

## RESEARCH ARTICLE

# Identification of inhibitors of an unconventional *Trypanosoma brucei* kinetochore kinase

Leah S. Torrie, Fabio Zuccotto, David A. Robinson, David W. Gray, Ian H. Gilbert\*, Manu De Rycker<sup>ID</sup>\*

Drug Discovery Unit, Wellcome Centre for Anti-Infectives Research, School of Life Sciences, University of Dundee, Dundee, United Kingdom

\* [I.H.Gilbert@dundee.ac.uk](mailto:I.H.Gilbert@dundee.ac.uk) (IHG); [M.DeRycker@dundee.ac.uk](mailto:M.DeRycker@dundee.ac.uk) (MDR)



## OPEN ACCESS

**Citation:** Torrie LS, Zuccotto F, Robinson DA, Gray DW, Gilbert IH, De Rycker M (2019) Identification of inhibitors of an unconventional *Trypanosoma brucei* kinetochore kinase. PLoS ONE 14(5): e0217828. <https://doi.org/10.1371/journal.pone.0217828>

**Editor:** Giovanni Maga, Istituto di Genetica Molecolare, ITALY

**Received:** April 3, 2019

**Accepted:** May 21, 2019

**Published:** May 31, 2019

**Copyright:** © 2019 Torrie et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** This work was funded by Wellcome Trust (<https://wellcome.ac.uk/>) awards 092340 and 204672. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

## Abstract

The discovery of 20 unconventional kinetochore proteins in *Trypanosoma brucei* has opened a new and interesting area of evolutionary research to study a biological process previously thought to be highly conserved in all eukaryotes. In addition, the discovery of novel proteins involved in a critical cellular process provides an opportunity to exploit differences between kinetoplastid and human kinetochore proteins to develop therapeutics for diseases caused by kinetoplastid parasites. Consequently, we identified two of the unconventional kinetochore proteins as key targets (the highly related kinases KKT10 and KKT19). Recombinant *T. brucei* KKT19 (*TbKKT19*) protein was produced, a peptide substrate phosphorylated by *TbKKT19* identified (KKLRRTLSVA), Michaelis constants for KKLRRTLSVA and ATP were determined (179  $\mu$ M and 102  $\mu$ M respectively) and a robust high-throughput compatible biochemical assay developed. This biochemical assay was validated pharmacologically with inhibition by staurosporine and hypothemycin (IC<sub>50</sub> values of 288 nM and 65 nM respectively). Surprisingly, a subsequent high-throughput screen of a kinase-relevant compound library (6,624 compounds) yielded few hits (8 hits; final hit rate 0.12%). The low hit rate observed was unusual for a kinase target, particularly when screened against a compound library enriched with kinase hinge binding scaffolds. In an attempt to understand the low hit rate a *TbKKT19* homology model, based on human cdc2-like kinase 1 (CLK1), was generated. Analysis of the *TbKKT19* sequence and structure revealed no obvious features that could explain the low hit rates. Further work will therefore be necessary to explore this unique kinetochore kinase as well as to assess whether the few hits identified can be developed into tool molecules or new drugs.

## Introduction

Kinetochores are multiprotein complexes that associate with the centromere of a chromosome during cell division, ensuring faithful transmission of genetic material to daughter cells [1, 2]. While kinetochores are evolutionary quite plastic, most eukaryotic kinetochores share the

same canonical kinetochore machinery [3]. The kinetochores of the protozoan parasite *Trypanosoma brucei*, the causative agent of human African trypanosomiasis [4], are an exception to this. The discovery of 20 novel kinetochore proteins in *T. brucei* [5–7], with no detectable homology to conventional eukaryotic kinetochore proteins, raises not only an interesting area of research to investigate the evolutionary history of these proteins, but also the possibility of exploiting these unique proteins in the search for new therapeutics, in particular as these proteins are conserved across the kinetoplastida order. This includes *Leishmania spp.* and *Trypanosoma cruzi*, the causative agents of leishmaniasis and Chagas disease [8, 9]. These neglected and devastating diseases affect some of the world's poorest populations with current drug treatments suffering from high cost, host toxicity and emerging resistance [10, 11]. In the search for new therapeutics for these diseases, targeting novel, but essential, enzymes within the parasite provides an attractive start point for drug discovery. The finding that four of the kinetoplastid kinetochore proteins (KKTs) are kinases offers an opportunity to exploit a known, druggable enzyme family, whilst also targeting proteins unique to the parasite.

Of the four kinetochore kinases previously described [5, 6] (KKT2 (Tb927.11.10520), KKT3 (Tb927.9.10920), KKT10 (Tb927.11.12410) and KKT19 (Tb927.11.12420)), the KKT2 and KKT3 proteins were found to be similar in terms of domain organisation and amino acid sequence (48% similarity), as were the KKT10 and KKT19 proteins (also referred to as CLK1 and CLK2 respectively [12, 13]) (85% similarity), with these pairs most probably arising from gene duplication events. Genome-wide RNAi studies in *T. brucei* reveal that knockdown of these kinetochore kinases is associated with loss-of-fitness of the parasite [14, 15]. Further studies confirm that knockdown of protein levels with either KKT2, KKT10 or dual KKT10 / KKT19 RNAi constructs affect cell growth [5, 12, 13]. Interestingly, despite a genome-wide RNAi screen showing loss-of-fitness when KKT19 alone is knocked down [14, 15], subsequent KKT19 RNAi studies in *T. brucei* have revealed no effect on parasite growth [13, 16].

Chemical validation of the KKT10 protein has been achieved using the natural product hypothemycin [13]. This compound was shown to inhibit both *T. brucei* KKT10 (*TbKKT10* / *TbCLK1*) activity ( $IC_{50} = 150$  nM) and *T. brucei* cell growth ( $EC_{50} = 170$  nM), with chemoproteomics confirming hypothemycin engages primarily with *TbKKT10* in *T. brucei* cell lysates at concentrations relevant to cellular efficacy [13]. Hypothemycin was also shown to reduce parasitemia in *T. brucei* infected mice, with prolonged survival of infected mice over 30 days and a 33% cure rate observed following 7 daily treatments with 10 mg/ml hypothemycin [13].

Based on these data, the KKT10 / KKT19 proteins were prioritised as targets for entry into a kinetoplastid drug discovery program. These kinases have been classified as members of the LAMMER subfamily of CMGC kinases [17] with 100% sequence identity in the active site. As such, any inhibitors identified would be expected to inhibit both kinases which, as highlighted in RNAi studies, confers a growth defect in *T. brucei* parasites [5]. Due to the successful production of recombinant *TbKKT19* protein, the current study focuses on the development of a high-throughput screening assay for *TbKKT19* and the subsequent high-throughput screen of a kinase-relevant compound library to identify ATP site inhibitors of this unique kinetoplastid kinetochore kinase.

## Materials and methods

### Protein expression and purification method for recombinant *TbKKT19*

A His-tagged, full-length *TbKKT19* expression construct (kindly provided by Bungo Akiyoshi) was transformed into Rosetta 2 competent cells for protein production. A 60 ml overnight culture was set up and grown at 37°C 200 rpm for 16 h and was used the next day to inoculate 6 litres of LB+AMP+Kan media. Cells were grown at 37°C until the

OD<sub>600</sub> = 0.704 then induced with 0.2 mM IPTG then grown 20°C for 16 hrs before harvesting by centrifugation at 3,500 g for 30 min and storage at -20°C. Lysis buffer (60 ml, 50 mM NaPhosphate / 500 mM NaCl / 10% Glycerol / 20 mM Imidazole pH 7.5 / protease inhibitor tablets / DNAase) was added and the pellets defrosted at 25°C in a water bath for approximately 20 min. The slurry was then passed through a Cell Disrupter (Constant Systems) set at 30 KPSI to lyse the cells. The sample was then centrifuged at 40,000 g for 30 min. The supernatant was then filtered using syringe filters to 0.2 µm. The supernatant was loaded onto a 5 ml HiTrap Ni HP column that had been equilibrated with Buffer A (50 mM NaPhosphate / 500 mM NaCl / 10% Glycerol / 20 mM Imidazole pH 7.5) at 5 ml/min using an AKTA Pure (GE). Once loaded the column was washed with 10 column volumes of buffer A followed by 5% Buffer B (50 mM NaPhosphate / 500 mM NaCl / 10% Glycerol / 500 mM Imidazole pH 7.5) to wash off His-rich contaminating proteins. A linear gradient of 5–50% B was used to elute the protein. Approximately 38 mg of protein (from 6 L culture) was present in the fractions containing the *Tb*KKT19 protein. The sample was then passed through a 0.2 µm filter, before loading onto a XK26/60 Superdex 75 column using an AKTA Pure at Room Temp at 2 ml/min.

### ***Tb*KKT19 substrate identification**

Following production of recombinant protein, *Tb*KKT19 was screened against a panel of 41 putative kinase peptide substrates (at a concentration of 300 µM or 100 µM for peptide KTF CGTPEY LAPEVRREPRLS EEEQEMFRDFDY IADWC) and 3 putative protein substrates (at a concentration of 1 mg/ml) using the kinase screening service provided by the MRC PPU International Centre for Kinase Profiling at the University of Dundee (S1 Table).

### ***Tb*KKT19 biochemical assay development**

The peptide substrate KKLRRTLSVA was custom synthesised by Pepceuticals Ltd. and used for further investigation in a non-radiometric high-throughput compatible assay. Activity of the *Tb*KKT19 enzyme was determined by monitoring levels of ADP released during the enzymatic reaction using the commercially available ADP Hunter Plus kit (DiscoverX).

Development of the *Tb*KKT19 assay was carried out in black, low-volume, 384-well plates (Greiner) at room temperature (~23°C). Enzyme linearity was assessed in a 10 minute endpoint assay with varying *Tb*KKT19 concentrations (0–20 nM), 200 µM ATP and 200 µM peptide substrate KKLRRTLSVA ('KKL' peptide).

Assays to determine the apparent Michaelis constants ( $K_m^{\text{app}}$ ) for the ATP and the 'KKL' peptide were carried out in 8 µl reaction volumes containing 5 nM recombinant *Tb*KKT19 prepared in the proprietary ADP Hunter buffer (pH 7.4) supplemented with 1 mM dithiothreitol. To determine the  $K_m^{\text{app}}$  of ATP, the ATP concentration was varied (0–800 µM), with the 'KKL' concentration fixed at a saturating concentration of 1 mM. To determine the  $K_m^{\text{app}}$  of the 'KKL' peptide, the 'KKL' concentration was varied (0–800 µM), with the ATP concentration fixed at a saturating concentration of 600 µM. Detection of ADP formation was carried out by the addition of 4 µl ADP Hunter kit reagent A, followed by the addition of 8 µl ADP Hunter kit reagent B to each assay well. The ADP Hunter signal was allowed to develop for 30 min before the fluorescence intensity (FI) of each well was read using a PheraStar plate reader (BMG) (Excitation 540 nm; Emission 590 nm). Assays were monitored over time (0–60 min), with linear reaction rates at each substrate concentration determined. Rate *versus* substrate concentration data were fitted to the Michaelis-Menten equation using GraFit (Erithacus Software).

### **TbKKT19 hit identification**

The *TbKKT19* high throughput screen was performed using our in-house kinase-relevant compound collection, containing 6,624 compounds [18]. All library compounds were solubilized in 100% DMSO to a concentration of 10 mM.

Single point inhibition assays were carried out at room temperature (~23 °C) in black, low-volume, 384-well plates (Greiner). Each assay was performed in an 8  $\mu$ l reaction volume containing ADP Hunter buffer (pH 7.4), 1 mM dithiothreitol, 5 nM *TbKKT19*, 80  $\mu$ M ATP, 100  $\mu$ M 'KKL' peptide, and 30  $\mu$ M test compound.

Test compounds (24 nl in DMSO) were transferred to assay plates using an ECHO 550 acoustic dispenser (Labcyte). DMSO was added to 0% inhibition control wells, with staurosporine added to 100% inhibition control wells (to a final assay concentration of 30  $\mu$ M). Assays were performed by adding 4  $\mu$ l buffer with enzyme to all assay plates before the reaction was initiated with the addition of a 4  $\mu$ l substrate mix containing ATP and 'KKL' peptide to all wells. Following a 30 min reaction at room temperature the assay was stopped with the addition of 4  $\mu$ l ADP Hunter kit reagent A, followed by the addition of 8  $\mu$ l ADP Hunter kit reagent B. The ADP Hunter signal was allowed to develop for 30 min before the FI of each well was read using a PheraStar plate reader (BMG) (Excitation 540 nm; Emission 590 nm). All liquid dispensing steps were carried out using a Thermo Scientific Multidrop dispenser (Matrix).

ActivityBase from IDBS was used for data processing and analysis. Test compound data were normalised to 0% inhibition and 100% inhibition control wells on each plate, with compounds designated as hits if the % inhibition at 30  $\mu$ M was >35%.

To generate IC<sub>50</sub> data for *TbKKT19* hit compounds, 10-point inhibitor curves were prepared in 384-well assay plates using an ECHO 550 acoustic dispenser (Labcyte). Following preparation of the inhibitor curves, assays were carried out using the ADP Hunter assay as described above.

ActivityBase from IDBS was again used for data processing and analysis. All IC<sub>50</sub> curve fitting was undertaken using ActivityBase XE from IDBS. A four-parameter logistic dose-response curve fit was utilized with prefit used for all four parameters.

### **TbKKT19 counterscreen**

As described above, 10-point inhibitor curves were prepared in 384-well plates using an ECHO 550 acoustic dispenser (Labcyte). Counterscreen assays, to identify any compounds inhibiting the ADP Hunter detection components rather than *TbKKT19*, were carried out by adding 8  $\mu$ l of 24  $\mu$ M ADP, prepared in ADP Hunter buffer (pH 7.4) containing 1 mM dithiothreitol, to all wells, with the exception of the 100% effect control wells, to which 8  $\mu$ l buffer only was added.

Following a 30 min incubation at room temperature (~23 °C), detection of ADP levels using the ADP Hunter kit was carried out as previously described. All liquid dispensing steps were carried out using a Thermo Scientific Multidrop dispenser (Matrix).

Data processing and analysis was performed using ActivityBase from IDBS as described above.

### **Kinase-related screening set**

The screening set contains 6,624 compounds containing a pharmacophore that could interact with the hinge in the ATP binding site based on principles laid out in [18]. For the compounds in the screening library a set of molecular properties comprising molecular weight (MW), hydrogen bond acceptors and donors count (HBA, HBD), rotatable bonds count (Rb), aromatic rings count (Ar), number of heavy atoms and AlogP were calculated using BIOVIA

Pipeline Pilot. To evaluate the scaffold distribution, compounds were clustered within Pipeline Pilot using their Bemis-Murcko assembly representation [19]. The clustering was carried out using FCFP4 as molecular descriptors and Tanimoto similarity 0.4 as maximum distance from cluster seed.

### TbKKT19 modelling

**Homology model.** The *TbKKT19* sequence was downloaded from the NCBI database (accession code: XP\_829304.1—see S1) and was used to query the PDB using “BLAST” as implemented in the NCBI blastp suite (<https://blast.ncbi.nlm.nih.gov/>) to identify suitable template structures to build the *TbKKT19* model. The structure of the human cdc2-like kinase 1 (*hCLK1*) was selected as template. An alignment between the target sequence *TbKKT19* and the sequence extracted from the *hCLK1* structure was generated using Schrödinger Prime and manually curated. The optimised sequence alignment was used to build the homology model using the knowledge-based method in Prime. Model refinement was performed by Molecular Dynamics using Schrödinger Desmond.

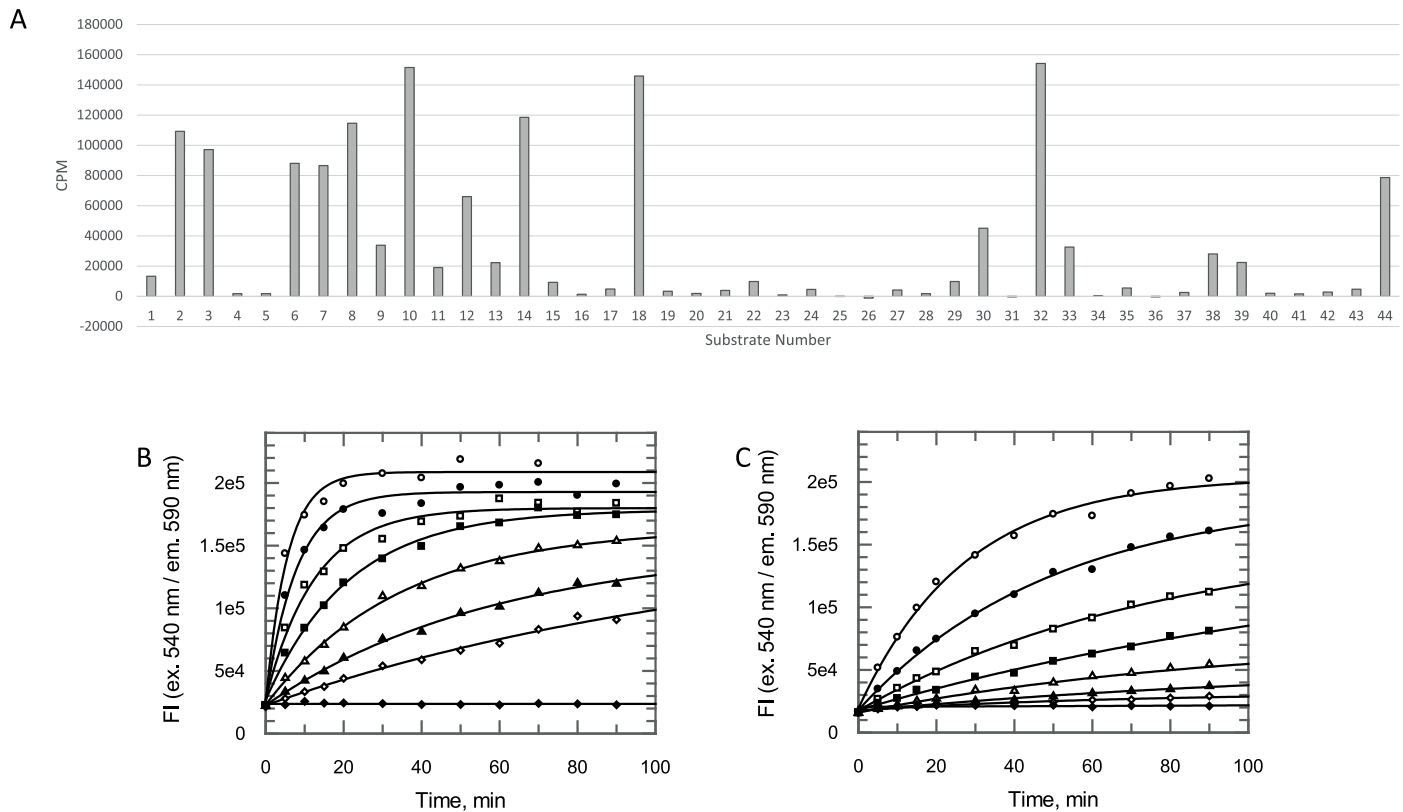
**Hit compounds docking.** The three-dimensional structures of the hits were built in Schrödinger Maestro, minimized with the OPLS3 force field [20] and docked in the catalytic site of the *TbKKT19* model using Schrödinger GLIDE-SP.

## Results

### Enzymatic characterisation of *TbKKT19*

Recombinant, His-tagged, full-length *TbKKT19* protein was produced (S1 Fig) and a panel of putative kinase substrates tested to identify a suitable peptide phosphorylated by *TbKKT19* (Fig 1A and S1 Table). On the basis of its availability in the laboratory, KKLRRTLSVA (‘KKL’ peptide) was selected for further investigation. Using an assay that measures ADP production, ‘KKL’ peptide was confirmed to be phosphorylated by *TbKKT19* (Fig 1B) and was used as a peptide substrate in biochemical *TbKKT19* assays. In addition, it was also noted that ADP formation could be detected in the absence of the ‘KKL’ peptide (Fig 1C), which may indicate autophosphorylation of the enzyme. Alternatively *TbKKT19* may display intrinsic ATPase activity in the absence of peptide substrate, a feature previously observed in other CMGC kinases [21, 22]. For future assay development a concentration of 5 nM *TbKKT19* was selected, as ADP formation in the absence of ‘KKL’ peptide could not be detected using this enzyme concentration, whereas phosphorylation of the ‘KKL’ peptide gave a robust and measurable assay signal.

Using the ‘KKL’ peptide as a substrate, the *TbKKT19* enzyme was characterised and assay conditions optimised to ensure appropriate inhibitor screening conditions were used. This involved determining the optimal enzyme concentration, assessing assay linearity with respect to time, and determining apparent Michaelis constants ( $K_m^{\text{app}}$ ) for the substrates (Fig 2). The ATP and ‘KKL’ peptide  $K_m^{\text{app}}$  values were determined to be 102  $\mu\text{M}$  (95% CI 87–117  $\mu\text{M}$ ) and 179  $\mu\text{M}$  (95% CI 150–208  $\mu\text{M}$ ) respectively. Pharmacological validation of the assay was achieved using both the pan-kinase inhibitor staurosporine and the *TbKKT10* (*TbCLK1*) inhibitor hypothemycin [13]. Staurosporine and hypothemycin inhibited *TbKKT19* with  $\text{IC}_{50}$  values of 288 nM (95% CI 190–386 nM) and 65 nM (95% CI 50–80 nM) respectively (Fig 3, closed circles). In addition, when the ATP concentration was increased 10-fold from 80  $\mu\text{M}$  to 800  $\mu\text{M}$ , 5 to 6-fold shifts in these  $\text{IC}_{50}$  values were observed (1.68  $\mu\text{M}$  (95% CI 0.99–2.37  $\mu\text{M}$ ) and 328 nM (95% CI 267–389 nM) for staurosporine and hypothemycin respectively), indicative of ATP competitive binding (Fig 3, open circles).

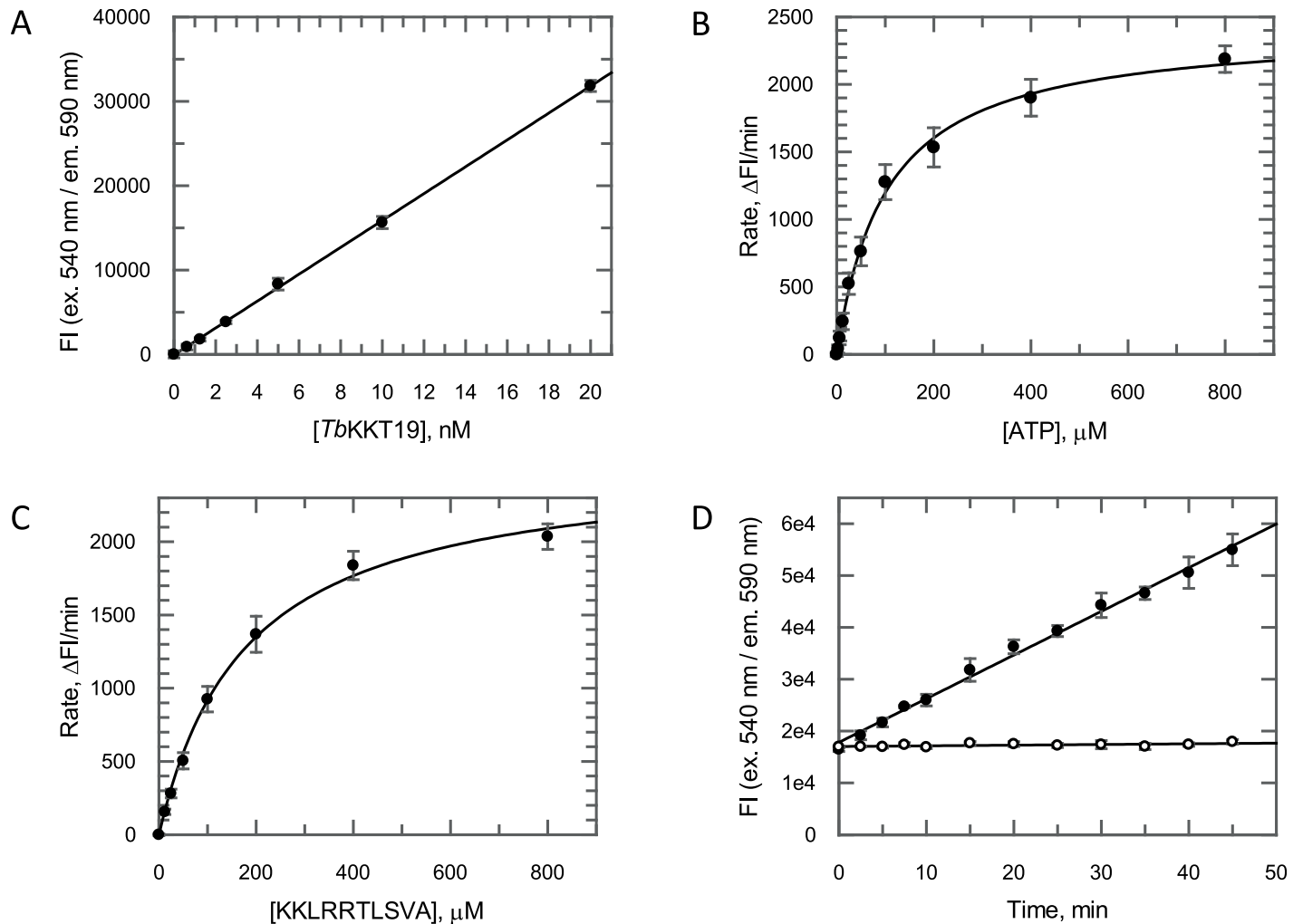


**Fig 1. *TbKKT19* substrate search.** (A) A panel of 41 peptides and 3 proteins (see [S1 Table](#) for details) were screened as possible substrates of *TbKKT19* in a radiometric kinase assay performed by the MRC PPU International Centre for Kinase Profiling at the University of Dundee (data shown as mean counts per minute above background (CPM) of two technical replicates). (B) Confirmation of peptide substrate activity using ADP Hunter Plus assay kit (DiscoverX). The activity of various *TbKKT19* concentrations monitored over time using 200 μM KKLRRRTLSVA (peptide 8 in panel A) and 200 μM ATP (data points are  $n = 1$ ). (C) Activity of various *TbKKT19* concentrations monitored over time in the presence of 200 μM ATP revealed ADP formation can be detected over time in the absence of 'KKL' peptide substrate (data points are  $n = 1$ ). In panels (B) and (C) *TbKKT19* concentrations are 175 nM (open circles), 88 nM (closed circles), 44 nM (open squares), 22 nM (closed squares), 11 nM (open triangles), 5.5 nM (closed triangles), 2.7 nM (open diamonds) and 0 nM (closed diamonds).

<https://doi.org/10.1371/journal.pone.0217828.g001>

### *TbKKT19* hit discovery

For high-throughput screening a non-biased approach was taken, with screening concentrations of 80 μM ATP and 100 μM 'KKL' peptide selected (*i.e.* substrate concentrations slightly below their respective  $K_m^{app}$ , allowing inhibitors with any inhibition modality to be identified). The screen was carried out with a kinase-relevant library of 6,624 compounds, characterised by the presence of a pharmacophore that could interact with the hinge motif in the ATP binding site. All compounds are Lipinski rule-of-five compliant [23] and [Table 1](#) shows the mean values of the calculated molecular parameters; the mean values are well below the Lipinski values, allowing for further development. The library is chemically diverse and includes 957 clusters and 969 singletons. The largest cluster contains 191 analogues whilst on average there are 6 molecules in each cluster. Examples of the most populated scaffolds are reported in [S2 Fig](#). The library was screened in single replicate at a concentration of 30 μM. To assess the robustness and reproducibility of the assay and the quality of the hit discovery campaign, various criteria were assessed following completion of the primary screen. These data reveal a high quality screening campaign, with a mean  $Z'$  [24] of  $0.88 \pm 0.03$  and a mean signal to background ratio of  $2.69 \pm 0.12$ .

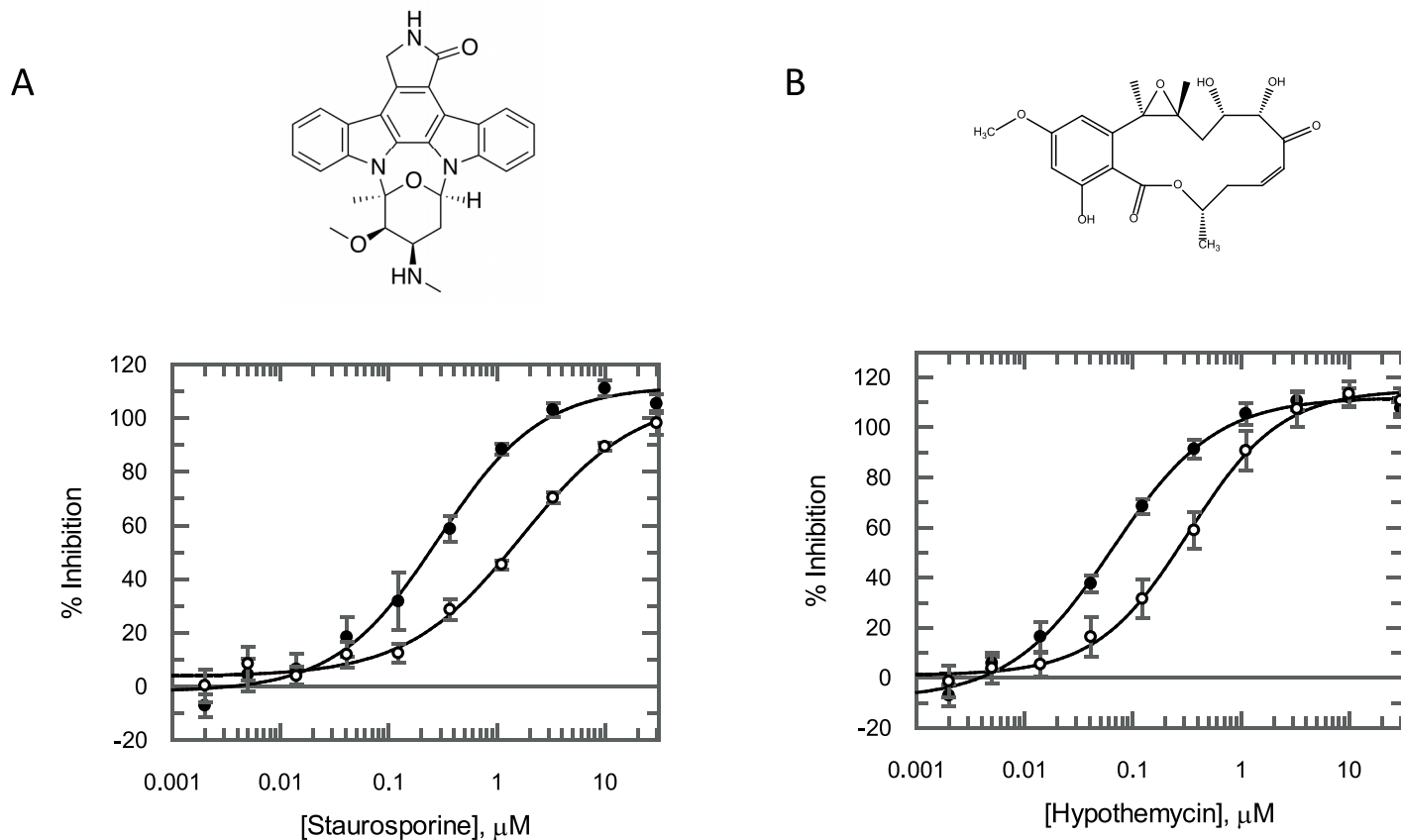


**Fig 2. *TbKKT19* assay development summary.** (A) Linearity of the *TbKKT19* assay with respect to enzyme concentration. Data are shown as mean FI  $\pm$  SD ( $n = 3$  technical replicates). (B) ATP  $K_m^{app}$  determination in the presence of a saturating concentration of 1 mM 'KKL' peptide. Data are shown as rate  $\pm$  SE of linear fit (rates determined from  $n = 2$  technical replicates). (C) 'KKL' peptide  $K_m^{app}$  determination in the presence of a saturating concentration of 600  $\mu$ M ATP. Data are shown as rate  $\pm$  SE of linear fit (rates determined from  $n = 2$  technical replicates). (D) Assay linearity with respect to time under the final assay screening conditions of 5 nM *TbKKT19*, 100  $\mu$ M 'KKL' peptide and 80  $\mu$ M ATP either with (open circles) or without (closed circles) 30  $\mu$ M staurosporine. Data are shown as mean FI  $\pm$  SD ( $n = 3$  technical replicates).

<https://doi.org/10.1371/journal.pone.0217828.g002>

Hit compounds were identified by applying an arbitrary threshold cut-off of 35% inhibition with 31 compounds from the kinase library meeting this criteria (0.46% hit rate) (Fig 4A). Hit compounds were cherry picked and retested as 10-point dose response curves in the *TbKKT19* assay to determine  $IC_{50}$  values (potencies in the range of 9.8–100  $\mu$ M determined) (Fig 4B). In addition, hit compounds were also tested in a counterscreen assay to exclude compounds that inhibit the ADP Hunter assay technology. Following this assessment only 8 confirmed kinase library hits remained (Fig 4C), corresponding to a final hit rate of 0.12%.

Two of the most potent hits (compound 1 and compound 2; Fig 5) returned  $IC_{50}$  values of 18.9  $\mu$ M (95% CI 16.5–21.6  $\mu$ M) and 14.9  $\mu$ M (95% CI 14.5–15.4  $\mu$ M) respectively. Both compounds were repurchased and reconfirmed as hits in the *TbKKT19* biochemical assay, however, due to their weak potency against the *TbKKT19* enzyme, no meaningful *T. brucei* cell-based data could be generated for these compounds.



**Fig 3. Staurosporine and hypothemycin inhibit *Tb*KKT19.** (A) Staurosporine inhibits *Tb*KKT19 with an  $IC_{50}$  of 288 nM (95% CI of best fit 190–386 nM) in the presence of 80  $\mu$ M ATP (closed circles) and 1.68  $\mu$ M (95% CI of best fit 0.99–2.37  $\mu$ M) in the presence of 800  $\mu$ M ATP (open circles). (B) Hypothemycin inhibits the *Tb*KKT19 enzyme with an  $IC_{50}$  of 65 nM (95% CI of best fit 50–80 nM) in the presence of 80  $\mu$ M ATP (closed circles) and 328 nM (95% CI of best fit 267–389 nM) in the presence of 800  $\mu$ M ATP (open circles). The 5.8-fold and 5.0-fold shifts in potency for staurosporine and hypothemycin respectively with increasing ATP concentrations are consistent with ATP competitive inhibition. All data points presented as mean % inhibition  $\pm$  SD ( $n = 3$  technical replicates).

<https://doi.org/10.1371/journal.pone.0217828.g003>

### *Tb*KKT19 homology model

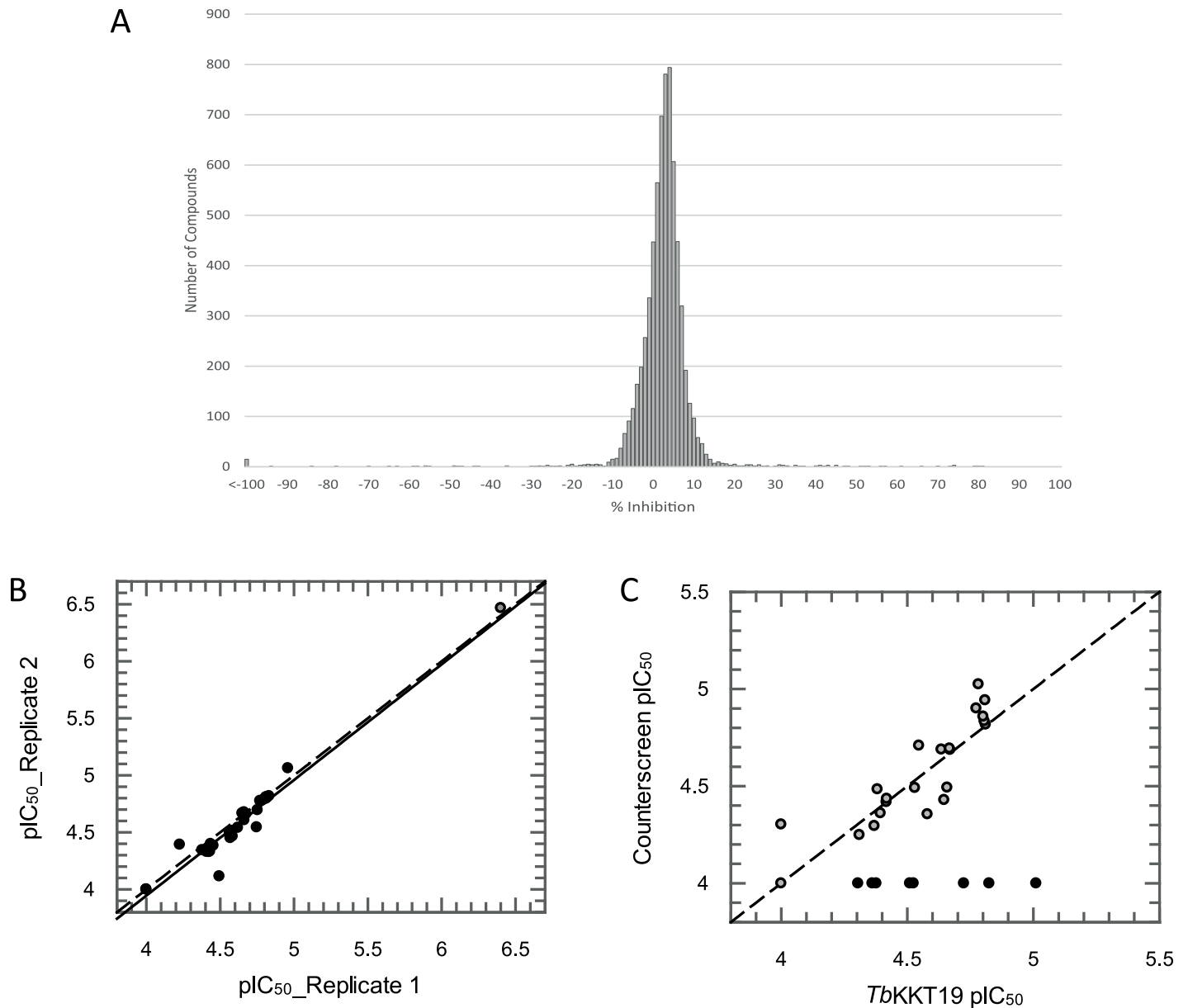
To understand the binding mode of our hit compounds and to try to account for the low hit rate observed for this kinetochore kinase, a *Tb*KKT19 homology model was generated. The *Tb*KKT19 sequence used to identify suitable 3D template structures using NCBI BLAST is shown in S3 Fig. Several human cdc2-like kinase (CLK) structures (e.g. *h*CLK3—PDB codes 2eu9; 2exe, 2wu6, 2wu7h—sequence identity 34%; *h*CLK1—PDB code: 1z57—sequence

**Table 1. Calculated properties for kinase-relevant compound library.**

Molecular property	Mean value
Molecular weight	320.35
HBA count	4.1
HBD count	1.2
Heavy atoms count	23
Rotatable bonds count	3.9
Number of aromatic rings	2.7
AlogP	2.67

<https://doi.org/10.1371/journal.pone.0217828.t001>

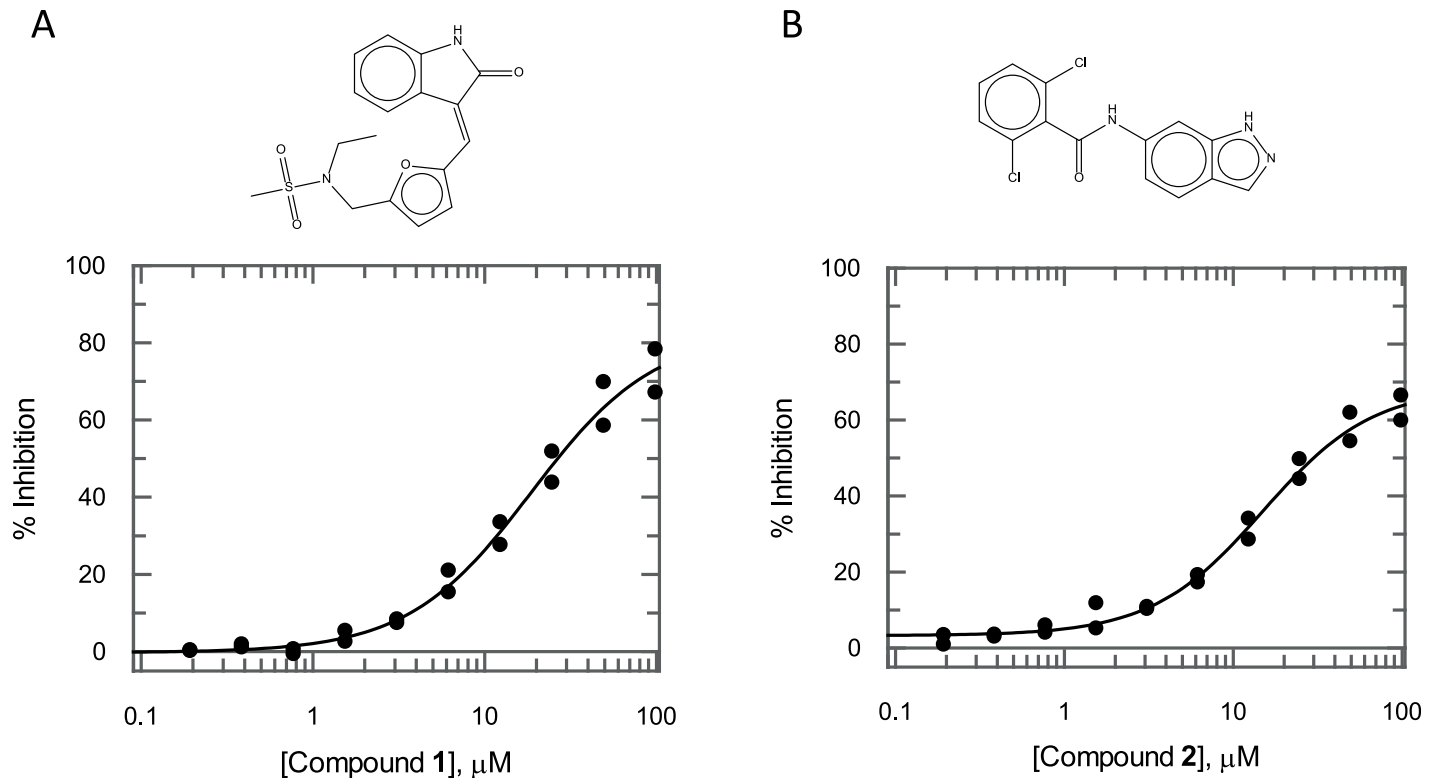




**Fig 4. *TbKKT19* hit discovery summary.** (A) Frequency distribution of compound percentage inhibition values. An arbitrary hit identification cut-off of 35% inhibition identified 31 putative hit molecules out of 6,624 molecules screened (0.46% hit rate). (B) Correlation between two independent  $pIC_{50}$  ( $-\log IC_{50}$  (M)) determinations for each of the 31 hit compounds. Black circles represent test compounds, grey circle is staurosporine control. Dashed line is line of equipotence; solid line is linear regression of these data with a correlation coefficient of 0.93. (C) Correlation between mean *TbKKT19*  $pIC_{50}$  from two independent determinations and the  $pIC_{50}$  measured in an ADP Hunter technology interference counterscreen. Black circles represent true *TbKKT19* hits (inactive in counterscreen assay), grey circles represent compounds displaying technology interference. Dashed line is line of equipotence.

<https://doi.org/10.1371/journal.pone.0217828.g004>

identity 34%; *hCLK2*—PDB code 3nr9—sequence identity 34%) were identified as suitable template structures for homology modelling. After inspection of the highly homologous structures only structures containing small molecule ligands were considered. Apo structures (e.g. PDB code 2eu9), structure with missing N lobe (PDB code 2exe) or structures with compounds not matching the hinge pharmacophore [25] (e.g. PDB code 2vag) were excluded.



**Fig 5. Compound 1 and compound 2 inhibit *TbKKT19*.** (A) Compound 1 and (B) compound 2 inhibit *TbKKT19* with  $IC_{50}$  values of 18.9  $\mu$ M (95% CI of best fit 16.5–21.6  $\mu$ M) and 14.9  $\mu$ M (95% CI of best fit 14.5–15.4  $\mu$ M) respectively when screened using final assay conditions of 5 nM *TbKKT19*, 100  $\mu$ M 'KKL' peptide and 80  $\mu$ M ATP. All data points shown ( $n = 2$  biological replicates).

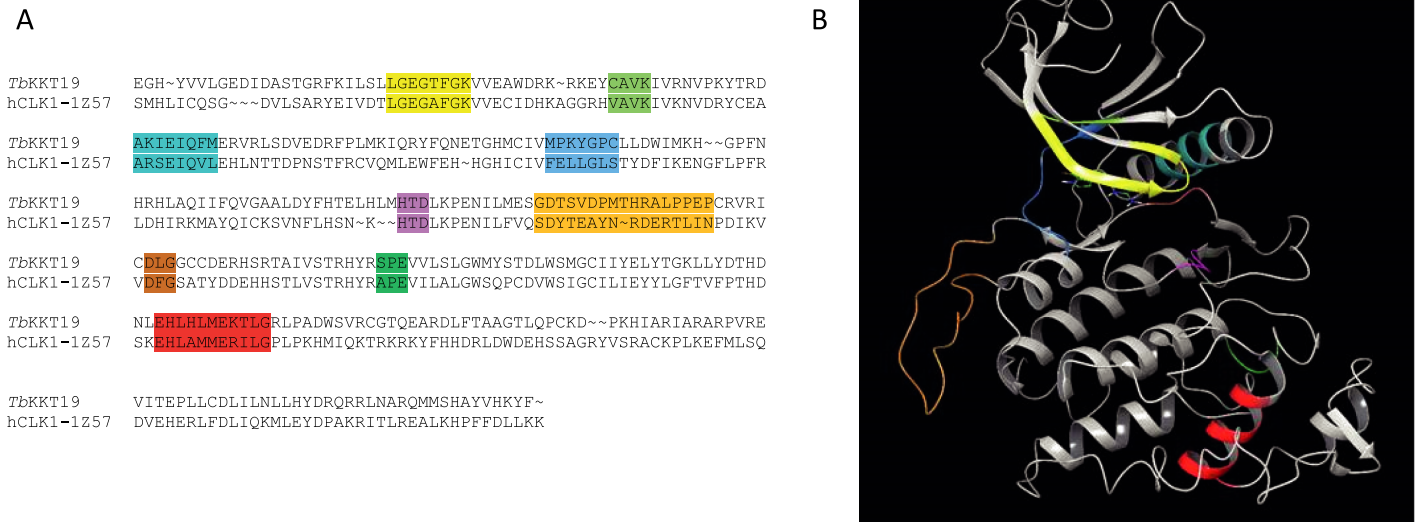
<https://doi.org/10.1371/journal.pone.0217828.g005>

Finally, the structure of the human *cdc2*-like kinase 1 (*hCLK1*) complexed with 10Z-Hymenialdisine [PDB code: 1z57] [26] was selected as a template to build the *TbKKT19* model.

An alignment between the target sequence *TbKKT19* and the sequence extracted from the *hCLK1* structure (Fig 6A) was generated. Despite a relatively low level of sequence conservation, kinases are characterised by a high degree of structural similarity. The alignment was manually curated to remove gaps / insertions from secondary structural elements and to ensure that the typical kinase sequence motifs important for catalysis and regulation were correctly aligned. A homology model for *TbKKT19* was then generated based on *hCLK1* structure using the optimised sequence alignment. The homology model was further optimised by energy minimisation and further validated by molecular dynamics (MD). An MD simulation of 100 ns showed that the system is well equilibrated and the template ligand is stably bound to the ATP pocket interacting with the hinge and the conserved catalytic Lys and Asp of the DLG motif (S4 Fig).

### Compound 1 and compound 2 binding mode

Both validated hit compounds (Fig 5) contain well-characterised hinge-binder chemical scaffolds. For instance, compound 2 contains an indazole ring that is also present in the kinase inhibitor entrectinib, currently in Phase II against NTRK/ROS1/ALK driven tumours, whereas the *trans*-3-ethylideneindolin-2-one group in compound 1 is also the hinge binding moiety in sunitinib, an inhibitor of receptor tyrosine kinases, currently on the market for the treatment of renal cell carcinoma and gastrointestinal stromal tumours.



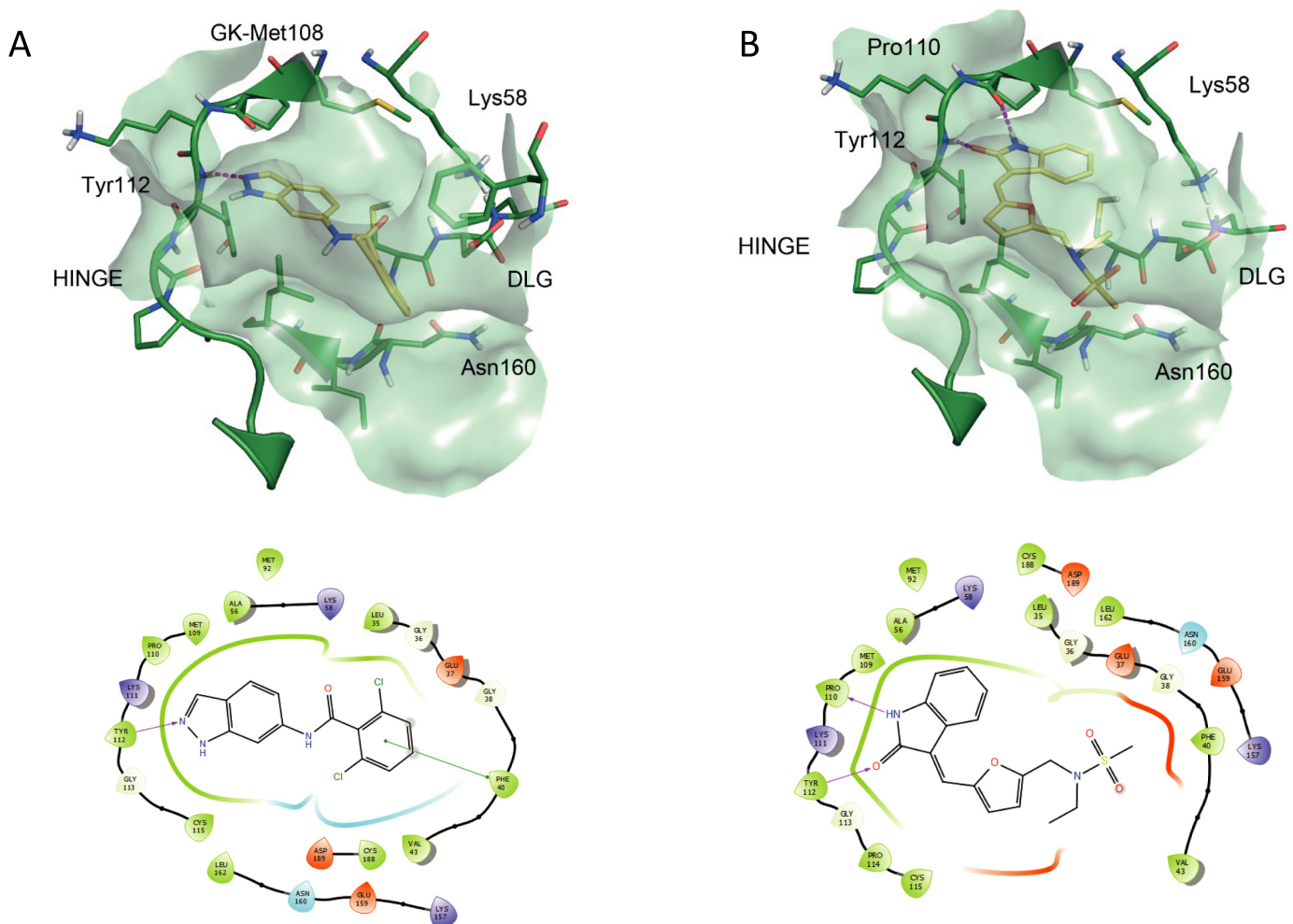
**Fig 6. Sequence alignment and *TbKKT19* homology model.** In yellow, the xGxGxxGx motif of the Gly-rich loop, in light green the xAVK motif of the highly conserved catalytic Lys, in teal the  $\alpha$ -C helix with the conserved catalytic Glu, in blue the hinge region starting with the gatekeeper residue, in magenta the HxD motif, in brown the DxG motif and in dark green the xPE motif. In orange the  $\beta$ -hairpin insertion and in red the LAMMER motif typical of the *cdc2*-like kinase family.

<https://doi.org/10.1371/journal.pone.0217828.g006>

The binding mode for both compounds in *TbKKT19* was investigated. The docking model for compound **1** (Fig 7A) is consistent with the mode of binding of sunitinib in different kinases (e.g. KIT, CDK2). The indolinone scaffold establishes two hydrogen bond interactions with the hinge motif of KKT19. One between the Tyr112 backbone NH and the carbonyl oxygen of the ligand and one between the Pro110 backbone carbonyl oxygen and the lactam NH. The phenyl ring of the indolinone moiety faces the gatekeeper residue Met 109. The double bond in compound **1** is trans whereas the equivalent bond in sunitinib is cis. That results in compound **1** developing into the ATP binding site following a different vector and placing the furan ring in the sugar subpocket. The indazole scaffold of compound **2** also occupies the adenine pocket of the ATP binding site and interacts with the Tyr112 backbone NH in the hinge (Fig 7B). The di-*o*-Cl-phenyl moiety is positioned underneath the Gly-rich loop establishing  $\pi$ -stacking with Phe40.

### Analysis of low hit rate

As the low hit-rate in the kinase-relevant compound library screen was unexpected we looked for possible explanations. An analysis of the *TbKKT19* sequence highlighted a number of unusual features (Fig 6). The beginning of the kinase activation loop is normally characterised by the DFG motif. The aspartate residue is crucial for the binding of divalent cations that coordinate the ATP phosphate groups whereas the phenylalanine of the DFG motif plays an important role in regulatory mechanisms and catalytic efficiency as it is part of the regulatory hydrophobic spine (R-spine), a highly conserved spatial motif comprising four non-consecutive hydrophobic residues [27]. The structural assembly of the R-spine is a requirement for activated kinases. Hence, formation of the R-spine is highly regulated and typically requires phosphorylation of the activation loop. Together with the catalytic hydrophobic spine (C-spine) and the gatekeeper residue, the R-spine is functionally relevant [28]. In *TbKKT19* the activation loop starts with a DLG motif where the phenylalanine residue is replaced by a leucine. We hypothesised that the change in the hydrophobic spine amino-acid composition



**Fig 7. Docking study results.** (A) Binding mode and 2D interaction diagram for compound 1 and (B) binding mode and 2D interaction diagram for compound 2. Ligand is characterised by yellow carbon atoms. Red arrows/dashed lines indicate hydrogen bonding interactions, green line indicate  $\pi$ -stacking.

<https://doi.org/10.1371/journal.pone.0217828.g007>

could have an impact on ATP  $K_m$ . Kinases characterised by a low ATP  $K_m$  value can be difficult to target as small molecule compounds will have to be very potent to be able to compete with cellular ATP. However, we determined that *Tb*KKT19 has a relatively high ATP  $K_m^{app}$  (102  $\mu$ M), indicating that it should be possible to identify compounds capable of competing for the ATP binding site. The end of the *Tb*KKT19 activation loop is characterised by an SPE motif instead of the more common APE motif. The APE motif is important for catalysis as it interacts with the  $\alpha$ F-helix stabilising the activation loop. The replacement of the alanine or proline residues of the APE motif is however not predicted to disrupt this interaction.

The *Tb*KKT19 sequence also presents a HTD motif instead of the common HRD motif. The HRD motif is a highly conserved motif in the catalytic loop, it is a key regulatory and substrate binding element as the histidine residue is also part of the hydrophobic spine and the aspartate is involved in the phospho transfer step [29]. The HRD replacement by a HTD motif is common in all members of the CLK family and does not seem to have an impact on the catalytic efficiency of the kinase [26].

The *Tb*KKT19 hinge region comprises the motif MPKYGPC. The presence of the two proline residues in position 1 and 5 relative to the gatekeeper residue Met109 (GK+1 and GK+5)

was investigated to verify the impact on the hinge pharmacophore. The presence of Pro residues can alter the recognition pattern between the hinge and the ligand as the unique Pro geometry changes the orientation of the backbone features interacting with the ligand. A typical example is PIM1 where the presence of a proline in position GK+3 removes the key hydrogen bond acceptor feature from the hinge resulting in an unconventional kinase hinge pharmacophore. A 100 ns dynamic simulation was carried out showing that the proline residues in the *TbKKT19* hinge do not alter the recognition motif at the hinge. This is also consistent with the hinge geometry observed in other kinase structures with one Pro residues either in position GK+1 (e.g. MET) or position GK+5 (e.g. CK1).

Finally, an analysis of the *TbKKT19* sequence in relation to the *hCLK1* structure used as template shows that the *TbKKT19* sequence matches a number of long insertions typical of the CLK family. In particular, residues 165–183 (Fig 6) are mapping to an extended  $\beta$ -hairpin at the top of the C-terminal lobe in the template structure. Residues 251–283 map onto a long insertion at the bottom of the C-terminal domain. In the CLK1 structure this insertion renders the  $\alpha$ G-helix that includes part of the LAMMER motif that characterises the CLK family, solvent inaccessible. The LAMMER motif seems to be only partially conserved in *TbKKT19* (the EHLAMMERILG in human is EHLHLMKTLG in *T. brucei*). These elements are important in the recognition of substrate but should not interfere with the binding of small molecules at the ATP site. Altogether, none of the *TbKKT19* specific sequence features appear to explain the unexpectedly low hit-rate.

## Discussion

KKT19 was previously identified as an unconventional kinetochore kinase in *T. brucei* [5]. Kinetic characterisation of this protein revealed it was enzymatically active as a kinase, with a substrate specificity profile similar to other reported CLK kinases (consensus sequence R-X-X-S) [26, 30, 31]. In addition, *TbKKT19* was inhibited by the pan-kinase inhibitor staurosporine as well as hypothemycin, a natural product known to inhibit kinases with a cysteine residue before the DXG motif [13], as seen for *TbKKT19*. Hypothemycin was previously shown to inhibit *TbKKT10* and the  $IC_{50}$  generated here for *TbKKT19*, in presence of 80  $\mu$ M ATP, as well as the ATP-competitive profile, are highly comparable to the data published for *TbKKT10* ( $IC_{50}$  = 150 nM when assayed using 100  $\mu$ M ATP [13]). This is not unexpected as both enzymes are closely related, and may well have redundant functions, it is therefore reassuring from a drug discovery perspective that dual inhibition of *TbKKT10* and *TbKKT19* is possible.

Despite appearing to function as a typical protein kinase and showing a pharmacological profile consistent with other kinases of the same family, a high-throughput screen of a kinase-relevant library yielded a very low hit rate. This result was unexpected as this library is composed of compounds with scaffolds designed to bind the ATP site of kinases and previous screens using a similar compound set with other kinases yielded higher hit rates (for example, the *T. brucei* GSK3 and *Leishmania* CRK3 kinases returned hit rates of 12.8% and 2.2% respectively) [32, 33]. The small number of *TbKKT19* hit molecules that we did identify contained known kinase hinge-binding motifs.

In the absence of a crystal structure, and to investigate the unexpectedly low hit rate further, a homology model of *TbKKT19* was generated, and docking of two key hits into the model was consistent with other known hinge-binding compounds. *TbKKT19* has a series of unusual amino acid sequences in conserved kinase motifs, but based on the homology model none of these should affect compound binding in the ATP-binding pocket. In conclusion, homology modelling of this kinetoplastid kinetochore kinase could not provide any explanation for the

low hit rates observed and the divergent pharmacology between *Tb*KKT19 and other kinases should be investigated further.

The few hit molecules that we identified had low potency against the enzyme. In order to develop tool molecules for target validation or potential new drug candidates for HAT a significant amount of work will be required not only to increase potency but also to generate compounds with optimal solubility, metabolic stability, oral bio-availability and CNS penetration. While the homology model will be useful to guide such efforts ideally a crystal structure is generated to facilitate this.

## Supporting information

### S1 Table. Substrate search peptide sequences and proteins.

(DOCX)

**S1 Fig. *Tb*KKT19 protein production.** SDS-PAGE gel showing production of 91.2% pure His-tagged, full-length *Tb*KKT19 protein.

(DOCX)

**S2 Fig. Scaffold representation in kinase-relevant compound library.** Compound structures exemplifying the chemical scaffolds in the 10 largest of the 957 clusters obtained from the 6,624 compounds in the kinase-relevant library.

(DOCX)

**S3 Fig. *T. brucei* KKT19 sequence.** The *Tb*KKT19 sequence is shown ([https://www.ncbi.nlm.nih.gov/protein/XP\\_829304.1](https://www.ncbi.nlm.nih.gov/protein/XP_829304.1)). In green are the residues that have been incorporated in the homology model.

(DOCX)

**S4 Fig. *Tb*KKT19 molecular dynamics simulation.** Interaction that the 10Z-Hymenialdisine compounds from the *h*CLK1 template structure establishes in the *Tb*KKT19 model. The percentage values indicate the proportion of time a specific interaction is present during an MD simulation of 100 ns.

(DOCX)

## Acknowledgments

We thank Bungo Akiyoshi for provision of the *Tb*KKT19 plasmid and helpful discussions. We also acknowledge the support of the Protein Production, Compound Management and Data Management Teams within the Wellcome Centre for Anti-Infectives Research (WCAIR).

## Author Contributions

**Conceptualization:** Leah S. Torrie, David A. Robinson, Ian H. Gilbert, Manu De Rycker.

**Data curation:** Leah S. Torrie.

**Formal analysis:** Leah S. Torrie, Fabio Zuccotto, Manu De Rycker.

**Funding acquisition:** David W. Gray, Ian H. Gilbert.

**Investigation:** Leah S. Torrie, Fabio Zuccotto.

**Methodology:** Leah S. Torrie, Fabio Zuccotto.

**Project administration:** Leah S. Torrie, David A. Robinson, Ian H. Gilbert, Manu De Rycker.

**Supervision:** David W. Gray, Ian H. Gilbert, Manu De Rycker.

**Validation:** Leah S. Torrie.

**Visualization:** Leah S. Torrie.

**Writing – original draft:** Leah S. Torrie, Fabio Zuccotto.

**Writing – review & editing:** David A. Robinson, David W. Gray, Ian H. Gilbert, Manu De Rycker.

## References

1. Monda JK, Cheeseman IM. The kinetochore-microtubule interface at a glance. *J Cell Sci.* 2018; 131(16).
2. Cheeseman IM. The kinetochore. *Cold Spring Harbor perspectives in biology.* 2014; 6(7):a015826. <https://doi.org/10.1101/cshperspect.a015826> PMID: 24984773
3. van Hooff JJ, Tromer E, van Wijk LM, Snel B, Kops GJ. Evolutionary dynamics of the kinetochore network in eukaryotes as revealed by comparative genomics. *EMBO Rep.* 2017; 18(9):1559–71. <https://doi.org/10.15252/embr.201744102> PMID: 28642229
4. Buscher P, Cecchi G, Jamonneau V, Priotto G. Human African trypanosomiasis. *Lancet.* 2017; 390(10110):2397–409. [https://doi.org/10.1016/S0140-6736\(17\)31510-6](https://doi.org/10.1016/S0140-6736(17)31510-6) PMID: 28673422
5. Akiyoshi B, Gull K. Discovery of unconventional kinetochores in kinetoplastids. *Cell.* 2014; 156(6):1247–58. <https://doi.org/10.1016/j.cell.2014.01.049> PMID: 24582333
6. Nerusheva OO, Akiyoshi B. Divergent polo box domains underpin the unique kinetoplastid kinetochore. *Open Biol.* 2016; 6(3).
7. Akiyoshi B. The unconventional kinetoplastid kinetochore: from discovery toward functional understanding. *Biochem Soc Trans.* 2016; 44(5):1201–17. <https://doi.org/10.1042/BST20160112> PMID: 27911702
8. Perez-Molina JA, Molina I. Chagas disease. *Lancet.* 2018; 391(10115):82–94. [https://doi.org/10.1016/S0140-6736\(17\)31612-4](https://doi.org/10.1016/S0140-6736(17)31612-4) PMID: 28673423
9. Burza S, Croft SL, Boelaert M. Leishmaniasis. *Lancet.* 2018; 392(10151):951–70. [https://doi.org/10.1016/S0140-6736\(18\)31204-2](https://doi.org/10.1016/S0140-6736(18)31204-2) PMID: 30126638
10. De Rycker M, Baragana B, Duce SL, Gilbert IH. Challenges and recent progress in drug discovery for tropical diseases. *Nature.* 2018; 559(7715):498–506. <https://doi.org/10.1038/s41586-018-0327-4> PMID: 30046073
11. Alves F, Bilbe G, Blesson S, Goyal V, Monnerat S, Mowbray C, et al. Recent Development of Visceral Leishmaniasis Treatments: Successes, Pitfalls, and Perspectives. *Clin Microbiol Rev.* 2018; 31(4).
12. Jones NG, Thomas EB, Brown E, Dickens NJ, Hammarton TC, Mottram JC. Regulators of *Trypanosoma brucei* cell cycle progression and differentiation identified using a kinome-wide RNAi screen. *PLoS pathogens.* 2014; 10(1):e1003886. <https://doi.org/10.1371/journal.ppat.1003886> PMID: 24453978
13. Nishino M, Choy JW, Gushwa NN, Osés-Prieto JA, Koupparis K, Burlingame AL, et al. Hypothemycin, a fungal natural product, identifies therapeutic targets in *Trypanosoma brucei* [corrected]. *Elife.* 2013; 2:e00712. PMID: 23853713
14. Alford S, Turner DJ, Obado SO, Sanchez-Flores A, Glover L, Berriman M, et al. High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res.* 2011; 21(6):915–24. <https://doi.org/10.1101/gr.115089.110> PMID: 21363968
15. Glover L, Alford S, Baker N, Turner DJ, Sanchez-Flores A, Hutchinson S, et al. Genome-scale RNAi screens for high-throughput phenotyping in bloodstream-form African trypanosomes. *Nat Protoc.* 2015; 10(1):106–33. <https://doi.org/10.1038/nprot.2015.005> PMID: 25502887
16. Mackey ZB, Koupparis K, Nishino M, McKerrow JH. High-throughput analysis of an RNAi library identifies novel kinase targets in *Trypanosoma brucei*. *Chem Biol Drug Des.* 2011; 78(3):454–63. PMID: 21668652
17. Parsons M, Worthey EA, Ward PN, Mottram JC. Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*. *BMC Genomics.* 2005; 6:127. <https://doi.org/10.1186/1471-2164-6-127> PMID: 16164760
18. Brenk R, Schipani A, James D, Krasowski A, Gilbert IH, Frearson J, et al. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. *ChemMedChem.* 2008; 3(3):435–44. PMID: 18064617

19. Bemis GW, Murcko MA. The properties of known drugs. 1. Molecular frameworks. *J Med Chem.* 1996; 39(15):2887–93. <https://doi.org/10.1021/jm9602928> PMID: 8709122
20. Harder E, Damm W, Maple J, Wu C, Reboul M, Xiang JY, et al. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J Chem Theory Comput.* 2016; 12(1):281–96. <https://doi.org/10.1021/acs.jctc.5b00864> PMID: 26584231
21. Chen G, Porter MD, Bristol JR, Fitzgibbon MJ, Pazhanisamy S. Kinetic mechanism of the p38-alpha MAP kinase: phosphoryl transfer to synthetic peptides. *Biochemistry.* 2000; 39(8):2079–87. PMID: 10684658
22. Prowse CN, Lew J. Mechanism of activation of ERK2 by dual phosphorylation. *J Biol Chem.* 2001; 276(1):99–103. <https://doi.org/10.1074/jbc.M008137200> PMID: 11016942
23. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews.* 2001; 46(1–3):3–26. PMID: 11259830
24. Zhang JH, Chung TD, Oldenburg KR. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J Biomol Screen.* 1999; 4(2):67–73. <https://doi.org/10.1177/108705719900400206> PMID: 10838414
25. Zuccotto F, Ardini E, Casale E, Angiolini M. Through the "gatekeeper door": exploiting the active kinase conformation. *J Med Chem.* 2010; 53(7):2681–94. <https://doi.org/10.1021/jm901443h> PMID: 20000735
26. Bullock AN, Das S, Debreczeni JE, Rellos P, Fedorov O, Niesen FH, et al. Kinase domain insertions define distinct roles of CLK kinases in SR protein phosphorylation. *Structure.* 2009; 17(3):352–62. <https://doi.org/10.1016/j.str.2008.12.023> PMID: 19278650
27. Taylor SS, Kornev AP. Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem Sci.* 2011; 36(2):65–77. <https://doi.org/10.1016/j.tibs.2010.09.006> PMID: 20971646
28. Kim J, Ahuja LG, Chao FA, Xia Y, McClendon CL, Kornev AP, et al. A dynamic hydrophobic core orchestrates allostery in protein kinases. *Science advances.* 2017; 3(4):e1600663. <https://doi.org/10.1126/sciadv.1600663> PMID: 28435869
29. Zhang L, Wang JC, Hou L, Cao PR, Wu L, Zhang QS, et al. Functional Role of Histidine in the Conserved His-x-Asp Motif in the Catalytic Core of Protein Kinases. *Sci Rep.* 2015; 5:10115. <https://doi.org/10.1038/srep10115> PMID: 25960268
30. Menegay HJ, Myers MP, Moeslein FM, Landreth GE. Biochemical characterization and localization of the dual specificity kinase CLK1. *J Cell Sci.* 2000; 113 (Pt 18):3241–53.
31. Moeslein FM, Myers MP, Landreth GE. The CLK family kinases, CLK1 and CLK2, phosphorylate and activate the tyrosine phosphatase, PTP-1B. *J Biol Chem.* 1999; 274(38):26697–704. <https://doi.org/10.1074/jbc.274.38.26697> PMID: 10480872
32. Woodland A, Grimaldi R, Luksch T, Cleghorn LA, Ojo KK, Van Voorhis WC, et al. From on-target to off-target activity: identification and optimisation of *Trypanosoma brucei* GSK3 inhibitors and their characterisation as anti-*Trypanosoma brucei* drug discovery lead molecules. *ChemMedChem.* 2013; 8(7):1127–37. PMID: 23776181
33. Cleghorn LA, Woodland A, Collie IT, Torrie LS, Norcross N, Luksch T, et al. Identification of inhibitors of the *Leishmania cdc2*-related protein kinase CRK3. *ChemMedChem.* 2011; 6(12):2214–24. PMID: 21913331