University of South Bohemia

Faculty of Science

A novel structurally characterized haloacid dehalogenase superfamily phosphatase from *Thermococcus thioreducens* with diverse substrate specificity

RNDr. thesis

Mgr. Petra Havlíčková

České Budějovice 2024

RNDr. thesis: Havlíčková, P., 2024: A novel structurally characterized haloacid dehalogenase superfamily phosphatase from *Thermococcus thioreducens* with diverse substrate specificity. RNDr. Thesis, in English – 10 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Annotation

This RNDr. thesis is focused on complex studies of a haloacid dehalogenase superfamily phosphatase member with an unknown substrate specificity, isolated from a hyperthermofilic archeon *Thermococcus thioreducens.* The HAD phosphatase, internally named as Tt82, was isolated, expressed and purified using standard techniques and the crystal structure was solved by X-ray crystallography method. The refined crystal structure displayed characteristic HAD superfamily elements, such as the HAD canonical fold with the HAD signature motifs in the active site. Further, possible substrates for the enzyme were proposed and examined by computational docking in the active site of the HAD phosphatase. The results of the docking studies served as a basis for an experimental measurement of Tt82 enzymatic activity against selected substrates at two different temperatures. HAD phosphatase Tt82 is a promising target for the biotechnological field, mainly for its wide substrate specificity and higher temperature optima.

Declaration:

I declare that I am the author of this qualification thesis and that I have used only the sources and literature indicated in the list of sources for the preparation of this thesis.

České Budějovice 22.1. 2024

.....

This thesis is based on following publication:

Havlickova, P., Brinsa, V., Brynda, J., Pachl, P., Prudnikova, T., Mesters, J. R., Kascakova, B., Kuty, M., Pusey, M. L., Ng, J. D., Rezacova, P., & Kuta Smatanova, I. (2019). A novel structurally characterized haloacid dehalogenase superfamily phosphatase from *Thermococcus thioreducens* with diverse substrate specificity. Acta crystallographica. Section D, Structural biology, 75(Pt 8), 743–752. https://doi.org/10.1107/S2059798319009586

I declare that I participated significantly in the publication by investigation, data curation, and writing – original draft.

My total share in this publication was 40 %.

STRUCTURAL BIOLOGY

ISSN 2059-7983

A novel structurally characterized haloacid dehalogenase superfamily phosphatase from *Thermococcus thioreducens* with diverse substrate specificity

Petra Havlickova,[a]‡ Vitezslav Brinsa,[b]‡ Jiri Brynda,[b,c] Petr Pachl,[b] Tatyana Prudnikova,[a,d] Jeroen R. Mesters,[e] Barbora Kascakova,[a] Michal Kuty,[a,d] Marc L. Pusey,[f] Joseph D. Ng,[f] Pavlina Rezacova[b,c] and Ivana Kuta Smatanova[a,d]*

[a]Institute of Chemistry, Faculty of Science, University of South Bohemia, Branisovska 1760, CZ-370 05 Ceske Budejovice, Czech Republic, [b]Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo namesti 2, CZ-166 10 Prague, Czech Republic, [c]Institute of Molecular Genetics of the Czech Academy of Sciences, Videnska 1083, CZ-142 10 Prague, Czech Republic, [d]Center of Nanobiology and Structural Biology, Institute of Microbiology of the Czech Academy of Sciences, Zamek 136, CZ-373 33 Nove Hrady, Czech Republic, [e]Institute of Biochemistry, University of Lübeck, Ratzeburger Allee 160, D-23538 Lübeck, Germany, and [f]Department of Biological Sciences, University of Alabama in Huntsville, 301 Sparkman Drive, Huntsville, AL 35899, USA. *Correspondence e-mail: ivanaks@seznam.cz

Received 29 November 2018
Accepted 3 July 2019

Edited by R. C. Garratt, University of São Paulo, Brazil

‡ These authors contributed equally to this work.

Keywords: HAD superfamily; hypothetical phosphatase; crystal structure; docking; phosphatase assay.

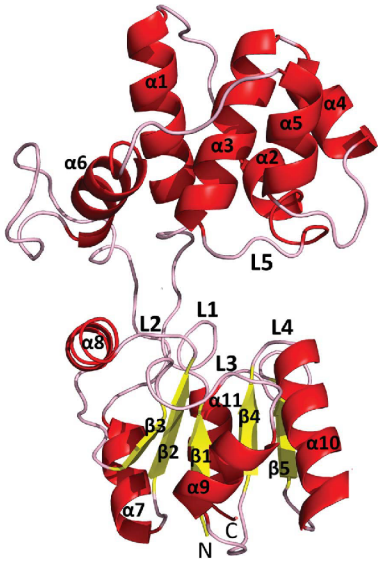
PDB reference: phosphatase Tt82 from *Thermococcus thioreducens*, 6iah

Supporting information: this article has supporting information at journals.iucr.org/d

The haloacid dehalogenase (HAD) superfamily is one of the largest known groups of enzymes and the majority of its members catalyze the hydrolysis of phosphoric acid monoesters into a phosphate ion and an alcohol. Despite the fact that sequence similarity between HAD phosphatases is generally very low, the members of the family possess some characteristic features, such as a Rossmann-like fold, HAD signature motifs or the requirement for Mg^{2+} ion as an obligatory cofactor. This study focuses on a new hypothetical HAD phosphatase from *Thermococcus thioreducens*. The protein crystallized in space group $P2_12_12$, with unit-cell parameters a = 66.3, b = 117.0, c = 33.8 Å, and the crystals contained one molecule in the asymmetric unit. The protein structure was determined by X-ray crystallography and was refined to 1.75 Å resolution. The structure revealed a putative active site common to all HAD members. Computational docking into the crystal structure was used to propose substrates of the enzyme. The activity of this thermophilic enzyme towards several of the selected substrates was confirmed at temperatures of 37°C as well as 60°C.

1. Introduction

Phosphatases in general are enzymes that are classified into a number of superfamilies. Owing to their diverse substrate specificity, understanding of their complete substrate profile and function is very limited (Fahs *et al.*, 2016). To gain insight into their biological functions and possible biotechnological applications, various biochemical and computational methods can be used. One of the desired biotechnological applications of phosphatases could be the ability to biotransform organophosphate molecules. Many of these compounds act as inhibitors of nervous-system enzymes (Kertesz *et al.*, 1994), and they are involved in several muscular and nerve diseases (Karpouzas & Singh, 2006). Nowadays, they are applied as insecticides and pesticides in agriculture (such as the commercially available herbicide glyphosate and the insecticide chlorpyrifos). In our study, we have focused on a hypothetical HAD phosphatase, internally designated Tt82, from *Thermococcus thioreducens*. The genome of *T. thioreducens*, a hyperthermophilic organism isolated from a hydrothermal vent, includes more than 300 open reading frames that



correspond to proteins with unknown function (Pikuta *et al.*, 2007; Hughes & Ng, 2007).

The haloacid dehalogenase (HAD) superfamily is a large group of enzymes that are involved in diverse cellular processes (Koonin & Tatusov, 1994). HAD superfamily members catalyze a wide range of reactions, including C—Cl bond cleavage (2-haloalkonoate dehalogenases), C—P bond cleavage (phosphoacetaldehyde hydrolase), CO—P bond cleavage (phosphate monoesterases), PO—P bond cleavage (ATPases) and P—OC bond cleavage, including intramolecular phosphoryl-group transfer (phosphomutases) (Allen & Dunaway-Mariano, 2004). The vast majority of HAD superfamily members have an unknown or a predicted function and, despite the family name, the majority of the enzymes catalyze the hydrolysis of phosphoric acid monoesters into a phosphate ion and an alcohol (Allen & Dunaway-Mariano, 2004). While HAD phosphatases share very low sequence identity in general (<15%), there are certain aspects that are characteristic of this group of enzymes. Based on the presence of these features, members of the HAD phosphatases can be identified as follows (Seifried *et al.*, 2013). (i) The overall structure of HAD phosphatases is generally divided into two domains linked by two peptide linkers (Allen & Dunaway-Mariano, 2004; Zhang *et al.*, 2004). (ii) The larger core catalytic domain is highly conserved in fold, and the characteristic structural arrangement of the catalytic domain is referred to as a Rossmann-like fold comprising repeating β–α units that form a three-layered α/β sandwich, specifically five central parallel β-strands surrounded by α-helices (Rao & Rossmann, 1973; Burroughs *et al.*, 2006; Seifried *et al.*, 2013). (iii) HAD phosphatase activity is connected with the N-terminal core domain of the protein, which contains highly conserved motifs also referred to as HAD signature motifs positioned on four loops forming the active site (Allen & Dunaway-Mariano, 2004; Zhang *et al.*, 2004; Parsons *et al.*, 2002). Motif I is defined as D*X*D*X*(T/V) (where *X* is any residue). The first highly conserved Asp residue serves as a nucleophile (Collet *et al.*, 1998) and the second Asp residue (located two residues C-terminal to the first Asp residue) serves as a general acid/base catalyst (Lahiri *et al.*, 2004). Initiation of the nucleophilic attack on the substrate (phosphomonoester) is carried out by the first Asp, while a water molecule is utilized for a nucleophilic attack on the phosphoaspartyl intermediate. In the second step of catalysis, the second Asp residue supports protonation of the substrate and deprotonation of the water molecule, leading to destruction of the intermediate form and the release of phosphate ion and substrate alcohol (Seifried *et al.*, 2013; Allen & Dunaway-Mariano, 2004). Motif II, which is poorly conserved compared with the other motifs, is defined as (S/T)G and possesses a conserved Ser or Thr that helps to orient the substrate for nucleophilic attack (Wang *et al.*, 2001). Motif III is defined as K-$(X)_n$-(G/S)(D/S)*XXX*(D/N) and comprises a conserved Lys or Arg that stabilizes the negative charge of the phosphoryl groups and also Ser/Thr from motif II (Wang *et al.*, 2001; Seifried *et al.*, 2013). Motif IV is defined as (G/S)(D/S)X_{3-4}(D/E) and comprises Asp residues that together with the Asp residues from motif I are responsible for coordination of an Mg^{2+} cofactor (Seifried *et al.*, 2013; Allen & Dunaway-Mariano, 2009). (iv) The HAD phosphatases are magnesium-dependent enzymes. Mg^{2+} ion is an essential cofactor that is required for catalysis. It is necessary for the proper positioning of the phosphoryl group and the Asp nucleophile and to electrostatically stabilize the reaction (Allen & Dunaway-Mariano, 2004). (v) The general proposed reaction for catalysis by HAD phosphatase includes two steps. The first step requires solvent exclusion, while the second step requires massive contact with bulk solvent. 'Squiggle' and 'flap' elements located in the core catalytic domain orchestrate this machinery. The 'squiggle' element forms an almost complete single-helical turn, whereas the 'flap' element constitutes a β-hairpin turn (Seifried *et al.*, 2013; Burroughs *et al.*, 2006). (vi) HAD phosphatases can be categorized into three general subclasses based on the arrangement, the topology and even the presence of the cap domain, which is diverse and plays a role in substrate specificity (Allen & Dunaway-Mariano, 2004; Peisach *et al.*, 2004; Seifried *et al.*, 2013). The C0 cap class lacks the cap domain. In the C1 cap class the cap domain is α-helical and is inserted between HAD motif I (loop 1) and HAD motif II (loop 2) (Burroughs *et al.*, 2006). The C2 cap consists of β-strands as well as α-helices and is inserted between HAD motif II (loop 2) and HAD motif III (loop 3) (Seifried *et al.*, 2013). The cap domain provides substrate specificity, and HAD family members are reported to be potentially promiscuous catalysts that are able to catalyze the enzymatic reaction of many different substrates (Pandya *et al.*, 2014). The cap domain also contains the so-called substrate-specificity loop 5 that provides the signature sequence motif V (a highly conserved Gly). This loop undergoes an open-to-closed conformational change upon substrate binding and defines the active site and specific chemistry (Caparrós-Martín *et al.*, 2013; Lahiri *et al.*, 2004).

Here, we report the expression, purification, crystallization and structure determination of the hypothetical HAD phosphatase Tt82 from *T. thioreducens*. The crystal structure confirmed the presence of the canonical HAD phosphatase fold, conserved motifs and active-site Mg^{2+} ion. The structure was used for computational docking and possible substrates were identified. Finally, activity assays with selected compounds were performed to identify substrates of Tt82.

2. Materials and methods

2.1. Macromolecule production

The gene encoding Tt82 was amplified by PCR from *T. thioreducens* genomic DNA using primers containing NdeI and BamHI restriction sites for cloning into pET-24a(+) (Novagen). The resulting construct was transformed into *Escherichia coli* BL21(DE3) cells for gene expression, which was performed in LB medium supplemented with 100 µg ml^{-1} carbenicillin and 35 µg ml^{-1} chloramphenicol at 37°C. Gene expression was induced with 0.5 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) when the optical density at 600 nm reached 0.6. The cells were cultured for 16 h before they were

Table 1
Macromolecule-production information.

Source organism	*T. thioreducens*
DNA source	*T. thioreducens* genomic DNA
Forward primer†	TTTGTTTAACTTTAAGAAGGAGATATACAT ATGTGGATAGTCTTTGACGT
Reverse primer‡	CTTCCTTTCGGGCTTTGTTAGCAGCCGGAT CCTCACAGGCCACCCTCCAGTG
Expression host	*E. coli*
Complete amino-acid sequence of the construct produced§	MWIVFDVDGVLIDVRESYDEATKLTAEYFL GLFGVEREIKPEWVRELRRKGSFGDDFK VSEALILFALSGRAEELVEEFPEGGTIE WVREKFGFQVFGGSIERVFNTFYLGREY PERLFDFPGLWKKERPIVRRGLLERASK HFKLGVVTGRSALEMELAERIIGFKFEN AVTREAYLKPDPRALWELVRGEPGVYIG DTINDELFVENYRGKYGDFDFVMVGRDV KDVNEFLENALEGGL

† The restriction site for NdeI is underlined. ‡ The restriction site for BamHI is underlined. § NCBI Reference Sequence WP_055429797.1.

Table 2
Crystallization.

Method	Sitting-drop vapour diffusion
Plate type	CombiClover 24-well plate
Temperature (°C)	22
Protein concentration (mg ml^{-1})	4.4
Buffer composition of protein solution	0.1 *M* MES pH 6.0
Composition of reservoir solution	0.2 *M* magnesium chloride hexahydrate, 25%(*w*/*v*) polyethylene glycol 3350, 0.1 *M* Tris pH 8.5
Volume and ratio of drop	2 µl, 1:1
Volume of reservoir (µl)	100

harvested by centrifugation, resuspended in buffer *A* (50 m*M* Tris–acetate, 50 m*M* ammonium sulfate pH 8.2) and lysed by sonication; cell debris was then removed by centrifugation (17 000*g*, 40 min). The supernatant was heated for 30 min at 75°C and the precipitate was removed by centrifugation (17 000*g*, 5 min). The supernatant was loaded onto an anion-exchange column (Bio-Rad Q Sepharose column; Bio-Rad, USA) equilibrated with buffer *A*. The protein was subsequently eluted with a linear gradient of 0.05–1 *M* ammonium sulfate in buffer *B* (50 m*M* Tris–acetate, 0.5 *M* ammonium sulfate pH 8.2) using a Bio-Rad low-pressure system (Bio-Rad, USA). Eluted fractions containing the protein band of expected molecular weight on an SDS–PAGE gel were pooled and further concentrated using a Millipore Ultra concentrator. The concentrated protein was then subjected to size-exclusion chromatography using a Sephacryl S-200 gel-filtration column (GE Healthcare) pre-equilibrated with buffer *C* (50 m*M* HEPES–NaOH pH 7.5, 100 m*M* NaCl). Fractions corresponding to the principal peak were run on an SDS–PAGE gel. Fractions with a clean single band at the expected molecular weight were collected and concentrated to 40 mg ml^{-1}. Macromolecule-production information is summarized in Table 1.

2.2. Crystallization

In order to determine the appropriate protein concentration for crystallization screening, PCT (Pre-Crystallization Test; Hampton Research, Aliso Viejo, California, USA) was applied. Initial crystallization trials were carried out with an Oryx-6 crystallization robot (Douglas Instruments, Hungerford, England) using the sitting-drop vapour-diffusion procedure in 96-well crystallization plates. The commercial Index screen kit (Hampton Research, Aliso Viejo, California, USA) was applied. Two protein:well solution ratios (0.5:1 and 1:1) were set up at a protein concentration of 4.4 mg ml^{-1} and were equilibrated against 100 µl reservoir solution at 22°C. Optimization of the preliminary crystallization conditions yielding crystals was performed manually by the sitting-drop vapour-diffusion technique in 24-well CombiClover crystallization plates (Rigaku Reagents, USA) at 22°C by varying the drop ratio and the protein concentration. Crystals that were suitable for X-ray diffraction measurements were obtained in condition No. 85, which consisted of 0.2 *M* magnesium chloride hexahydrate, 25%(*w*/*v*) polyethylene glycol 3350, 0.1 *M* Tris pH 8.5. Crystallization information is summarized in Table 2.

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BL14.2, BESSY II
Wavelength (Å)	0.9184
Temperature (K)	100
Detector	PILATUS3 S 2M
Crystal-to-detector distance (mm)	266.6
Rotation range per image (°)	0.1
Total rotation range (°)	200
Exposure time per image (s)	2.5
Space group	$P2_12_12$
a, *b*, *c* (Å)	66.3, 117.0, 33.8
α, β, γ (°)	90.00, 90.00, 90.00
Mosaicity (°)	0.15
Resolution range (Å)	50.00–1.75 (1.80–1.75)
Total No. of reflections	196687 (14247)
No. of unique reflections	27374 (1997)
Completeness (%)	99.9 (96.4)
Multiplicity	7.2 (7.1)
$\langle I/\sigma(I)\rangle$	18.4 (3.6)
$R_{\rm r.i.m.}$ (%)	6.1 (51.0)
Overall *B* factor from Wilson plot (Å^2)	29.4

2.3. Data collection and processing

A single crystal was fished out directly from the crystallization drop and was flashed-cooled in liquid nitrogen without any additional cryoprotection. X-ray diffraction data were collected on BL14.2 at the BESSY II electron-storage ring operated by the Helmholz-Zentrum Berlin (Mueller *et al.*, 2015). The data set was processed using *XDS* (Kabsch, 2010) with the *XDSAPP* graphical user interface (Sparta *et al.*, 2016). The crystal belonged to space group $P2_12_12$, with unit-cell parameters *a* = 66.3, *b* = 117.0, *c* = 33.8 Å, and contained one molecule in the asymmetric unit. Data-collection and processing statistics are given in Table 3.

2.4. Structure solution and refinement

Initial attempts to solve the structure of the hypothetical HAD phosphatase Tt82 by a molecular-replacement method

using the closest sequence homolog from the PDB, a putative phosphoglycolate phosphatase from *Lactobacillus plantarum* (PDB entry 2hdo, 22% sequence identity; Joint Center for Structural Genomics, unpublished work), as a model with *MOLREP* (Vagin & Teplyakov, 2010) and *Phaser* (McCoy *et al.*, 2007) from the *CCP*4 suite (Winn *et al.*, 2011) failed. The crystal structure was solved using *ARCIMBOLDO_LITE* (Sammito *et al.*, 2015; Millán, 2015) from the *CCP*4 suite (Winn *et al.*, 2011). The structure was refined using *REFMAC*5 (Murshudov *et al.*, 2011) and manually rebuilt using *Coot* (Emsley *et al.*, 2010). The refinement statistics are shown in Table 4. The quality of the final model was validated with *MolProbity* (Chen *et al.*, 2010). All figures showing structural representations were prepared with *PyMOL* (DeLano, 2002). The atomic coordinates and experimental structure factors were deposited in the Protein Data Bank under accession code 6iah.

2.5. Docking studies

Computational docking studies were performed using *AutoDock Vina* v.1.1.2 (Trott & Olson, 2010) as implemented in *UCSF Chimera* 1.3.1 (Pettersen *et al.*, 2004). The crystal structure of Tt82 with a magnesium ion bound in the putative active site was used as the receptor. Thirty different ligands were chosen for docking according to the list of the most commonly used substrates of HAD phosphatases (see Table 5; Huang *et al.*, 2015). The ligands were obtained from the Pubchem database (Kim *et al.*, 2016) or built manually and the geometry was optimized using *UCSF Chimera* (Pettersen *et al.*, 2004). Hydrogens and charges were added to the ligands and the receptor. The docking searches were performed with an exhaustiveness of eight, ten modes and energy ranges of 3 kcal mol^{-1}. Searches were carried out over the whole molecule, allowing ligands to be flexible. The results were examined based on the involvement of magnesium ion in the ligand interactions and were visualized in *PyMOL* (DeLano, 2002).

2.6. Substrate screening

To find potential substrates of Tt82, commercially available molecules that scored well in docking studies were selected and tested using the EnzChek Phosphate Assay Kit (Invitrogen). This kit enables the continuous detection of inorganic phosphate generated by enzymatic reactions (Webb, 1992). The set of tested compounds consisted of acetylphosphate, AMP, D-erythritol 4-phosphate, D-glucose 1-phosphate, D-glucose 6-phosphate, glycerol 1-phosphate and glycerol 2-phosphate, all of which were purchased from Sigma–Aldrich. Enzymatic reactions were performed at 26°C in a 96-microwell plate (MediSorp Nunc-Immuno 96 MicroWell plate, Fisher Scientific) and the reaction mixture consisted of 25 µ*M* Tt82, 1 U purine nucleoside phosphorylase and 0.2 m*M* 2-amino-6-mercapto-7-methylpurine riboside (MESG) in 50 m*M* Tris–HCl pH 7.5 supplemented with 1 m*M* $MgCl_2$ and 0.1 m*M* sodium azide. Reactions were started by the addition of potential substrates to the reaction mixture, resulting in a 10 m*M* final concentration of the compound in a total volume of 100 µl. The absorbance of the end product, 2-amino-6-mercapto-7-methylpurine, was measured at 360 nm for 120 min in 30 d intervals using a Tecan Infinite M200 microplate reader (Tecan Life Sciences). Each reaction was performed in duplicate. Corresponding negative controls lacking the Tt82 enzyme component were also performed.

The absorbance of the corresponding negative control was then subtracted from the absorbance of each sample during the reaction time of 30 min. A calibration curve for inorganic phosphate was used to calculate the conversion rates of the particular substrates in the enzymatic reactions.

Table 4
Structure refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	43.85–1.70 (1.80–1.75)
Completeness (%)	99.9 (99.6)
σ Cutoff	None
No. of reflections, working set	26033 (1898)
No. of reflections, test set	1340 (97)
Final R_{cryst} (%)	18.5 (34.0)
Final R_{free} (%)	22.5 (36.8)
Cruickshank DPI	0.112
No. of non-H atoms	
Protein	2078
Ion	10
Water	231
Total	2319
R.m.s. deviations	
Bonds (Å)	0.017
Angles (°)	1.78
Average *B* factors (Å^2)	
Overall	31.5
Protein	31.0
Ion	40.6
Water	41.4
Ramachandran plot†	
Most favoured (%)	97.1
Allowed (%)	2.1

† Determined by the program *MolProbity* (Chen *et al.*, 2010).

2.7. Steady-state kinetics

To determine the steady-state kinetics, the phosphatase activity of Tt82 towards AMP and 4-nitrophenyl phosphate (*p*NPP) was measured with 5 µ*M* Tt82 in 50 m*M* Tris–HCl pH 7.5 supplemented with 1 m*M* $MgCl_2$. Reactions were started by the addition of various amounts of substrate to the reaction mixture of total volume 100 µl, resulting in various concentrations of substrate ranging from 0.1 to 300 m*M*. Enzyme assays with AMP were performed for 120 min at 37°C and for 20 min at 60°C, while enzyme assays with *p*NPP were performed for 60 min at 37°C and for 10 min at 60°C. Each reaction was then chilled on ice and stopped by the addition of 0.14 *M* ethylenediaminetetraacetic acid (EDTA). Each reaction was performed in duplicate.

The hydrolysis of substrates to products was detected using reverse-phase high-performance liquid chromatography with spectrophotometric detection at λ = 260 and 310 nm for AMP and *p*NPP, respectively. A 5 µl aliquot of reaction mixture was applied onto a Gemini 5 µm C18 110 Å column (LC Column

Table 5
Results of docking studies.

Substrate	Free energy of binding (kcal mol^{-1})	Asp–PO_4 distance† (Å)	H-bnd‡	H-bnd lig§	H-bnd rec¶
2-Deoxy-6-phosphoglucitol	−5.3	3.55	4	3	3
2-Deoxy-D-glucose 6-phosphate	−5.7	3.61	2	2	2
2-Deoxy-D-ribitol 5-phosphate	−4.2	3.53	5	4	5
2-Deoxy-D-ribonate 5-phosphate	−5.2	3.54	4	4	4
2-Deoxy-D-ribose 5-phosphate	−5.1	3.68	1	1	1
3-Deoxy-D-gluconate 6-phosphate	−5.1	3.53	5	5	4
3-Deoxy-D-glucose 6-phosphate	−6.1	3.68	4	4	4
3-Deoxy-D-sorbitol 6-phosphate	−4.7	3.56	4	4	3
Acetylphosphate	−3.4	3.67	1	1	1
Adenosine monophosphate (AMP)	−7.0	3.58	3	3	3
Arabinose 5-phosphate	−4.8	3.51	3	3	3
D-Allitol 6-phosphate	−5.7	3.55	4	3	3
D-Glucose 1-phosphate	−5.6	3.60	1	1	1
D-Glucose 6-phosphate	−6.0	3.51	3	3	3
D-Lyxose 5-phosphate	−5.1	3.55	2	2	2
D-Mannitol 6-phosphate	−5.4	3.60	6	5	4
D-Ribitol 5-phosphate	−5.4	3.52	7	6	5
D-Ribonate 5-phosphate	−5.4	3.53	4	4	3
D-Sorbitol 1-phosphate	−5.4	3.51	7	6	6
D-Sorbitol 6-phosphate	−5.3	3.56	6	5	5
Glycerol 1-phosphate	−4.3	3.90	2	2	1
Glycerol 2-phosphate	−4.3	3.70	3	3	2
Imidodiphosphate	−4.6	3.84	4	4	2
L-Arabitol 1-phosphate	−5.2	3.81	3	3	3
L-Xylitol 5-phosphate	−5.4	3.64	3	3	3
L-Xylonate 5-phosphate	−5.2	3.56	5	5	5
meso-Erythrol 4-phosphate	−4.9	3.52	4	4	4
para-Nitrophenyl phosphate (*p*NPP)	−5.5	3.62	3	3	3
Pyridoxal 5-phosphate	−5.9	3.74	2	2	1
Riboflavin 5-phosphate	−7.5	3.51	1	1	1

† Distance between the $O^{\delta 2}$ atom of Asp6 and the P atom of the phosphate group in the ligand. ‡ No. of ligand–protein hydrogen bonds. § No. of ligand atoms involved in hydrogen bonds. ¶ No. of protein atoms involved in hydrogen bonds.

100 × 3 mm, Phenomenex) connected to an Agilent 1100 system and separated by isocratic elution in 0.1 *M* KH_2PO_4 pH 6.0 buffer supplemented with 5% or 20% methanol for the separation of samples containing AMP or *p*NPP, respectively. Isocratic elution was followed by washing the column with a solution of pure methanol containing 0.05%(*v*/*v*) trifluoroacetic acid (TFA) in each run.

Conversion rates and the respective initial reaction velocities in each sample were calculated from the areas of the peak belonging to the product and the peak with mobility corresponding to the standard for the substrate, where the resulting value of the substrate conversion was always below 10%. Using *GraphPad Prism* 5 (GraphPad Software), initial reaction velocities were plotted against respective initial substrate concentrations and the steady-state kinetics parameters K_m and k_{cat} (and K_i in the corresponding cases) were then obtained from nonlinear regression of the plots with either the classical Michaelis–Menten equation (1) or the substrate-inhibition equation (2) (Copeland, 2000):

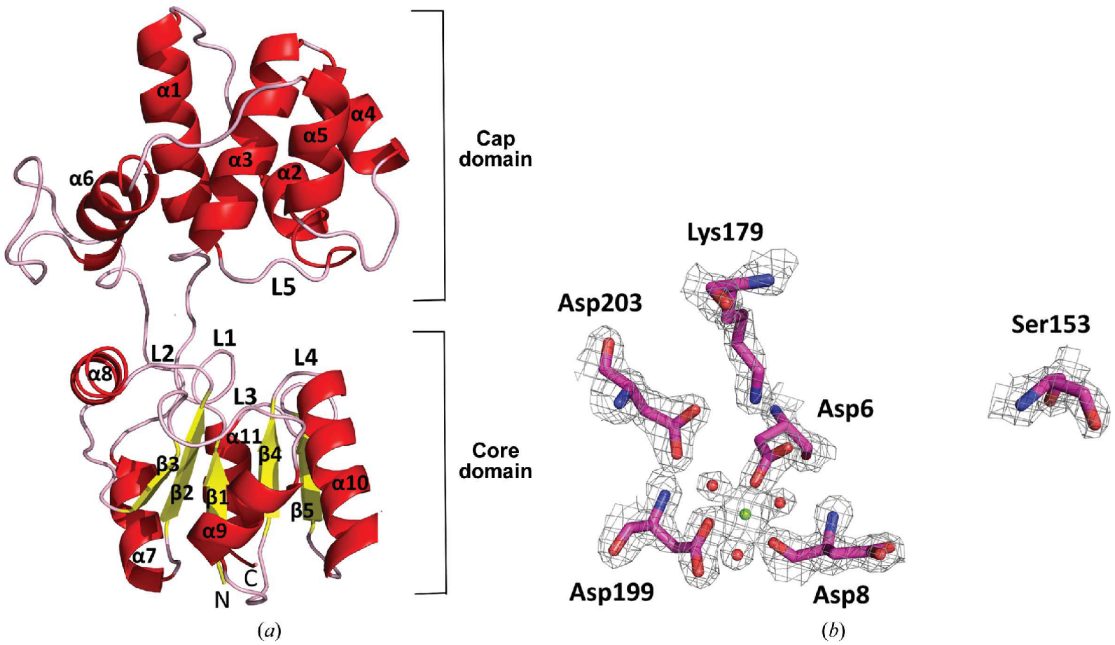


Figure 1
(*a*) The overall structure of the Tt82 protein is shown in cartoon representation; the N- and C-termini and secondary-structure elements are labelled. Four loops from the core domain that form the active site are labelled L1–L4 and the cap domain specificity loop is labelled L5. (*b*) Detail of the predicted active site. The catalytic residues are represented by sticks; the magnesium ion is shown as a green sphere with coordinating water molecules as red spheres. The $2F_o - F_c$ electron-density map for ions and interacting residues is contoured at 2.0σ above the mean.

$$v_0 = \frac{V_{max} \cdot [S]}{K_m + [S]}, \quad (1)$$

$$v_0 = \frac{V_{max} \cdot [S]}{K_m + [S] \cdot \left(1 + \frac{[S]}{K_i}\right)}. \quad (2)$$

3. Results and discussion

3.1. Crystal structure of Tt82

The crystal structure of HAD phosphatase Tt82 originating from the hyperthermophilic archeon *T. thioreducens* was refined to 1.75 Å resolution. Tt82 is a 241-residue monomeric protein with a molecular weight of approximately 27.9 kDa. The protein crystallized in space group $P2_12_12$ and contained one molecule in the asymmetric unit, with a solvent content of 48%. The crystallographic model contained 240 residues in one polypeptide chain. Clear electron density for the C-terminal leucine residue was not present and this residue is thus missing in the crystallographic model. Nonprotein positive $2F_o - F_c$ electron density was attributed to three magnesium ions, seven chloride ions and 231 water molecules.

The native crystal structure of Tt82 is composed of 11 α-helices and five β-strands sequentially arranged in the order $\beta1$–$\alpha1$–$\alpha2$–$\alpha3$–$\alpha4$–$\alpha5$–$\alpha6$–$\alpha7$–$\beta2$–$\alpha8$–$\beta3$–$\alpha9$–$\beta4$–$\alpha10$–$\beta5$–$\alpha11$ [Fig. 1(*a*)]. The protein structure is divided into two domains: the core and cap domains. Residues 1–12 and 128–240 form a core domain, which is arranged into a three-layered β/α sandwich known as the Rossmann-like fold (Rao & Rossmann, 1973). Residues 19–122 form the cap domain. Two linkers, residues 13–18 and 123–127, connect the two domains. The predicted active site is formed by signature motifs located on five loops, similar to other HAD members [Fig. 1(*a*)]. Loop 1 (connecting $\beta1$ to $\alpha1$) contains Asp6 and Asp8. Loop 2 (connecting $\beta2$ to $\alpha8$) includes Ser153, and loop 3 (connecting $\beta3$ to $\alpha9$) includes Lys179. Loop 4 (connecting $\beta4$ to $\alpha10$) contains two conserved Asp residues, namely Asp199 and Asp203. Cap domain specificity loop 5 contains the conserved Gly54 and is labelled L5 [Fig. 1(*a*)]. The putative active site is occupied by a magnesium ion coordinated by three water molecules (W144, W145 and W146), the $O^{\delta1}$ atom of Asp199, the $O^{\delta2}$ atom of Asp6 and the main-chain carbonyl of Asp8 [Fig. 1(*b*)]. The cap domain is a C1 cap type as it is helical and is inserted between motif I and motif II (Burroughs *et al.*, 2006).

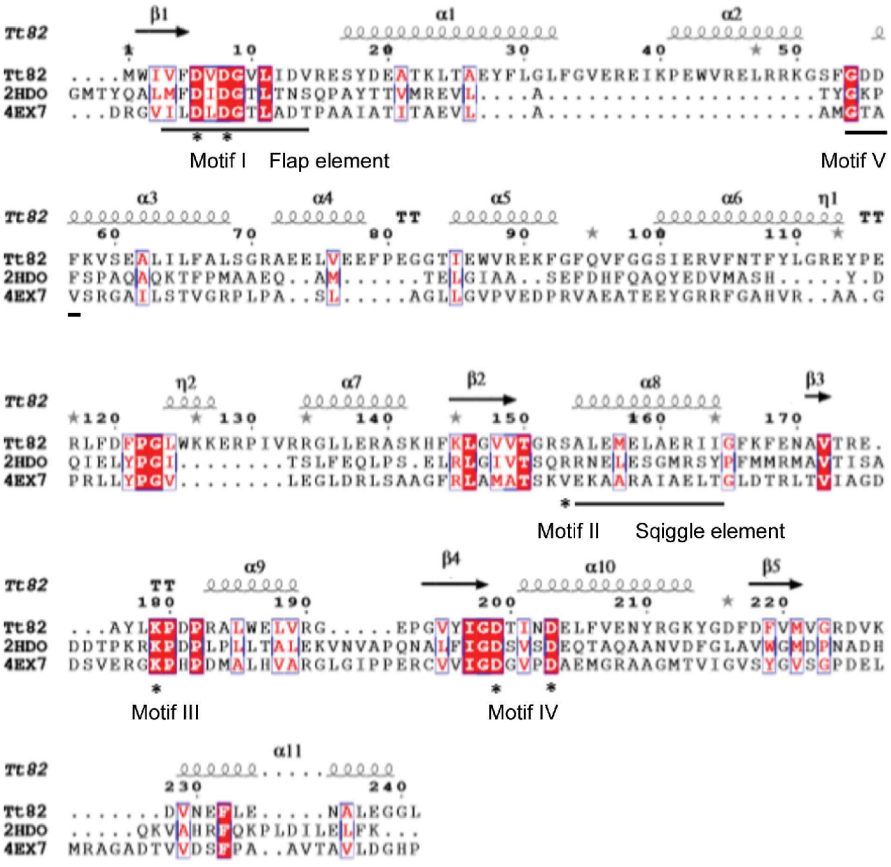


Figure 2
Sequence alignment and secondary-structure elements of Tt82. Conserved motifs are also highlighted. The putative phosphoglycolate phosphatase from *L. plantarum* (PDB entry 2hdo) and the alnumycin P phosphatase AlnB from *Streptomyces* sp. (PDB entry 4ex7) were chosen for sequence-alignment analysis. The multiple sequence alignment was performed using *ClustalW* (Larkin *et al.*, 2007) and displayed using *ESPript* (Robert & Gouet, 2014).

3.2. Comparison with structural homologs

A sequence search in the PDB was performed to identify protein homologs with known three-dimensional structures. The search revealed two structures with an identity of greater than 20%, namely the crystal structure of the putative phosphoglycolate phosphatase from *L. plantarum* (PDB entry 2hdo; 22% sequence identity; Joint Center for Structural Genomics, unpublished work) and the crystal structure of an *O*-methyltransferase from *Legionella pneumophila* (PDB entry 3r3h; 27% sequence identity; New York Structural Genomics Research Consortium, unpublished work). These structures were used in unsuccessful attempts to solve the crystal structure by molecular replacement.

Once the structure of Tt82 had been determined *ab initio* using *ARCIMBOLDO_LITE* (Millán *et al.*, 2015), the search for structural homologs was repeated using the *PDBeFold* (Krissinel & Henrick, 2004) and *DALI* (Holm & Sander, 1997) servers. As expected, numerous HAD enzymes were found

among the structural homologs of Tt82. The closest structural homologs were a predicted phosphoglycolate phosphatase (PDB entry 2hsz; 16% sequence identity; *Z*-score of 8.651 and an r.m.s.d. of 0.52 Å for the superposition of 138 C^{α} atoms) and the hypothetical protein Q8NW41 from *Staphylococcus aureus* (PDB entry 1qyi; 20% sequence identity; *Z*-score of 17.7 and an r.m.s.d. of 0.34 Å for the superposition of 63 C^{α} atoms).

Two homologs were selected for structural comparison: (i) the putative phosphoglycolate phosphatase from *L. plantarum* (PDB entry 2hdo), with an r.m.s.d. value of 0.31 Å for the superposition for 106 C^{α} atoms, and (ii) the closest structurally

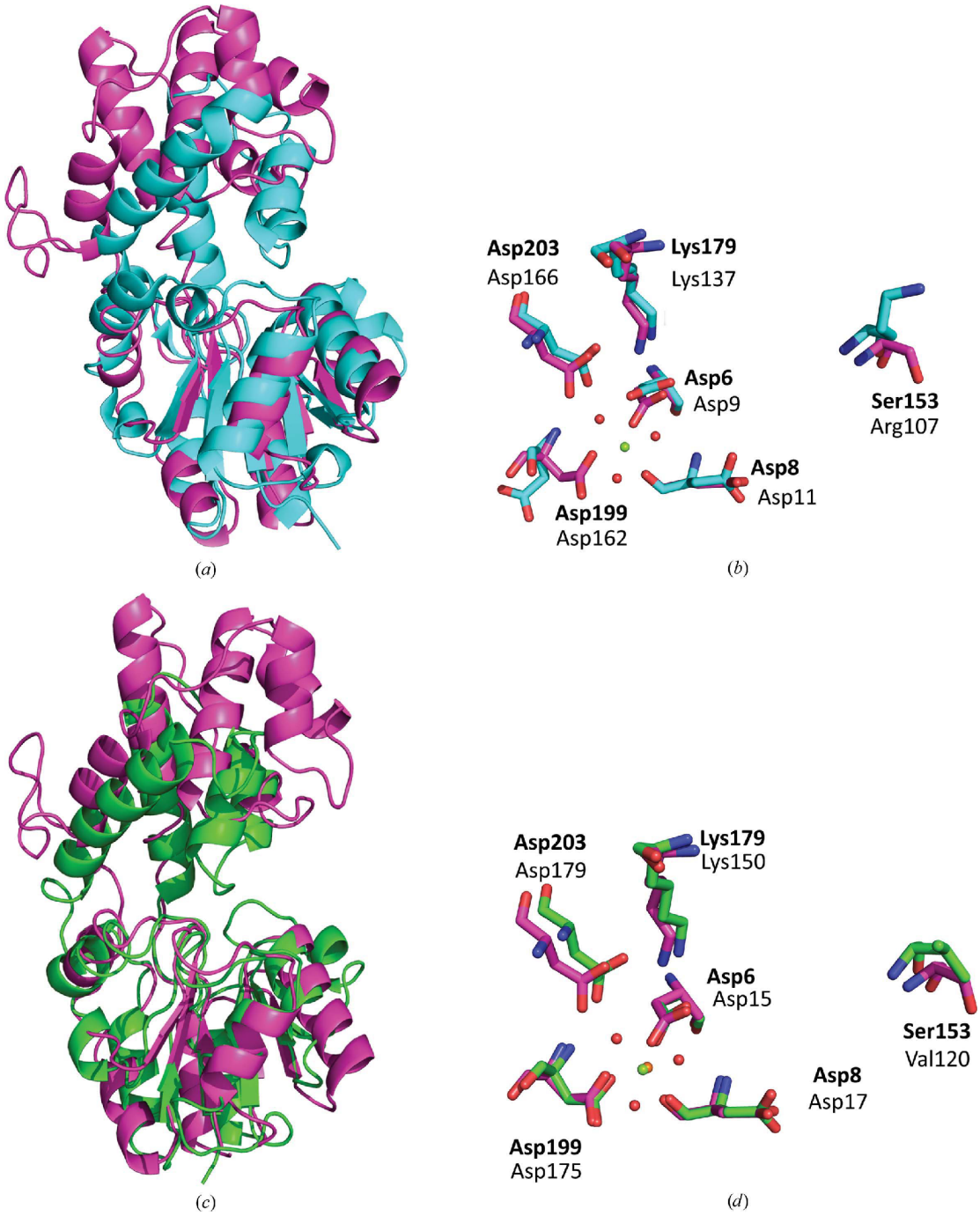


Figure 3
Superposition of Tt82 with structural homologs. (*a*) Overall superposition of Tt82 (magenta) with PDB entry 2hdo (cyan). (*b*) Superposition of the active site of Tt82 (magenta) with the active site of PDB entry 2hdo; Arg107 is incomplete (cyan). (*c*) Overall superposition of Tt82 (magenta) with PDB entry 4ex7 (green). (*d*) Superposition of the active site of Tt82 (magenta) with the active site of PDB entry 4ex7 (green). Tt82 residues are labelled in bold. The magnesium ion in Tt82 is shown as a green sphere and coordinating water molecules are shown as red spheres. The magnesium ion in PDB entry 4ex7 is shown as an orange sphere.

and functionally characterized enzyme alnumycin P phosphatase AlnB from *Streptomyces* sp. CM020 (PDB entry 4ex7; Oja *et al.*, 2012), with an r.m.s.d. value of 0.37 Å for the superposition of 100 C^{α} atoms (Figs. 2 and 3).

The sequence and structural alignment revealed that the HAD signature motifs (motif I, motif II and motif IV) are well conserved in all analyzed proteins, while motif III is poorly conserved. Squiggle and flap elements are also present in all three structures (Fig. 2). Detailed comparison of the active site of the structure shows the conserved position of the magnesium ion and coordinating water molecules. The active site of the AlnB phosphatase is also occupied by a magnesium ion that is in an octahedral coordination geometry involving three water molecules and three essential Asp residues [Asp15, Asp17 and Asp175; Fig. 3(*d*)].

Despite the fact that all cap domains are C1 cap type, the most significant differences are observed in the cap domains, which differ not only in sequence but also in the arrangement, number and positions of α-helices [Figs. 3(*a*) and 3(*c*)].

3.3. Docking studies

As the true physiological substrate of Tt82 is presently unknown, we carried out computational docking of potential substrates into the rigid crystal structure of Tt82. The structure of Tt82 displays the cap-closed conformation, in which the cap substrate specificity loop 5 (residues 53–56) enters the active site, and it is able to participate in substrate binding and catalysis (Huang *et al.*, 2011; Lahiri *et al.*, 2004). The inter-domain movement that is commonly observed during the catalytic cycle for the C1 HAD members (Burroughs *et al.*, 2006) was not involved in our docking studies because the active site of the Tt82 crystal structure is in an appropriate conformation for substrate binding.

The ligands/substrates were selected based on a list of the most commonly used substrates of HAD phosphatases (Huang *et al.*, 2015). The docking study was performed over the whole molecule and all selected substrates docked into the predicted active site. The results of the docking studies are summarized in Table 5, which shows the list of ligands and the corresponding values representing ligand–protein binding energy, the distance between the $O^{\delta 2}$ atom of Asp6 and the P atom of the phosphate group of the ligand, the number of ligand–protein hydrogen bonds, the number of ligand atoms involved in hydrogen bonds and, finally, the number of protein atoms involved in hydrogen bonds. For each ligand only one pose was selected (from the total of nine poses found in the docking process) based on the best binding position in the Tt82 active site and the lowest binding energy, respectively.

We examined the results of our docking studies according to the previous study (Krachtus *et al.*, 2018), in which the involvement of magnesium ion and the Asp nucleophile in ligand interaction was also observed.

The results of the docking simulations suggest that Tt82 is likely to dephosphorylate most of the examined substrates. The phosphate group of the docked compounds acquires a position in the vicinity of the catalytic residues and magnesium ion is also involved in the interaction (Fig. 4). We assume that the most important residues for the enzymatic reaction are two conserved Asp residues (Asp6 and Asp8), which is supported by mutagenesis of these Asp residues in the homolog AlnB phosphatase from *Streptomyces* sp., which leads to a loss of activity (Oja *et al.*, 2012).

3.4. Enzymatic activity of Tt82

The predicted phosphatase activity of Tt82 was tested with the EnzChek Phosphate Assay Kit (Invitrogen) at a temperature of 26°C using a set of molecules which were

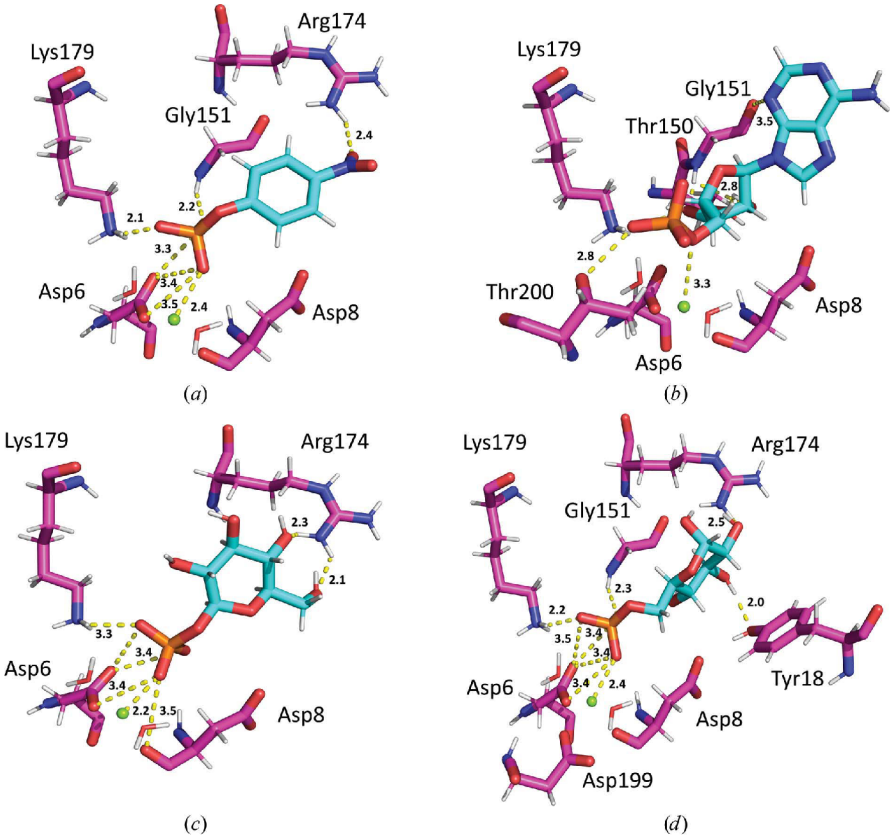


Figure 4
Predicted substrates docked in the active site of Tt82: (*a*) *para*-nitrophenyl phosphate (*p*NPP), (*b*) adenosine monophosphate (AMP), (*c*) D-glucose 1-phosphate and (*d*) D-glucose 6-phosphate. C atoms of Tt82 residues are shown in magenta and C atoms of ligands are in cyan. O atoms are depicted in red, N atoms in blue and P atoms in orange. The magnesium ion in Tt82 is shown as a green sphere. Distances between hydrogen-bond donors and acceptors are shown in Å.

successfully docked into the enzyme active site and are commercially available: acetyl phosphate, AMP, D-erythritol 4-phosphate, D-glucose 1-phosphate, D-glucose 6-phosphate, glycerol 1-phosphate and glycerol 2-phosphate.

The phosphatase activity of Tt82 towards AMP, D-glucose 1-phosphate and D-glucose 6-phosphate was confirmed, while the other tested compounds appeared to be rather poor substrates or were not substrates at all [Fig. 5(*a*)].

The activity of Tt82 towards AMP was further characterized by steady-state kinetics using reverse-phase high-performance liquid chromatography with spectrophotometric detection. *para*-Nitrophenyl phosphate (*p*NPP), a commonly used chromogenic substrate for phosphatases (Lorenz, 2011), was also included in the experiment, from which the steady-state kinetics parameters K_m and k_{cat} were obtained at two different temperatures [37 and 60°C; Figs. 5(*b*)–5(*e*)].

Activity towards both of them was detected at a temperature of 37°C and was found to be markedly greater at 60°C, consistent with the hyperthermophilic origin of Tt82 [Figs. 5(*b*)–5(*e*)]. At both temperatures Tt82 exhibited a higher affinity towards *p*NPP than towards AMP, as documented by the Michaelis–Menten constants [Figs. 5(*b*)–5(*e*)]. Nevertheless, the turnover number k_{cat} was almost the same for both *p*NPP and AMP at a given temperature [37 or 60°C; Figs. 5(*b*)–5(*e*)]. Interestingly, the phenomenon of inhibition by substrate was observed at high substrate concentrations (Copeland, 2000). The activity of Tt28 was inhibited by AMP, with an estimated K_i value of 27.72 ± 3.54 m*M* at 37°C and 13.82 ± 3.60 m*M* at 60°C. In the tested range of substrate concentrations, inhibition by *p*NPP was only observed at 60°C, with a K_i value of 142.9 ± 7.0 m*M*.

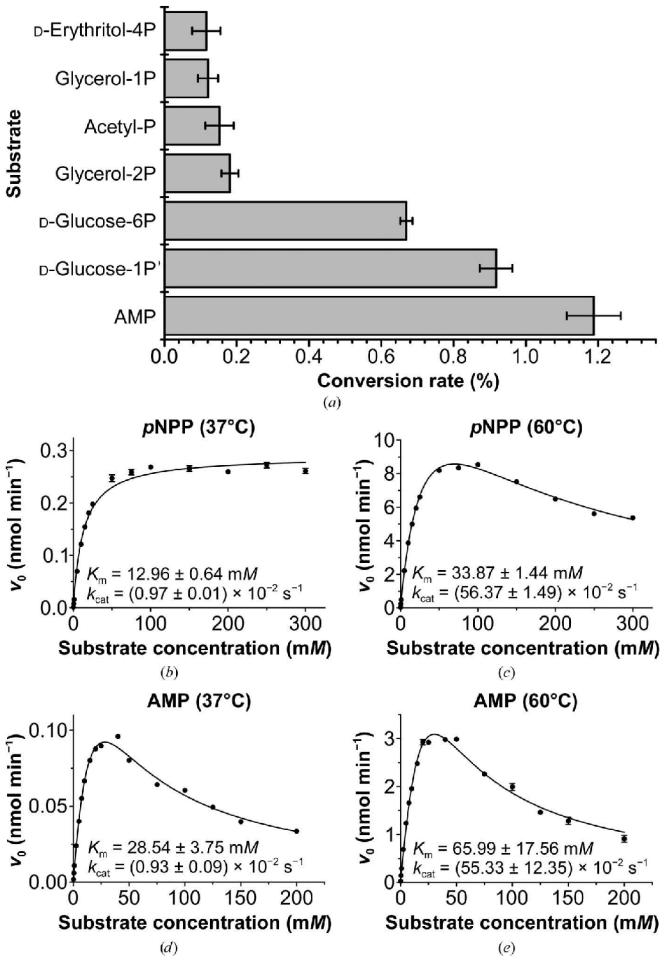


Figure 5
Phosphatase activity of Tt82. (*a*) Conversion rates of selected compounds assayed by the EnzChek Phosphate Assay Kit. (*b*)–(*e*) Steady-state kinetics of *para*-nitrophenyl phosphate (*p*NPP) and AMP at two different temperatures (37 and 60°C). At higher substrate concentrations inhibition by substrate was observed (see the text for details).

4. Conclusions

We have expressed, purified and crystallized the hypothetical HAD phosphatase Tt82 from *T. thioreducens* and have determined its structure at 1.75 Å resolution. Based on the structural information, we have identified the putative active site with the presence of the HAD signature motifs. We have predicted potential substrates of Tt82 by molecular-docking studies and have confirmed some of them experimentally. Specifically, AMP, D-glucose 1-phosphate, D-glucose 6-phosphate and *p*NPP were confirmed as possible substrates of Tt82. Consistent with the hyperthermophilic origin of Tt82, a higher activity at elevated temperature (60°C) was observed. The catalytic efficiency for the predicted substrates was actually very low. The reason for this might be that the true physiological substrates for this enzyme were not identified in our study, as only a limited number of commercially available compounds were tested. Alternatively, we might speculate that the low catalytic efficiency is the consequence of a wide substrate specificity. The promiscuity of Tt82 might play a physiological role, although the true biological function of this enzyme in the lifecycle of *T. thioreducens*, a specialized extreme thermophile, remains unclear. Its higher temperature optima and wide substrate specificity, which have yet to be fully explored, make Tt82 an attractive candidate for biotechnological applications.

Acknowledgements

Diffraction data have been collected on BL14.2 at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin. We would particularly like to acknowledge the help and support of Dr Manfred S. Weiss and Dr Christian Feiler during the diffraction experiment. We would also like to thank Dr Zdenek Franta for his help and inspirational conversations, and Rostislav Simunek for his contribution to the docking studies. We would specifically like to acknowledge Crissy L. Tarver for helping to make things possible.

Funding information

Funding for this research was provided by: Grantová Agentura České Republiky (grant No. 17-24321S); Jihočeská Univerzita v Českých Budějovicích (grant No. GAJU 04-149/2016/P); European Regional Development Fund (award No.

CZ.02.1.01/0.0/0.0/15_003/0000441); Deutscher Akademischer Austauschdienst (award No. DAAD-16-09); Rijksdienst voor Ondernemend Nederland (award Nos. RVO 68378050 and RVO 61388963); Ministry of Education of the Czech Republic (programme NPU I, project LO1304).

References

Allen, K. N. & Dunaway-Mariano, D. (2004). *Trends Biochem. Sci.* **29**, 495–503.

Allen, K. N. & Dunaway-Mariano, D. (2009). *Curr. Opin. Struct. Biol.* **19**, 658–665.

Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D. & Aravind, L. (2006). *J. Mol. Biol.* **361**, 1003–1034.

Caparrós-Martín, J. A., McCarthy-Suárez, I. & Culiáñez-Macià, F. A. (2013). *Planta*, **237**, 943–954.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Collet, J. F., Stroobant, V., Pirard, M., Delpierre, G. & Van Schaftingen, E. A. (1998). *J. Biol. Chem.* **273**, 14107–14112.

Copeland, R. A. (2000). *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*. New York: Wiley.

DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Fahs, S., Lujan, P. & Kohn, M. (2016). *ACS Chem. Biol.* **11**, 2944–2961.

Holm, L. & Sander, C. (1997). *Nucleic Acids Res.* **25**, 231–234.

Huang, H., Pandya, C., Liu, C., Al-Obaidi, N. F., Wang, M., Zheng, L., Keating, S. T., Aono, M., Love, J. D., Evans, B., Seidel, R. D., Hillerich, B. S., Garforth, S., Almo, S. C., Mariano, P. S., Dunaway-Mariano, D., Allen, K. N. & Farelli, J. D. (2015). *Proc. Natl Acad. Sci. USA*, **112**, E1974–E1983.

Huang, H., Patskovsky, Y., Toro, R., Farelli, J. D., Pandya, C., Almo, S. C., Allen, K. N. & Dunaway-Mariano, D. (2011). *Biochemistry*, **50**, 8937–8949.

Hughes, R. C. & Ng, J. D. (2007). *Cryst. Growth Des.* **7**, 2226–2238.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Karpouzas, D. G. & Singh, B. K. (2006). *Adv. Microb. Physiol.* **51**, 119–185.

Kertesz, M. A., Cook, A. M. & Leisinger, T. (1994). *FEMS Microbiol. Rev.* **15**, 195–215.

Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B. A., Wang, J., Yu, B., Zhang, J. & Bryant, S. H. (2016). *Nucleic Acids Res.* **44**, D1202–D1213.

Koonin, E. V. & Tatusov, R. L. (1994). *J. Mol. Biol.* **244**, 125–132.

Krachtus, D., Smith, J. C. & Imhof, P. (2018). *Molecules*, **23**, 3342.

Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.

Lahiri, S. D., Zhang, G., Dai, J., Dunaway-Mariano, D. & Allen, K. N. (2004). *Biochemistry*, **43**, 2812–2820.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947–2948.

Lorenz, U. (2011). *Curr. Protoc. Immunol.* **91**, 11.7.1–11.7.12.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Millán, C., Sammito, M. & Usón, I. (2015). *IUCrJ*, **2**, 95–105.

Mueller, U., Förster, R., Hellmig, M., Huschmann, F. U., Kastner, A., Malecki, P., Pühringer, S., Röwer, M., Sparta, K., Steffien, M., Ühlein, M., Wilk, P. & Weiss, M. S. (2015). *Eur. Phys. J. Plus*, **130**, 141–150.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Oja, T., Niiranen, L., Sandalova, T., Klika, K. D., Niemi, J., Mantsala, P., Schneider, G. & Metsa-Ketela, M. (2012). *Proc. Natl Acad. Sci. USA*, **110**, 1291–1296.

Pandya, C., Farelli, J. D., Dunaway-Mariano, D. & Allen, K. N. (2014). *J. Biol. Chem.* **289**, 30229–30236.

Parsons, J. F., Lim, K., Tempczyk, A., Krajewski, W., Eisenstein, E. & Herzberg, O. (2002). *Proteins*, **46**, 393–404.

Peisach, E., Selengut, J. D., Dunaway-Mariano, D. & Allen, K. N. (2004). *Biochemistry*, **43**, 12770–12779.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.

Pikuta, E. V., Marsic, D., Itoh, T., Bej, A. K., Tang, J., Whitman, W. B., Ng, J. D., Garriott, O. K. & Hoover, R. B. (2007). *Int. J. Syst. Evol. Microbiol.* **57**, 1612–1618.

Rao, S. T. & Rossmann, M. G. (1973). *J. Mol. Biol.* **76**, 241–256.

Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.

Sammito, M., Millán, C., Frieske, D., Rodríguez-Freire, E., Borges, R. J. & Usón, I. (2015). *Acta Cryst.* D**71**, 1921–1930.

Seifried, A., Schultz, J. & Gohla, A. (2013). *FEBS J.* **280**, 549–571.

Sparta, K. M., Krug, M., Heinemann, U., Mueller, U. & Weiss, M. S. (2016). *J. Appl. Cryst.* **49**, 1085–1092.

Trott, O. & Olson, A. J. (2010). *J. Comput. Chem.* **31**, 455–461.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.

Wang, W., Kim, R., Jancarik, J., Yokota, H. & Kim, S.-H. (2001). *Structure*, **9**, 65–71.

Webb, M. R. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 4884–4887.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* D**67**, 235–242.

Zhang, G., Morais, M. C., Dai, J., Zhang, W., Dunaway-Mariano, D. & Allen, K. N. (2004). *Biochemistry*, **43**, 4990–4997.