: [10.1016/j.actatropica.2006.05.014](https://doi.org/10.1016/j.actatropica.2006.05.014) PMID: [17055992](https://pubmed.ncbi.nlm.nih.gov/17055992/)
47. Neves RF, Fernandes AC, Meyer-Fernandes JR, Souto-Pradon T (2014). *Trypanosoma cruzi*-secreted vesicles have acid and alkaline phosphatase activities capable of increasing parasite adhesion and infection. *Parasitol Res*. 113:2961–2972. doi: [10.1007/s00436-014-3958-x](https://doi.org/10.1007/s00436-014-3958-x) PMID: [24906990](https://pubmed.ncbi.nlm.nih.gov/24906990/)
48. de Almeida-Amaral EE, Belmont-Firpo R, Vannier-Santos MA, Meyer-Fernandes JR (2006). *Leishmania amazonensis*: characterization of an ecto-phosphatase activity. *Exp Parasitol* 114: 334–340. doi: [10.1016/j.exppara.2006.04.011](https://doi.org/10.1016/j.exppara.2006.04.011) PMID: [16814776](https://pubmed.ncbi.nlm.nih.gov/16814776/)
49. Pinheiro AAD, Amazonas JN, Barros FD, De Menezes LF, Batista EJO, Silva EF, et al. (2007). *Entamoeba histolytica*: an ecto-phosphatase activity regulated by oxidation-reduction reactions. *Exp Parasitol* 115: 352–358. doi: [10.1016/j.exppara.2006.09.019](https://doi.org/10.1016/j.exppara.2006.09.019) PMID: [17113080](https://pubmed.ncbi.nlm.nih.gov/17113080/)

50. Amazonas JN, Cosentino-Gomes D, Werneck-Lacerda A, Pinheiro AAD, Lanfredi-Rangel A, De Souza W, et al. (2009). *Giardia lamblia*: Characterization of ecto-phosphatase activities. *Exp Parasitol* 121: 15–21. doi: [10.1016/j.exppara.2008.09.010](https://doi.org/10.1016/j.exppara.2008.09.010) PMID: [18845148](https://pubmed.ncbi.nlm.nih.gov/18845148/)
51. de Jesus JB, Podlyska TM, Hampshire A, Lopes CS, Vannier-Santos MA, Meyer-Fernandes JR (2002). Characterization of an ecto-phosphatase activity in the human parasite *Trichomonas vaginalis*. *Parasitol Res* 88: 991–997. doi: [10.1007/s00436-001-0583-2](https://doi.org/10.1007/s00436-001-0583-2) PMID: [12375165](https://pubmed.ncbi.nlm.nih.gov/12375165/)
52. De Jesus JB, Ferreira MA, Cuervo P, Britto C, Silva-Filho FC, Meyer-Fernandes JR (2006). Iron modulates ecto-phosphohydrolase activities in pathogenic trichomonads. *Parasitol Int* 55: 285–290. doi: [10.1016/j.parint.2006.08.002](https://doi.org/10.1016/j.parint.2006.08.002) PMID: [17010660](https://pubmed.ncbi.nlm.nih.gov/17010660/)
53. Gottlieb M, Dwyer DM (1981) Protozoan parasite of humans: surface membrane with externally disposed acid phosphatase. *Science* 212: 939–941. PMID: [7233189](https://pubmed.ncbi.nlm.nih.gov/7233189/)
54. Shakarian AM, Joshi MB, Ghedin E, Dwyer DM (2002) Molecular dissection of the functional domains of a unique, tartrate-resistant, surface membrane acid phosphatase in the primitive human pathogen *Leishmania donovani*. *J Biol Chem* 277: 17994–18001. doi: [10.1074/jbc.M200114200](https://doi.org/10.1074/jbc.M200114200) PMID: [11882658](https://pubmed.ncbi.nlm.nih.gov/11882658/)
55. Aguirre-Garcia MM, Anaya-Ruiz M, Talamas-Rohana P (2003) Membrane-bound acid phosphatase (MAP) from *Entamoeba histolytica* has phosphotyrosine phosphatase activity and disrupts the actin cytoskeleton of host cells. *Parasitology* 126: 195–202. PMID: [12666878](https://pubmed.ncbi.nlm.nih.gov/12666878/)