Published in final edited form as:

Mol Biochem Parasitol. 2013 January ; 187(1): 1–8. doi:10.1016/j.molbiopara.2012.10.006.

Structure of the *Trypanosoma cruzi* protein tyrosine phosphatase TcPTP1, a potential therapeutic target for Chagas' disease

George T. Lountos^{1,2,*}, Joseph E. Tropea², and David S. Waugh^{2,*}

¹Basic Science Program, SAIC-Frederick, Inc., Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA

²Macromolecular Crystallography Laboratory, Center for Cancer Research, Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA

Abstract

Chagas' disease, a neglected tropical affliction transmitted by the flagellated protozoan *Trypanosoma cruzi*, is prevalent in Latin America and affects nearly 18 million people worldwide, yet few approved drugs are available to treat the disease. Moreover, the currently available drugs exhibit severe toxicity or are poorly effective in the chronic phase of the disease. This limitation, along with the large population at risk, underscores the urgent need to discover new molecular targets and novel therapeutic agents. Recently, the *T. cruzi* protein tyrosine phosphatase TcPTP1 has been implicated in the cellular differentiation and infectivity of the parasite and is therefore a promising target for the design of novel anti-parasitic drugs. Here, we report the X-ray crystal structure of TcPTP1 refined to a resolution of 2.18 Å, which provides structural insights into the active site environment that can be used to initiate structure-based drug design efforts to develop specific TcPTP1 inhibitors. Potential strategies to develop such inhibitors are also discussed.

Keywords

Trypanosoma cruzi; Chagas's disease; phosphatase; drug design; crystal structure; inhibitors

1. Introduction

Chagas' disease is a neglected tropical disease that is endemic to Latin America and affects nearly 18 million people worldwide[1, 2]. The etiological agent of Chagas' disease is the flagellated protozoan parasite *Trypanosoma cruzi* [3, 4]. *T. cruzi* infection poses a serious global health and economic problem that places millions of people at risk, particularly in the tropical and subtropical regions of the world[5, 6]. Transmittal to humans occurs mainly via

*Corresponding authors: Macromolecular Crystallography Laboratory Frederick National Laboratory for Cancer Research P.O. Box B Frederick, MD 21702-1201 Tel: (301) 846-5345; Fax: (301) 846-7148 lountosg@mail.nih.gov (GTL) Tel: (301) 846-1842; Fax (301) 846-7148; waughd@mail.nih.gov (DSW).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

the feces of infected blood-feeding triatomine insects, commonly called kissing bugs, which serve as vectors for the *T. cruzi* parasite[5]. Infection can also occur through blood transfusion[7], congenital transmission[8], organ transplant[9], and consumption of contaminated food and drink[10]. Clinical Chagas' disease manifests itself in patients as life-threatening cardiovascular disease or pathological gut enlargement due to chronic infection by *T. cruzi*[11-13].

Because Chagas' disease kills more people in Latin America than any other parasite-borne illness, including malaria, there is an urgent need for effective drugs to treat this affliction[2]. Nifurtimox and benznidazole are the only currently approved drugs. Although effective in the acute phase of infection, both of them can exhibit severe side effects and exhibit lower efficacy in the chronic phase[14-16]. Therefore, improved therapeutic agents against Chagas' disease with higher efficacy and lower toxicity are needed, especially because past efforts to target the disease have been limited[15, 17-19]. The identification of valid drug targets, coupled with rational, structure-based design of specific small molecule inhibitors, can be an effective strategy for discovering new drugs[20, 21].

Protein phosphorylation is an important post-translational modification that regulates a variety of signaling processes in living organisms[22]. Protein kinases and phosphatases play pivotal and interconnected roles in regulating the phosphorylation and dephosphorylation of tyrosine, serine and threonine in many signaling proteins[23]. Reversible protein phosphorylation results in the modulation of the structure and functions of many proteins that can affect protein-protein interactions, enzymatic activity, protein stability, and subcellular localization[24-26]. The phosphorylation status of proteins can have a pronounced effect on many signaling cascades that regulate the cell cycle, differentiation, cell motility, immune responses, metabolism and other biological processes[22, 27-29]. Therefore, the pharmacological targeting of phosphatases involved in these processes can be an effective therapeutic strategy[30].

Several trypanosomatid-specific protein phosphatases have been identified recently[4, 31, 32], but their molecular functions and the signaling pathways in which they are involved within the trypanosomatid life cycle are not well understood[33-35]. The TcPTP1 protein is one of only two protein tyrosine phosphatases (PTPases) encoded by the *T. cruzi* genome[32]. Studies have suggested that TcPTP1 participates in the differentiation process and infectivity of the parasite[36]. Indeed, inhibition of TcPTP1 by a broad spectrum PTPase inhibitor resulted in reduced infectivity of *T. cruzi* trypomastigotes, suggesting that it is a potential molecular target for the design of novel anti-parasitic agents[36]. Although PTPases are generally considered to be highly challenging targets, recent work has demonstrated that structure-based methods can be an effective strategy for the development of new phosphatase-specific inhibitors[37-40]. Therefore, we set out to crystallize and solve the X-ray crystal structure of the TcPTP1 as an important first step in initiating drug design efforts.

2. Materials and methods

2.1. Molecular cloning and protein expression

The *T. cruzi* TcPTP1 gene encoding Met1 to Arg302 was amplified from genomic DNA (Tulahuen strain, ATCC30266D, American Type Culture Collection) by PCR using the following primers: 5'-GAG-AAC-CTG-TAC-TTC-CAG-ATG-AAT-GAT-TCGAAC-TGC-TTC-CCT-TTC-3' and 5'-GGG-GAC-CAC-TTT-GTA-CAA-GAA-AGCTGG-GTT-ATT-ACC-TGC-TAT-TCA-ACA-GAC-CGA-CAA-GTG-C-3' (primer R). The resulting PCR amplicon was used as the template for a second PCR with the following primers: 5'-GGG-GAC-AAG-TTT-GTA-CAA-AAA-AGC-AGG-CTC-GGAGAA-CCT-GTA-CTT-CCA-G-3' and primer R. The amplicon from the second PCR was inserted by recombinational cloning into the entry vector pDONR221 (Life Technologies), and the nucleotide sequence confirmed experimentally. The open reading frame of PTP1 (Met1-Arg302), now with a recognition site (ENLYFQ/M) for tobacco etch virus (TEV) protease added to its N terminus, was moved by recombinational cloning into the heptahistidine-maltose-binding protein (MBP) fusion vector pMT28 (M. Tropea, Clinical Center, NIH) to produce pJT206. This plasmid directs the expression of *T. cruzi* TcPTP1 as a His₇-MBP-TcPTP1 fusion protein with a TEV protease recognition site between His₇-MBP and TcPTP1. The fusion protein was expressed in the *E. coli* strain Rosetta2(DE3) (Novagen EMD Millipore). Cells were grown to mid-log phase (OD₆₀₀ ~ 0.5) at 37°C in LB broth containing 100 µg mL⁻¹ ampicillin, 30 µg mL⁻¹ chloramphenicol and 0.2% glucose. Overproduction of fusion protein was induced with isopropyl-β-D-thiogalactopyranoside at a final concentration of 1 mM for 20 h at 18°C. The cells were pelleted by centrifugation and stored at -80°C.

2.2. Protein purification

All procedures were performed at 4-8°C. Twenty grams of *E. coli* cell paste was suspended in 200 ml ice-cold buffer A (50 mM HEPES, 200 mM NaCl, 25 mM imidazole, pH 7.5) containing 5 mM benzamidine HCl (Sigma Chemical Company) and Complete EDTA-free protease inhibitor cocktail tablets (Roche Molecular Biochemicals). The cells were lysed with an APV-1000 homogenizer (Invensys APV Products) at 10,000 psi, and centrifuged at 30,000 × *g* for 30 min. The supernatant was filtered through a 0.22 µm polyethersulfone membrane and applied to a 12 ml Ni-NTA superflow column (Qiagen) equilibrated in buffer A. The column was washed to baseline with buffer A and eluted with a linear gradient of imidazole to 250 mM. Fractions containing recombinant His₇-MBP-TcPTP1 (Met1-Arg302) fusion protein were pooled, concentrated using an Amicon YM30 membrane (Millipore Corporation), diluted with 50 mM HEPES, 200 mM NaCl, pH 7.5 buffer to reduce the imidazole concentration to about 25 mM, and digested overnight at 4°C with His₆-tagged TEV protease[41]. The digest was applied to the 12 ml Ni-NTA superflow column equilibrated in buffer A and recombinant TcPTP1 (Met1-Arg302) emerged in the column effluent. The effluent was incubated overnight at 4°C with 10 mM dithiothreitol, concentrated using an Amicon YM10 membrane, and applied to a HiPrep 26/60 Sephacryl S-200 HR column (GE Healthcare Bio-Sciences Corporation) equilibrated in 25 mM HEPES, 150 mM NaCl, 2 mM tris(2-carboxyethyl) phosphine, pH 7.2 buffer. The peak fractions containing TcPTP1 (Met1-Arg302) were pooled and concentrated to 15-25 mg mL⁻¹ (estimated at 280 nm using a molar extinction coefficient of 31,860 M⁻¹ cm⁻¹ derived

using the Expasy ProtPram tool)[42]. Aliquots were flash-frozen in liquid nitrogen and stored at -80°C. The final product was judged to be >95% pure by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The molecular weight was confirmed by electrospray ionization mass spectroscopy.

2.3 Crystallization and data collection

Recombinant TcPTP1 (15 mg/mL buffered in 25 mM Tris-HCl pH 7.2, 150 mM NaCl, and 2 mM tris(2-carboxyethyl) phosphine) was screened for crystals using a Phoenix crystallization robot (Art Robbins Instruments) and several commercial screening kits from Hampton Research, Qiagen, Jena Biosciences, and Molecular Dimensions. Initial crystals appeared after 2 days in condition C1 (3.5 M sodium formate pH 7.0) of the Qiagen Classics II screen. Optimization trials were performed by setting up grid screens with varying concentrations of protein and sodium formate and also screening additives from the Hampton Research Additive Screen. Crystals used in data collection were finally obtained by mixing 2 μ L of protein (7 mg/mL) with 2 μ L of well solution consisting of 3.2 M sodium formate and incubating the trays at 18° C. After one week, a single crystal was transferred to a cryoprotectant solution consisting of 3.2 M sodium formate and 20% (v/v) ethylene glycol, soaked for 1 minute, and flash-frozen by plunging into liquid nitrogen using a LithoLoop (Molecular Dimensions). X-ray diffraction data were collected remotely at beamline 22-BM of the SER-CAT facilities at the Advanced Photon Source, Argonne National Laboratory, using a MAR 225 CCD detector. For data collection, a 1.0 Å wavelength was used and images were collected with a 10 second exposure time and 1.0° oscillation angle. The data were indexed, integrated, and scaled using HKL3000[43]. The crystal belonged to space group R 3 2 with unit cell lengths $a=b=154.3$, $c=195.7$ Å. The Matthews coefficient of 3.32 Å³ Da⁻¹ and solvent content of 62.9% suggested the presence of two molecules in the asymmetric unit[44, 45]. The quality of the diffraction data was further assessed by *phenix.xtriage* to detect any crystal defects or potential twinning[46].

2.4 Structure determination

The structure was solved using the automated molecular replacement option with *Phaser* in *PHENIX*[46, 47]. A monomer of the *Trypanosoma brucei* protein tyrosine phosphatase TbPTP1 (PDB code: 3m4u) was selected as a search model based on 64% sequence identity with TcPTP1 and prepared for molecular replacement by removing all solvent molecules and side chains of non-identical residues[48]. The search model was used to locate two molecules in the asymmetric unit with *Phaser* and data extending to 3.0 Å resolution. The solution was then subjected to rigid-body refinement using *phenix.refine*[49]. The side chains and model corrections were built into the electron density maps after iterative rounds of model rebuilding and refinement using *phenix.refine* and *Coot*[50] while gradually increasing the resolution of the data to the maximum of 2.18 Å. After the protein model was completed, the positions of water and solvent molecules were located with *Coot*, manually inspected, and refined with *phenix.refine*. During the final stages of refinement, TLS (translational/libration/screw) refinement was performed with *phenix.refine*[51]. Model quality was assessed by *PROCHECK*[52] from the CCP4 suite of programs[53] and *MolProbity*[54]. All data collection and refinement statistics are presented in Table 1.

Structural alignments were performed with *PyMOL* (<http://www.pymol.org>) and *PDBefold*[55]. Electrostatic surface potentials were calculated with *PyMOL*.

3. Results and Discussion

3.1 Overall structure and comparison

The crystal structure of *T. cruzi* protein tyrosine phosphatase, TcPTP1, was determined at 2.18 Å resolution by molecular replacement. The asymmetric unit consists of two molecules that superimpose with an average r.m.s. deviation of 0.6 Å. Chain A consists of residues Asp3-Asn300 and chain B consists of Ser4-Leu295. In chain A, there is a lack of clear electron density for residues 68-71 found on a surface loop and so they were not included in the model. However, there is electron density visible for these residues in chain B. Furthermore, although there is continuous electron density for the backbone atoms of the N-terminal residues 4-13 in chain B, some of the side chains are disordered. Although the crystal structure consists of two molecules in the asymmetric unit, there are only limited interactions between the two molecules consisting mainly of two salt bridges (Glu A136-His B47 and His A47-Asp B33) and a few other hydrogen bonds. Analysis of the protein interfaces in the crystal lattice using PDBePISA[56] did not reveal any strong indications that the structures form a dimer in solution. Indeed, the elution profile from size exclusion chromatography demonstrated that TcPTP1 migrates as a monomer (data not shown), as was also observed for the classical PTPase from *Trypanosoma brucei*, TbPTP1b[48]. Since the structures of chain A and chain B are very similar, the coordinates of chain A were used for structural analysis.

The overall fold of TcPTP1 is very similar to those of other classical PTPases, consisting of a central, mixed β -sheet that adopts a twisted conformation with several α -helices surrounding it (Fig. 1)[57]. The catalytic, phosphate-binding loop (P loop) is located at the center of the protein (residues H₂₂₅CSAGIGRT₂₃₃) and includes the PTPase signature motif HCXXGXXR(S/T). Adjacent to the P loop, the WPD-loop containing the general acid/base Asp196 residue adopts a closed conformation that is commonly observed in PTPase structures with bound ligands[58, 59].

Sequence analysis of the trypanosomal PTPase subfamily reveals that they share many of the conserved structural motifs found in human PTPases such as the WPD loop harboring the catalytic aspartic acid, the catalytic P loop containing the invariant arginine and nucleophilic cysteine residues, the Q loop component of the active site, and several other highly conserved structural motifs located within the core of the PTPase catalytic domain[60]. In addition to these conserved structural motifs, the trypanosomal PTPases also contain several trypanosome-specific motifs. These distinct motifs consist of two regions *a* (R(M/L)QREFXQLQ) and *b* (ENPRXI(D/N)FTTSL) identified as “pre-catalytic” along with four other trypanosome-specific regions *c* (LANEXTIYP), *d* (EV(D/E)), *e* (LIGAYA), and *f* (RLGV(D/S)(I/V))[60]. The sequences found in these motifs appear to be unique to the trypanosomal PTPase family and are postulated to function as molecular recognition sites for cellular targets. Although the sequences in these motifs are distinct, they are found within the classical PTPase fold and do not give rise to any new structural features. Indeed, these regions are all located on the surface of the TcPTP1 molecule with the exception of

region *e*, which is located behind the active site P-loop and is mostly buried in the core of the protein (Fig. 2). Motifs *a*, *b*, and *c* form an extended solvent exposed region that surrounds an extended groove adjacent to the active site pocket which may be a binding site for the phosphorylated substrate.

3.2 The active site

The active site of PTPases is formed primarily by the catalytic P loop which harbors the active-site motif HCXXGXXR(S/T), where X is a variable residue, and residues from the Q- and WPD loops[57]. In TcPTP1, the P loop is composed of residues H₂₂₅CSAGIGRT₂₃₃, which are identical to the ones found in the human PTP1B P loop[61]. The catalytic P loop forms a positively charged, shallow pocket that is approximately 9 Å deep (Fig. 3). The side chain of Arg232, which is a structural determinant that provides binding interactions to the incoming substrate phosphate moiety[59], adopts a cradle-like conformation around the edge of the pocket. The His225 imidazole side chain contributes hydrogen-bonding interactions to the backbone amide (3.1 Å) of Cys226 and the OH of the side chain of Tyr133, a conserved residue among PTPases, thereby providing stabilizing interactions for the conformation of the active site cysteine located at the base of the pocket. The depth of the pocket is an important structural feature that confers phosphotyrosine specificity to the phosphatase. PTPases that have deeper pockets are typically unable to position the shorter phosphothreonine/phosphoserine residues properly for catalysis and therefore exhibit selectivity for phosphotyrosine[59]. Phosphatases that have shallower pockets typically exhibit dual-specificity and therefore are able to dephosphorylate phosphoserine and phosphothreonine residues as well[62, 63].

The removal of phosphate groups by PTPases proceeds by a two-step mechanism[30, 57, 64]. The conserved cysteine, which has been observed to possess a low pK_a in PTPases, initiates catalysis through a cysteine thiolate anion intermediate that attacks the tyrosine phosphate to form a cysteinyl-phosphate intermediate[65-67]. The tyrosine is then released via the donation of a proton from an invariant aspartic acid, (Asp196 in TcPTP1), which then serves as a general base by removing a proton from a water molecule that attacks the phospho-enzyme intermediate to eliminate phosphate and regenerate the active enzyme[67]. In the TcPTP1 structure, the WPD loop adopts a closed conformation that orients the catalytic Asp196 residue approximately 7.2 Å from the nucleophilic Cys226 residue and 4.6 Å from the Arg232 side chain. This conserved aspartic acid has been observed to undergo a large conformational shift from the apo- to substrate-bound enzyme in many PTPases, highlighting the requirement for the proper orientation of this residue in order to function as a general acid/base during the dephosphorylation reaction[58, 59, 68]. A single formate molecule from the crystallization solution is bound to the active site in chain A and forms hydrogen bonds between the carbonyl oxygens and the side chain of NH₂ of Arg232 and the backbone amides of Ser227 and Ala228. In chain B, the active site is occupied only by water molecules. There are two water molecules, Wa2123 and Wa2135, in the active site that are highly conserved in other structures of PTPases such as PTP1B and YopH. Water 2123 is coordinated by hydrogen bonds to the side chain of Gln272 (2.9 Å), located on the Q loop, and to the catalytic Asp196 (2.9 Å). This water molecule aligns well with the analogous conserved water in crystal structures of PTP1B complexed with vanadate, a transition state

analogue, and YopH complexed with nitrate, and therefore likely corresponds to the nucleophilic water that participates in the dephosphorylation reaction[58, 69]. Water 2135, which forms hydrogen bonds to the Arg232 side chain and the carbonyl oxygen of formate, is another highly conserved water molecule found in many crystal structures of PTPases. This water molecule is displaced upon binding of the phosphotyrosine substrate and its position has been used in past efforts to aid in PTPase inhibitor design[70, 71].

3.3 Comparison with structural homologs

A search of the Protein Data Bank for related structural homologs using the PDBfold server[55] indicated the closest structural homologs to TcPTP1 were the *T. brucei* protein tyrosine phosphatase TbPTP1 (r.m.s.d. =1.0 Å over 274 aligned residues, 64% sequence identity, PDB code: 3m4u)[48] and human PTP1B (1.8 Å over 251 aligned atoms, 27% sequence identity, PDB code: 2hnq)[61]. Structural alignments of TcPTP1 and these two homologs reveal that the secondary structure elements align well with most of the structural deviations existing in the loops that connect them (Fig. 4). The WPD loop in the PTP1B structure with bound tungstate (PDB code: 2hnq) is found in the open “out” position while in TcPTP1 and TbPTP1 (which has a phosphate ion bound in the active site), it adopts the closed conformation. The WPD loop in PTP1B adopts the closed conformation when substrates or small molecule inhibitors are bound in the active site[59, 69]. Although a formate molecule is found bound only in chain A of TcPTP1, the WPD loop in chain B, which only has water molecules in the active site, also adopts the closed conformation.

3.4 Insights into drug design

Although an effort to inhibit the activity of TcPTP1 using the broad spectrum PTPase inhibitor BZ3 resulted in reduced infectivity of *T. cruzi* [36], the possibility that the observed phenotype was due to the inhibition of other phosphatases such as TcPTP2, which shares 97% sequence identity with TcPTP1, could not be ruled out. Indeed, considering the very high degree of sequence identity, it is highly likely that BZ3 inhibits both enzymes. Therefore, further validation of TcPTP1 as a drug target is still needed. Yet inhibition of TcPTP2 may also have therapeutic benefits. Its ortholog in *Trypanosoma rangeli* (TrPTP2) was shown to be associated with the parasite's flagellum and may play a role in modulating the phosphorylation state of proteins that regulate its structure [72]. In any case, the development of highly specific inhibitors, facilitated by the crystal structure of TcPTP1, will ultimately resolve any lingering uncertainty about the validity of this enzyme as a molecular target for the treatment of Chagas' disease.

Close inspection of the active sites of TcPTP1 and human PTP1B indicates that the identity and conformations of the catalytic P loop residues are essentially the same (Fig. 5), which presents a significant hurdle for the design of specific inhibitors of TcPTP1. However, in such cases specificity can often be achieved through the use of bidentate or multidentate inhibitors[40, 73-75]. In this strategy, secondary, noncatalytic pockets near the active site that exhibit sequence and structural deviations from other PTPases are targeted with the bidentate inhibitors[76]. Another major challenge that is frequently encountered during the design of PTPase inhibitors is the high rate of “false positives” that can arise through inhibition of enzyme function by “promiscuous” mechanisms that can be attributed to

nonspecific factors such as protein aggregation and oxidation[77, 78]. Such promiscuous inhibitors are generally not valid leads and therefore should be avoided in the inhibitor design process[79]. The application of “substrate-activity screening” (SAS), a technique first applied to protease and PTPase targets, has shown to be highly effective at avoiding promiscuous inhibitors[80, 81]. In this approach, high affinity *substrates* are first identified and then converted to *inhibitors* by replacement of the phosphoryl group with a nonhydrolyzable phosphoryl mimetic. This process avoids promiscuous inhibition because of the requirement for proper binding of the substrate in order for turnover to occur. The potency and selectivity of the inhibitor then can be modified through further structural variations to target nearby secondary pockets[40, 82]. The combination of SAS using arylphosphate fragments, coupled with their subsequent derivatization by oxime-based click chemistry to target nearby secondary pockets on the enzyme's surface, has been successfully utilized to design potent and non-promiscuous inhibitors of the *Yersinia pestis* protein tyrosine phosphatase YopH[40].

Fig. 6 (A-C) compares the electrostatic surface potentials of the YopH, TcPTP1, and human PTP1B X-ray structures, respectively. The position occupied by an aminooxy-containing platform ((4'-((aminooxy)methyl)-[1,1'-biphenyl]-3-yl)difluoromethyl)phosphonic acid) used in the design of bidentate inhibitors of YopH is shown in its crystal structure (Fig. 6A) [40]. The crystal structures of TcPTP1 and PTP1B were aligned with that of YopH in order to model the position of the aminooxy platform in these homologs. A successful effort to design specific and potent YopH inhibitors involved the identification of tight binding arylphosphate fragment platforms to the active site followed by derivitization with additional fragments (tethering) to target a secondary pocket, P1[40]. The analogous P1 pockets in TcPTP1 and PTP1B differ in their shape, depth, and electrostatic potential, which makes them attractive prospects for achieving PTPase selectivity. The P1 pocket in TcPTP1 is both deeper and wider than that of PTP1B. Although the P1 pocket of PTP1B functions as a secondary aryl phosphate-binding site [83], it is highly unlikely that it also does so in TcPTP1 because of its substantially greater negative charge. In addition to these differences, the region around the P1 pocket in TcPTP1 is bordered by three trypanosomatid-specific motifs (Fig 2B). Therefore, the P1 pocket in TcPTP1 should be an excellent specificity determinant for inhibitor design. Comparison of the three crystal structures suggests that, as in the case of YopH[40], the application of SAS combined with fragment tethering to target secondary pockets (e.g., P1) is likely to lead to the successful development of potent and specific TcPTP1 inhibitors.

4. Conclusions

The crystal structure of TcPTP1, determined at 2.18 Å resolution, closely resembles those of the *T. brucei* TbPTP1 and human PTP1B. Although the active site of TcPTP1 is very similar to that of human PTP1B and other PTPases, a pocket situated adjacent to the active site exhibits structural deviations that may be exploited to design bidentate inhibitors that are specific for TcPTP1. Since there are very few drugs available to treat Chagas' disease, structure-based approaches to designing inhibitors of TcPTP1 may help fulfill the urgent need for new anti-parasitic drugs.

Acknowledgments

We thank Scott Cherry for technical assistance. This project has been funded in whole or in part with Federal funds from the Frederick National Laboratory for Cancer Research, National Institutes of Health, under contract HHSN261200800001E and the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does the mention of tradenames, commercial products or organizations imply endorsement by the US Government. We thank the Biophysics Resource in the Structural Biophysics Laboratory, Frederick National Lab, for use of the LC/ESMS instrument. X-ray diffraction data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22-BM of the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at <http://www.ser-cat.org/members.html>. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. W-31-109-Eng-38.

REFERENCES

1. Gurtler RE, Diotaiuti L, Kitron U. Commentary: Chagas disease: 100 years since discovery and lessons for the future. *Int J Epidemiol*. 2008; 37:698–701. [PubMed: 18653505]
2. Chagas disease: a neglected emergency. *Lancet*. 2009; 373:1820.
3. De Souza W. Basic cell biology of *Trypanosoma cruzi*. *Curr Pharm Des*. 2002; 8:269–85. [PubMed: 11860366]
4. El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, et al. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science*. 2005; 309:409–15. [PubMed: 16020725]
5. Stuart K, Brun R, Croft S, Fairlamb A, Gurtler RE, McKerrow J, et al. Kinetoplastids: related protozoan pathogens, different diseases. *J Clin Invest*. 2008; 118:1301–10. [PubMed: 18382742]
6. Schmunis GA, Yadon ZE. Chagas disease: a Latin American health problem becoming a world health problem. *Acta Trop*. 2010; 115:14–21. [PubMed: 19932071]
7. Young C, Losikoff P, Chawla A, Glasser L, Forman E. Transfusion-acquired *Trypanosoma cruzi* infection. *Transfusion*. 2007; 47:540–4. [PubMed: 17319837]
8. Carlier Y, Truyens C, Deloron P, Peyron F. Congenital parasitic infections: a review. *Acta Trop*. 2012; 121:55–70. [PubMed: 22085916]
9. Barsoum RS. Parasitic infections in transplant recipients. *Nat Clin Pract Nephrol*. 2006; 2:490–503. [PubMed: 16941042]
10. Shikanai-Yasuda MA, Carvalho NB. Oral transmission of Chagas disease. *Clin Infect Dis*. 2012; 54:845–52. [PubMed: 22238161]
11. Prata A. Clinical and epidemiological aspects of Chagas disease. *Lancet Infect Dis*. 2001; 1:92–100. [PubMed: 11871482]
12. Marin-Neto JA, Cunha-Neto E, Maciel BC, Simoes MV. Pathogenesis of chronic Chagas heart disease. *Circulation*. 2007; 115:1109–23. [PubMed: 17339569]
13. Matsuda NM, Miller SM, Evora PR. The chronic gastrointestinal manifestations of Chagas disease. *Clinics (Sao Paulo)*. 2009; 64:1219–24. [PubMed: 20037711]
14. Viotti R, Vigliano C, Lococo B, Alvarez MG, Petti M, Bertocchi G, et al. Side effects of benznidazole as treatment in chronic Chagas disease: fears and realities. *Expert Rev Anti Infect Ther*. 2009; 7:157–63. [PubMed: 19254164]
15. Buckner FS. Experimental chemotherapy and approaches to drug discovery for *Trypanosoma cruzi* infection. *Adv Parasitol*. 2011; 75:89–119. [PubMed: 21820553]
16. Guedes PM, Silva GK, Gutierrez FR, Silva JS. Current status of Chagas disease chemotherapy. *Expert Rev Anti Infect Ther*. 2011; 9:609–20. [PubMed: 21609270]
17. Tarleton RL, Reithinger R, Urbina JA, Kitron U, Gurtler RE. The challenges of Chagas Disease--grim outlook or glimmer of hope. *PLoS Med*. 2007; 4:e332. [PubMed: 18162039]
18. Lima FM, Oliveira P, Mortara RA, Silveira JF, Bahia D. The challenge of Chagas' disease: has the human pathogen, *Trypanosoma cruzi*, learned how to modulate signaling events to subvert host cells? *N Biotechnol*. 2010; 27:837–43. [PubMed: 20172059]

19. Croft SL, Barrett MP, Urbina JA. Chemotherapy of trypanosomiasis and leishmaniasis. *Trends Parasitol.* 2005; 21:508–12. [PubMed: 16150644]
20. Schaffhausen J. Advances in structure-based drug design. *Trends Pharmacol Sci.* 2012; 33:223. [PubMed: 22503451]
21. Maddaford SP. A medicinal chemistry perspective on structure-based drug design and development. *Methods Mol Biol.* 2012; 841:351–81. [PubMed: 22222460]
22. Pawson T. Protein modules and signalling networks. *Nature.* 1995; 373:573–80. [PubMed: 7531822]
23. Hunter T. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell.* 1995; 80:225–36. [PubMed: 7834742]
24. Van Vactor D, O'Reilly AM, Neel BG. Genetic analysis of protein tyrosine phosphatases. *Curr Opin Genet Dev.* 1998; 8:112–26. [PubMed: 9529614]
25. Pawson T, Nash P. Protein-protein interactions define specificity in signal transduction. *Genes Dev.* 2000; 14:1027–47. [PubMed: 10809663]
26. Alonso A, Sasin J, Bottini N, Friedberg I, Osterman A, Godzik A, et al. Protein tyrosine phosphatases in the human genome. *Cell.* 2004; 117:699–711. [PubMed: 15186772]
27. Tonks NK. Protein tyrosine phosphatases: from genes, to function, to disease. *Nat Rev Mol Cell Biol.* 2006; 7:833–46. [PubMed: 17057753]
28. Soulsby M, Bennett AM. Physiological signaling specificity by protein tyrosine phosphatases. *Physiology (Bethesda).* 2009; 24:281–9. [PubMed: 19815854]
29. Tarrant MK, Cole PA. The chemical biology of protein phosphorylation. *Annu Rev Biochem.* 2009; 78:797–825. [PubMed: 19489734]
30. Zhang ZY. Protein tyrosine phosphatases: structure and function, substrate specificity, and inhibitor development. *Annu Rev Pharmacol Toxicol.* 2002; 42:209–34. [PubMed: 11807171]
31. El-Sayed NM, Myler PJ, Blandin G, Berriman M, Crabtree J, Aggarwal G, et al. Comparative genomics of trypanosomatid parasitic protozoa. *Science.* 2005; 309:404–9. [PubMed: 16020724]
32. Brenchley R, Tariq H, McElhinney H, Szoor B, Huxley-Jones J, Stevens R, et al. The TriTryp phosphatome: analysis of the protein phosphatase catalytic domains. *BMC Genomics.* 2007; 8:434. [PubMed: 18039372]
33. Morales-Neto R, Hulshof L, Ferreira CV, Gadelha FR. Distinct phosphatase activity profiles in two strains of *Trypanosoma cruzi*. *J Parasitol.* 2009; 95:1525–31. [PubMed: 19476404]
34. Szoor B. Trypanosomatid protein phosphatases. *Mol Biochem Parasitol.* 2010; 173:53–63. [PubMed: 20594956]
35. Huang H. Signal transduction in *Trypanosoma cruzi*. *Adv Parasitol.* 2011; 75:325–44. [PubMed: 21820563]
36. Gallo G, Ramos TC, Tavares F, Rocha AA, Machi E, Schenkman S, et al. Biochemical characterization of a protein tyrosine phosphatase from *Trypanosoma cruzi* involved in metacyclogenesis and cell invasion. *Biochem Biophys Res Commun.* 2011; 408:427–31. [PubMed: 21514274]
37. Combs AP. Structure-based drug design of new leads for phosphatase research. *IDrugs.* 2007; 10:112–5. [PubMed: 17285463]
38. Zhang S, Zhang ZY. PTP1B as a drug target: recent developments in PTP1B inhibitor discovery. *Drug Discov Today.* 2007; 12:373–81. [PubMed: 17467573]
39. Blaskovich MA. Drug discovery and protein tyrosine phosphatases. *Curr Med Chem.* 2009; 16:2095–176. [PubMed: 19519384]
40. Bahta M, Lountos GT, Dyas B, Kim SE, Ulrich RG, Waugh DS, et al. Utilization of nitrophenylphosphates and oxime-based ligation for the development of nanomolar affinity inhibitors of the *Yersinia pestis* outer protein H (YopH) phosphatase. *J Med Chem.* 2011; 54:2933–43. [PubMed: 21443195]
41. Kapust RB, Tozser J, Fox JD, Anderson DE, Cherry S, Copeland TD, et al. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* 2001; 14:993–1000. [PubMed: 11809930]

42. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 2003; 31:3784–8. [PubMed: 12824418]
43. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr.* 2006; 62:859–66. [PubMed: 16855301]
44. Matthews BW. Solvent content of protein crystals. *J Mol Biol.* 1968; 33:491–7. [PubMed: 5700707]
45. Kantardjieff KA, Rupp B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.* 2003; 12:1865–71. [PubMed: 12930986]
46. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Echols N, Headd JJ, et al. The Phenix software for automated determination of macromolecular structures. *Methods.* 2011
47. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *Journal of Applied Crystallography.* 2007; 40:658–74. [PubMed: 19461840]
48. Chou S, Jensen BC, Parsons M, Alber T, Grundner C. The *Trypanosoma brucei* life cycle switch TbPTP1 is structurally conserved and dephosphorylates the nucleolar protein NOPP44/46. *J Biol Chem.* 2010; 285:22075–81. [PubMed: 20444707]
49. Afonine PV, Grosse-Kunstleve RW, Adams PD. A robust bulk-solvent correction and anisotropic scaling procedure. *Acta Crystallogr D Biol Crystallogr.* 2005; 61:850–5. [PubMed: 15983406]
50. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:486–501. [PubMed: 20383002]
51. Painter J, Merritt EA. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr.* 2006; 62:439–50. [PubMed: 16552146]
52. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. Procheck -a Program to Check the Stereochemical Quality of Protein Structures. *Journal of Applied Crystallography.* 1993; 26:283–91.
53. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr.* 2011; 67:235–42. [PubMed: 21460441]
54. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:12–21. [PubMed: 20057044]
55. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr.* 2004; 60:2256–68. [PubMed: 15572779]
56. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol.* 2007; 372:774–97. [PubMed: 17681537]
57. Zhang ZY. Protein-tyrosine phosphatases: biological function, structural characteristics, and mechanism of catalysis. *Crit Rev Biochem Mol Biol.* 1998; 33:1–52. [PubMed: 9543627]
58. Fauman EB, Yuvaniyama C, Schubert HL, Stuckey JA, Saper MA. The X-ray crystal structures of *Yersinia* tyrosine phosphatase with bound tungstate and nitrate. Mechanistic implications. *J Biol Chem.* 1996; 271:18780–8. [PubMed: 8702535]
59. Jia Z, Barford D, Flint AJ, Tonks NK. Structural basis for phosphotyrosine peptide recognition by protein tyrosine phosphatase 1B. *Science.* 1995; 268:1754–8. [PubMed: 7540771]
60. Szoor B, Wilson J, McElhinney H, Tabernero L, Matthews KR. Protein tyrosine phosphatase TbPTP1: A molecular switch controlling life cycle differentiation in trypanosomes. *J Cell Biol.* 2006; 175:293–303. [PubMed: 17043136]
61. Barford D, Flint AJ, Tonks NK. Crystal structure of human protein tyrosine phosphatase 1B. *Science.* 1994; 263:1397–404. [PubMed: 8128219]
62. Yuvaniyama J, Denu JM, Dixon JE, Saper MA. Crystal structure of the dual specificity protein phosphatase VHR. *Science.* 1996; 272:1328–31. [PubMed: 8650541]

63. Schumacher MA, Todd JL, Rice AE, Tanner KG, Denu JM. Structural basis for the recognition of a bisphosphorylated MAP kinase peptide by human VHR protein Phosphatase. *Biochemistry*. 2002; 41:3009–17. [PubMed: 11863439]
64. Zhang ZY. Mechanistic studies on protein tyrosine phosphatases. *Prog Nucleic Acid Res Mol Biol*. 2003; 73:171–220. [PubMed: 12882518]
65. Guan KL, Dixon JE. Evidence for protein-tyrosine-phosphatase catalysis proceeding via a cysteine-phosphate intermediate. *J Biol Chem*. 1991; 266:17026–30. [PubMed: 1654322]
66. Cho HJ, Krishnaraj R, Kitas E, Bannwarth W, Walsh CT, Anderson KS. Isolation and Structural Elucidation of a Novel Phosphocysteine Intermediate in the Lar Protein Tyrosine Phosphatase Enzymatic Pathway. *J Am Chem Soc*. 1992; 114:7296–8.
67. Zhang ZY, Dixon JE. Active site labeling of the Yersinia protein tyrosine phosphatase: the determination of the pKa of the active site cysteine and the function of the conserved histidine 402. *Biochemistry*. 1993; 32:9340–5. [PubMed: 8369304]
68. Denu JM, Lohse DL, Vijayalakshmi J, Saper MA, Dixon JE. Visualization of intermediate and transition-state structures in protein-tyrosine phosphatase catalysis. *Proc Natl Acad Sci U S A*. 1996; 93:2493–8. [PubMed: 8637902]
69. Brandao TA, Hengge AC, Johnson SJ. Insights into the reaction of protein-tyrosine phosphatase 1B: crystal structures for transition state analogs of both catalytic steps. *J Biol Chem*. 2010; 285:15874–83. [PubMed: 20236928]
70. Kim SE, Bahta M, Lountos GT, Ulrich RG, Burke TR Jr. Waugh DS. Isothiazolidinone (IZD) as a phosphoryl mimetic in inhibitors of the Yersinia pestis protein tyrosine phosphatase YopH. *Acta Crystallogr D Biol Crystallogr*. 2011; 67:639–45. [PubMed: 21697602]
71. Douty B, Wayland B, Ala PJ, Bower MJ, Pruitt J, Bostrom L, et al. Isothiazolidinone inhibitors of PTP1B containing imidazoles and imidazolines. *Bioorg Med Chem Lett*. 2008; 18:66–71. [PubMed: 18037290]
72. Prestes EB, Bayer-Santos E, Hermes Stoco P, Sincero TC, Wagner G, Umaki A, et al. Trypanosoma rangeli protein tyrosine phosphatase is associated with the parasite's flagellum. *Mem Inst Oswaldo Cruz*. 2012; 107:713–9. [PubMed: 22990958]
73. Xie J, Seto CT. Investigations of linker structure on the potency of a series of bidentate protein tyrosine phosphatase inhibitors. *Bioorg Med Chem*. 2005; 13:2981–91. [PubMed: 15781408]
74. Sun JP, Fedorov AA, Lee SY, Guo XL, Shen K, Lawrence DS, et al. Crystal structure of PTP1B complexed with a potent and selective bidentate inhibitor. *J Biol Chem*. 2003; 278:12406–14. [PubMed: 12547827]
75. Leone M, Barile E, Vazquez J, Mei A, Guiney D, Dahl R, et al. NMR-based design and evaluation of novel bidentate inhibitors of the protein tyrosine phosphatase YopH. *Chem Biol Drug Des*. 2010; 76:10–6. [PubMed: 20456369]
76. Guo XL, Shen K, Wang F, Lawrence DS, Zhang ZY. Probing the molecular basis for potent and selective protein-tyrosine phosphatase 1B inhibition. *J Biol Chem*. 2002; 277:41014–22. [PubMed: 12193602]
77. McGovern SL, Caselli E, Grigorieff N, Shoichet BK. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J Med Chem*. 2002; 45:1712–22. [PubMed: 11931626]
78. McGovern SL, Helfand BT, Feng B, Shoichet BK. A specific mechanism of nonspecific inhibition. *J Med Chem*. 2003; 46:4265–72. [PubMed: 13678405]
79. Seidler J, McGovern SL, Doman TN, Shoichet BK. Identification and prediction of promiscuous aggregating inhibitors among known drugs. *J Med Chem*. 2003; 46:4477–86. [PubMed: 14521410]
80. Wood WJL, Patterson AW, Tsuruoka H, Jain RK, Ellman JA. Substrate activity screening: A fragment-based method for the rapid identification of nonpeptidic protease inhibitors. *J Am Chem Soc*. 2005; 127:15521–7. [PubMed: 16262416]
81. Soellner MB, Rawls KA, Grundner C, Alber T, Ellman JA. Fragment-based substrate activity screening method for the identification of potent inhibitors of the Mycobacterium tuberculosis phosphatase PtpB. *J Am Chem Soc*. 2007; 129:9613. [PubMed: 17636914]

82. Bahta M, Liu F, Kim SE, Stephen AG, Fisher RJ, Burke TR Jr. Oxime-based linker libraries as a general approach for the rapid generation and screening of multidentate inhibitors. *Nat Protoc.* 2012; 7:686–702. [PubMed: 22422315]
83. Puius YA, Zhao Y, Sullivan M, Lawrence DS, Almo SC, Zhang ZY. Identification of a second aryl phosphate-binding site in protein-tyrosine phosphatase 1B: a paradigm for inhibitor design. *Proc Natl Acad Sci U S A.* 1997; 94:13420–5. [PubMed: 9391040]

Highlights

- The structure of TcPTP1 was solved at 2.18 Å resolution
- The overall fold closely resembles those of the *T. brucei* TbPTP1 and human PTP1B homologs
- A secondary pocket situated adjacent to the active site could be targeted to develop specific bidentate inhibitors of TcPTP1.

Figure 1.

Stereo view of the three-dimensional structure of *T. cruzi* TcPTP1 with secondary structure elements colored as follows: helix (cyan), loop (magenta), sheet (red).

Figure 2. Ribbon model (A) and space-filling model (B) of TcPTP1. Trypanosomatid-specific motifs are colored in red. Residues found in the catalytic P loop, the Q loop and the WPD loop are colored blue.

Figure 3.

The final $2F_o - F_c$ electron density map (blue, 2.18 Å resolution, contoured at 1.0 σ level) superimposed onto the TcPTP1 active site residues (carbon in gray, nitrogen in blue, oxygen in red, and sulfur in yellow). A formate molecule (carbon in cyan), from the crystallization solution, was found in the active site of chain A.

Figure 4.

Stereo views of the crystal structure of TcPTP1 (cyan) superimposed onto the structures of (A) *T. brucei* TbPTP1 (red, PDB code: 3m4u) and (B) human PTP1B (wheat, PDB code: 2hnq).

Figure 5.

The active site residues of TcPTP1 (carbon in gray, nitrogen in blue, oxygen in red, and sulfur in yellow) are aligned with the corresponding residues of human PTP1B (carbon in cyan, PDB code: 2hnq).

Figure 6.

The electrostatic surface potentials (positive in blue, negative in red) of (A) *Y. pestis* YopH (PDB code: 2y2f), (B) *T. cruzi* TcPTP1, and (C) human PTP1B (PDB code: 2hnq). An aminoxy-containing platform compound used to develop bidentate inhibitors of YopH is shown in the crystal structure of YopH (A) and modeled in the TcPTP1 and hPTP1B active sites (B-C). Secondary pockets P1 and P2 located adjacent to the active sites of the PTPases are indicated by the arrows.

Table 1**X-ray Data Collection & Refinement Statistics**

	TcPTPI
Data collection:	
X-ray Source	22-BM, SER-CAT
Wavelength (Å)	1.0
Resolution (Å) ^a	50-2.18 (2.22-2.18)
Space group	R 3 2
Unit cell parameters a=b, c= (Å)	154.3, 195.7
Total Reflections	289993
Completeness (%)	99.6 (100)
R _{sym} (%) ^b	0.061 (0.660)
Mean I/σ(I)	37.0 (2.7)
Redundancy	6.2 (6.2)
Refinement statistics	
Resolution (Å)	30.94-2.18 (2.22-2.18)
No. of Unique reflections used	46377
R _{work} ^c	0.190
R _{free}	0.232
No. of atoms/Mean B factor (Å) ²	
Protein chain A	2327/48.4
Protein chain B	2248/60.5
Water	261/51.7
Ethylene glycol	4/60.4
Formate	15/20.8
R.m.s. deviations from ideal geometry	
Bond lengths (Å)	0.008
Bond angles (°)	1.1
Ramachandran plot	
Most favored (%)	89.1
Additionally allowed (%)	10.2
Generously allowed (%)	0.6
Disallowed (%)	0.2
Molprobity Analysis	
Molprobity clash score	8.01 (95 th percentile)
Molprobity protein geometry score	2.06 (82 nd percentile)
Cβ deviations >0.25Å	0
Residues with bad bonds	0

	TcPTPI
Residues with bad angles	0
PDB code	4a1k

^aValues in parenthesis are for reflections in the highest resolution shell.

^b $R_{\text{sym}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where $\langle I(hkl) \rangle$ is the mean intensity of multiply recorded reflections.

^c $R = \frac{\sum |F_{\text{obs}}(hkl) - F_{\text{calc}}(hkl)|}{\sum F_{\text{obs}}(hkl)}$. R_{free} is the R value calculated for 5% of the data set not included in the refinement.