
Article:
Valente, Maria, Timm, Jennifer, Castillo-Acosta, Víctor M et al. (7 more authors) (2016) Cell cycle regulation and novel structural features of thymidine kinase, an essential enzyme in Trypanosoma brucei. Molecular Microbiology. pp. 365-385. ISSN: 0950-382X

https://doi.org/10.1111/mmi.13467

Cell cycle regulation and novel structural features of thymidine kinase, an essential enzyme in *Trypanosoma brucei*

Maria Valente[1§], Jennifer Timm[2§], Víctor M. Castillo-Acosta[1], Luis M. Ruiz-Pérez[1], Tom Balzarini[1], Joanne E. Nettleship[3], Louise E Bird[3], Heather Rada[3], Keith S. Wilson[2]* & Dolores González-Pacanowska[1]*

[1] Instituto de Parasitología y Biomedicina "López-Neyra", Consejo Superior de Investigaciones Científicas, Granada, Spain

[2] York Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5DD, UK

[3] The Oxford Protein Production Facility, Research Complex at Harwell, Rutherford Appleton Laboratory, R92 Harwell, Didcot, Oxfordshire OX11 0FA, UK

*Correspondence to: Dolores González-Pacanowska; dgonzalez@ipb.csic.es and Keith S. Wilson; keith.wilson@york.ac.uk

Running title: Cell cycle regulation and structure of *Tb*TK

Keywords: *Trypanosoma brucei*, thymidine kinase, cell cycle, protein structure, pyrimidines

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1111/mmi.13467

Abstract

Thymidine kinase (TK) is a key enzyme in the pyrimidine salvage pathway which catalyzes the transfer of the γ-phosphate of ATP to 2'-deoxythymidine (dThd) forming thymidine monophosphate (dTMP). Unlike other type II TKs, the *Trypanosoma brucei* enzyme (*Tb*TK) is a tandem protein with two TK homolog domains of which only the C-terminal one is active. In this study, we establish that *Tb*TK is essential for parasite viability and cell cycle progression, independently of extracellular pyrimidine concentrations. We show that expression of *Tb*TK is cell cycle regulated and that depletion of *Tb*TK leads to strongly diminished dTTP pools and DNA damage indicating intracellular dThd to be an essential intermediate metabolite for the synthesis of thymine-derived nucleotides. In addition, we report the X-ray structure of the catalytically active domain of *Tb*TK in complex with dThd and dTMP at resolutions up to 2.2 Å. In spite of the high conservation of the active site residues, the structures reveal a widened active site cavity near the nucleobase moiety compared to the human enzyme. Our findings strongly support *Tb*TK as a crucial enzyme in dTTP homeostasis and identify structural differences within the active site that could be exploited in the process of rational drug design.

Introduction

Trypanosoma brucei is a unicellular parasite of the class Kinetoplastida and the causative agent of Human African trypanosomiasis (HAT, (World Health Organisation, 2014)). In the last decade *Trypanosoma brucei* has been intensely studied in order to find new drugs and new drug targets that could improve parasitic chemotherapy and overcome the deficiencies of existing treatments. The identification of major differences in the biochemical pathways between the parasites and their human hosts would help reduce toxicity of novel drug candidates.

Deoxynucleotide metabolism provides drug targets for many diseases as a balanced and well-controlled pool of deoxyribonucleotides (dNTPs) is essential for DNA replication and repair. In most species, two metabolic routes contribute to the intracellular dNTP pool, the *de novo* and the salvage pathway (Reichard, 1988). While deoxythymidine triphosphate (dTTP) biosynthesis via deoxyuridine monophosphate (dUMP) formation has been proven to be necessary for parasite survival *in vitro* (Sienkiewicz *et al.*, 2008), very little is known about the relevance of dTTP formation via thymidine salvage. Thymidine kinases (TKs, EC 2.7.1.21) are key enzymes in the pyrimidine salvage pathway, catalyzing the magnesium-dependent transfer of the γ-phosphate of ATP to thymidine (dThd), forming thymidine monophosphate (dTMP).

$$ATP + dThd \overset{TK}{\leftrightarrow} ADP + dTMP$$

Subsequently, dTMP is phosphorylated further to dTTP, a substrate for DNA synthesis and therefore essential for replication (Al-Madhoun *et al.*, 2004; Birringer *et al.*, 2005; Birringer *et al.*, 2006). There are two types of TKs, I and II, with quite different 3D

folds and substrate specificities. Type I TKs include the herpes simplex virus TK (HSV-TK) and human mitochondrial TK2, and their broad substrate specificity is exploited in antiviral therapy. Type II TKs (TKIIs) are present in many organisms, for example protozoan parasites like *Trypanosoma brucei* and *Leishmania* spp., plants, mammals (cytosolic), bacteria and certain viruses (Ranjbarian *et al.*, 2012; Timm *et al.*, 2014; Lee & Cheng, 1976).

The crystal structures of many TKIIs have been solved, including those from human (*Hs*TK1, (Welin *et al.*, 2004; Birringer *et al.*, 2005)), *Ureaplasma urealyticum* (*Uu*TK, (Kosinska *et al.*, 2005)), *Bacillus anthracis* and *B. cereus* (*Ba*TK and *Bc*TK, (Kosinska *et al.*, 2007)), *Thermotoga maritima* (*Tm*TK, (Segura-Pena *et al.*, 2007a; Segura-Pena *et al.*, 2007b)), *Clostridium acetobutylicum* and *L. major* (*Lm*TK, (Timm *et al.*, 2014)). All these structures have the enzyme packed as a tetramer in the crystal with similar quaternary interactions: most forming a "closed" tetramer with no ligand in the phosphate donor site, and a smaller number with an "open" tetramer with the site occupied by a nucleobase. The majority of TKIIs are reported to be either dimers or tetramers in solution, or an equilibrium between both (Eriksson *et al.*, 2002; Mutahir *et al.*, 2013). Most are highly specific, phosphorylating only dThd and deoxyuridine (dUrd), and tolerate only minor changes in the 5-position of the pyrimidine base and/or in the 3'-position of the 2'-deoxyribose sugar (Johansson & Eriksson, 1996).

Unlike other TKIIs, *Tb*TK deviates from the usual homodimer/homotetramer form. Instead, it is a tandem protein with two homologous *Hs*TK1-like domains in a single chain both with the conserved TKII fold, assumed to have evolved as a result of gene duplication and subsequent fusion. The full length *Tb*TK, has been reported to be a pseudo-dimer, with its two domains corresponding to two adjacent subunits of a typical TKII (Ranjbarian *et al.*, 2012). The two domains have 67% sequence identity on the

amino acid level. Ranjbarian *et al* furthermore showed that only the C-terminal domain 2 is active, while in the N-terminal domain the catalytically important D286 is replaced by an asparagine and two residues proposed to be involved in substrate binding (D245) and transition state stabilization (R247) are also different. In addition, *Tb*TK was shown to have broader substrate specificity compared to *Hs*TK1, which provides the potential for identifying inhibitors that specifically target the parasitic enzyme.

The essential character of *Tb*TK has not been fully analyzed to date, yet the existence of the *de novo* biosynthesis would preclude the requirement of thymidine salvage for parasite survival in *T. brucei*. However, in the closely related species *L. major*, *Lm*TK was shown to be important for infectivity of the parasites (Thiel *et al.*, 2008). On the other hand, cellular production of dTTP is tightly modulated throughout the cell cycle (Reichard, 1988). In human cells, cell cycle-dependent regulation of enzymes involved in dTTP formation has been extensively described (Navalgund *et al.*, 1980; Bradshaw, 1983; Sherley & Kelly, 1988) and *Hs*TK1 has been shown to have an important role in this process. During mitosis, the enzyme is phosphorylated on Ser13 (Chang *et al.*, 1998), which interferes with active tetramerization and gives rise to a reduction of catalytic efficiency (Li *et al.*, 2004). *Hs*TK1 function is practically undetectable at the entry to the G1 phase due to anaphase promoting complex/cyclosome (APC/C)-dependent degradation at the end of mitosis (Ke & Chang, 2004).

Here we assess the importance of *Tb*TK for parasite growth and correct cell cycle progression of *T. brucei* bloodstream forms (BF) *in vitro* as well as its fundamental role in the maintenance of the dTTP pool. Our results showed that *Tb*TK depletion resulted in a decreased cell proliferation rate and cell cycle abnormalities together with a depletion of the dTTP pools revealing an essential role of thymidine salvage in the cell cycle regulation of pyrimidine nucleotide homeostasis. The enzyme exhibits a cell cycle

dependent regulation related to both the levels and relative intracellular localization. Additionally, we report the crystal structure of the catalytically active C-terminal domain of *Tb*TK solved in complex with dThd and dTMP to 2.2 Å resolution, highlighting features of the active site which might be exploited in the process of rational drug design.

Results

*Tb*TK is essential for parasite survival and salvage of intracellular thymidine

A correct balance of the deoxyribonucleotide pool is essential for DNA replication and consequent parasite duplication. Hence, we explored the role of TbTK in parasite growth by RNAi mediated depletion of the enzyme in BFs of *T. brucei*. TK knockdown impaired dramatically the proliferation at 48 hours, which correlated with a 90% reduction of *TbTK* mRNA levels compared to the parental line (Fig. 1A) and an almost complete elimination of the protein (Fig. 1B). In particular, a time course analysis showed a rapid decline in the growth of parasites in a TK-depleted culture, reaching a strong reduction at 16 hours post-induction (70%) and a maximum of 95% at 48 hours (Fig. 1C). A second RNAi construct containing a 437 bp fragment of the 3'UTR of the gene was obtained in order to corroborate the phenotype and rule out possible off-target effects. A second cell line transfected with this construct also showed a dramatic reduction in growth upon TK depletion (80% at 24 hours).

To analyze whether the growth defect could be affected by pyrimidine availability, parasites were grown for several days in an HMI-9 pyrimidine-free medium (Fig. 1D) and in HMI-9 medium supplemented with the nucleosides uridine (Urd), dUrd, deoxycytidine (dCyd) or dThd (Fig 1E). These nucleosides were tested at a high concentration (1 mM) in order to overcome their inefficient uptake by the U3 or T1 transporters (Ali *et al.*, 2013). While *T. brucei* cannot perform salvage of extracellular dCyd via phosphorylation, this nucleoside was also assessed since there is evidence for a dCyd deaminase activity (Ali *et al.*, 2013) which could give rise to dUrd from dCyd,

which can then be converted to uracil via uridine phosphorylase (UPase) or to dUMP via TK and therefore sustain synthesis of dTMP via synthesis *de novo*. However, the absence of an external source of pyrimidines had no additional impact on proliferation of TK-deficient parasites. On the other hand, none of the nucleosides tested were capable of reversing the proliferation defect caused by the depletion of TK. Next, growth was monitored in the presence of uracil, a major substrate for pyrimidine salvage which is efficiently transported in BFs (Ali *et al.*, 2013). Since uracil may be cytotoxic at high concentrations (Ong *et al.*, 2013), and is already present in the fetal bovine serum, the final uracil concentration in the medium was adjusted by using a pyrimidine-free HMI-9 medium with 10% dialyzed FBS supplemented with 100 μM or 1 mM of uracil. Neither low nor high uracil supplementation promoted growth of TK-depleted parasites (Ong *et al.*, 2013) (Fig. 1F). Thus, the data suggest that the essential role of TK is independent of the utilization of extracellular pyrimidines and appears to be exclusively related to the salvage of intracellular dThd most probably located in specific intracellular compartments.

TK-depleted cells exhibit reduced dTTP levels and accumulate genomic DNA breaks

In order to establish the role of TK in the homeostasis of dTTP, measurements of intracellular deoxyribonucleotides were performed at 16, 24 and 48 hours post-induction in *TbTK* RNAi cell lines. For this purpose, we used a DNA polymerase based assay previously described (Castillo-Acosta *et al.*, 2008; Requena *et al.*, 2014) with minor modifications. In agreement with previously reported data (Vodnala *et al.*, 2008), the amount of dTTP in BFs was 0.81 ± 0.06 pmol/10^6 cells. After TK depletion, a statistically significant decrease in dTTP levels was detected at 16 hours (0.57 ± 0.12

pmol/10^6 cells) (Fig. 2A), that was further enhanced at 24 and 48 hours post-induction (0.21 ± 0.03 pmol/10^6 cells and 0.20 ± 0.05 pmol/10^6 cells respectively) (Fig. 2A). Conversely, the amount of dGTP, that was equal to 0.21 ± 0.06 pmol/10^6 cells in the parental cell line, did not undergo significant changes upon TK depletion. Previous studies have reported dTTP as an allosteric effector of GDP reduction by *Trypanosoma* ribonucleotide reductase (Hofer *et al.*, 1998) which would lead to a concomitant decrease in dGTP upon dTTP depletion. This effect was not observed in the present study and an alternative methodology for determining nucleotide levels would be necessary to confirm this observation.

The depletion of cellular dTTP leads to the so-called thymineless death that has been described to occur in replicating cells (Cummings & Kusy, 1970), probably because of massive DNA damage sustained during replication in the absence of thymine nucleotides. Misincorporation of genomic uracil and formation of DNA double strand breaks, which involve an activation of the DNA damage response and repair (Khodursky *et al.*, 2015), are known consequences of thymineless death (Ahmad *et al.*, 1998). We therefore proceeded to first quantify the accumulation of DNA damage by using the TUNEL assay which specifically marks DNA strand breaks (Fig. 2 panels B and C) and then determined the presence of DNA repair foci upon TK depletion at16 and 24 hours post-induction (Fig. 3).

To investigate the presence of repair foci, both parental and *TbTK* RNAi cell lines were subjected to an immunofluorescence analysis using the anti-*Tb*γH2A antibody, which has been described as prominent repair marker (Glover & Horn, 2012). The percentage of immunofluorescent γH2A-positive cells was first determined by flow cytometry (Fig. 3A) which showed a significant increase in the γH2A-positive population at 16 and 24 hours after TK depletion (15.7% and 22.2% respectively versus 9.3% in the parental cell

line). These results were further confirmed by immunofluorescence microscopy. The percentage of positive cells was determined by counting γH2A nuclear foci of more than 300 cells per sample. An almost three-fold increase of γH2A-positive cells was found in *TbTK* RNAi cells at both 16 and 24 hours after induction (Fig. 3B) and whole nuclear staining was typically observed in these cells (Fig. 3C). Hence, TK depletion triggers the nuclear accumulation of γH2A in *T. brucei* BFs, which is a marker of DNA damage with a recognized role in the DNA damage signaling cascade (Canman, 2003, Lukas *et al.*, 2011). While an increase of γH2A-positive cells was observed at 24 hours after *TbTK* RNAi induction, no significant changes in TUNEL-positive cells number appeared at this time point (Fig. 2 panels B and C). However, a two-fold increase of the TUNEL-positive population was observed at 48 hours after TK knockdown compared to parental line, thus indicating that TK depletion induces accumulation of DNA strand breaks in BFs of *T. brucei*.

As previously mentioned, inhibition of *de novo* synthesis of dTTP triggers thymineless death phenomena (Ahmad *et al.*, 1998, Khodursky *et al.*, 2015). In *T. brucei* the *de novo* synthesis of dTMP is catalyzed by the bifunctional enzyme dihydrofolate reductase–thymidylate synthase (DHFR–TS) which is essential for parasite survival in the absence of dThd (Sienkiewicz *et al.*, 2008). Since thymidine salvage is able to counteract the lack of DHFR-TS *in vitro* (Sienkiewicz *et al.*, 2008), it was of interest to investigate how TK deficiency could affect the sensitivity to inhibitors of *de novo* thymidylate (dTMP) synthesis (Ahmad *et al.*, 1998; Van Triest *et al.*, 2000). Parental and *TbTK* RNAi cell lines were treated with methotrexate, a classical DHFR inhibitor, for 24 hours and the relative population doubling (RPD) was calculated (O'Donovan, 2012). As indicated in Figure 4, TK-depleted cells were hypersensitive to methotrexate compared to the parental line with a statistically significant RPD reduction ($p<0.001$,

ANOVA test) of more than 25% at concentrations above 25 μM. These observations further support that TK has an important role in the thymineless death induced by the interruption of *de novo* biosynthesis of dTTP in *T. brucei* BFs.

TK expression is cell-cycle-regulated and plays an important role in cell cycle progression in *T. brucei* bloodstream forms

In the majority of eukaryotes, including human cells and the Kinetoplastida *Leishmania donovani*, type II TKs are localized primarily in the cytosol. However *Hs*TK has been shown to accumulate in the nucleus under specific conditions such as in the case of response to genotoxic insult (Chen *et al.*, 2010b). In order to evaluate the localization of TK in *T. brucei*, an immunofluorescence analysis using an affinity-purified polyclonal anti-*Tb*TK antibody was performed. Additionally, an expression construct encoding a *Tb*TK-myc fusion protein was stably integrated in procyclic cells for visualization by immunofluorescence using a monoclonal antibody against the c-myc epitope. The results show that TK appears to be located predominantly within the nucleus but it was also detected in the cytosol (Fig. 5). Nuclear localization was determined by the colocalization between the FITC and DAPI signals with a Pearson's Coefficient of 0.6 ± 0.06. We next analyzed the expression levels and the cell cycle distribution of *Tb*TK. Different staining intensities were denoted and the relative distribution of the enzyme changed through the cell cycle. We therefore defined and quantified by microscopy BFs with nuclei that exhibit intense, moderate or low levels of TK and classified them according to the nuclei and kinetoplast content. More than 500 BFs stained with an anti-*Tb*TK antibody were analyzed under the same conditions of light exposure. The analysis showed that more than 90% of the cells with 1N1K* (S phase) and almost 70% of the 1N2K population (G2 phase) displayed intense fluorescence, while a moderate FITC

signal was observed in the majority of cells with 1N*2K (M phase) and in 40% of the 1N1K population (G1 phase) (Fig. 6A). Very low or no fluorescence was observed in cells undergoing cytokinesis (2N2K). Thus, intense staining was preferentially observed in cells with an elongated kinetoplast that occur prior to kinetoplast division and correspond to cells in the S phase of the cell cycle (1N1K*). *Tb*TK expression declined during the onset of mitosis (1N*2K) and appeared severely reduced in cells where mitosis has been completed (2N2K). Cytosol staining intensity was also higher in 1N1K* (S) and lower in 1N1K cells (G1), yet the relative amount of *Tb*TK in the nucleus versus the cytosol declined from 88% in 1N1K* cells to 55% in 1N1K cells (Fig. 6B). In addition, the cell cycle was analyzed by measuring the DNA content by flow cytometry. This method, which does not allow for such a clear distinction of the different phases, revealed a 2.5-fold increase in TK content in cells in S phase and G2/M relative to G1 (Fig. 6C). Thus, *Tb*TK levels and its relative distribution are regulated in a cell cycle-dependent manner, reaching a maximum during the S phase and declining thereafter.

A similar cell distribution and cell cycle regulation was obtained for a procyclic cell line overexpressing *Tb*TK-myc, a c-myc-tagged version of the enzyme (Fig. 6D). Since in this case the *Tb*TK-myc coding sequence was inserted in non-transcribed spacer region of the rRNA locus, we conclude that cell cycle regulation of *Tb*TK expression is independent of the flanking untranslated regions of the endogenous *TK* locus.

The modulation of *Tb*TK expression along the cell cycle suggests stage-specific functions for this enzyme. In order to explore a possible role of TK in the S phase, we determined the effect of RNAi-mediated knockdown of *Tb*TK on cell cycle progression. The depletion of *Tb*TK protein for 24 hours resulted in a significant reduction of the population in G1 phase (1N1K) (23% versus 63% in the parental line), a two-fold

increase in the populations of cells in early S (1N1K*) and a seven-fold increase in cells in G2 phase (1N*2K) (Fig. 7A). The time course analysis of the cell cycle progression of *TbTK* RNAi-induced cells performed by flow cytometry upon TK knockdown showed a gradual decrease of the percentage of cells in G1 phase, a concomitant increase in the populations in G2/M and the accumulation of cells with an abnormal DNA content (post-G2/M) (Fig. 7B). Indeed, *TbTK* silencing provoked a significant perturbation of cytokinesis, likely as a consequence of mitotic arrest (Tu & Wang, 2005), giving rise to an aberrant population with several nuclei and kinetoplasts (XNXK) (Fig. 7C). The data indicate that *Tb*TK is essential for the correct cell cycle progression and completion of mitosis with the knockdown of TK expression resulting in an accumulation of cells in the S and early G2/M phases. To investigate further if kinetoplast replication was specifically affected by *Tb*TK depletion, a flow cytometry analysis was performed using dihydroethidium (DHE) as a kDNA marker of living cells at 24 and 48 hours after protein knockdown. The fluorescence caused by DHE staining was significantly increased in *TbTK* RNAi-induced cells relative to parental or non-induced parasites (Fig. 7D) showing that an aberrant number of kinetoplasts is present in *Tb*TK depleted cells. No differences in kinetoplast shape or size was observed.

In summary, the results show that progression through the cell cycle is highly dependent on TK and it is likely that the maintenance of an adequate dTTP pool via thymidine salvage is required for replication and completion of mitosis.

TK-deficient parasites exhibit reduced infectivity in mice

In order to gain insight into the role of *Tb*TK in parasite survival within the mammalian host, C57BL/6 mice were infected with the *TbTK* RNAi and parental strains and down-regulation *in vivo* was induced by doxycycline administration. While the parental and

TbTK RNAi cell lines without doxycycline gave rise to high levels of parasitaemia (10^9 parasites ml^{-1}) and consequently produced mice death within 6 days post-infection, mice infected with *TbTK* RNAi and given doxycycline showed no parasites during the period monitored and all survived the infection (Fig. 8) thus indicating that *Tb*TK is also important for parasite viability *in vivo*.

Overexpression, purification and crystallization of *Tb*TK

The observation that *Tb*TK is essential for parasite survival together with the profound differences in the properties of the *T. brucei* enzyme with regard to the human counterpart suggest the possibility of considering *Tb*TK as an exploitable therapeutic target. In order to obtain detailed structural information exploitable in inhibitor design, crystallization experiments were performed. Expression of a variety of constructs, full-length and truncations, in *E. coli* did not yield any soluble and/or stable protein and *Tb*TK was therefore expressed using the baculovirus system. None of the constructs of the full-length enzyme or the inactive TK-homology domain gave sufficiently soluble and stable protein for structural studies. Purification using nickel affinity and size exclusion chromatography was successful for several constructs of the catalytic domain, but only one formed crystals. Addition of dTMP, AppNHp and $MgCl_2$ to the protein during purification was necessary for stabilization. With an overall polydispersity of 19.9 %, as determined by DLS, the protein was considered suitably monodisperse for crystallization screening.

*Tb*TK preferentially formed crystals in space group $P4_322$, in which the ATP-binding region is partially blocked by an adjacent *Tb*TK molecule. The *Tb*TK-dThd/dTMP crystal shows clear electron density for the dThd from the dTMP with which it was co-crystallized, as well as additional density most probably resulting from a partially

occupied PO_4^{2-} from the dTMP. The *Tb*TK-dTMP-PO_4^{2-}-a crystal shows a well-resolved dTMP molecule in the active site. Additionally, there is well-defined electron density for a phosphate where the β-phosphate of the AppNHp was expected to bind. The *Tb*TK-dThd-glycerol crystal was soaked in cryo-protectant containing dThd and AppNHp with the aim of replacing the co-crystallized dTMP. This structure shows clear electron density for dThd, but none for AppNHp, and has a glycerol molecule bound from the cryo-protectant. In order to crystallize *Tb*TK in a different crystal form which might allow the AppNHp to bind, cross seeding experiments were carried out using seed stocks prepared from *Lm*TK-dThd crystals in space group $P6_522$ (Timm *et al.*, 2014). These led to the *Tb*TK-dTMP-PO_4^{2-}-b crystal in space group C222. However, as in the previous crystals, *Tb*TK formed closed tetramers and there was no electron density for the AppNHp moiety, but instead density for bound dTMP and phosphate in the β-position as seen in *Tb*TK-dTMP-PO_4^{2-}-a.

Structures of *Tb*TK

Overall fold and assembly

Only the catalytically active of the two TK homology domains expressed in soluble form, so that the solved structure represents less than half of the full-length *Tb*TK. The structure shows close similarity to other TKIIs with an overall RMSD of 0.31 Å to *Hs*TK1 (PDB ID: 1w4r, (Welin *et al.*, 2004)). Each protomer has the two-domain fold typical for TKIIs with the active site sitting between the domains (Fig. 9A). The larger N-terminal α/β-domain consists of a central six-stranded parallel β-sheet (β1, β2, β3, β4, β5, β8), flanked by α-helices (α1, α2 and α3) and loop regions. The P-loop,

essential for the kinase activity, is located between β1 and α1. There is a flexible loop region between residues 245 and 260 in *Tb*TK, which is similarly disordered in most published TK structures. Like in other TKIIs, the smaller C-terminal domain contains three β-strands (β9-11) and a short α-helix (α4) at the very C-terminus of the structure. It comprises a structural Zn^{2+} ion tetrahedrally coordinated by four cysteine residues C341, C344, C374 and C377. A lasso structure between β10 and β11 covers the thymidine-binding site.

The catalytic domain of *Tb*TK forms tetramers like other TKIIs (Fig. 9B). The crystals in space group $P4_322$ contain one dimer in the asymmetric unit (AU), with the tetramers generated by symmetry. The *Tb*TK-dTMP-PO_4^{2-}-b crystal in space group C222 has one tetramer and a dimer in the AU. Analysis using PISA (Krissinel & Henrick, 2007), indicated that the *Tb*TK construct was likely to be a dimer or a tetramer in solution, depending on the inclusion or exclusion of phosphates in the P-loop from the analysis. Inclusion of phosphate led to prediction of a tetramer, exclusion of all ligands suggested a dimer. There are two protomer-protomer interfaces in the tetramer, one defined by the central β-sheets of two protomers forming a single continuous β-sheet, the other by the interaction of the α1-helices of two adjacent subunits. The first interaction results in the formation of a stable dimer, the latter in a tetramer.

Two different types of tetramer are seen in TK crystal structures in the PDB depending on the presence or absence of phosphate donor. In the absence of ligand in the phosphate donor site, TKIIs form a closed tetramer with the α1-helices interacting directly with each other. In the presence of a phosphate donor the α1-helices are ~ 3 Å apart forming more open tetramers leaving space for the nucleobase to bind. In both tetramers, the dimer interfaces remain identical. In our crystals, *Tb*TK forms closed

tetramers, not showing electron density for the AppNHp with which they were co-crystallized.

Substrate and Inhibitor binding

Most active site residues forming the dThd-binding site have very similar conformations to those in other TKIIs in the PDB (Fig. 10A). A detailed description of the active site was given previously for *Hs*TK1 (Welin *et al.*, 2004) and *Ba*TK and *Bc*TK (Kosinska *et al.*, 2007). In brief for *Tb*TK, the dThd molecule is coordinated by hydrogen bonds to the main chain of residues, R361, L363 and G365, located in the lasso domain, and Y316 from the α/β-domain. Stacking interactions with Y370 from the lasso structure and hydrophobic interactions of F289, L312 and F321 on the α/β-domain hold the thymidine ring in place. Residue E286, essential for catalysis, forms a hydrogen bond to the 5'-oxygen of the ribose ring of the dThd. It is positioned to serve as a catalytic base, abstracting a proton from the oxygen, which is then enabled to perform a nucleophilic attack on the γ-phosphate of the phosphate donor.

A unique feature in the active site of *Tb*TK is the position of residue F321, which is pulled outwards due to the bending of the α3-helix (Fig. 10B), thus widening the dThd binding cavity. This might explain the lower substrate specificity of *Tb*TK compared to other TKIIs.

None of the *Tb*TK structures contains the phosphate donor mimic AppNHp, in keeping with the formation of the closed tetramer. However, three of the structures show a phosphate coordinated by the P-loop in the position expected for the β-phosphate of the donor, as seen in the structure of *Tm*TK-dThd-AppNHp (PDB: 2qq0) (Fig. 10C). The expected ATP-binding site, determined by homology to the structures of *Ca*TK and

*Tm*TK with occupied phosphate donor sites, is blocked in our structures by residues of the α1-helix of an adjacent tetramer.

The TbTK pseudo-dimer

*Tb*TK is a tandem enzyme, containing two fused TK-homology domains. When expressed separately in *E. coli* the catalytically active domain does not dimerize efficiently in solution perhaps due to the fact that lower protein concentrations were used (Ranjbarian *et al.*, 2012). The full-length enzyme in solution dimerizes to some extent to form a pseudo-tetramer with the main species remaining the monomer (pseudo-dimer). Within the pseudo-dimer an 18 amino acid residue linker region (aa 181-197) bridges the gap between the C-terminus of the N-terminal inactive domain and the N-terminus of the catalytic domain. The high similarity of the two domains led us to the hypothesis that the pseudo-dimer assembly involves comparable protomer-protomer interfaces to those found in the typical TKII tetramers. On the basis of this proposal, the sequence of the N-terminal TK-homology domain was modelled onto the structure of one of the two protomers making up a catalytically active dimer (Fig. 11). As there are two different protomer-protomer interfaces in a TK tetramer, two alternative full-length *Tb*TK pseudo-dimers were constructed. In both of them (Fig. 11 A and B), the termini of the two protomers are close enough to be connected by the 18 amino acid linker (shown as a dashed gray line in the Figure). Highlighting the sequence identity on the structure reveals the conservation of the central β-sheet and the α1-helix, while most amino acid replacements are located on the outward facing parts of helices α2 and α3 and the region between the β-strands β5 and β6. Analysis with PISA suggests the pseudo-dimer in Figure 11 A to be the more probable biological entity. We propose that

the full-length *Tb*TK forms a pseudo-dimer similar to that shown in Figure 11 A, which then assembles into a pseudo-tetramer.

Comparison of the active site residues of structures of catalytically active TKs with the modelled inactive TK-homology domain highlights the high sequence identity with all residues in direct interaction with the substrate being fully conserved except the amino acid substitution D286N, which renders the protein inactive. The two other replaced residues (D247N and R245, deletion mutant,) involved in substrate binding and transition state stabilization, are located in the disordered loop between residues 245 and 260.

Discussion

In this study, we provide evidence that TK plays an important role in cell survival in *Trypanosoma brucei*. Our experiments of media supplementation with different nucleobases and nucleosides further show that this function is independent of extracellular pyrimidines and therefore phosphorylation of intracellular-derived dThd emerges as an essential requirement for efficient DNA replication. This observation establishes a unique essential character of pyrimidine salvage which would be linked to its role in the maintenance of the dTTP pools and cell cycle control of DNA replication in *T. brucei*.

The essential character of TK has been explored in different cell types. In human cells the expression of TK1 is dispensable for proliferation and, neither the steady-state level of dTTP nor cell proliferation under normal growth conditions was affected by TK1 depletion (Chen *et al.*, 2010b). Rather TK is required for repair of DNA damage (Chen *et al.*, 2010b; al-Nabulsi *et al.*, 1994; Wakazono *et al.*, 1996). Indeed, checkpoint activation after DNA damage allows tumor cells to integrate *de novo* synthesis of dNTP and salvage synthesis of dTTP to achieve an appropriate level of nucleotides for efficient DNA repair and survival. Consequently it has been suggested that inhibition of targets such as TK1 would be non-toxic to normal cycling cells but lead to chemosensitization and elimination of tumor cells (Chen *et al.*, 2010b). While in cell lines, TK1 is non-essential for survival (Chen *et al.*, 2010b), homozygous knockout (TK1−/−) mice display sclerosis of kidney glomeruli and spleen abnormalities and the animals die before 1 year of age showing that the enzyme is required for development (Dobrovolsky *et al.*, 2003).

Here we show that down-regulation of *Tb*TK results in growth arrest, depletion of dTTP pools and early DNA damage. It has been fully demonstrated that complete synthesis of pyrimidines through the *de novo* pathway occurs in this parasite and that this process is essential in the absence of extracellular pyrimidines (Ali *et al.*, 2013; Ong *et al.*, 2013). Likewise, DHFR-TS null mutants are dThd auxotrophs (Sienkiewicz *et al.*, 2008). On the other hand, metabolic labeling studies in *T. brucei* have indicated that dThd salvage mediated by the action of TK is a preferential pathway for dTMP synthesis when dThd is available in the medium (Creek *et al.*, 2015). The phenotype obtained here after *Tb*TK depletion, despite the existence of the *de novo* pathway, reveals that dTTP levels are highly dependent on TK function and that both the *de novo* and salvage pathways are essential for dTMP formation in BF trypanosomes in the absence of extracellular dThd. An explanation for the strong dependence of dTTP pools on dThd involves the presence of a preexisting pool which, in the absence of extracellular dThd, would be derived from synthesis *de novo* and in the presence of the nucleoside from both synthesis and salvage. In support of these observations the labeling and metabolomic studies performed by Creek et al. have shown the existence of a pool of dThd in *T. brucei* BFs which is partially derived from synthesis *de novo* (Creek et al., 2015). The origin of a pool of dThd is unclear but would implicate the action of a nucleotidase acting on thymidine nucleotides. While canonical 5' nucleotidases acting on pyrimidine deoxynucleoside mono- and diphosphates in *T. brucei* remain to be characterized, two orthologues of a putative dNTP triphosphohydrolase have been identified which are homologues of human sterile alpha motif (SAM) and histidine/aspartate (HD)-domain containing protein 1 (SAMHD1). These enzymes could provide the nucleoside substrates for dTMP formation via TK.

Curiously, the labeling studies have also shown the presence of a pool of *de novo* formed deoxyuridine (dUrd). Canonical dCMP/dCTP deaminases are apparently absent from the *Trypanosoma* genome and thus dUMP for dTMP synthesis would be generated via deoxyuridine 5'-triphosphate nucleotidohydrolase from dUTP or dUDP (Hidalgo-Zarco *et al.*, 2001, Hemsworth *et al.*, 2013). However an additional role of TK may be the phosphorylation of dUrd (generated through dCyd deamination) which is also a substrate of the enzyme (Ranjbarian *et al.*, 2012) providing this way dUMP for DHFR-TS, thus contributing to the dTTP pool although this latter possibility requires further investigation.

In summary, we propose that TK utilizes both extracellular and *de novo*-derived dThd to synthesize dTMP. Due to the strict requirement of TK for dTTP formation, knock-down of *Tb*TK results in the activation of DNA damage signaling, an accumulation of DNA lesions and, cell cycle arrest demonstrating that dTMP formation via TK is absolutely required for DNA replication/repair and cannot be compensated by synthesis *de novo* per se. The presence and origin of different pools of dThd, their relative distribution between the cytosol and the nucleus, and the existence of a flux of dThd between different cellular compartments all remain to be established.

In human cells, cytosolic dThd is salvaged to dTMP by TK1 and then phosphorylated by thymidylate kinase (TMPK) and nucleoside-diphosphate kinase (NDPK) to dTTP. The expression of the first two enzymes increases in the G1/S transition and diminishes in the mitotic phase via proteolytic destruction, while an increase of the relative mRNA levels in S-phase has been shown for NDPK (Caligo *et al.*, 1995). In *Trypanosoma* a similar TK expression pattern was observed: high in S-phase, decreasing in cells close to completion of mitosis and very low prior to cytokinesis and G1. *Hs*TK1 is phosphorylated on Ser13 in the mitotic phase (Chang *et al.*, 1998), which perturbs its

active tetramerization status and diminishes its catalytic efficiency (Li *et al.*, 2004). On the basis of our structural data we propose that the full-length *Tb*TK forms a pseudo-dimer, which then assembles into a pseudo-tetramer. Whether this process is regulated by phosphorylation, as in the case of the human enzyme, is yet unknown.

Our observations during the cell cycle showing that *Tb*TK appears to be differentially expressed and present in the nucleus may provide clues for explaining the essential character of thymidine salvage. We propose that dThd is a major intermediary metabolite required for the synthesis of thymine-derived nucleotides during DNA replication and repair. Enzymes involved in the subsequent steps of dTTP formation such as nucleoside diphosphate kinase have also been described to be located in the nucleus in *T. brucei* (Hunger-Glaser *et al.*, 2000), thus supporting the notion that dTTP is synthesized from dThd *in situ* within the nucleus and has a major contribution to the pool of dTTP.

Unfortunately, only heavily truncated versions of *Tb*TK are expressed in soluble and stable form, so structural work was limited to the core protein closely similar to those of other species (*Hs*TK1, (Welin *et al.*, 2004; Birringer *et al.*, 2005); *Uu*TK, (Kosinska *et al.*, 2005); *Ba*TK and *Bc*TK,(Kosinska *et al.*, 2007) ; *Tm*TK, (Segura-Pena *et al.*, 2007a; Segura-Pena *et al.*, 2007b); *Ca*TK. All these *Tb*TK structures form closed tetramers with the blocked phosphate donor site.

The active site residues whose side chains coordinate the dThd are highly conserved, with one important difference: the dThd-binding site in *Tb*TK has a widened cavity beside the C5 position of the dThd ring, where the main chain is bent outwards pulling F321 ~1.1 Å further away from the dThd than seen in *Hs*TK1. While the disordered loop region *Tb*TK (aa 245-260) lies close to the 3'-oxygen of the ribose ring and seems to leave more room for the substrate compared to *Hs*TK1, this is presumed to be an

artefact of the crystal packing. Sequence alignment indicated the relevant residues (ordered in *Hs*TK1 and disordered in *Tb*TK) to be conserved in all TKs except the inactive TK-homolog domain of *Tb*TK. The structural differences in the active site support the wider substrate specificity of *Tb*TK compared to *Hs*TK1 demonstrated previously (Ranjbarian *et al.*, 2012) which showed *Tb*TK to readily phosphorylate deoxyinosine and deoxyguanosine, while the activity of *Hs*TK1 against these substrates was poorer. This leads to the possibility of developing nucleoside analogues and other ligands that selectively bind to *Tb*TK either as cytotoxic substrate analogues or inhibitors.

In summary the information here provided establishes that TK has a critical role in the maintenance of a correct dTTP pool, which is required for DNA replication and cell cycle progression. The essential character of the unique thymidine kinase from *T. brucei* together with the structural information on the catalytically active domain may contribute to the exploitation of this enzyme as a novel drug target.

Experimental procedures

Generation of *TbTK* RNAi and *Tb*TK-myc lines

In order to down-regulate the expression of *Tb*TK by RNA interference in BFs of *T. brucei,* two fragments of 505 bp (position 810-1315 of the coding region) and 437 bp (position 298-735 of the 3' UTR) of *TbTK* were amplified by PCR. The 5' primers 5'-GCGGATCCAAGCTTGACGTGGGTGATGAGTGGCG-3' and 5'-GCGGATCCAAGCTTAATTTTCGAAACGCGGTATG-3', for the amplification of the first and second fragment respectively, were designed containing *Bam*H I and *Hind* III restriction sites (underlined), while the 3' primers 5'-GCGTTAACGGGCCCTAGCTTCTGTGTTCGGTTCA-3' and 5'-GCGTTAACGGGCCCACCTTGACAAGCGTTTCACC-3' contained *Hpa* I and *Apa* I restriction sites (underlined). All the primers used were designed from the sequence found in the GeneDB database (Tb927.10.880). The fragments were first cloned into the *Hind* III and *Apa* I sites of plasmid pGR19 (Clayton *et al.*, 2005), then digested with *Bam*H I and *Hpa* I and cloned in the *Bam*H I and *Hpa* I restriction sites of the previous construct in an antisense orientation, yielding the pGRMV8 and pGRMV23 RNAi constructs. In addition, a construct for overexpressing TK fused to a c-myc tag was generated by PCR using the following primers: 5'-GGCATATGCACGACGGAGATGGCAAT-3' (*Nde*I restriction site underlined) and 5'-CGGTTAACAGTAGTATCAACGGCCATTTG-3' (*Hpa* I restriction site underlined). The *Tb*TK-myc construct obtained by PCR was cloned into the pGRV33 plasmid (Castillo-Acosta *et al.*, 2012) yielding pGRMV10.

Trypanosome growth and transfection

T. brucei single-marker bloodstream form (BF) parasites (Wirtz *et al.*, 1999) were used to generate the *TK*RNAi cell lines. The parasites were cultured in HMI-9 supplemented with 10% (v/v) fetal bovine serum (FBS) at 37ºC and 5% CO_2 or in the case of pyrimidine-free experiments in HMI-9 without dThd and supplemented with 10% (v/v) dialyzed fetal bovine serum (FBS). The constructs pGRMV8, pGRMV10 and pGRMV23 were used for transfection. Transfected cell lines were obtained following protocols previously described (Wirtz *et al.*, 1998; Wirtz *et al.*, 1999). Briefly, 2.4×10^7 BF parasites maintained in logarithmic phase were harvested, washed with Modified Cytomix Buffer, optimized from the original Cytomix buffer (van den Hoff *et al.*, 1992), mixed with 10 μg of *Not* I -linearized plasmidic DNA in a total volume of 400 μL. Transfection was performed by electroporation using a BTX ECM 630 electroporator. Parasites were recovered at 37ºC and 5% CO_2 for around 18 h before adding the corresponding drug used as selectable marker (hygromycin in the case of pGRMV8 and pGRMV23 and puromycin in the case of pGRMV10)

Measurement of intracellular nucleotides

The nucleotides dGTP and dTTP were quantified using a modified DNA polymerase assay. The oligonucleotides 5′-TTTATTTATTTATTTATTTAGGCGGTGGAGGCGG-3′ and 5′-TTTCTTTCTTTCTTTCTTTCGGCGGTGGAGGCGG-3′ were employed as template for dTTP and dGTP measurements respectively, and the oligonucleotide 5′-CCGCCTCCACCGCC-3′ as primer. Parasites (2×10^6 parasites per sample) were harvested by centrifugation and washed once in 1× PBS. The cells were extracted with 250 μL of cold 60% methanol overnight at −20 °C. The suspension was heated for 5 min in boiling water bath, followed by centrifugation for 20 min at 17,000 × *g*. The

supernatant was transferred to a fresh tube and dried under vacuum. The residues were dissolved in 40 µL either of dUTPase buffer (34 mM Tris–HCl, pH 7.8, 5 mM $MgCl_2$) or dUTPase buffer plus 30 ng of *human* dUTPase and incubated for 20 min at 37 °C. To stop the reaction 60 µL of 100% methanol were added and the samples were incubated for 1 hour at −20 °C, followed by centrifugation for 20 min at 17,000 × *g*. The supernatant was again dried under vacuum and the residues were then dissolved in 100 µL of reaction mixture containing 32 nM DNA template, 32 nM DNA primer, 1× NEBuffer 2 (New England BioLabs®), 0.3 units of DNA polymerase I Klenow fragment (New England Biolabs®) and 0.0032 µCi μL^{-1} [3H]dATP (PerkinElmer®). A standard curve was obtained by using different amounts of the corresponding dNTP instead of dry extract. The polymerase reaction was then performed for 15 min at 25 °C and stopped by addition of 10 mM EDTA and heating at 75 °C for 20 min. Samples were incubated with 10% (v/v) TCA (trichloroacetic acid) for 30 min at 4 °C to precipitate the DNA. The solution was blotted on Glass Microfibre Filters GF/C (Whatman) and each filter was washed under vacuum with 30 ml of 5% TCA. Filters were then rinsed with 3 ml of ethanol, dried and the radioactivity was measured in a LS 6500 Multi-Purpose Scintillation counter (Beckman Coulter). Standard curves are indicated in Fig. S1.

TUNEL assay

Parasites (1.5×10^7 cells) were subjected to terminal deoxynucleotidyltransferase enzyme-mediated dUTP end labelling (TUNEL) in order to study DNA fragmentation mediated by TK depletion. The cells were first harvested by centrifugation at 7000 × *g* at 4 ºC for 2 minutes, then washed twice with cold 1X PBS and fixed by incubation in 2% p-formaldehyde for 45 minutes. After fixation the cells were washed twice with cold

1× PBS and permeabilized with 0.1% (v/v) Triton X-100 in 0.1% (w/v) sodium citrate for 2 minutes on ice.

Fluorescence activated cell sorting (FACS) and morphological phenotypic analysis

Cells (10^7 parasites per sample) were harvested by centrifugation fixed with ethanol and subjected to FACS analysis with propidium iodide staining as previously described (Castillo-Acosta *et al.*, 2008). For morphological analysis, parasites (10^6 cells per sample) were harvested and after paraformaldehyde fixation mounted with Vectashield-DAPI (Vector Laboratories, Inc.). The microscopy and digital image acquisition was carried out with a Zeiss Axiophot microscope (Carl Zeiss, Inc.). The subsequent analysis was performed by quantification of the number of nuclei and kinetoplasts in individual cells from two independent experiments in which more than 500 cells were analyzed. Dihydroethidium staining and FACS analysis of live cells was used to analyze the kDNA content.

Sensitivity of *T. brucei TbTK* RNAi cell lines to methotrexate

Cell proliferation of parental and TK-depleted line was measured at 24 hours after exposure to methotrexate (MTX) (Sigma). Cells were cultured in 96-well microplates starting from an initial concentration of 10^3 cells ml^{-1} per well, and exposed to different MTX concentrations (0–50 μM). After 24 hours the number of viable cells was measured by colorimetric Resazurin redox assay. Briefly, cells were incubated with 0.01 mg mL^{-1} Resazurin (Sigma) for 1 hour in the dark. After incubation, the fluorescence was measured by using a SpectraMax® GEMINI EM microplate reader (Molecular

Devices, Sunnyvale, CA, USA). The relative population doubling (RPD) was then calculated as in the formula (O'Donovan, 2012):

$$RPD = \frac{PD_{treated}}{PD_{untreated}} \times 100\%$$

Where $$PD = \{\log\ (N_{24h} \div N_0)\} \div log\ 2$$

Antibody generation

Denatured and purified recombinant *Tb*TK was used to immunize a rabbit to obtain a polyclonal anti-TK antibody. Around 450 µg of pure protein, resuspended in PBS and mixed with Freund's adjuvant at 1:1 ratio, were used in each injection. The anti-*Tb*TK serum was collected after four injections. The antibody was then affinity-purified using the recombinant protein coupled to Affi-Gel® 15 Gel (Bio-Rad) resin, according to the manufacturer's instructions. The specificity of the anti-TbTK antibody is shown in Figure S2.

3D-Immunofluorescence studies

Immunofluorescence was performed as previously described (Landeira *et al.*, 2009). Briefly, after 15 minutes incubation with MitoTraker Red, parasites were fixed in a suspension of TDB-glucose with 2% PFA for 1 hour and then permeabilized with 1% NP40 for 30 min. The IF was performed by using 0.5% blocking reagent (Roche) in PBS (Sigma) using either polyclonal anti-*Tb*TK antibody (1:100) and FITC-conjugated anti-rabbit secondary antibody (Sigma, 1:500) , monoclonal anti-c-myc (Sigma, 1:100) and Alexa Fluor® 488 goat anti-mouse secondary antibody (Sigma, 1:500) or the polyclonal anti-*Tb*γH2A antibody (1:100, gift from David Horn) (Glover & Horn, 2012)

and Alexa Fluor® 488 goat anti-rabbit secondary antibody (Sigma, 1:100). The cells were then placed on poly-l-lysine coated slides and stained and mounted with Vectashield-DAPI (Vector Laboratories, Inc.). Vertical stacks of 30-40 slices (0.2 μm steps) were captured using either an Olympus wide-field microscope and Cell R IX81software or confocal Leica SP5 microscope. Images were deconvolved and pseudo-colored by using Huygens Essential software (version 3.3; Scientific Volume Imaging) and Fiji software (version 1.5e; ImageJ) (Schindelin *et al.*, 2012), respectively. Colocalization analysis was performed by using JACoP plugin (Bolte & Cordelieres, 2006). Aliquots of the same samples were subjected to flow cytometry analysis with BD FACSAria™ III Cell Sorter and results interpreted by Flow Jo software (version 7.6). γH2A foci counting was performed with Fiji software following the Duke University Light Microscopy Core Facility protocol (microscopy.duke.edu).

In vivo studies

Four female C57BL/6J mice per group (8 weeks old) were infected via intraperitoneal injection with 1000 monomorphic *T. brucei* parasites of either parental (BF) or *TbTK* RNAi cell lines. Doxycycline (1mg ml-1) was given in the water of the group infected with the *TbTK* RNAi cell line to induce *Tb*TK knockdown. The animal research described in this manuscript complied with Spanish (Ley 32/2007) and European Union Legislation (2010/63/UE). The protocols used were approved by the Animal Care Committee of the Instituto de Parasitología y Biomedicina 'López-Neyra', CSIC (protocol CEEA/2014/DGP/1).

Construction of baculoviruses and protein expression in Sf9 cells

*Tb*TK is a tandem enzyme with two TK-homology domains, and its gene (GeneDB: Tb427.10.880) was cloned from genomic DNA of *T. brucei* Lister427. All constructs tested contained either N-terminal or C-terminal His-tags with or without a HRV 3C protease recognition site for removal of the tag. Beside the full-length *Tb*TK, we tested a range of truncations, designed according to homology to other TKs, removing disordered regions predicted using RONN (Yang *et al.*, 2005), as well as constructs separating the two TK-homology domains. Expression tests of a variety of constructs in *E. coli* failed to yield sufficient quantities of stable protein for crystallization, so we changed to the baculovirus expression system. Baculovirus generation and protein expression in Sf9 cells were carried out as described before (Timm *et al.*, 2014), following standard procedures of the Oxford Protein Production Facility (OPPF, (Berrow *et al.*, 2007)).

Purification and Characterization

Cells of 1 L Sf9 culture expressing the construct of the catalytic TK-homology domain of *Tb*TK fused to a cleavable N-terminal Histag were thawed and resuspended in 35 mL lysis buffer (20 mM Tris pH 8.0, 300 mM NaCl, 30 mM imidazole, 1 mM DTT and 100 µM PMSF) and lysed by sonication. To degrade nucleic acids, 25-29 U benzonase was added and incubated for 10 min at room temperature. Cell debris was removed by centrifugation for 30 min at 5000 rcf and 4°C and the supernatant was loaded onto an equilibrated 1 mL HisTrap crude FF column for 1 h in a loop on ice. After washing with 10 mL lysis buffer the column was connected to an ÄKTA purifier (GE Healthcare), washed with another 10 mL of lysis buffer and eluted with an imidazole gradient from 30 mM to 500 mM. The protein containing fractions were pooled and concentrated up

to 4 mL total volume. To prevent precipitation and degradation, the protein was diluted in 1 mL SEC buffer (10 mM Tris pH 8.0, 200 mM NaCl, 1 mM DTT) containing 1 mM dTMP, 1mM AppNHp and 3 mM $MgCl_2$. Due to a mistake, dTMP was added instead of dThd. The stabilized protein was loaded onto an equilibrated Superdex 200 column (GE Healthcare) in SEC buffer. The purity was assessed by SDS-PAGE and the protein-containing fractions were pooled and concentrated to 35 mg mL^{-1} in presence of 1 mM dTMP, 1mM AppNHp and 3 mM $MgCl_2$. Aliquots were flash frozen in liquid nitrogen and stored at -80°C. From 1 L Sf9 cell culture ~ 6 mg of pure protein were obtained.

Electrospray ionization mass spectrometry (ESI-MS) and tandem mass spectrometry (MS/MS, using a Bruker autoflex III MALDI-TOF/TOF) after trypsin digest, were used to confirm the identity of the purified protein. The polydispersity was determined by Dynamic light scattering (DLS) at 20°C. For this the protein was diluted in SEC buffer to a concentration of 3 mg mL^{-1} with a total sample volume of 20 µL.

Crystallization

Initial crystallization screens (Index, PACT and JCSG) were set up in sitting-drop vapor diffusion format with 150 nL protein solution including additives plus an equal volume of precipitant, using a Mosquito robot (TTP Labtech, UK). Crystal optimization was carried out in hanging-drop vapor diffusion with 1 µL protein solution plus an equal volume of precipitant. The crystals obtained in the original crystal form did not contain the phosphate donor AppNHp and the formation of a closed dimer in the crystals blocked the binding site. With the aim of obtaining a different crystal form with an open tetramer, cross seeding experiments were carried out using an Oryx 4 robot (Douglas Instruments, UK), using the Index screen in sitting-drop vapor diffusion with 250 nL protein solution plus 170 nL precipitant and 80 nL seed stock. The seed stock used was

made up using the seed bead kit (Hampton Research) from *Lm*TK crystals (Timm *et al.*, 2014) in 1.4 M Na/K PO_4 pH 5.4 with 1 mM dThd, 1 mM AppNHp and 3 mM $MgCl_2$ added to the protein. 1 mM dThd was added to replace the dTMP incorrectly used for stabilization during the preparation. The conditions leading to the best crystals are listed in Table 1. The *Tb*TK crystals grown without seeding were cryo-protected with 10 % glycerol added to the precipitant solution prior to vitrification. The crystal obtained after seeding was not cryo-protected.

Data collection and Structure solution

Diffraction data were collected at the Diamond Light Source (DLS) and the crystallographic data and statistics are summarized in Table 2. All computations were carried our using the CCP4 software package (Winn *et al.*, 2011). The diffraction images were processed with MOSFLM (Battye *et al.*, 2011) and the data reduced and merged with AIMLESS (Evans & Murshudov, 2013).

The structures were determined by molecular replacement using PHASER (McCoy, 2007). At the time the *Hs*TK1 (PDB code: 1w4r) showed the highest sequence identity of 49% and was selected as search model. The structure of *Tb*TK-dThd/dTMP was used as search model for subsequent structures. Structure refinement was performed using REFMAC5 (Murshudov *et al.*, 2011) with one TLS group per chain and automatically generated local NCS. The last rounds of refinement were carried out using the jelly body refinement option (with sigma 0.02). The manual fitting of the electron-density maps, model building and fitting of the ligands were carried out in COOT (Emsley *et al.*, 2010). The models were validated in COOT and *MolProbity* (Chen *et al.*, 2010a). The X-ray coordinates and data have been deposited in the Protein Data Bank with the codes 5fuv, 5fuw, 5fux and 5fuy.

Statistics

The Student's t-test was used for comparison of sets of data. The variances in normally distributed data sets were analyzed with two-way ANOVA and Dunnett's post-hoc tests. Normality was assessed using Kolmogorov-Smirnov and Shapiro-Wilk tests. Statistics were calculated with either Graphpad QuickCalcs (GraphPad Software) or IBM SPSS Statistics. * $p<0.05$, ** $p<0.01$ and *** $p<0.001$.

Acknowledgements

We thank Dr. Jenny Campos-Salinas for qualified support on the digital image analysis. We acknowledge Dr. Antonio E. Vidal for critical reading of the manuscript. We thank David Horn for kindly providing the anti-TbγH2A antibody. This work was supported by the Junta de Andalucía (BIO-199, P12-BIO-2059), the Plan Nacional de Investigación Científica, Instituto de Salud Carlos III-Subdirección General de Redes y Centros de Investigación Cooperativa-Red de Investigación Cooperativa en Enfermedades Tropicales (RICET FIS Network: RD12/0018/0017), the Plan Nacional (SAF2013-48999-R), the FEDER funds from the EU, the PARAMET network (FP7-PEOPLE-2011-ITN. GA290080) and the EU project TRYPOBASE (Grant agreement no: 223238).

Author Contributions

MV, JT, VMC-A, KW, DG-P contributed to the conception and design of the study. MV, JT, VMC-A, LMR-P, TB, JEN, LEB, HR, KW, DG-P performed the acquisition,

analysis and interpretation of the data. MV, JT, KW, DG-P contributed to the writing of the manuscript.

References

Ahmad, S.I., S.H. Kirk & A. Eisenstark, (1998) Thymine metabolism and thymineless death in prokaryotes and eukaryotes. *Annu Rev Microbiol* **52**: 591-625.

Al-Madhoun, A.S., W. Tjarks & S. Eriksson, (2004) The role of thymidine kinases in the activation of pyrimidine nucleoside analogues. *Mini Reviews in Medicinal Chemistry* **4**: 341-350.

al-Nabulsi, I., Y. Takamiya, Y. Voloshin, A. Dritschilo, R.L. Martuza & T.J. Jorgensen, (1994) Expression of thymidine kinase is essential to low dose radiation resistance of rat glioma cells. *Cancer Research* **54**: 5614-5617.

Ali, J.A., D.J. Creek, K. Burgess, H.C. Allison, M.C. Field, P. Maser & H.P. De Koning, (2013) Pyrimidine salvage in *Trypanosoma brucei* bloodstream forms and the trypanocidal action of halogenated pyrimidines. *Mol Pharmacol* **83**: 439-453.

Amie, S.M., R.A. Bambara & B. Kim, (2013) GTP is the primary activator of the anti-HIV restriction factor SAMHD1. *J Biol Chem* 288: 25001-25006.

Battye, T.G., L. Kontogiannis, O. Johnson, H.R. Powell & A.G. Leslie, (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica. Section D, Biological Crystallography* **67**: 271-281.

Berrow, N.S., D. Alderton, S. Sainsbury, J. Nettleship, R. Assenberg, N. Rahman, D.I. Stuart & R.J. Owens, (2007) A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Research* **35**: e45.

Birringer, M.S., M.T. Claus, G. Folkers, D.P. Kloer, G.E. Schulz & L. Scapozza, (2005) Structure of a type II thymidine kinase with bound dTTP. *FEBS letters* **579**: 1376-1382.

Birringer, M.S., R. Perozzo, E. Kut, C. Stillhart, W. Surber, L. Scapozza & G. Folkers, (2006) High-level expression and purification of human thymidine kinase 1: quaternary structure, stability, and kinetics. *Protein Expression and Purification* **47**: 506-515.

Bolte, S. & F.P. Cordelieres, (2006) A guided tour into subcellular colocalization analysis in light microscopy. *Journal of Microscopy* **224**: 213-232.

Bradshaw, H.D., Jr., (1983) Molecular cloning and cell cycle-specific regulation of a functional human thymidine kinase gene. *Proc Natl Acad Sci U S A* **80**: 5588-5591.

Caligo, M.A., G. Cipollini, L. Fiore, S. Calvo, F. Basolo, P. Collecchi, F. Ciardiello, S. Pepe, M. Petrini & G. Bevilacqua, (1995) NM23 gene expression correlates with cell growth rate and S-phase. *International journal of cancer* **60**: 837-842.

Canman, C.E., (2003) Checkpoint mediators: relaying signals from DNA strand breaks. *Current biology* : CB **13**: R488-490.

Castillo-Acosta, V.M., F. Aguilar-Pereyra, A.E. Vidal, M. Navarro, L.M. Ruiz-Perez & D. Gonzalez-Pacanowska, (2012) Trypanosomes lacking uracil-DNA glycosylase are hypersensitive to antifolates and present a mutator phenotype. *Int J Biochem Cell Biol* **44**: 1555-1568.

Castillo-Acosta, V.M., A.M. Estevez, A.E. Vidal, L.M. Ruiz-Perez & D. Gonzalez-Pacanowska, (2008) Depletion of dimeric all-alpha dUTPase induces DNA strand breaks and impairs cell cycle progression in *Trypanosoma brucei*. *Int J Biochem Cell Biol* **40**: 2901-2913.

Chang, Z.F., D.Y. Huang & L.M. Chi, (1998) Serine 13 is the site of mitotic phosphorylation of human thymidine kinase. *J Biol Chem* **273**: 12095-12100.

Chen, V.B., W.B. Arendall, 3rd, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson & D.C. Richardson, (2010a) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica. Section D, Biological Crystallography* **66**: 12-21.

Chen, Y.L., S. Eriksson & Z.F. Chang, (2010b) Regulation and functional contribution of thymidine kinase 1 in repair of DNA damage. *J Biol Chem* **285**: 27327-27335.

Clayton, C.E., A.M. Estevez, C. Hartmann, V.P. Alibu, M. Field & D. Horn, (2005) Down-regulating gene expression by RNA interference in *Trypanosoma brucei*. *Methods in Molecular Biology* **309**: 39-60.

Creek, D.J., M. Mazet, F. Achcar, J. Anderson, D.H. Kim, R. Kamour, P. Morand, Y. Millerioux, M. Biran, E.J. Kerkhoven, A. Chokkathukalam, S.K. Weidt, K.E. Burgess, R. Breitling, D.G. Watson, F. Bringaud & M.P. Barrett, (2015) Probing the metabolic network in bloodstream-form *Trypanosoma brucei* using untargeted metabolomics with stable isotope labelled glucose. *PLoS Pathogens* **11**: e1004689.

Cummings, D.J. & A.R. Kusy, (1970) Thymineless death in *Escherichia coli*: deoxyribonucleic acid replication and the immune state. *Journal of Bacteriology* **102**: 106-117.

Dobrovolsky, V.N., T. Bucci, R.H. Heflich, J. Desjardins & F.C. Richardson, (2003) Mice deficient for cytosolic thymidine kinase gene develop fatal kidney disease. *Molecular Genetics and Metabolism* **78**: 1-10.

Emsley, P., B. Lohkamp, W.G. Scott & K. Cowtan, (2010) Features and development of Coot. *Acta Crystallographica. Section D, Biological Crystallography* **66**: 486-501.

Eriksson, S., B. Munch-Petersen, K. Johansson & H. Eklund, (2002) Structure and function of cellular deoxyribonucleoside kinases. *Cell Mol Life Sci* **59**: 1327-1346.

Evans, P.R. & G.N. Murshudov, (2013) How good are my data and what is the resolution? *Acta Crystallographica. Section D, Biological Crystallography* **69**: 1204-1214.

Glover, L. & D. Horn, (2012) Trypanosomal histone gammaH2A and the DNA damage response. *Molecular and biochemical parasitology* **183**: 78-83.

Guerra-Giraldez, C., L. Quijada & C.E. Clayton, (2002) Compartmentation of enzymes in a microbody, the glycosome, is essential in *Trypanosoma brucei*. *J Cell Sci* **115**: 2651-2658.

Hemsworth, G.R., D. Gonzalez-Pacanowska & K.S. Wilson, (2013) On the catalytic mechanism of dimeric dUTPases. *The Biochemical journal* **456**: 81-88.

Hidalgo-Zarco, F., A.G. Camacho, V. Bernier-Villamor, J. Nord, L.M. Ruiz-Perez & D. Gonzalez-Pacanowska, (2001) Kinetic properties and inhibition of the dimeric dUTPase-dUDPase from Leishmania major. *Protein science : a publication of the Protein Society* **10**: 1426-1433.

Hofer, A., J.T. Ekanem & L. Thelander, (1998) Allosteric regulation of Trypanosoma brucei ribonucleotide reductase studied in vitro and in vivo. *J Biol Chem* **273**: 34098-34104.

Hunger-Glaser, I., A. Hemphill, T. Shalaby, M. Hanni & T. Seebeck, (2000) Nucleoside diphosphate kinase of *Trypanosoma brucei*. *Gene* **257**: 251-257.

Johansson, N.G. & S. Eriksson, (1996) Structure-activity relationships for phosphorylation of nucleoside analogs to monophosphates by nucleoside kinases. *Acta Biochimica Polonica* **43**: 143-160.

Ke, P.Y. & Z.F. Chang, (2004) Mitotic degradation of human thymidine kinase 1 is dependent on the anaphase-promoting complex/cyclosome-CDH1-mediated pathway. *Molecular and Cellular Biology* **24**: 514-526.

Khodursky, A., E.C. Guzman & P.C. Hanawalt, (2015) Thymineless Death Lives On: New Insights into a Classic Phenomenon. *Annu Rev Microbiol* **69**: 247-263.

Kosinska, U., C. Carnrot, S. Eriksson, L. Wang & H. Eklund, (2005) Structure of the substrate complex of thymidine kinase from *Ureaplasma urealyticum* and investigations of possible drug targets for the enzyme. *FEBS J* **272**: 6365-6372.

Kosinska, U., C. Carnrot, M.P. Sandrini, A.R. Clausen, L. Wang, J. Piskur, S. Eriksson & H. Eklund, (2007) Structural studies of thymidine kinases from *Bacillus anthracis* and *Bacillus cereus* provide insights into quaternary structure and conformational changes upon substrate binding. *FEBS J* **274**: 727-737.

Krissinel, E. & K. Henrick, (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica. Section D, Biological Crystallography* **60**: 2256-2268.

Krissinel, E. & K. Henrick, (2007) Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* **372**: 774-797.

Landeira, D., J.M. Bart, D. Van Tyne & M. Navarro, (2009) Cohesin regulates VSG monoallelic expression in trypanosomes. *J Cell Biol* **186**: 243-254.

Lee, L.S. & Y.C. Cheng, (1976) Human deoxythymidine kinase. I. Purification and general properties of the cytoplasmic and mitochondrial isozymes derived from blast cells of acute myelocytic leukemia. *J Biol Chem* **251**: 2600-2604.

Li, C.L., C.Y. Lu, P.Y. Ke & Z.F. Chang, (2004) Perturbation of ATP-induced tetramerization of human cytosolic thymidine kinase by substitution of serine-13 with aspartic acid at the mitotic phosphorylation site. *Biochem Biophys Res Commun* **313**: 587-593.

Lukas, J., C. Lukas & J. Bartek, (2011) More than just a focus: The chromatin response to DNA damage and its role in genome integrity maintenance. *Nature cell biology* **13**: 1161-1169.

McCoy, A.J., (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallographica. Section D, Biological Crystallography* **63**: 32-41.

McNicholas, S., E. Potterton, K.S. Wilson & M.E. Noble, (2011) Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallographica. Section D, Biological Crystallography* **67**: 386-394.

Murshudov, G.N., P. Skubak, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long & A.A. Vagin, (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica. Section D, Biological Crystallography* **67**: 355-367.

Mutahir, Z., A.R. Clausen, K.M. Andersson, S.M. Wisen, B. Munch-Petersen & J. Piskur, (2013) Thymidine kinase 1 regulatory fine-tuning through tetramer formation. *Febs J* **280**: 1531-1541.

Navalgund, L.G., C. Rossana, A.J. Muench & L.F. Johnson, (1980) Cell cycle regulation of thymidylate synthetase gene expression in cultured mouse fibroblasts. *J Biol Chem* **255**: 7386-7390.

O'Donovan, M., (2012) A critique of methods to measure cytotoxicity in mammalian cell genotoxicity assays. *Mutagenesis* **27**: 615-621.

Ong, H.B., N. Sienkiewicz, S. Wyllie, S. Patterson & A.H. Fairlamb, (2013) *Trypanosoma brucei* (UMP synthase null mutants) are avirulent in mice, but recover virulence upon prolonged culture in vitro while retaining pyrimidine auxotrophy. *Mol Microbiol* **90**: 443-455.

Organisation, W.H., (2014) Human African trypanosomiasis In., pp.

Ranjbarian, F., M. Vodnala, S.M. Vodnala, R. Rofougaran, L. Thelander & A. Hofer, (2012) *Trypanosoma brucei* thymidine kinase is tandem protein consisting of two homologous parts, which together enable efficient substrate binding. *J Biol Chem* **287**: 17628-17636.

Reichard, P., (1988) Interactions between deoxyribonucleotide and DNA synthesis. *Annual Review of Biochemistry* **57**: 349-374.

Requena, C.E., G. Perez-Moreno, L.M. Ruiz-Perez, A.E. Vidal & D. Gonzalez-Pacanowska, (2014) The NTP pyrophosphatase DCTPP1 contributes to the homoeostasis and cleansing of the dNTP pool in human cells. *The Biochemical Journal* **459**: 171-180.

Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak & A. Cardona, (2012) Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**: 676-682.

Segura-Pena, D., J. Lichter, M. Trani, M. Konrad, A. Lavie & S. Lutz, (2007a) Quaternary structure change as a mechanism for the regulation of thymidine kinase 1-like enzymes. *Structure* **15**: 1555-1566.

Segura-Pena, D., S. Lutz, C. Monnerjahn, M. Konrad & A. Lavie, (2007b) Binding of ATP to TK1-like enzymes is associated with a conformational change in the quaternary structure. *Journal of Molecular Biology* **369**: 129-141.

Sherley, J.L. & T.J. Kelly, (1988) Regulation of human thymidine kinase during the cell cycle. *J Biol Chem* **263**: 8350-8358.

Sienkiewicz, N., S. Jaroslawski, S. Wyllie & A.H. Fairlamb, (2008) Chemical and genetic validation of dihydrofolate reductase-thymidylate synthase as a drug target in African trypanosomes. *Mol Microbiol* **69**: 520-533.

Thiel, M., S. Harder, M. Wiese, M. Kroemer & I. Bruchhaus, (2008) Involvement of a *Leishmania* thymidine kinase in flagellum formation, promastigote shape and growth as well as virulence. *Molecular and Biochemical Parasitology* **158**: 152-162.

Timm, J., C. Bosch-Navarrete, E. Recio, J.E. Nettleship, H. Rada, D. Gonzalez-Pacanowska & K.S. Wilson, (2014) Structural and kinetic characterization of thymidine kinase from *Leishmania major*. *PLoS Negl Trop Dis* **9**: e0003781.

Tu, X. & C.C. Wang, (2005) Pairwise knockdowns of cdc2-related kinases (CRKs) in *Trypanosoma brucei* identified the CRKs for G1/S and G2/M transitions and demonstrated distinctive cytokinetic regulations between two developmental stages of the organism. *Eukaryot Cell* **4**: 755-764.

van den Hoff, M.J., A.F. Moorman & W.H. Lamers, (1992) Electroporation in 'intracellular' buffer increases cell survival. *Nucleic Acids Research* **20**: 2902.

Van Triest, B., H.M. Pinedo, G. Giaccone & G.J. Peters, (2000) Downstream molecular determinants of response to 5-fluorouracil and antifolate thymidylate synthase inhibitors. *Annals of Oncology: Official Journal of the European Society for Medical Oncology / ESMO* **11**: 385-391.

Vodnala, M., A. Fijolek, R. Rofougaran, M. Mosimann, P. Maser & A. Hofer, (2008) Adenosine kinase mediates high affinity adenosine salvage in *Trypanosoma brucei*. *J Biol Chem* **283**: 5380-5388.

Wakazono, Y., M. Kubota, K. Furusho, L. Liu & S.L. Gerson, (1996) Thymidine kinase deficient cells with decreased TTP pools are hypersensitive to DNA alkylating agents. *Mutation Research* **362**: 119-125.

Welin, M., U. Kosinska, N.E. Mikkelsen, C. Carnrot, C. Zhu, L. Wang, S. Eriksson, B. Munch-Petersen & H. Eklund, (2004) Structures of thymidine kinase 1 of human and mycoplasmic origin. *Proc Natl Acad Sci U S A* **101**: 17970-17975.

Winn, M.D., C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin & K.S. Wilson, (2011) Overview of the CCP4 suite and current developments. *Acta Crystallographica. Section D, Biological Crystallography* **67**: 235-242.

Wirtz, E., M. Hoek & G.A. Cross, (1998) Regulated processive transcription of chromatin by T7 RNA polymerase in *Trypanosoma brucei*. *Nucleic Acids Research* **26**: 4626-4634.

Wirtz, E., S. Leal, C. Ochatt & G.A. Cross, (1999) A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **99**: 89-101.

Yang, Z.R., R. Thomson, P. McNeil & R.M. Esnouf, (2005) RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics* **21**: 3369-3376.

Tables

Table 1: Crystallization conditions for the *Tb*TK structures, including additives to the crystallization solutions and composition of soaking solutions used for cryo-protection.

Crystal name	Additives in protein solution	Crystallization condition	Ligands seen in the active site in the structure
TbTK- dThd/dTMP	1 mM dTMP, 1 mM AppNHp, 3 mM $MgCl_2$	0.1 M HEPES pH 7.0, 0.9 M Succinic acid	dThd/dTMP
TbTK- dTMP-PO_4^{2-}-a	1 mM dTMP, 1 mM AppNHp, 3 mM $MgCl_2$	0.1 M HEPES pH 7.0, 0.9 M Succinic acid, 1% (w/v) PEG 2K MME	dTMP, PO_4^{2-}
TbTK- dThd-glycerol	1 mM dTMP, 1 mM AppNHp, 3 mM $MgCl_2$	0.1 M HEPES pH 7.0, 0.8 M Succinic acid	dThd, glycerol
TbTK- dTMP-PO_4^{2-}-b	1 mM dTMP, 1 mM AppNHp, 3 mM $MgCl_2$, Seed stock of *Lm*TK-dThd crystals	0.15 M DL-Malic acid pH 7.0, 20% (w/v) PEG 3350	dTMP, PO_4^{2-}

Table 2: Crystallographic data collection and refinement statistics.

	***Tb*TK-dThd-glycerol**	***Tb*TK-dThd/dTMP**	***Tb*TK-dTMP-PO_4^{2-}-a**	***Tb*TK-dTMP-PO_4^{2-}-b**
Data collection				
Diffraction Source	DLS beamline I04	DLS beamline I02	DLS beamline I02	DLS beamline I04
Wavelength (Å)	0.9795	0.9795	0.9795	
Temperature (K)	100	100	100	100
Detector	Pilatus 6M	Pilatus 6M-F	Pilatus 6M-F	Pilatus 6M
Rotation range per image (°)	0.2	0.2	0.2	0.2
Total Rotation range (°)	220	180	180	220
Crystal data				
Space group	$P4_322$	$P4_322$	$P4_322$	C222
a, b, c (Å)	114.7, 114.7,	115, 115, 104.8	114.8, 114.8, 104.5	87.61, 97.0,
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 89.99
Resolution (Å)	44.25-2.30 (2.39-2.30)	64.24-2.20 (2.27-2.20)	43.92-2.20 (2.27-2.20)	65.02-2.80 (2.95-2.80)
Total Reflections	456390	422851 (33374)	441889 (37765)	173354
Unique Reflections	31864 (3298)	36299 (3079)	36045 (3273)	31168 (4599)
Completeness (%)	99.9 (99.9)	100 (100)	100 (100)	95.9 (98.7)
Redundancy	14.3 (14.3)	11.6 (10.8)	12.3 (11.5)	5.6 (4.4)
R_{merge}	0.085 (1.024)	0.105 (0.885)	0.071 (1.215)	0.098 (0.373)
Mean [I/σ (I)]	17.0 (2.6)	11.8 (2.6)	18.9 (2.3)	8.5 (2.7)
Matthews' coefficient (Å3/Da)	4.01	3.81	3.97	2.34
Refinement statistics				
Reflections, working set	30146	34470	34261	29477
Reflections, test set	1708	1757	1749	1570
Resolution Range (Å)	44.25-2.30	64.24-2.20	43.92-2.20	65.02-2.80
R-factor	0.1680	0.1800	0.1810	0.2062
R_{free}	0.1888	0.1889	0.1887	0.2344
No. of non-H atoms				
Protein	2598	2588	2582	7420
Ligand	59	36	54	162
Water	94	123	76	32
Mean B factors (Å^2)				
Protein	42.2	39.3	43.1	48.6
Ligands	56.7	52.8	60.3	60.5
Water	55.5	59.0	60.2	47.5
RMS deviation from				
Bond length (Å)	0.020	0.020	0.020	0.013
Bond angles (°)	2.08	1.91	2.00	1.78
Ramachandran Plot				
Residues in favoured	97.3	97.0	96.6	96.0
Residues in allowed	2.4	3.0	3.4	4.0
Outliers	0.3	0	0	0
Molprobity score	1.44	1.36	1.43	1.39
Poor Rotamers (%)	0.36	0	0.36	0.26
PDB ID				

Figures

Fig 1. *Tb*TK depletion dramatically impairs parasite proliferation.

(A) Percentage of *TbTK* mRNA levels in parental and *TbTK* RNAi-induced and non-induced cells from a 48h culture, as resulted by RT-qPCR. (B) Expression profile of *Tb*TK in parental, *TbTK* RNAi plus and minus doxycycline as shown by Western blotting using affinity-purified polyclonal anti-*Tb*TK antibody. Anti-*Tb*CSM was used as loading control (Guerra-Giraldez *et al.*, 2002). (C) Time course analysis showing relative growth of *TbTK* RNAi line minus or plus doxycycline calculated as percentage of the parental line. Measurements of the parasite growth were taken at 8, 16, 24 and 48 hours after induction. (D) Cumulative growth of parental and *TbTK* RNAi-induced and non-induced parasites grown in standard HMI-9 medium supplemented with 10% FBS or pyrimidine-free HMI medium with 10% dialyzed FBS. (E) Relative growth of *TbTK* RNAi-induced line grown in normal medium or after supplementation with 1 mM uridine (Urd), deoxyuridine (dUrd), thymidine (dThd) or deoxycytidine (dCyd). Growth is expressed as percentage of the parental line's growth at the same time point in normal medium with no supplementation. (F) Cumulative growth of *TbTK* RNAi non-induced or induced cells by 1 μM doxycycline after supplementation with uracil (Ura) at the concentration of 100 μM or 1 mM.

Fig 2. *Tb*TK knockdown causes thymineless death in *Trypanosoma brucei*.

(A) dGTP and dTTP levels (± S.E.) in parental and *TbTK* RNAi cells at 16, 24 and 48 hours after induction. *p* represents Student's t-test. (B) Histograms of TUNEL-labeled cells of the parental and *TbTK* RNAi lines (plus and minus doxycycline) measured by flow cytometry. The negative control shown as gray area represents the counts of cells incubated with Label Solution in the absence of terminal transferase, while the

histogram demarcated by a black line overlay indicates TUNEL positive cells for each cell line subjected to the study. (C) Plot showing the percentage of TUNEL positive cells in parental and *TbTK* RNAi lines where TK knockdown was induced with doxycycline. The plot shows measurements taken at 24 and 48 hours after induction. *p* represents Student's t-test. * $p<0.05$, ** $p<0.01$ and *** $p<0.001$.

Fig 3. TbTK-depleted cells exhibit higher levels of the DNA repair marker γH2A.

(A) Flow cytometry histograms of γH2A -labeled cells, obtained by immunofluorescence with an anti-*Tb*γH2A antibody and an Alexa Fluor® 488 -conjugated anti-rabbit secondary antibody. The gray area represents the parental line, while the black dotted and continuous marks represent the *TbTK* RNAi line at 16 and 24 hours after RNAi induction respectively. The red area represents the positive sub-population for each cell line. (B) Percentage of γH2A positive cells, as determined by immunofluorescence microscopy of parental and *TbTK* RNAi cells at 16 and 24 hours after RNAi induction. Positive cells were considered as cells showing at least one nuclear γH2A focus. $n \geq 300$ for each sample. Error bars represent the standard error. (C) Representative immunofluorescence microscopy images showing γH2A positive and negative cells in the parental and *TbTK* RNAi cell lines plus and minus doxycycline. N, negative; P, positive; Bar, 5 μm.

Fig 4. Sensitivity of *TbTK* RNAi cells to methotrexate.

Relative population doubling (RPD) of *TbTK* RNAi and parental cell lines treated with increasing concentrations of methotrexate for 24 hours. The RPD was calculated as PDtreated/ PDnot-treated x100% were PD = {log [N24h ÷ N0]} ÷ log 2. The plot represents the mean RPD (± S.E.) of two independent experiments where RPD for each

point was the result of four independent wells of a 96-well plate culture. *P* represents an ANOVA test, *** p<0.001.

Fig 5. *Tb*TK intracellular localization.

Immunofluorescence analysis of the intracellular localization of TK from the parental line in bloodstream forms (BF) and in procyclics (PF) expressing a myc tagged fusion protein (*Tb*TK-myc). TK signal (green) was detected by either anti-*Tb*TK and FITC-conjugated anti-rabbit, or anti-c-myc and Alexa Fluor® 488 conjugated goat anti-mouse antibodies. Mitochondria were stained with MitoTracker® Red CMXRos (red) while nuclei and kinetoplasts were stained with DAPI (blue). Images were collected with a confocal Leica SP5 microscope, deconvolved using Huygens Professional from Scientific Volume Imaging (version 3.3) and pseudo-colored by using Fiji software (Schindelin *et al.*, 2012). Bar, 5 μm.

Fig 6. *Tb*TK expression is cell cycle-regulated in *T. brucei.*

(A) Left panel, percentage of cells of the parental line showing intense, moderate or low *Tb*TK expression according to the number of nuclei (N) and kinetoplasts (K). TK expression was determined by immunofluorescence analysis with an anti-*Tb*TK antibody and FITC-conjugated anti-rabbit secondary antibody. Right panel, an example of different FITC signal intensity (I= Intense, N= No). Nuclei and kinetoplasts were detected by DAPI staining. The differential interference contrast (DIC) image is shown as a grayscale inset.

(B) Percentage of bloodstream forms (BF) that express *Tb*TK in the nucleus (violet) or cytosol (orange), classified according to the number of nuclei (N) and kinetoplasts (K). The quantification of *Tb*TK in each compartment was determined by using Fiji software (Schindelin *et al.*, 2012). (C) Fold changes in *Tb*TK expression during the cell cycle

with regard to the G1 phase as shown by FACS analysis with propidium iodide. (D) Left panel, percentage of procyclic *Tb*TK-myc cells showing intense, moderate or low *Tb*TK expression according to the number of nuclei (N) and kinetoplasts (K), as determined by immunofluorescence analysis with anti-c-myc antibody and Alexa Fluor® 488 -conjugated anti-rabbit secondary antibody. Right panel, an example of cells with different Alexa Fluor® 488 signal intensity (I= Intense, L= Low). Nuclei and kinetoplasts were detected by DAPI staining. The differential interference contrast (DIC) image is shown as a grayscale inset.

Images were collected with an Olympus microscope and Cell R1X81 software. Deconvolution and pseudo-coloring of images was carried out using Huygens Essential software (version 3.3; Scientific Volume Imaging) and Fiji software (version 1.5e; ImageJ) (Schindelin *et al.*, 2012), respectively.

Fig 7. *Tb*TK depletion highly impairs cell cycle progression in *Trypanosoma brucei* bloodstream forms.

(A) Quantification of parental and *TbTK* RNAi cells plus and minus doxycycline after 24 hours of induction according to the number of nuclei (N) and kinetoplasts (K). Data are shown as the mean percentage (± S.E.) of total cells (>500) coming from two independent experiments. The asterisks show significant differences calculated by the Student's t-test: * $p<0.05$, ** $p<0.01$ and *** $p<0.001$. (B) Cell cycle progression in *TbTK* RNAi induced cells as shown by FACS analysis. The plot shows the percentage of cells in each cell cycle stage (G1, S, G2/M, and post-G2/M) and at different time points (8, 16 and 24 hours). Each point in the plot represents the mean (± S.E.) of at least two samples from two independent experiments. (C) Fluorescence microscopy of *T. brucei* upon *Tb*TK depletion after 24 hours of RNAi induction. Parasite morphology

(phase) and nuclei (N) and kinetoplasts (K) number and morphology (DAPI staining) are shown. (D) kDNA signal plot and histogram of parental and *TbTK* RNAi lines at 24 and 48 hours after induction as revealed by FACS analysis of live cells stained with dihydroethidium.

Fig 8. Survival profile analysis of mice infected with *Tb*TK- deficient *T. brucei*.

Kaplan-Meier survival analysis for mice infected with parental bloodstream forms (BF) and *TbTK* RNAi line in the absence or presence of doxycycline to induce *Tb*TK knockdown.

Fig 9. Overall structure and assembly of *Tb*TK.

(A) *Tb*TK-dThd-glycerol protomer in ribbon representation with dThd shown as spheres colored by atom type. The disordered loop between amino acid residues 245-260 is represented by a dotted line, but this must not be taken to indicate its true position and is only included to show that while the protein is intact, there is no electron density for this loop. (B) *Tb*TK tetramer in worm-tube representation colored by chain. Figure made with CCP4*mg* (McNicholas *et al.*, 2011).

Fig 10. *Tb*TK substrate binding site.

(A) Superposition of the dThd binding site and P-loop of the four *Tb*TK structures in cylinder representation; *Tb*TK-dThd-glycerol in light green, *Tb*TK-dThd/dTMP in ice blue, *Tb*TK-dTMP-PO_4^{2-}-a in lawn green, and *Tb*TK-dTMP-PO_4^{2-}-b in light blue. Hydrogen bonds are represented as dashed lines. (B) dThd binding site with focus on residue F321. Phenylalanines and substrate shown as cylinders with *Tb*TK in lawn green and *Hs*TK1 (PDB ID: 1w4r) in lemon. (C) Superposition of the substrates of

*Tb*TK-dTMP-PO_4^{2-}-a (lawn green) and *Tm*TK-dThd-AppNHp (PDB ID: 2qq0, orange). The position of the P-loop is identical in both structures and indicated for *Tb*TK. Figure made with CCP4*mg* (McNicholas *et al.*, 2011) with the superposition carried out using SSM (Krissinel & Henrick, 2004).

Fig 11. Model of *Tb*TK pseudo-dimers.

Two chains are shown, one with the modeled sequence of the inactive N-terminal domain (chain A), the other with the sequence of the catalytically active C-terminal one (chain B), colored by sequence identity: identical residues (green), similar residues (yellow), amino acid replacements with amino acids harboring different properties (red). (A) A *Tb*TK pseudo-dimer forming at the dimer interface with the first residue of chain A (modeled TK-homology domain) and last residue of chain B (catalytically active domain, crystal structure of *Tb*TK-dThd-glycerol) are shown as spheres; dashed lines representing the 18 amino acid residue linker region connecting the termini of the two TK-homology domains in full-length enzyme. (B) Alternative pseudo-dimer formed by full-length *Tb*TK along the tetramer interface. Figure made with CCP4*mg* (McNicholas *et al.*, 2011).

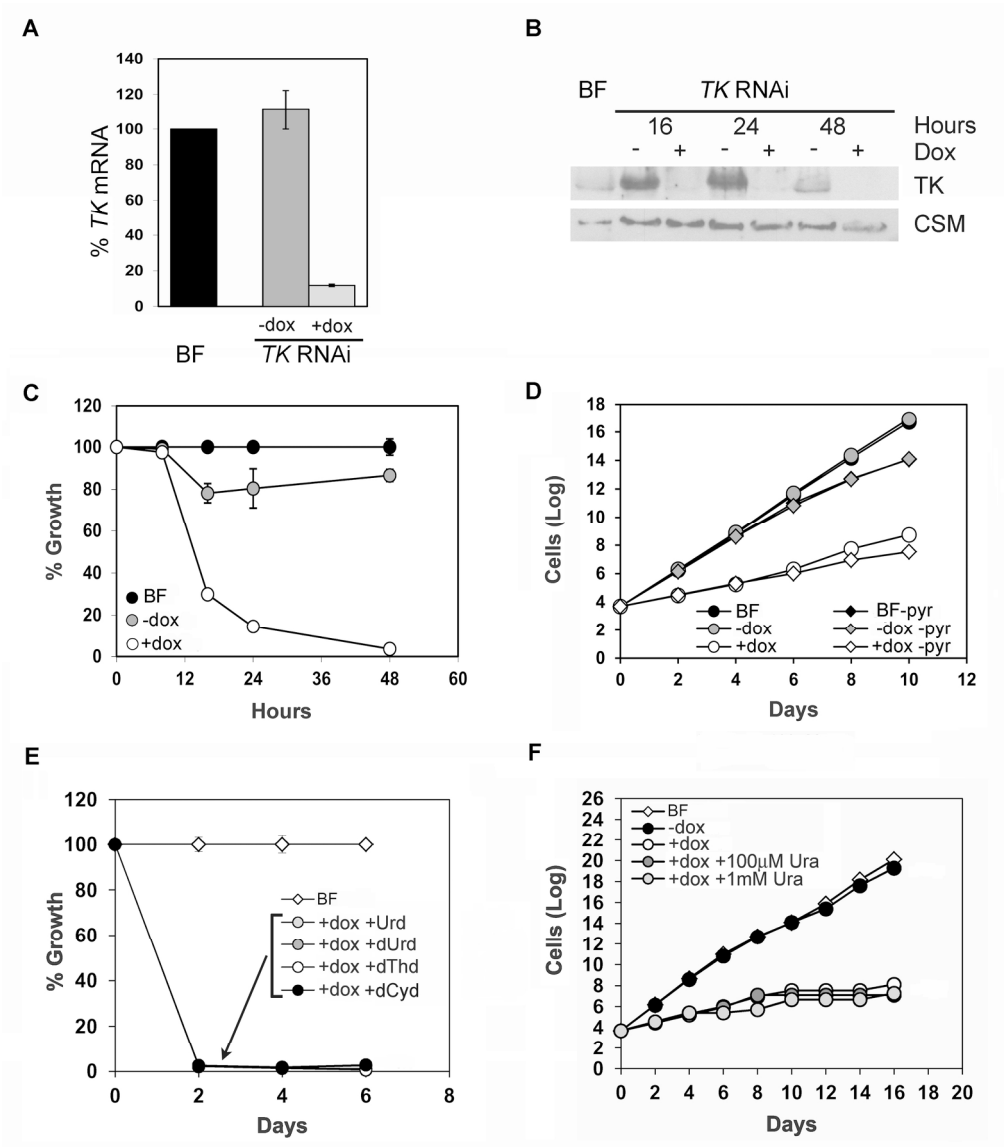


Figure 1. TbTK depletion dramatically impairs parasite proliferation.
Fig. 1
192x220mm (300 x 300 DPI)

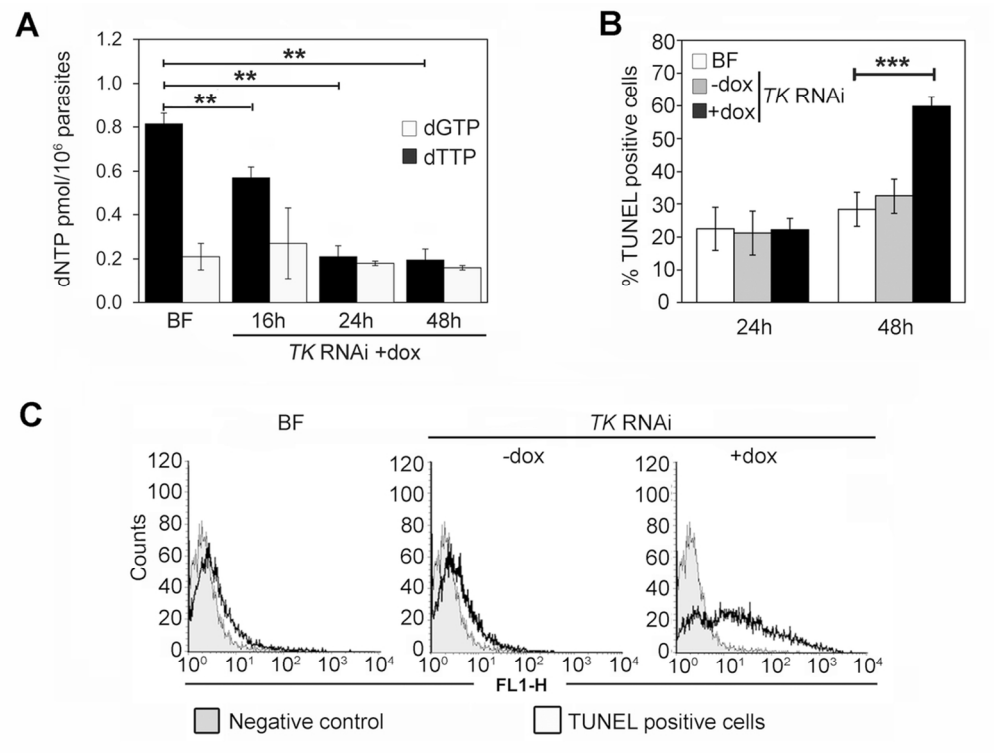


Figure 2. TbTK knockdown causes thymineless death in Trypanosoma brucei.
Fig. 2
105x79mm (300 x 300 DPI)

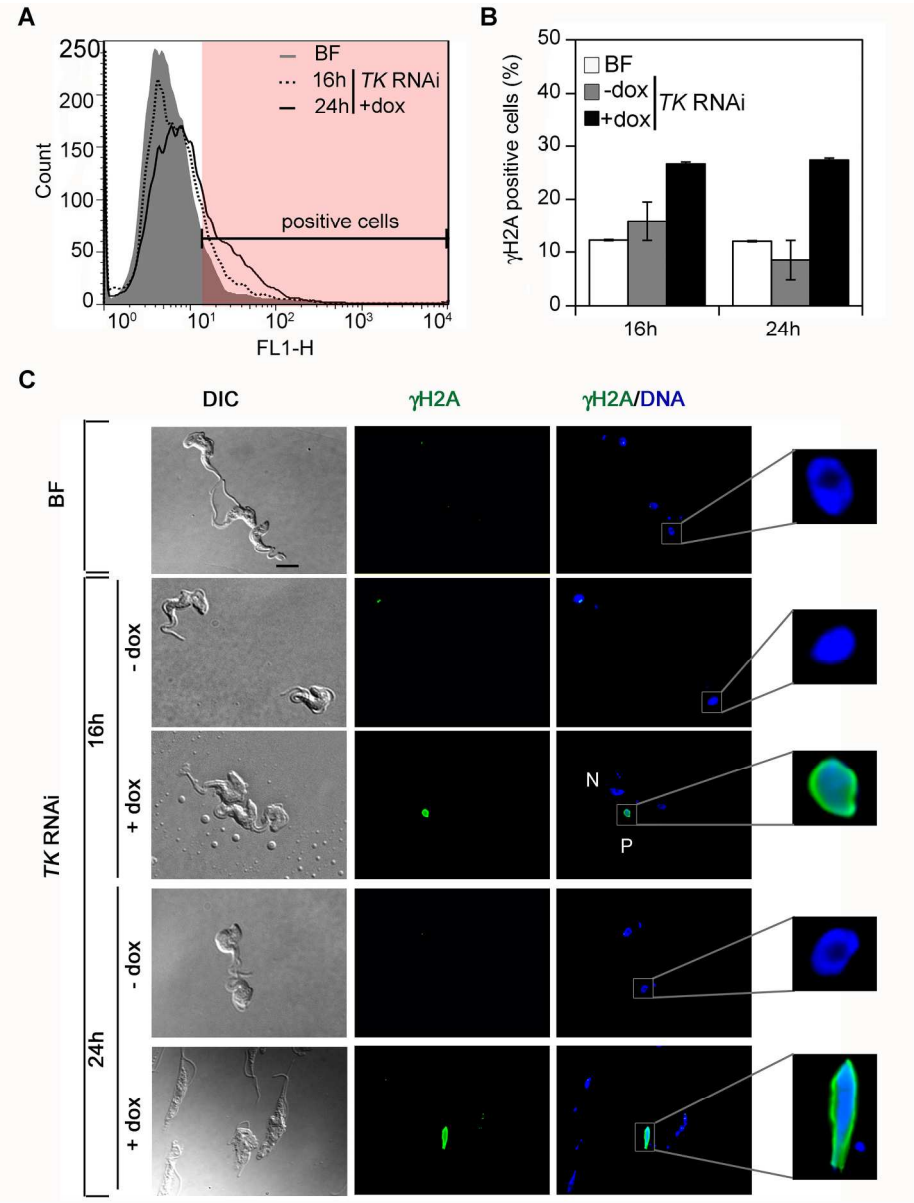


Figure 3. TbTK-depleted cells exhibit higher levels of the DNA repair marker γH2A.
Fig. 3
196x257mm (300 x 300 DPI)

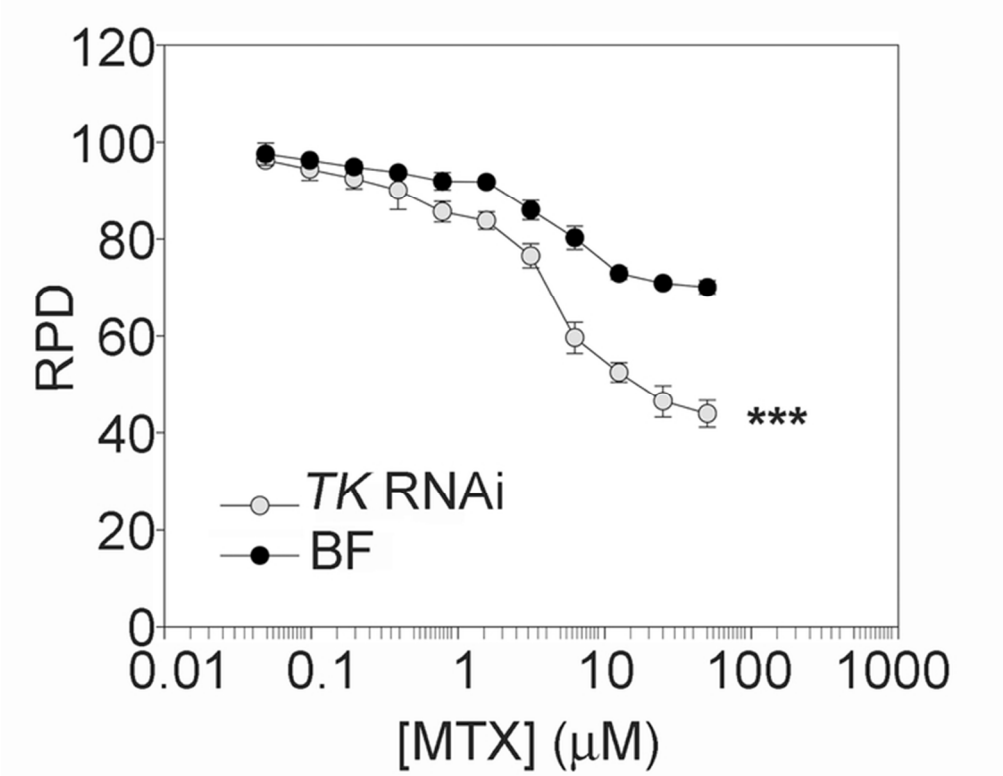


Figure 4. Sensitivity of TbTK RNAi cells to methotrexate.
Fig. 4
63x49mm (300 x 300 DPI)

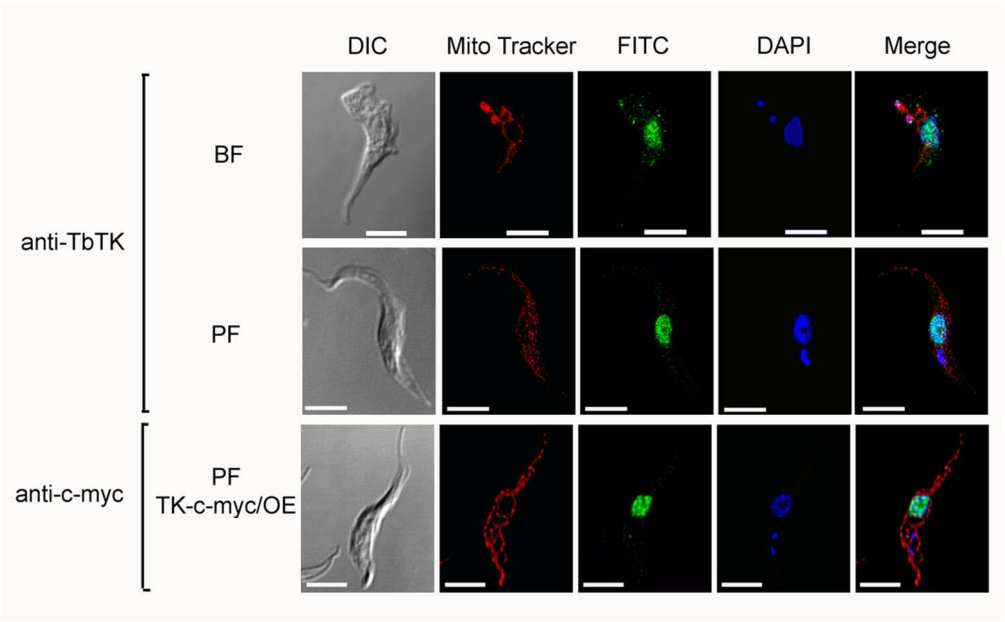


Figure 5. TbTK intracellular localization.
Fig. 5
104x64mm (300 x 300 DPI)

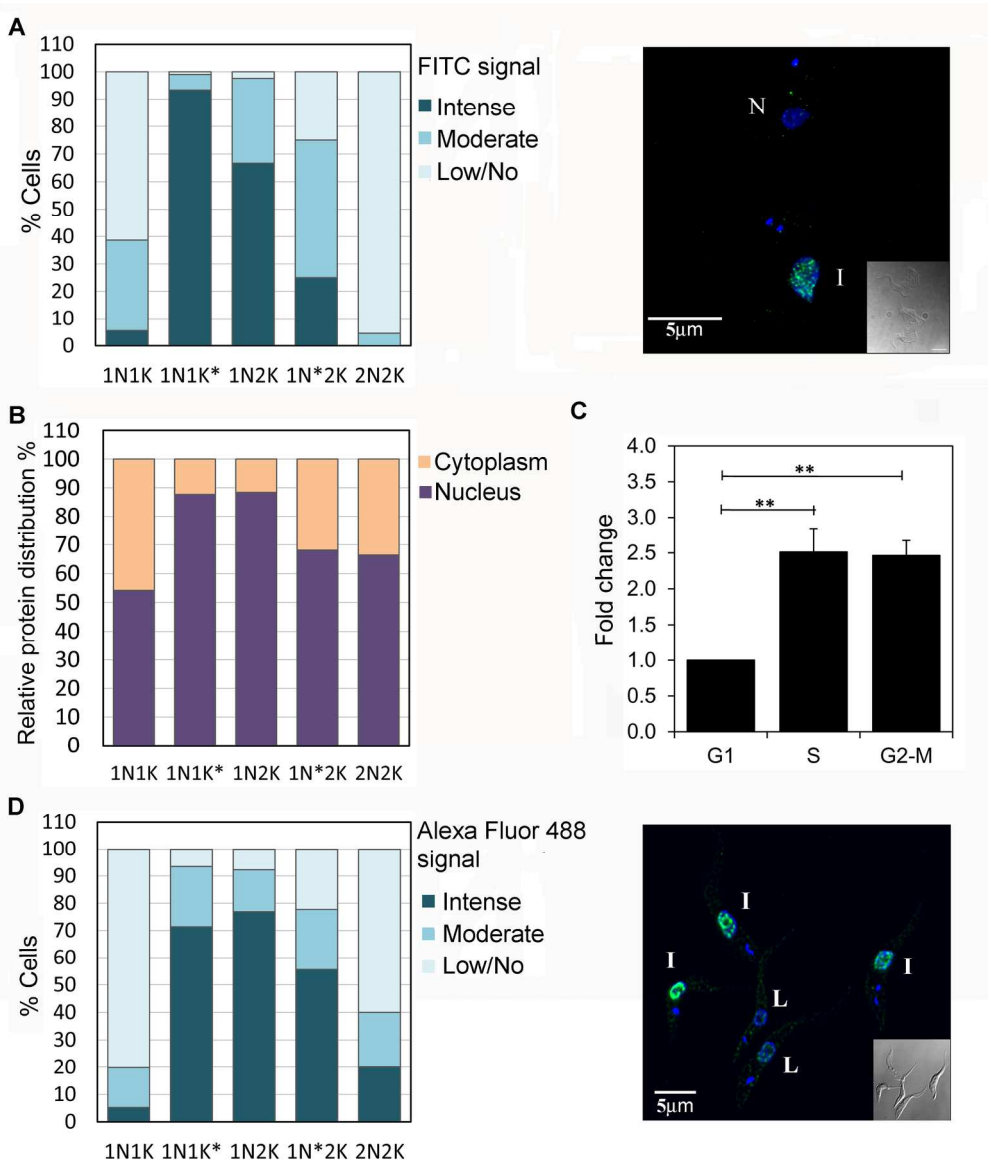


Figure 6. TbTK expression is cell cycle-regulated in T. brucei.
Fig. 6
197x229mm (300 x 300 DPI)

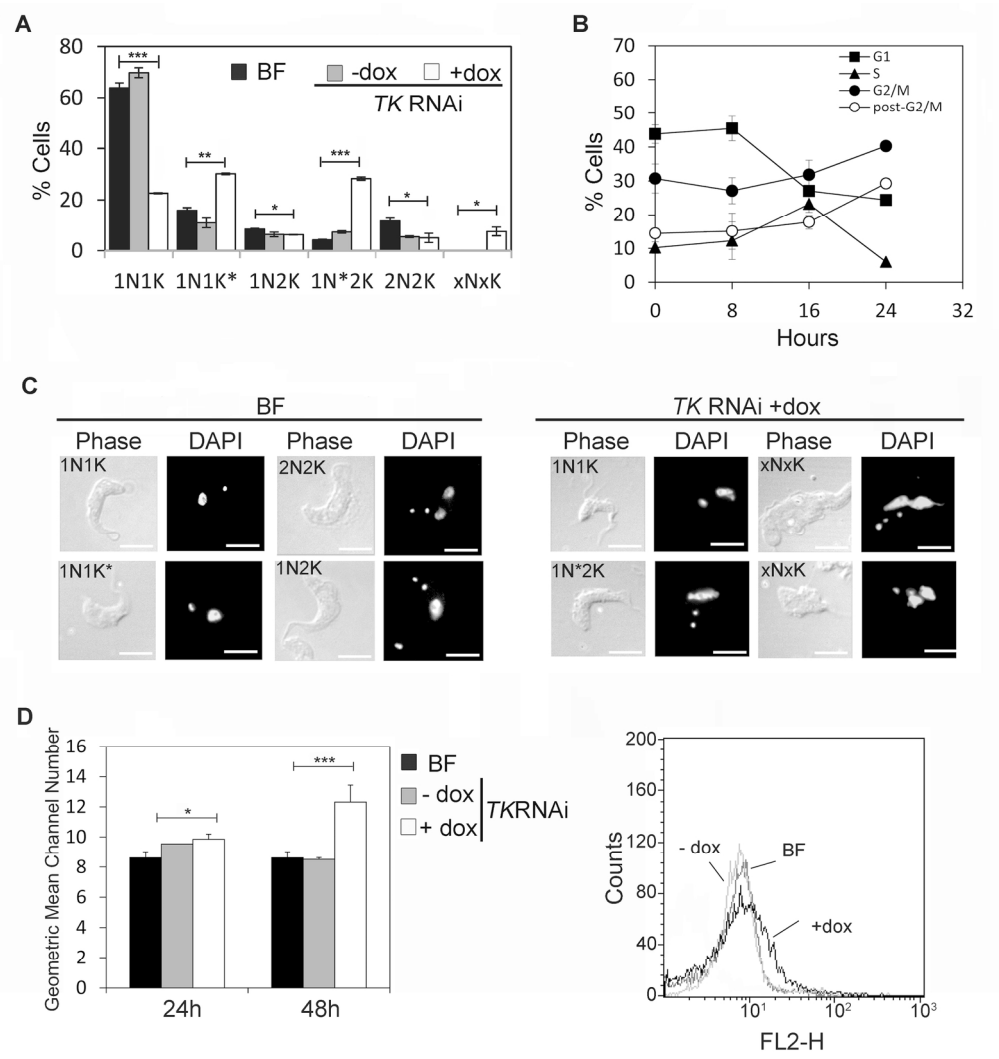


Figure 7. TbTK depletion highly impairs cell cycle progression in Trypanosoma brucei bloodstream forms.
Fig. 7
169x180mm (300 x 300 DPI)

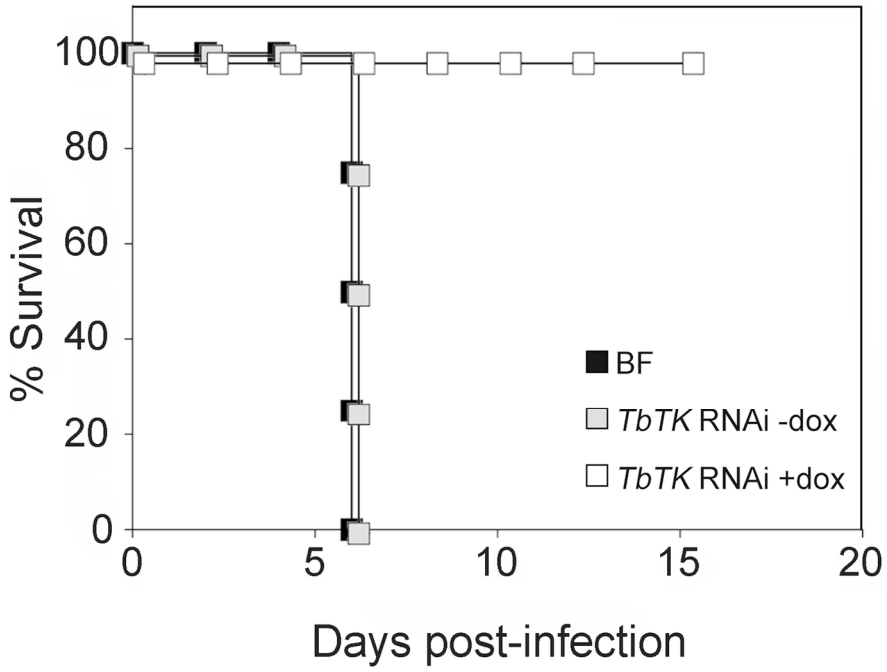


Figure 8. Survival profile analysis of mice infected with TbTK- deficient T. b. brucei.
Fig. 8
154x119mm (300 x 300 DPI)

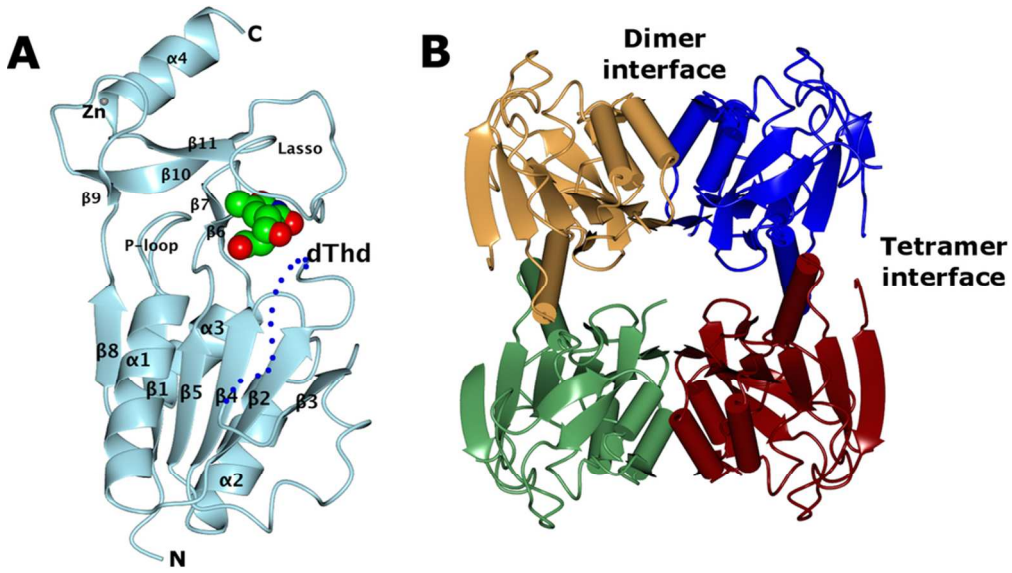


Figure 9. Overall structure and assembly of TbTK.
Fig. 9
90x50mm (300 x 300 DPI)

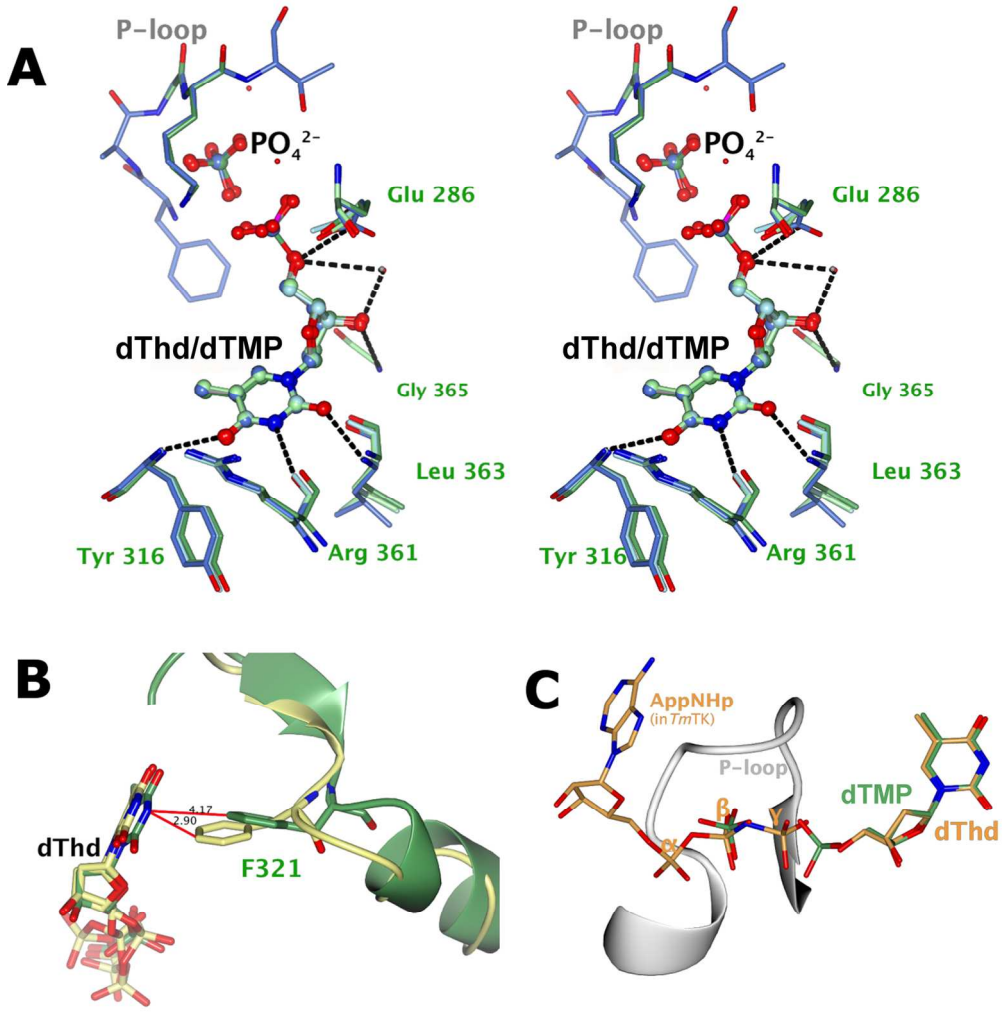


Figure 10. TbTK substrate binding site.
Fig. 10
165x171mm (300 x 300 DPI)

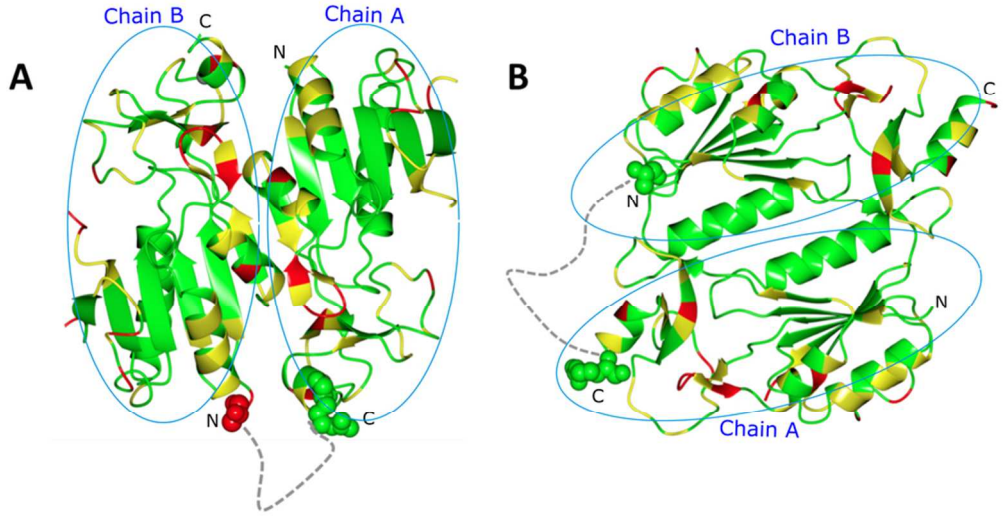
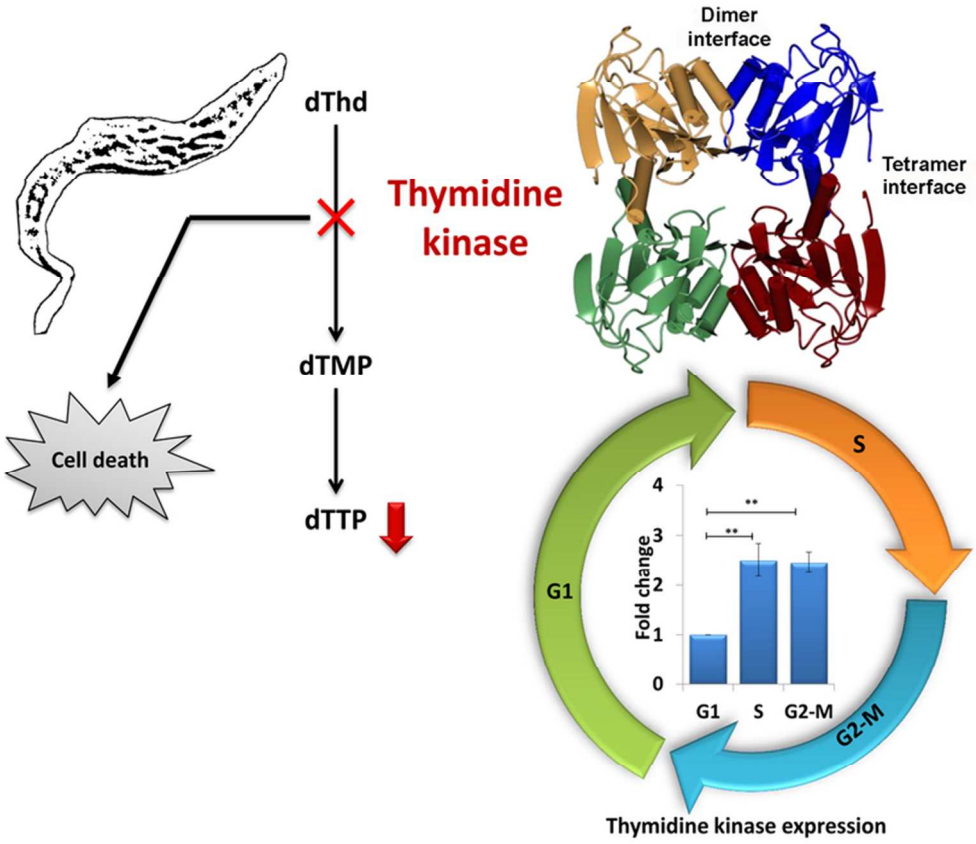


Figure 11. Model of TbTK pseudo-dimers.
Fig. 11
84x44mm (300 x 300 DPI)



70x61mm (300 x 300 DPI)

Abbreviated Summary

In this study, we establish that the depletion of *Trypanosoma brucei* thymidine kinase (*Tb*TK) leads to strongly diminished dTTP pools, DNA damage and parasite death. We show that *Tb*TK expression is cell cycle regulated and essential for cell cycle progression. In addition, we report the X-ray structure of the catalytically active domain of *Tb*TK and identify structural peculiarities within the active site that could be exploited in the process of rational drug design.