Identification of a proteasome-targeting arylsulfonamide with potential for the treatment of Chagas' disease

Lima, Marta; Tulloch, Lindsay; Corpas Lopez, Victoriano; Carvalho, Sandra; Wall, Richard; Milne, Rachel

Published in:
Antimicrobial Agents and Chemotherapy

DOI:
10.1128/AAC.01535-21

Publication date:
2021

Identification of a proteasome-targeting arylsulfonamide with potential for the treatment of Chagas' disease

Marta Lima[1], Lindsay B, Tulloch[1], Victoriano Corpas-Lopez[1], Sandra Carvalho[1], Richard J. Wall[1] Rachel Milne[1], Eva Rico[1], Stephen Patterson[1], Ian H. Gilbert, Sonia Moniz[1], Lorna MacLean[2], Leah S. Torrie[2], Carmine Morgillo[2], David Horn[1], Fabio Zuccotto[2] and Susan Wyllie[1]*

[1] Division of Biological Chemistry and Drug Discovery, Wellcome Centre for Anti-Infectives Research, School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, United Kingdom.

[2] Drug Discovery Unit, Wellcome Centre for Anti-Infectives Research, School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, United Kingdom.

Running title: Anti-chagasic compound targets the proteasome

Keywords: Chagas disease, drug discovery, proteasome, malic enzyme, drug target, mechanism of action

***Corresponding author:** Dr Susan Wyllie, Division of Biological Chemistry and Drug Discovery, Wellcome Centre for Anti-Infectives Research, School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, United Kingdom.

Tel: (44)1382 38 5761; Email: s.wyllie@dundee.ac.uk

Abstract

Phenotypic screening identified an arylsulfonamide compound with activity against *Trypanosoma cruzi*, the causative agent of Chagas' disease. Comprehensive mode of action studies revealed that this compound primarily targets the *T. cruzi* proteasome, binding at the interface between β4 and β5 subunits that catalyse chymotrypsin-like activity. A mutation in the β5 subunit of the proteasome was associated with resistance to compound **1**, while overexpression of this mutated subunit also reduced susceptibility to compound **1**. Further genetically engineered and *in vitro* selected clones resistant to proteasome inhibitors known to bind at the β4/β5 interface were cross-resistant to compound **1**. Ubiquitinylated proteins were additionally found to accumulate in compound **1**-treated epimastigotes. Finally, thermal proteome profiling identified malic enzyme as a secondary target of compound **1**, although malic enzyme inhibition was not found to drive potency. These studies identify a novel pharmacophore capable of inhibiting the *T. cruzi* proteasome that may be exploitable for anti-chagasic drug discovery.

Introduction

The protozoan parasite *Trypanosoma cruzi* is the etiological agent of Chagas' disease, also known as American trypanosomiasis. This zoonotic disease is endemic in Latin American countries where an estimated 6–7 million individuals across 21 countries are infected. Due to migration from endemic countries, Chagas' disease is now a worldwide problem, with hundreds of thousands of infected individuals now residing in the United States and Europe. The acute stage of Chagas' disease has very mild and non-specific symptoms that occur 4–8 weeks post-infection. As a result, very few infections are diagnosed at this stage. However, in ~30% of individuals, infection manifests as a symptomatic chronic condition, although this can take many years to emerge (1). Most commonly, chronic disease is associated with cardiac dysfunction, and to a lesser extent, digestive tract pathologies. These sequelae result in the death of ~12,500 people each year (2).

To date, benznidazole (2-nitroimidazole) and nifurtimox (5-nitrofuran) are the only approved drugs available for the treatment of Chagas' disease. Prolonged treatment with these nitroimidazoles during the acute stage cures up to 70% of individuals; however, their efficacy decreases significantly in the chronic stage (3). Both therapies are associated with severe toxic side effects that can lead to the interruption or discontinuation of treatment in as many as 30% of cases (4, 5). It is clear that new, safe, effective, oral drugs that are suitable for short-course regimens are urgently required.

No new drugs have been developed for Chagas' disease for over 30 years. Recent clinical trials with posaconazole and the ravuconazole prodrug E1224, were disappointing with relapse in between 70-90% of patients (6, 7), compared with to 6-30% failure for the benznidazole-treated arm of the study. The failure of both azoles, known to act via inhibition of lanosterol C_{14}α-demethylase (CYP51), has led to a "root and branch" overhaul of the screening

cascade and drug discovery approach for Chagas (8, 9). The principal goal of this process will be to vastly improve translation from *in vitro* and *in vivo* models for Chagas' disease to the clinic.

Successful treatment of Chagas' disease is now believed to require removal of every viable parasite within the infected patient. To complicate matters further, transiently dormant or persister forms of *T. cruzi* that are refractory to drugs acting via certain mechanisms of action (MoA) have recently been identified (10, 11). An additional barrier to the development of new drugs is the relative lack of robustly validated targets in *T. cruzi*. This has limited target-focused screening programs and led to a reliance upon phenotypic screening to identify start points for drug discovery. Phenotypic approaches have proven effective, however, a lack of information regarding the MoA or specific molecular target(s) of active compounds can hinder their downstream optimization in order to overcome pharmacokinetic and/or toxicity issues and to derive selectivity compared to human homologues. Furthermore, a comprehensive understanding of MoA can facilitate the deprioritization of compounds with unattractive or failed targets, such as CYP51, or those unable to clear all parasites and/or kill persister forms, as well as allowing the triaging of compounds targeting the same, promising targets.

Here, we use a range of genetic and chemical proteomic approaches to determine the MoA of an arylsulfonamide compound demonstrating promising *in vitro* activity against *T. cruzi*. Our comprehensive studies reveal that this compound principally targets the *T. cruzi* proteasome, binding at the interface between the β4 and β5 subunits that catalyse chymotrypsin-like peptidase activity. Using thermal proteome profiling, we also confirm that this compound interacts with a secondary target, malic enzyme, albeit this interaction does not appear to drive potency. The implications of developing compounds with this MoA as anti-chagasic drugs in the future is discussed.

Methods

Compounds

Compound **1** was kindly provided by GlaxoSmithKline. The malic enzyme inhibitor ATR-073 was purchased from MolPort. DDD01012248 (12) and GNF6702 (13) were kindly provided by the Drug Discovery Unit, University of Dundee. Bortezomib was purchased from Sigma-Aldrich. The structures of all compounds used in this study are shown in Figure **1**.

Cell lines and culture conditions

The clonal *Leishmania donovani* cell line LdBOB (derived from MHOM/SD/62/1S-CL2D) was grown as either promastigotes or axenic amastigotes in media-specific for each developmental stage, as previously described (14). *T. cruzi* epimastigotes from the Silvio strain (MHOM/BR/78/Silvio; clone X10/7A, (15)) were grown at 28°C in RTH/FBS [RPMI 1640 medium supplemented with trypticase, haemin, Hepes and 10% heat-inactivated FBS (Fisher Scientific) (16). Bloodstream form *T. brucei* "single marker" S427 (T7RPOL TETR NEO) were cultured in the presence of G418 (15 µg mL^{-1}) at 37 °C in HMI9-T media in the presence of 5% CO_2. Hep G2 cells (ECACC 85011430) were obtained from European Collection of Authenticated Cell Culture (ECACC). Cells were maintained in full growth medium (MEM with Glutamax (ThermoFisher), supplemented with 1% MEM non-essential amino acids (Sigma) and 10% heat-inactivated FBS and cultured at 37°C in the presence of 5% CO_2. Cells were passaged twice weekly by detaching adherent cells with 0.05% trypsin/EDTA (Sigma) and diluted into fresh media. Cells were never grown beyond 80% confluency. Vero cells (African green monkey kidney cells, ECCAC 84113001) were maintained at 37°C, 5% CO_2 in Dulbecco's modified Eagles's medium (DMEM, Lonza) supplemented with 10% FCS and sub-cultured every 2 days. *T. cruzi* metacyclic trypomastigotes were obtained from late-log epimastigotes (initial inoculum

10^6. mL^{-1}) cultures after ~7 -10 days at 28°C in RTH/FBS. Trypomastigote-rich cultures were incubated with Vero cells overnight at 37°C 5% CO_2 in DMEM/10% FCS. The following day, extracellular parasites were removed by washing the Vero cell monolayer three times. Infected monolayers were maintained at 37°C 5% CO_2 and DMEM/10% FCS replaced every 48 h until trypomastigotes re-emerged from Vero cells (9).

Drug sensitivity assays

To examine the effects of test compounds on the growth of *T. cruzi* epimastigotes, mid-log parasites were seeded into 96 well plates at a cell density of 5×10^5 cells·mL^{-1}. Cells were exposed to test compounds over a range of concentrations (two-fold serial dilutions). Cells were incubated for 4 days, after which 20 µL 2.5 mM resazurin was added to each well, before measuring fluorescence (excitation of 528 nm and emission of 590 nm), after a further 24 h incubation. Data were processed using GRAFIT (Erithacus software) and fitted to a 2-parameter equation, where the data are corrected for background fluorescence, to obtain the effective concentration inhibiting growth by 50% (EC_{50}):

$$y = \frac{100}{1 + \left(\frac{[I]}{EC_{50}} \right)^m}$$

In this equation, [I] represents inhibitor concentration and m is the slope factor. Experiments were repeated at least two times, and the data is presented as the mean plus standard deviation.

L. donovani promastigote (17), axenic amastigote (18) and *T. brucei* bloodstream form (19) drug sensitivity assays were carried out as previously described. HepG2 monolayers were washed twice in PBS, then detached with 0.05% trypsin/EDTA, and diluted with fully

supplemented MEM growth medium. Cells were pelleted at 80 ×*g* for 5 min and resuspended in fully-supplemented growth medium. Cells were seeded into 96-well plates (5 × 10^4/well) and allowed to adhere prior to exposure to test compounds. Plates were incubated at 37°C in the presence of 5% CO_2 for 72 h. Following incubation, resazurin (20 µL of a 2.5 mg/mL^{-1} stock solution) was added to each well, incubated for a further two hours prior to fluorescence being read as described.

Drug sensitivity assays against *T. cruzi*-infected Vero cells were carried out as previously described (20), however in this instance assays were carried out in 96-well plates. Data were processed using GRAFIT (Erithacus software) and fitted to a 2-parameter equation, as described above.

Cosmid library screening

The construction of our cosmid-based genome-wide overexpression library in *L. donovani* and strategy used to screen the library have been described in detail previously (18). For compound **2** (Figure **1**), the library was selected for 2 days at 8 nM, 2 days at 16 nM and a further 12 days at 30 nM prior to harvesting and analysis. For compound **1** (Figure **1**), the library was selected for 7 days at 300 nM and for a further 14 days at 600 nM. The associated data sets have been deposited with the European Nucleotide Archive under the following accession number: **PRJEB39157**.

Resistance generation

Compound-resistant cell lines were generated by subculturing a clone of wild-type *T. cruzi* epimastigotes in the continuous presence of compound **1**. Starting at sublethal concentrations,

drug concentrations in 5 independent cultures were increased in a stepwise manner. When parasites were able to survive and grow in concentrations of compound **1** equivalent to >20x the established EC_{50} value, the resulting cell lines were cloned by limiting dilution in the presence of compound. Five clones (RES 1-5) were selected for further biological study.

Whole genome sequencing analysis

Genomic DNA was isolated from WT and resistant clones using a standard alkaline lysis protocol. DNA was sequencing on an Illumina 4000 machine by the Beijing Genomics Institute (BGI). Sequences reads were aligned to the *T. cruzi* Sylvio x10 (v39) or *T. cruzi* Dm28c 2018 genome sequence (v46; tritrypdb.org) alongside the maxi-circle sequence (FJ203996.1, NCBI). Reads were aligned using Bowtie2 using the settings "--very-sensitive" and Samtools software. Single nucleotide polymorphisms (SNPs) and indels were called using Samtools (mpileup) and BCFtools (21) where overall quality score (QUAL) was >100 when compared with the wild type starter clone. Chromosome and gene copy number variation (CNV) analysis, as well as visualisations, was performed using Artemis. Median read counts of the wild type and resistant clones were used to normalise copy number. The associated data sets have been deposited with the European Nucleotide Archive under the following accession number: **PRJEB39157**.

Lysate production for thermal proteome profiling (TPP)

T. cruzi (X10/7 strain) mid-log epimastigotes (~1×10^{10}) were harvested by centrifugation (1912 ×*g*, 15 min, 4ºC), washed with ice-cold PBS (1912 ×*g*, 5 min, 4 ºC), and finally the cell pellet was resuspended in 8 mL ice-cold lysis buffer (1 mM EDTA, 1 mM DTT, 100 µM TLCK and 1× Roche EDTA-free COmplete protease inhibitor cocktail in 50 mM potassium phosphate buffer, pH7.4). The cell suspension was submitted to 3 freeze-thaw cycles in a dry ice/ethanol bath to

biologically inactivate the parasites and then submitted to cell disruption (Constant Systems, UK) at 30 kpsi. The resulting lysate was centrifuged (100,000 ×*g*, 20 min, 4ºC), supernatant was collected, and the protein concentration was determined using the Bio-Rad Protein Assay.

TPP assays

The lysate concentration was adjusted to 2.5 mg mL^{-1} with lysis buffer and then 2 × 2 mL aliquots were incubated at room temperature for 30 min in the presence of test compound at 20 µM (equivalent to 10×EC_{50}) or vehicle (0.1% DMSO). Each 2-mL aliquot (drug and vehicle treated) was divided into 10 × 100 µL aliquots in 0.5-mL thin–walled PCR tubes and incubated at a designated temperature (33, 37, 41, 45, 49, 53, 57, 61, 65 or 69ºC) for 3 min followed by incubation at RT for 3 min before each sample was placed on ice. Each aliquot was centrifuged (100,000 ×*g*, 20 min, 4 ºC), supernatants were harvested, and the protein concentration was assessed.

TPP sample processing, analyses and data processing

All aspects of sample processing, peptide and protein identification and quantitation; and target identification were carried out as previously described (22). However, in this instance, proteins were identified by searching the MS and MS/MS data for the peptides against *T. cruzi* proteome Dm28c 2018 version 50 (https://tritrypdb.org/tritrypdb).

Generation of overexpression constructs

Malic enzyme (*TcME*, C4B63_28g106) β5 ($Tc\beta5^{WT}$, C4B63_48g131) and $\beta5^{D225N}$ ($Tc\beta5^{D225N}$) overexpression constructs were assembled by inserting synthetic versions of each gene

(GeneArt, Invitrogen) into the pTREX vector via EcoRI and XhoI sites (23). The *L. donovani* malic enzyme (LdBPK_240780.1) overexpression construct was assembled by inserting a synthetic version of the gene (GeneArt) into the pIR1SAT vector via BglII sites. All overexpression constructs were sequenced in-house to confirm their accuracy. *T. cruzi* overexpression constructs were linearised with NheI prior to transfection.

Generation of CRISPR-cas9 edited *T. cruzi* cell line

T. cruzi proteasome β4 subunit base editing was achieved by mixing a Cas9 expression plasmid (10 µg), a specific sgRNA template, a repair template (40 µg) and T7 RNA polymerase; then transfecting the mixture into *T. cruzi* epimastigotes. Briefly, pRPa^{T7Cas9} was assembled by replacing the *rDNA* promoter in pRPa^{Cas9} (24) with a T7 promoter. pRPa^{Cas9} was digested with NheI and HindIII to remove the *rDNA* promoter and replaced by the following sequence containing the T7 promoter:

<u>GCTAGC</u>**TAATACGACTCACTATAGGG**CCCTGCACGCGCCTTCGAGTTTTTTTTCCTTTTCC CCATTTTTTTCAACTTGAAGACTTCAATTACACCAAAAAGTAAAATTCAC<u>AAGCTT.</u>

Restriction sites are underlined, and T7 promoter sequence is bold. The remaining sequence corresponds to an untranslated region upstream of the *procyclin* gene that was removed with the *rDNA* promoter and needs to be reinstated for the correct processing of Cas9 mRNA. The sgRNA template was generated by annealing and end-filling the FTcProtB4g and R-uni-scaf oligonucleotides: FTcProtB4g (<u>TAATACGACTCACTATAGGG</u>**CATCAAGATCATGGACACGG***GTTTTAGAGCTAGAAATAGC AAG*), the T7 promoter is underlined, the specific gRNA target sequence is bold and the partial gRNA scaffold sequence is in italics; R-uni-scaf (GCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATT

TCTAGCTCTAAAAC), the full sgRNA scaffold sequence. These oligonucleotides (both at 2 μM) were annealed at 50°C and end-filled at 72°C for 15 sec (5 cycles) in the presence of HiFi polymerase (Roche). The repair template was FTcProtB4g: (TAGCAGCAGCAGGGCTGAATGCCTT**A**TA<u>C</u>TA<u>T</u>AT<u>T</u>AA<u>A</u>AT**G**ATGGA<u>T</u>AC<u>A</u>GA<u>A</u>GATAAGGT CACGCAGTTGGATTCCC); non-synonymous edits are bold and synonymous changes are underlined, one of which disrupts the Cas9 protospacer-adjacent motif to prevent further DNA breaks. The three DNA components were combined, ethanol precipitated, resuspended in 10 µL of dH_2O and, following addition of 5 µL of T7 RNA polymerase, electroporated into *T. cruzi* epimastigotes, as described below. Cells were allowed to recover for 24 h, then selected with GNF6702 at 1.5 µM. Resistant cells were subcloned, DNA was extracted from independent subclones, and a specific portion of the *T. cruzi* proteasome β4 gene (TCSYLVIO_007432) encompassing the edited region, was amplified using the following PCR primers: FB4-PCR (atgtcggagacaaccattgcttttc) and RB3-PCR (ccatgtagtacaagtgtggtcc). The PCR products were Sanger sequenced in-house.

Transfection of *L. donovani* and *T. cruzi* transgenic cell lines

Mid-log-epimastigotes (2×10^7 cells in total) were transfected with 5 – 10 µg of overexpression constructs using the Human T-Cell Nucleofector kit and Amaxa Nucleofector electroporator (program U-033). Following transfection, cells were allowed to recover for 16-24 h, before the appropriated drug selection (200 µg mL^{-1} G418). *L. donovani* transgenic cell lines were generated as previously described (25) and selected with nourseothricin (100 µg mL^{-1}). In all cases, cloned cell lines were generated by limiting dilution, maintained in selective medium and removed from drug selection for one passage prior to experiments.

RT-qPCR

RNA was harvested from mid-log epimastigotes (1×10^8 cells total) using the RNeasy Mini Kit (Qiagen), as per the manufacturer's instructions. The remaining DNA was degraded from samples using the RNase-Free DNase Set (Qiagen). Quantitative RT-PCR was performed with 100 ng of total RNA using the Luna Universal One-Step RT-qPCR Kit (New England Biolabs) with the following reaction conditions: 10 min at 55°C for the reverse transcription step, followed by a denaturation step of 1 min at 95°C and then by 40 cycles of 10 s at 95°C; finally extension for 30 s at 60°C. Relative quantification was established using the reference gene glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Primers (listed below), were designed using the Primer3Plus website. The levels of each transcript in the overexpression cell lines were normalised to wild-type using the $\Delta\Delta C_t$ method. Two independently transfected clones for each construct were used, and statistical significance was measured using a student's unpaired *t* test.

Label-free quantification

Relative protein abundance in WT versus overexpressing cell lines was established as previously described (18). In this instance proteins were identified by searching protein sequence database containing *L. donovani* BPK282A1 or *T. cruzi* Dm28c annotated proteins (downloaded from TriTrypDB 46, http://www.tritrypdb.org).

Proteasome assays

Proteasome assay – luminescence

The effect of inhibitors on the chymotrypsin-like activity of the *T. cruzi* proteasome was assessed via a luminescence-based assay, as previously described (26).

Proteasome assay - treatment with proteasome inhibitors and lysate preparation

T. cruzi epimastigotes in the logarithmic growth phase (3×10^6 cells mL^{-1}) were incubated for 12h with bortezomib (1.8 µM), GNF6702 (2.9 µM) (Figure **1**) or compound **1** (24 µM), equivalent to 8× the EC_{50} values of each compound. Controls were incubated in the presence of diluent (DMSO). Cells were harvested by centrifugation (1912 ×*g*, 15 min, 4 °C) and washed with ice-cold PBS (1912 ×*g*, 5 min, 4 °C), and finally, the cell pellets were resuspended in 1.5 mL of ice-cold lysis buffer (1 mM EDTA, 1 mM DTT, 100 µM TLCK, and 1× Roche EDTA-free cOmplete protease inhibitor cocktail in 50 mM potassium phosphate buffer, pH 7.4). Cell suspensions were submitted to 3 freeze–thaw cycles in a dry ice/ethanol bath to biologically inactivate the parasites and then lysed using the One Shot™ Cell disruptor (Constant Systems, UK) at 30 kpsi.

Proteasome assay - sample processing and enrichment

Cell lysates were centrifuged (100,000 ×*g*, 20 min, 4°C), supernatants were collected, and the protein concentrations were determined using a standard Bio-Rad protein assay. Aliquots (1.1 mg) were reduced by incubating with 25 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) for 10 min at 37 °C and alkylated by incubating with 25 mM iodoacetamide (IAA) for 1 h at RT in the dark. Samples were then precipitated by incubating with 10% (v/v) trichloroacetic acid (TCA) for 3 h at -20 ºC then washed ×3 with ice-cold acetone. Protein pellets were resuspended in 100 mM triethylammonium bicarbonate (TEAB) and digested with 40 µg Lys-C for 6 h followed by 40 µg trypsin overnight (25:1 protein:enzyme ratio). Protein digests were

dried via evaporation, and the digestion efficiency was checked by mass spectrometry. Small aliquots of each sample (9% of the total sample) were kept for total proteome analysis, with the remainder submitted to enrichment. Ubiquitinated proteins were enriched using the PTMScan HS Ubiquitin/SUMO Remnant Motif (K-ε- GG) Kit (Cell Signalling Technologies) following the manufacturer's recommendations. This kit contains antibodies conjugated to magnetic beads that specifically recognise the remnant of ubiquitinated lysines following digestion with trypsin and/or LysC. This remnant consists of a Gly-Gly (diGly) motif bound to the ε-amine of lysine through an isopeptide bond. Lysine ubiquitination results in a mis-cleavage, as tryptic enzymes are not able cut after ubiquitinated lysines. Briefly, dried digests were resuspended in HS IAP Bind Buffer and incubated with magnetic beads conjugated to antibodies recognising the anti-K-ε- GG motif for 2h at 2°C. Then, the beads were washed with HS IAP Wash buffer and water to remove unbound peptides. Bound peptides were eluted by incubating in agitation with IAP elution buffer (0.15 % trifluoroacetic acid, TFA) for 10 min. Eluates were then dried under vacuum and labelled with TMTs using the TMT 10plex isobaric tagging kit (Thermo Scientific) as follows: 126, 127N, 127C and 128N, were used to label control, GNF6702, bortezomib and compound 1 samples, respectively, in parallel with their respective (total) proteome samples. After 1h, the reaction was quenched by the addition of 5% hydroxylamine for 15 min, and then the four samples were pooled and vacuum-dried. The two pooled samples (total proteome and enriched fraction) were desalted using the Pierce™ Peptide Desalting Spin Columns (Thermo) and the eluates vacuum-dried.

Proteasome assay - LC-MS/MS

Analysis of peptides was performed on a Q-Exactive-HF (Thermo Scientific) mass spectrometer coupled to a Dionex Ultimate 3000 RS (Thermo Scientific). LC buffers used were as follows: Buffer A (0.1% formic acid in Milli-Q water (v/v)) and Buffer B (80% acetonitrile and 0.08%

formic acid in Milli-Q water (v/v)). Aliquots of each sample (1 µL) were loaded at 5 µL/min onto a trap column (100 µm x 2 cm, PepMap nanoViper C18 column, 5 µm, 100 Å, Thermo Scientific) equilibrated in 5% Buffer B. The trap column was washed for 5 min at the same flow rate and then switched in-line with a Thermo Scientific, resolving C18 column (75 µm x 50 cm, PepMap RSLC C18 column, 2 µm, 100 Å). Peptides were eluted from the column at a constant flow rate of 300 nL/min^{-1} with a linear gradient from 5% Buffer B (for fractions 1-10, 7% for fractions 11-20) to 35% Buffer B in 130 min, and then to 98% Buffer B at 132 min. The column was then washed with 98% Buffer B for 20 min and re-equilibrated in 5% buffer B for 17 min. Q-Exactive HF was used in data-dependent mode. A scan cycle comprised MS1 scan [m/z range from 335-1800, with a maximum ion injection time of 50 ms, a resolution of 120,000 and automatic gain control (AGC) value of 3 × 106] followed by 15 sequential dependant MS2 scans (with an isolation window set to 0.7 Da, resolution at 60,000, maximum ion injection time at 200 ms and AGC 1 × 105). To ensure mass accuracy, the mass spectrometer was calibrated on the first day that the runs are performed.

Protein search and data analysis

MS data were analysed using the software MaxQuant (http:// https://maxquant.net/maxquant/, version 2.0.1.0). For the enriched fractions, Reporter ion MS2 mode was selected using N-terminus TMT10plex and Carbamidomethyl (C) as fixed modifications while oxidation (M), acetyl (Protein N-term), Lysine TMT10plex, DiGly and a set of DiGly-Lysine TMT10plex modifications were set as a variable modifications. Proteins were identified by searching the MS and MS/MS data for the peptides against *Trypanosoma cruzi* Dm28c proteome (TriTrypDB version 50, tritrypdb.org). Trypsin/P and LysC/P were selected as the digestive enzymes. For the total proteome samples, Reporter ion MS2 mode was selected using the TMT-10plex labels on N-terminus and lysine; Carbamidomethyl (C) was set as fixed modification while oxidation (M),

acetyl (Protein N-term) were set as variable modifications. Protein abundance was calculated according to the normalized reporter ion intensities, which for the enriched fractions were calculated using only DiGly-modified peptides. The FDR threshold for peptides and proteins was 0.01. Two missed tryptic cleavages were allowed in the global proteome samples while three in the enriched fractions, FTMS MS/MS mass tolerance was set to 10 ppm and ITMS MS/MS mass tolerance was 0.06 Da. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier **PXD027524**. Data were analysed using the Perseus software (https://maxquant.net/perseus/, version 1.6.15.0) and RStudio (version 1.2.5033).

Malic enzyme assays

Logarithmic *T. cruzi* epimastigotes were harvested (1,690 ×*g*, 10 min, 4°C), washed once in ice cold PBS and resuspended at 2×10^9 cells mL^{-1} in lysis buffer (10 mM phosphate buffer, pH 7.2, 10 mM EDTA, 5 mM benzamidine, 5 mM phenanthroline, 0.1 mM PMSF) containing 1 mg mL^{-1} digitonin and incubated at 28 °C for 10 min. The resulting lysate was then centrifuged (13,000 ×*g*, 10 min, 4°C) supernatant harvested and stored on ice. The protein concentration of this clarified lysate was determined via a standard Bio-Rad protein assay.

The activity of malic enzyme was assayed as previously described (27), with minor modifications. *T. cruzi* clarified lysate (30 µL) was added to a reaction mixture (500 µL final volume) containing 50 mM Tris-HCl buffer pH 7.6, 1 mM $MnCl_2$ and 0.12 mM $NADP^+$. Reactions were initiated by the addition of 5 mM L-malate. The reduction of $NADP^+$ was monitored at 340 nm using a UV-2401 spectrophotometer (Shimadzu). To monitor the effect of test compounds on the activity of malic enzyme, compounds (1 and 20 µM) were pre-incubated for 10 min with

the lysate in the reaction mixture. DMSO (0.2%) was used as control. All enzymatic activities were calculated as Δabs min^{-1} mg^{-1} protein.

Homology model

A model of the subunits β4 and β5 of the *T. cruzi* proteasome was generated using Modeller [version 9.24, https://www.salilab.org/modeller/, (28)] based on K and L chains of the *L. tarentolae* proteasome co-crystallized with compound GSK3494245 (PDB: 6QM7). *T. cruzi* β4 (C4B63_13g138) and β5 (C4B63_48g131) sequences were aligned to the *L. tarentolae* template using the alignment 2D function in Modeller (Figure **S2**). A set of 5 models were generated with the GSK3494245 compound in the binding site, and the best ranked, based on Modeller scores, was chosen for docking calculations.

Docking

Ionization states and tautomers for compound **1** were assigned using LigPrep (Schrödinger) at the default pH range (7 ± 2). Pyrazole tautomerism indicated two equally populated tautomers and both were used for subsequent docking calculations. The *T. cruzi* model was refined and optimized using Protein preparation wizard[i]: hydrogen bond networks were optimized for hydroxyls, thiols and sidechain amide groups of the protein residues (Ser, Tyr, Cys, Asn and Gln, respectively). Tautomers were evaluated for imidazole rings (His) and optimised. Protonation states of charged residues (His, Asp, Glu, Arg and Lys) were evaluated. Optimisation was carried out using OPLS3e forcefield and VSGB implicit solvation model. A grid centred on the GSK3494245 ligand was generated for Glide XP docking (Schrödinger).

Molecular electrostatic potential

Electrostatic potential (ESP) maps were generated on the ligand binding pose using DFT method in Jaguar at B3LYP-D3/6-31+G(d,p) level of theory. ESP surfaces were represented using 0.001 isovalue, and surfaces were mapped on the electrostatic potential using rainbow colour scale. Protein ESP was generated using the Electrostatic potential surface panel in Maestro by solving the Poisson-Boltzmann equations using the atomic partial charges of the protein residues (Schrödinger). The docking protocol XP was used using default settings which include a post docking minimisation step.

Results

An arylsulfonamide compound demonstrating promising activity against *T. cruzi*.

High-throughput screening of GSK's 1.8M compound library against *L. donovani*, *T. cruzi* and *T. brucei* resulted in the identification of a significant number of compounds active against these parasites (29). Among these hits, TCMDC-143194 was found to be moderately active against all three kinetoplastids (compound **1**, Figure **1**), with EC_{50} values ranging between 0.1 and 8 µM against the mammalian stages of these parasites (Table **1**). Bearing in mind the paucity of well-validated molecular targets for Chagas drug discovery and the fact that compound **1** does not act through inhibition of CYP51 [CYP51 pIC_{50} – 4.4, (29)], we proceeded with target identification studies predominantly in *T. cruzi*.

Resistance generation followed by whole genome sequencing (WGS)

Our first step towards determining the MoA of compound **1** was to select *T. cruzi* epimastigote cell lines resistant to this arylsulfonamide. Starting at 3 µM (~1 × EC_{50}), five independent clonal lines of compound-susceptible parasites were exposed to stepwise increasing levels of compound **1** for 40-60 days, until they were routinely growing at concentrations equivalent to 10-20× the established EC_{50} value (Figure **2A**). The five independently generated resistant cell lines were cloned by limiting dilution, and clones were assessed for susceptibility to compound **1**. The resulting clones were between 9 -21-fold less sensitive to compound **1** than the wild-type parental clone (Figure **2B** and Table **2**). In each case, the resistance demonstrated by each clone was stable over 20 passages in culture in the absence of compound.

Genomic DNA recovered from the five resistant clones was analysed by WGS. Sequence reads were aligned to both *T. cruzi* Dm28c or Sylvio X10 reference genomes and compared to the wild-type parental clone (data summarised in Tables **S2** and **S3**). Analysis of

single nucleotide polymorphisms (SNPs) of resistant lines RES 1-3 identified a homozygous mutation (D225N) in the β5 subunit of the proteasome (C4B63_48g131). In addition, pre-existing heterozygosity at position 54 (I/T) within β5 in the parental cell line became homozygous for Thr in all resistant lines. In line RES 5, a homozygous SNP (I27T) was also identified on the β4 subunit of the proteasome. Of the 5 compound **1**-resistant clones, only RES 4, the least resistant of the 5 lines (9-fold), bore no mutations in any subunits of the proteasome. Instead, RES 4 maintains a homozygous SNP (V253L) in an ABCG-like transporter (C4B63_63g86), a homologue of a *Leishmania* transporter with an established link to drug resistance (30, 31), as well as a heterozygous SNP (H605Q) on mitochondrial DNA polymerase I protein D. Relatively few copy number variations (CNV) were observed in resistant clones in comparison to the parental. Extra copies of the proteasome regulatory ATPase subunit 5 (C4B63_76g37) and subunit 1 (C4B63_76g43) were identified in RES 2, 3 and 4, with both genes encoded on the same contig.

Screening of compound 1 against a genome-wide overexpression library

As a parallel approach to identify the molecular target, compound **1** was screened against our genome-wide overexpression library in the closely related kinetoplastid parasite *Leishmania donovani* (22). The principle under-pinning this gain-of-function screen is that elevated levels of a drug target can result in resistance to the corresponding drug by increasing the pool of functional protein or by reducing free drug through binding. *L. donovani* promastigotes were transfected with a pooled population of cosmids containing genomic DNA fragments of between 35 and 45 kb. The final transfected library provides a >15-fold genome coverage with 99% of *Leishmania* genes represented. The library was selected with 300 nM compound **1** (equivalent to the EC_{99} in promastigotes) for 7 days and for a further 14 days at 600 nM. Following compound selection, cosmids maintained by the “resistant” parasite population were harvested

and analysed by next-generation sequencing. Mapping of overexpressed inserts to the *L. donovani* LV9 and BPK281 assembled genomes revealed that 80% of all mapped reads aligned to a single region on chromosome 35 (Figures **2C**, Tables **S4** and **S5**). This 63.3 kb region encodes 14 designated open reading frames (ORFs) in total. However, only two genes were flanked by all opposing barcodes: a CBS domain-containing protein (LdBPK.35.2.000780), also annotated as the γ subunit of an AMP-activated protein kinase, and the proteasome-activating protein PA28 (PA28; LdBPK.35.2.000770). Genome-wide overexpression library screening with the established proteasome inhibitor compound **2** [compound **7** in (12)], a close analogue of the clinical candidate for visceral leishmaniasis GSK3494245 (12, 32), confirmed that parasites "resistant" to this proteasome inhibitor were also found to overexpress PA28 (Figure **2C** and Table **S5**). Indeed, compound **2**-resistant *L. donovani* promastigotes, generated through *in vitro* selections, and subsequently found to bear a $G^{197}S$ mutation within the gene encoding the β5 proteasome subunit, demonstrate considerable (260-fold) cross-resistance to compound **1** (Table **S6**). These data, alongside our WGS analysis, strongly suggest that, like compound **2**, compound **1** is a proteasome inhibitor in both *L. donovani* and *T. cruzi*.

Target validation

The proteasome is a key component of the ubiquitin-proteasome protein degradation system and plays a crucial role in numerous cellular processes, including protein turnover and cell signalling (33). In eukaryotes, the proteasome consists of a central 20S cylindrical core flanked by two regulatory complexes (19S). The canonical 20S unit is comprised of two outer (α) and two inner (β) polypeptide rings. Three of the β-type subunits are responsible for chymotrypsin-, trypsin-, and caspase-like catalytic activities. The proteasome is a well-exploited target in drug discovery for a variety of indications, including cancer, inflammation, and a number of infectious diseases (34). A number of recent studies have illustrated the utility of the proteasome as a

viable drug target in kinetoplastids (12, 13, 26, 35). Indeed, GSK3494245 and LXE408 (12, 35), inhibitors of the chymotrypsin-like activity of the *Leishmania* proteasome are being clinically assessed for use in the treatment of visceral leishmaniasis. These studies confirm the feasibility of selectively inhibiting of the kinetoplastid proteasome and the value of this molecular target for drug discovery.

We next sought to interrogate the role of the proteasome in the MoA of compound **1**. Like compound **1**, compound **2** is active against both *T. cruzi* and *L. donovani*. Virtually all of our compound **1**-resistant clones, with the exception of RES 4 (2.5-fold resistant), demonstrated considerable cross-resistance to this established proteasome inhibitor (20-127-fold, Table **2**). In addition, the broad-spectrum anti-kinetoplastid proteasome inhibitor GNF6702 (13) demonstrated similar levels of cross-resistance against RES 1-5 clones while there was no evidence of cross-resistance to the classical proteasome inhibitor bortezomib, used in the treatment of multiple myeloma, mantle cell lymphoma and a number of other cancers (36) (Table **2**). GNF6702, GSK3494245 and analogues are known to target the same allosteric binding site at the interface of the β4/β5 subunits of the proteasome resulting in the inhibition of chymotrypsin-like activity. Collectively, our data suggests that compound **1** targets this same allosteric binding site rather than the bortezomib binding pocket in the active site of β5.

The impact of the β5 $D^{225}N$ mutation, identified in the majority of our resistant clones, was examined further. *T. cruzi* epimastigotes overexpressing the mutated version of the β5 subunit were generated, with elevated levels of this mutated protein confirmed by quantitative RT-PCR (Figure **S1**). Epimastigotes overexpressing β5 $D^{225}N$ were found to be 8-fold less sensitive to compound **1** than wild-type parasites. These transgenic parasites also demonstrated considerable resistance to both compound **2** and GNF6702 (Table **2**). Overexpression of $\beta5^{WT}$ in the RES 1 clone, which bears the $D^{225}N$ mutation, partially reverted the resistance phenotype of this cell line to all three compounds. CRISPR-Cas9 was used to

engineer specific mutations in the β4 subunit ($F^{24}L$ and $I^{29}M$) previously shown to confer resistance to GNF6702 (13). CRISPR-edited epimastigotes were refractory to compound **1**, compound **2** and GNF6702 at all the concentrations tested, once again linking the mechanism(s) of action/resistance of compound **1** with that of established proteasome inhibitors. Furthermore, these data strengthen our hypothesis that compound **1** exploits the same allosteric binding site at the interface of the β4/β5 subunits of the proteasome.

The *in vitro*-selected, CRISPR-edited and overexpressing *T. cruzi* cell lines were also assessed against compound **1,** GNF6702 and fexinidazole (control compound) as amastigotes within Vero cells (Table **S7**). The response to compound treatment and resistance/cross-resistance profiles of these intracellular amastigotes closely mimicked that seen with their respective epimastigote cell lines. These data are consistent with compound **1** inhibiting the function of the proteasome in the clinically-relevant, mammalian stage of *T. cruzi*.

Inhibition of proteasome activity

Initially, the impact of compound **1** on the chymotrypsin-like proteolytic activity of the proteasome was assessed using a commercially available indirect enzyme-based luminescent assay. In this assay, activity is monitored in proteasome-enriched *T. cruzi* epimastigote lysates using Suc-LLVY-aminoluciferin as a substrate (12). Unfortunately, the data produced by this assay was unreliable with regard to compound **1,** with IC_{50} values ranging from 0.9 to >15 μM reported. Based on these data and the statistics associated with the assay, it became clear that compound **1** directly interferes with the assay. Thus, an alternative route to determine the impact of this arylsulfonamide compound on *T. cruzi* proteasome function was required.

In cells where the function of the proteasome has been compromised, there is a concomitant build-up of ubiquitinylated proteins earmarked for degradation. Here,

ubiquitinylated peptides were recovered from the lysates of epimastigotes pre-treated with bortezomib, compound **1** or compound **2** over 8h at concentrations equivalent to 3× their respective EC_{50} values. Control cultures were treated for the same period of time in the presence of DMSO. Enrichment of ubiquitinylated proteins was achieved by immunoprecipitation using magnetic beads conjugated to an antibody specific for the remnant of ubiquitinated lysines following digestion with trypsin and/or LysC. Using LC-MS/MS, the ubiquitination ratio of each sample was calculated by dividing the reporter intensity of ubiquitinated proteins by the reporter intensity of total protein (Table **S8**). As expected, the accumulation of ubiquitinylated proteins was highest in epimastigotes treated with bortezomib (2.7-fold higher than in the DMSO-treated control), known to inhibit all three catalytic activities of the proteasome (Figure **3**). The build-up of ubiquitylated proteins in cells treated with GNF6702, an established inhibitor of the chymotrypsin-like activity of the proteasome, was clearly evident but more modest (1.45-fold higher than control). Similarly, ubiquitinylated proteins accumulated to levels 1.9-fold higher in compound **1**-treated parasites compared to those recovered from DMSO-treated control cells. These data are entirely consistent with our hypothesis that compound **1** acts principally as an inhibitor of the *T. cruzi* proteasome.

Identification of a secondary target

Thermal proteome profiling (TPP) can be used as an effective and unbiased approach to demonstrate compound-target engagement. It is based on the principle that binding of a drug to its protein target can significantly alter the thermal stability of that protein (37). Here, *T. cruzi* epimastigotes lysates were treated with compound **1** (10× established EC_{50} value) or DMSO vehicle. Aliquots of each lysate were then incubated at designated temperatures (33-69°C), and for each temperature, insoluble (denatured) proteins were removed. The resulting soluble protein samples were reduced, alkylated, and digested with trypsin prior to derivatization with

tandem mass tags. Pooled peptides were fractionated by HPLC and analyzed by LC/MS-MS prior to identification and quantitation. The melting points of identified proteins were then established using the TPP software package. Full melt curves were established for 6771 proteins, representing 39.4% coverage of the *T. cruzi* proteome. The top 20 proteins demonstrating thermal shift in the presence of compound **1** in two separate replicate experiments are summarised in Tables **S9** and **S10**.

While TPP has proven effective in a number of our MoA studies to date (18, 22), experience indicates that it is less effective in cases where the molecular target is part of a large multi-subunit complex such as the proteasome. Indeed, analysis of our two independently generated TPP datasets failed to identify any subunits of the proteasome as targets of compound **1** (Tables **S9** and **S10**). The only target candidate with significantly increased thermal stability in the presence of compound **1** across both datasets was the cytosolic *T. cruzi* malic enzyme (c*Tc*ME, C4B63_28g106) (38). Individual melting curves revealed that the thermal stability of c*Tc*ME increased by 8.8 °C (mean ΔT_m) in experiment 1 and 5.5°C in experiment 2 (Figure **4A,** Tables **S9 and S10**). In contrast, compound **1** had no impact on the thermal stability of the mitochondrial version of this enzyme (m*Tc*ME, Figure **4B**).

ME catalyses the oxidative decarboxylation of malate to pyruvate with the concomitant reduction of $NAD(P)^+$ to NAD(P)H. In light of our TPP studies, we next sought to determine if the binding of compound **1** to c*Tc*ME inhibits enzymatic activity. The ME-dependent production of NADPH was monitored in clarified lysates of *T. cruzi* epimastigotes at 340 nm following the addition of malate. In the first instance, this assay was validated using ATR-073, an established inhibitor of c*Tc*ME (39). In keeping with previous studies (39), pre-incubation of lysates with 20 µM ATR-073 resulted in a 96% reduction in ME activity while pre-treatment with 1 µM reduced activity by 63%. Similarly, pre-incubation of lysates with 20 µM and 1 µM of compound **1** inhibited ME activity by 96 and 68%, respectively (Figure **4C**). These data confirm compound **1**

as an inhibitor of c*Tc*ME. It should be noted that ATR-073 and compound **1** are structurally similar, both being substituted *N*-phenyl aryl sulfonamides, this common moiety constituting the pharmacophore of this chemical series.

In order to probe the role of c*Tc*ME inhibition in the MoA of compound **1**, a clonal cell line overexpressing this enzyme was generated. Elevated levels of c*Tc*ME in transgenic parasites relative to wild-type were confirmed by label-free MS quantitation (Figure **S2**). In addition, using the previously described spectrophotometric assay, ME activity in c*Tc*ME^{OE} lysates was shown to be 4.4-fold higher than in comparable wild-type lysates (Figure **4D)** confirming that overexpressed ME is enzymatically active. However, overexpression of c*Tc*ME had little or no effect on the potency of compound **1**, and this was also the case for c*Tc*ME-overexpressing amastigotes in Vero cell assays (Table **S7**) as well as c*Ld*ME-overexpressing *L. donovani* promastigotes (Table **S6**). Epimastigotes overexpressing c*Tc*ME also remained just as susceptible to ATR-073 as wild-type perhaps suggesting that ME is not the primary target of this compound (Table **S11**). It should also be noted that careful analysis of WGS data from our *T. cruzi* compound **1**-resistant clones identified no CNV or SNP associated with c*Tc*ME, nor did screening of our genome-wide overexpression identify c*Tc*ME as a hit. We next investigated the possibility that ATR-073 may actually inhibit the *T. cruzi* proteasome. Using an established luciferase-based biochemical assay (26), this compound had no effect on the chymotrypsin-like activity of the proteasome at concentrations up to and including 3.3 µM (Figure **S3A**). At concentrations above this threshold ATR-073 began to interfere directly with the assay. In addition, cell lines resistant to compound **1** and bearing mutations in the β4 and β5 subunits of the proteasome demonstrated no cross-reactivity to ATR-073 (Figure **S3B**). Collectively, these data suggest that despite the structural similarities between ATR-073 and compound **1** that they likely interact with different molecular targets with *T. cruzi*.

Docking studies

With the aim of defining the binding site of compound **1** and understanding the role of mutations in compound **1**-resistance, a homology model of the *T. cruzi* β4 and β5 proteasome subunits was generated. This model was based on the *L. tarentolae* orthologue structure (PDB:6QM7) complexed with GSK34944245 (12). The *L. tarentolae* β4/β5 proteasome subunits share 78% overall sequence identity with their counterparts in *T. cruzi* (Figure **S4**). Indeed, the sequence identity of the GSK3494245 binding site is even greater, with only one of the 26 amino acids within 5Å of the ligand differing (S^{132} of the β5 subunit in *T. cruzi* is T^{122} in *L. tarentolae*). As there are two tautomeric forms of the compound **1** pyrazole, both were docked into the model. The best docking poses of both tautomers occupy the same region occupied by GSK3494245 in the *L. tarentolae* cryoEM structure (12). The pyrazole moiety of compound **1** occupies the hydrophobic pocket formed by F^{24}, I^{27} and I^{29} residues of the β4 subunit and Y^{223}, V^{238} and Y^{246} from β5. It also establishes a hydrogen bond with the backbone nitrogen of G^{239} in β5 (or S^{242} side chain, depending on the tautomer). In this binding pose, compound **1** stacks with F^{24} from the β4 subunit, and the central phenyl ring is in close proximity to the β5 side chains of D^{225} and D^{226} while the sulphonamide points towards solvent, directing thedi-metoxy phenyl moiety into a groove defined by β4 residues Y^{25}, Y^{26} and I^{27} side chains (Figure **5**). Previous studies have shown that β5 residues D^{225} and D^{226} play an important role on the recognition of GSK3494245 by establishing long-range electrostatic interactions with a positively charged patch resulting from the unevenly distributed electrons of the ligand (32). Analysis of the electrostatic potential surface (ESP) of compound **1** suggests a similar scenario where a moderately electron-deficient area on the side of the pyrazole and central phenyl rings, possibly accentuated by the electron withdrawing effect of the sulfonyl amide, establishes a favourable electrostatic interaction with the side chains of the β5 residues Y^{223}, D^{225} and D^{226} (Figure **S5**). This is consistent with the reduction in affinity observed for compound **1** and GSK3494245 in the presence of the mutation

$D^{225}N$ in the β5 subunit, where the removal of the negatively charged side chain partially disrupts this favourable interaction. Based on the suggested mode of binding, I^{27} from β4 plays a critical role in defining the sub-pockets where the pyrazole and di-methoxy phenyl moieties of compound **1** bind, and its mutation would likely disrupt the binding of the ligand as evidenced by the resistance-conferring $I^{27}T$ mutation.

Discussion

Multiple orthogonal genetic, biochemical and proteomics approaches identify the proteasome as the primary molecular target of an arylsulfonamide compound with potential for anti-chagasic drug discovery. The proteasome has long been considered a viable, theoretical drug target in trypanosomatids (40), largely based on the enhanced susceptibility of these parasites to established proteasome inhibitors relative to mammalian cells (26, 41). Several recent studies have validated this assumption, with two proteasome inhibitors now in clinical trials for use in the treatment of visceral leishmaniasis (GSK3494245 and LXE408). Both compounds are structurally related and exploit the same binding pocket at the interface of the β4/β5 subunits of the *L. donovani* proteasome (12, 13, 35). Binding to this site inhibits the chymotrypsin- but not the trypsin- and caspase-like activities of the *Leishmania* proteasome. Inhibitors that target this pocket demonstrate impressive selective inhibition of the parasite versus the human proteasome, and this has been attributed to sequence variability in the narrow hydrophobic pocket (12) while the corresponding pocket in the human proteasome is also more open, shallow and solvent-exposed. Despite sharing no structural similarity with either LXE408, GSK3494245 or analogues, in the current study we present multiple lines of evidence that compound **1** targets the equivalent binding pocket in the *T. cruzi* proteasome (Figures **5B and C**). Thus, compound **1** represents a new scaffold capable of binding at the proteasome β4/β5 subunit interface that may represent a suitable starting point for future anti-chagasic drug discovery.

The structure of compound **1** closely resembles a series of compounds found to inhibit c*Tc*ME in a target-based high-throughput screen (39). While our data indicate that compound **1** directly binds to c*Tc*ME and inhibits enzymatic activity, this inhibition does not appear to drive potency in epimastigotes or intracellular amastigotes. However, we cannot rule out the possibility that c*Tc*ME inhibition in parasites where the function of the proteasome is also

compromised may contribute to compound **1** potency. To our knowledge, there is no direct evidence that either ME isoforms are essential in *T. cruzi*. In the related kinetoplastid *T. brucei*, RNAi studies in the insect procyclic stage of the parasite indicate that, while mME is essential in standard cell culture conditions, the cytosolic isoform is dispensable (42). Establishing whether cME is essential in *T. cruzi* will be important in understanding its precise role, if any, in the MoA of compound **1** and will also determine if this enzyme has any value for future Chagas' disease drug discovery.

Sub-populations of non-dividing intracellular amastigotes have been observed in both *in vitro* and *in vivo T. cruzi* infections. These "persister" parasites retain the ability to differentiate into infectious trypomastigote forms that can reactivate infection. Importantly, nonreplicating amastigotes are refractory to treatment with existing trypanocidal compounds, including the frontline therapy benznidazole (10). These observations have led many to associate dormancy with treatment failure and to suggest that only compounds capable of killing dormant forms should be pursued (11). Based on their reduced ability to express reporter proteins (10), it is assumed that persister amastigotes exist in a significantly reduced metabolic state concurrent with reduced protein synthesis. It is tempting to hypothesise that alongside reduced protein synthesis, the requirement to turnover proteins via the proteasome will be similarly reduced in persisters, leading to a decreased susceptibility to proteasome inhibitors. However, the proteasome inhibitor GNF6702 dosed twice-daily at $10\,\mathrm{mg\,kg^{-1}}$ matched the efficacy of benznidazole in a mouse model of infection (13). All but one of the eight treated mice had no detectable parasites in blood, colon or heart tissue, even after 4 weeks of immunosuppression. It should also be noted that a modified treatment regimen where higher doses of benznidazole were given in pulses over prolonged periods resulted in sterile parasitological cure of multiple mouse models of infection (43). This high-dose, extended-time protocol is believed to improve benznidazole efficacy by challenging the stochastic, time-limited nature of dormant parasites in

T. cruzi infections. These exemplary studies provide a template for the effective use of promising compounds that potentially have a modest impact on persister parasites. Assessing the efficacy of known proteasome inhibitors, including compound **1**, via this modified protocol should be a priority.

In conclusion, here we have identified a novel pharmacophore capable of inhibiting the *T. cruzi* proteasome. Future studies should focus on optimising the drug-like properties of this promising compound in order to assess its efficacy in mouse models of Chagas' disease.

Data availability

Genomics data sets from this study have been deposited with the European Nucleotide Archive under the following accession number: **PRJEB39157**.Proteomics data from this study have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier **PXD027524**. All additional information data is available upon request from the corresponding author (s.wyllie@dundee.ac.uk).

Acknowledgements

We would like to take this opportunity to thank the Fingerprints Proteomics Facility at the University of Dundee for their invaluable support throughout these studies. We also thank GlaxoSmithKline for provision of compound **1**. This work was supported by the following funding from the Wellcome Trust: Centre Award [203134/Z/16/Z], Strategic Award [105021] and Innovations Award [218448/Z/19/Z]. Dr Marta Lima is supported by a Newton International Fellowship [NIF\R1\181222] from The Royal Society. For the purpose of open access, the

Figure legends

Figure 1: **Structures of compounds used in this study.** These compounds include: TCMDC-143194, an *N*-aryl arylsulfonamides that originates from GSK's kinetobox library (29); the anti-cancer drug bortezomib, a peptidic boronic acid; the xxxxxxx DDD01012248, a close analogue of the Leishmania clinical candidate and proteasome inhibitor GSK3494245 (12); the imidazotriazine GNF6702, a pan-active inhibitor of the kinetoplastid proteasome developed by Novartis (13); and the triazolopyrimidine ATR-073, a proposed inhibitor of *T. cruzi* malic enzyme (cytosolic) (39).

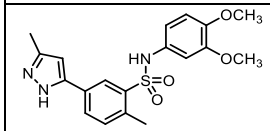
Figure 2: Target deconvolution studies with compound 1. (A) Schematic representation of the generation of compound **1**-resistant cell lines in *T. cruzi* epimastigotes. Each passage of cells in culture (circles) is indicated with cell lines 1-5 indicated in black, grey, blue, red, and green, respectively. (B) EC_{50} values for compound **1** were determined for WT (white circles) and cloned resistant cell lines I-V (black, red, blue, grey and green circles, respectively). These curves are the nonlinear fits of data using a two-parameter EC_{50} equation provided by GraFit. An EC_{50} value of 0.7 ± 0.01 µM was determined for compound **1** against WT promastigotes. EC_{50} values for resistant clones I-V were 23 ± 4, 16 ± 0.3, 13 ± 1, 11 ± 1 and 14 ± 6 µM, respectively. These EC_{50} curves and values are from one biological replicate, comprised of two technical replicates. Collated data sets reporting the weighted mean ± SD of multiple biological replicates are summarised in Table **S4**. (C) Genome-wide map indicating cosmid library hits from screening of compound **2** (upper panel) and compound **1** (lower panel). Primary "hits" on chromosome 35 indicated in green (LdBPK.35.2.000770 and LdBPK.35.2.0007800) with surrounding genes indicated in yellow. The blue/pink and black/green peaks indicate independent cosmid inserts in different orientations. Focus on barcodes flanking LdBPK.35.2.000770 and LdBPK.35.2.0007800 (lower panel inset).

Figure 3: **Relative levels of ubiquitinylated proteins in compound-treated and untreated *T. cruzi* epimastigotes**. LC-MS/MS was used to quantify ubiquitinylated proteins recovered from *T. cruzi* epimastigotes pre-treated (8 h) with bortezomib, GNF6702 or compound **1** at concentrations equivalent to 3× their respective EC_{50} values. Abundance of ubiquitinylated proteins relative to the levels in control cultures exposed to DMSO for 8 h are shown. Data is from one biological replicate.

Figure 4: Interrogation of *T. cruzi* ME as a potential target of compound 1. TPP melt curves for c*Tc*ME (A) and m*Tc*ME (B) following incubation with compound **1** (red) or vehicle (0.1% DMSO, black). Data from technical replicates (circles and squares) are shown, and the mean shift in melting temperature (ΔT_m) for c*Tc*ME in this biological replicate was 5.5°C. (C) Monitoring the effects ATR-073 and compound **1** pre-incubation on the ME-dependent production of NADPH in clarified lysates of *T. cruzi* epimastigotes following the addition of malate. Data represent enzyme activity relative to DMSO-treated control lysates and is the mean ± SD of >3 biological replicates. (D) Relative ME activity in wild-type and c*Tc*ME^{OE} cell lysates. Data represent enzyme activity relative to wild-type and is the mean ± SD of >3 biological replicates.

Figure 5: Compound 1 binding hypothesis. (A) Best scoring binding pose for compound **1** (green) in the *T. cruzi* homology model of the GSK3494245 binding site at the interface of the β4/β5 subunits of the proteasome. (B) Best scoring binding pose for compound **2**, consistent with the pose observed in the cryoEM structure of *L. tarentolae* proteasome with the close analogue GSK3494245 (PDB 6QM7). (C) 2D ligand interaction diagram based on the best-scoring docking pose for compounds **1** and **2**.

Table 1 – Potency

Compound	EC_{50} values, μM						
	T. cruzi		***L. donovani***			***T. brucei***	**Mammalian**
	Epimastigotes	**Intra-Vero**	**Promastigote**	**Axenic amastigote**	**Intra-macrophage***	**BSF**	**Hep G2**
	3 ± 0.08	1.3 ± 0.2	0.1 ± 0.005	0.7 ± 0.1	8*	0.3 ± 0.013	23 ± 3

EC_{50} values represent the weighted mean ± standard deviation of at least three biological replicates (n ≥ 3) with each biological replicate comprised of two technical replicates. * Intra-macrophage data from (29).

Table 2 - Collated EC_{50} data for WT, resistant and transgenic *T. cruzi* epimastigote cell lines

Cell line	EC_{50} values, µM (fold change versus WT)			
	Compound **1**	Compound **2**	GNF6702*	Bortezomib
Wild-type	3.3 ± 0.08	0.04 ± 0.001	0.3 ± 0.02	0.2 ± 0.004
RES I	61 ± 1 (18)	4 ± 0.08 (100)	>10	0.2 ± 0.004 (1)
RES 2	54 ± 1 (16)	5 ± 0.2 (127)	>10	0.3 ± 0.01 (1)
RES 3	61 ± 0.7 (18)	5 ± 0.4 (120)	>10	0.2 ± 0.08 (1)
RES 4	31 ± 2 (9)	0.1 ± 0.01 (2.5)	>10	0.6 ± 0.05 (2.5)
RES 5	70 ± 2 (21)	0.8 ± 0.03 (20)	10 ± 0.2 (33)	0.2 ± 0.08 (1)
$\beta5^{D225N\text{-}OE}$	27 ± 1 (8)	2 ± 0.1 (58)	> 10 (>28)	0.2 ± 0.008 (1)

Res 1-β5^{OE}	13 ± 0.8 (4)	0.3 ± 0.01 (6.5)	2 ± 0.3 (6)	0.2 ± 0.005 (1)
β4$^{F24L/I29M}$	> 50 (>15)	> 10 (>250)	> 10 (>28)	0.2 ± 0.005 (1)
ME^{OE}	5 ± 0.11 (1)	-	-	-

All EC_{50} values represent the weighted mean ± standard deviation of at least three biological replicates (n ≥ 3) with each biological replicate comprised of two technical replicates. *Solubility issues >10 µM.

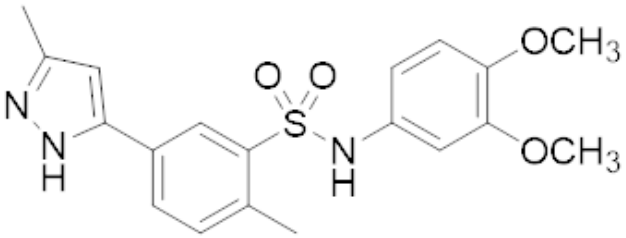
References

1. WHO. 2021. Chagas Disease Fact Sheet.
2. WHO. 2012. Research priorities for Chagas disease, human African trypanosomiasis and leishmaniasis. World Health Organ Tech Rep Ser:v-xii, 1-100.
3. Kratz JM, Garcia Bournissen F, Forsyth CJ, Sosa-Estani S. 2018. Clinical and pharmacological profile of benznidazole for treatment of Chagas disease. Expert Rev Clin Pharmacol 11:943-957.
4. Bern C. 2011. Antitrypanosomal therapy for chronic Chagas' disease. N Engl J Med 364:2527-34.
5. Molina I, Salvador F, Sánchez-Montalvá A, Treviño B, Serre N, Sao Avilés A, Almirante B. 2015. Toxic Profile of Benznidazole in Patients with Chronic Chagas Disease: Risk Factors and Comparison of the Product from Two Different Manufacturers. Antimicrob Agents Chemother 59:6125-31.
6. Molina I, Salvador F, Sánchez-Montalvá A. 2014. Posaconazole versus benznidazole for chronic Chagas' disease. N Engl J Med 371:966.
7. Torrico F, Gascon J, Ortiz L, Alonso-Vega C, Pinazo MJ, Schijman A, Almeida IC, Alves F, Strub-Wourgaft N, Ribeiro I. 2018. Treatment of adult chronic indeterminate Chagas disease with benznidazole and three E1224 dosing regimens: a proof-of-concept, randomised, placebo-controlled trial. Lancet Infect Dis 18:419-430.
8. Chatelain E. 2015. Chagas disease drug discovery: toward a new era. J Biomol Screen 20:22-35.
9. MacLean LM, Thomas J, Lewis MD, Cotillo I, Gray DW, De Rycker M. 2018. Development of Trypanosoma cruzi in vitro assays to identify compounds suitable for progression in Chagas' disease drug discovery. PLoS Negl Trop Dis 12:e0006612.
10. Sánchez-Valdéz FJ, Padilla A, Wang W, Orr D, Tarleton RL. 2018. Spontaneous dormancy protects Trypanosoma cruzi during extended drug exposure. Elife 7.
11. Barrett MP, Kyle DE, Sibley LD, Radke JB, Tarleton RL. 2019. Protozoan persister-like cells and drug treatment failure. Nat Rev Microbiol 17:607-620.
12. Wyllie S, Brand S, Thomas M, De Rycker M, Chung CW, Pena I, Bingham RP, Bueren-Calabuig JA, Cantizani J, Cebrian D, Craggs PD, Ferguson L, Goswami P, Hobrath J, Howe J, Jeacock L, Ko EJ, Korczynska J, MacLean L, Manthri S, Martinez MS, Mata-Cantero L, Moniz S, Nühs A, Osuna-Cabello M, Pinto E, Riley J, Robinson S, Rowland P, Simeons FRC, Shishikura Y, Spinks D, Stojanovski L, Thomas J, Thompson S, Viayna Gaza E, Wall RJ, Zuccotto F, Horn D, Ferguson MAJ, Fairlamb AH, Fiandor JM, Martin J, Gray DW, Miles TJ, Gilbert IH, Read KD, Marco M, Wyatt PG. 2019. Preclinical candidate for the treatment of visceral leishmaniasis that acts through proteasome inhibition. Proc Natl Acad Sci U S A 116:9318-9323.
13. Khare S, Nagle AS, Biggart A, Lai YH, Liang F, Davis LC, Barnes SW, Mathison CJ, Myburgh E, Gao MY, Gillespie JR, Liu X, Tan JL, Stinson M, Rivera IC, Ballard J, Yeh V, Groessl T, Federe G, Koh HX, Venable JD, Bursulaya B, Shapiro M, Mishra PK, Spraggon G, Brock A, Mottram JC, Buckner FS, Rao SP, Wen BG, Walker JR, Tuntland T, Molteni V, Glynne RJ, Supek F. 2016. Proteasome inhibition for treatment of leishmaniasis, Chagas disease and sleeping sickness. Nature 537:229-233.
14. Goyard S, Segawa H, Gordon J, Showalter M, Duncan R, Turco SJ, Beverley SM. 2003. An in vitro system for developmental and genetic studies of Leishmania donovani phosphoglycans. Mol Biochem Parasitol 130:31-42.
15. Roberts AJ, Torrie LS, Wyllie S, Fairlamb AH. 2014. Biochemical and genetic characterization of Trypanosoma cruzi N-myristoyltransferase. Biochem J 459:323-32.

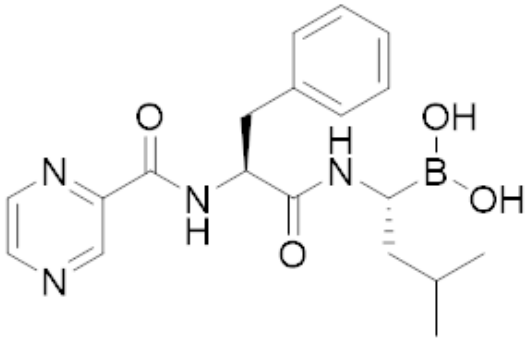
16. Hunter KJ, Le Quesne SA, Fairlamb AH. 1994. Identification and biosynthesis of N1,N9-bis(glutathionyl)aminopropylcadaverine (homotrypanothione) in Trypanosoma cruzi. Eur J Biochem 226:1019-27.
17. Wyllie S, Patterson S, Stojanovski L, Simeons FR, Norval S, Kime R, Read KD, Fairlamb AH. 2012. The anti-trypanosome drug fexinidazole shows potential for treating visceral leishmaniasis. Sci Transl Med 4:119re1.
18. Paradela LS, Wall RJ, Carvalho S, Chemi G, Corpas-Lopez V, Moynihan E, Bello D, Patterson S, Güther MLS, Fairlamb AH, Ferguson MAJ, Zuccotto F, Martin J, Gilbert IH, Wyllie S. 2021. Multiple unbiased approaches identify oxidosqualene cyclase as the molecular target of a promising anti-leishmanial. Cell Chem Biol 28:711-721.e8.
19. Jones DC, Hallyburton I, Stojanovski L, Read KD, Frearson JA, Fairlamb AH. 2010. Identification of a κ-opioid agonist as a potent and selective lead for drug development against human African trypanosomiasis. Biochem Pharmacol 80:1478-86.
20. Wyllie S, Thomas M, Patterson S, Crouch S, De Rycker M, Lowe R, Gresham S, Urbaniak MD, Otto TD, Stojanovski L, Simeons FRC, Manthri S, MacLean LM, Zuccotto F, Homeyer N, Pflaumer H, Boesche M, Sastry L, Connolly P, Albrecht S, Berriman M, Drewes G, Gray DW, Ghidelli-Disse S, Dixon S, Fiandor JM, Wyatt PG, Ferguson MAJ, Fairlamb AH, Miles TJ, Read KD, Gilbert IH. 2018. Cyclin-dependent kinase 12 is a drug target for visceral leishmaniasis. Nature 560:192-197.
21. Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987-93.
22. Corpas-Lopez V, Moniz S, Thomas M, Wall RJ, Torrie LS, Zander-Dinse D, Tinti M, Brand S, Stojanovski L, Manthri S, Hallyburton I, Zuccotto F, Wyatt PG, De Rycker M, Horn D, Ferguson MAJ, Clos J, Read KD, Fairlamb AH, Gilbert IH, Wyllie S. 2019. Pharmacological Validation of N-Myristoyltransferase as a Drug Target in Leishmania donovani. ACS Infect Dis 5:111-122.
23. Vazquez MP, Levin MJ. 1999. Functional analysis of the intergenic regions of TcP2beta gene loci allowed the construction of an improved Trypanosoma cruzi expression vector. Gene 239:217-25.
24. Rico E, Jeacock L, Kovářová J, Horn D. 2018. Inducible high-efficiency CRISPR-Cas9-targeted gene editing and precision base editing in African trypanosomes. Sci Rep 8:7960.
25. Wyllie S, Roberts AJ, Norval S, Patterson S, Foth BJ, Berriman M, Read KD, Fairlamb AH. 2016. Activation of Bicyclic Nitro-drugs by a Novel Nitroreductase (NTR2) in Leishmania. PLoS Pathog 12:e1005971.
26. Zmuda F, Sastry L, Shepherd SM, Jones D, Scott A, Craggs PD, Cortes A, Gray DW, Torrie LS, De Rycker M. 2019. Identification of Novel Trypanosoma cruzi Proteasome Inhibitors Using a Luminescence-Based High-Throughput Screening Assay. Antimicrob Agents Chemother 63.
27. Cazzulo JJ, Juan SM, Segura EL. 1977. The malic enzyme from Trypanosoma cruzi. J Gen Microbiol 99:237-41.
28. Webb B, Sali A. 2016. Comparative Protein Structure Modeling Using MODELLER. Curr Protoc Bioinformatics 54:5.6.1-5.6.37.
29. Peña I, Pilar Manzano M, Cantizani J, Kessler A, Alonso-Padilla J, Bardera AI, Alvarez E, Colmenarejo G, Cotillo I, Roquero I, de Dios-Anton F, Barroso V, Rodriguez A, Gray DW, Navarro M, Kumar V, Sherstnev A, Drewry DH, Brown JR, Fiandor JM, Julio Martin J. 2015. New compound sets identified from high throughput phenotypic screening against three kinetoplastid parasites: an open resource. Sci Rep 5:8771.

30. Castanys-Muñoz E, Pérez-Victoria JM, Gamarro F, Castanys S. 2008. Characterization of an ABCG-like transporter from the protozoan parasite Leishmania with a role in drug resistance and transbilayer lipid movement. Antimicrob Agents Chemother 52:3573-9.
31. BoseDasgupta S, Ganguly A, Roy A, Mukherjee T, Majumder HK. 2008. A novel ATP-binding cassette transporter, ABCG6 is involved in chemoresistance of Leishmania. Mol Biochem Parasitol 158:176-88.
32. Thomas M, Brand S, De Rycker M, Zuccotto F, Lukac I, Dodd PG, Ko EJ, Manthri S, McGonagle K, Osuna-Cabello M, Riley J, Pont C, Simeons F, Stojanovski L, Thomas J, Thompson S, Viayna E, Fiandor JM, Martin J, Wyatt PG, Miles TJ, Read KD, Marco M, Gilbert IH. 2021. Scaffold-Hopping Strategy on a Series of Proteasome Inhibitors Led to a Preclinical Candidate for the Treatment of Visceral Leishmaniasis. J Med Chem 64:5905-5930.
33. Ciechanover A. 2017. Intracellular protein degradation: From a vague idea thru the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. Best Pract Res Clin Haematol 30:341-355.
34. Tanaka K. 2009. The proteasome: overview of structure and functions. Proc Jpn Acad Ser B Phys Biol Sci 85:12-36.
35. Nagle A, Biggart A, Be C, Srinivas H, Hein A, Caridha D, Sciotti RJ, Pybus B, Kreishman-Deitrick M, Bursulaya B, Lai YH, Gao MY, Liang F, Mathison CJN, Liu X, Yeh V, Smith J, Lerario I, Xie Y, Chianelli D, Gibney M, Berman A, Chen YL, Jiricek J, Davis LC, Liu X, Ballard J, Khare S, Eggimann FK, Luneau A, Groessl T, Shapiro M, Richmond W, Johnson K, Rudewicz PJ, Rao SPS, Thompson C, Tuntland T, Spraggon G, Glynne RJ, Supek F, Wiesmann C, Molteni V. 2020. Discovery and Characterization of Clinical Candidate LXE408 as a Kinetoplastid-Selective Proteasome Inhibitor for the Treatment of Leishmaniases. J Med Chem 63:10773-10781.
36. Vora PA, Patel R, Dharamsi A. 2020. Bortezomib - First Therapeutic Proteasome Inhibitor for Cancer Therapy: A Review of Patent Literature. Recent Pat Anticancer Drug Discov 15:113-131.
37. Jafari R, Almqvist H, Axelsson H, Ignatushchenko M, Lundbäck T, Nordlund P, Martinez Molina D. 2014. The cellular thermal shift assay for evaluating drug target interactions in cells. Nat Protoc 9:2100-22.
38. Cannata JJ, Frasch AC, Cataldi de Flombaum MA, Segura EL, Cazzulo JJ. 1979. Two forms of 'malic' enzyme with different regulatory properties in Trypanosoma cruzi. Biochem J 184:409-19.
39. Ranzani AT, Nowicki C, Wilkinson SR, Cordeiro AT. 2017. Identification of Specific Inhibitors of Trypanosoma cruzi Malic Enzyme Isoforms by Target-Based HTS. SLAS Discov 22:1150-1161.
40. Kourbeli V, Chontzopoulou E, Moschovou K, Pavlos D, Mavromoustakos T, Papanastasiou IP. 2021. An Overview on Target-Based Drug Design against Kinetoplastid Protozoan Infections: Human African Trypanosomiasis, Chagas Disease and Leishmaniases. Molecules 26.
41. Nkemngu NJ, Rosenkranz V, Wink M, Steverding D. 2002. Antitrypanosomal activities of proteasome inhibitors. Antimicrob Agents Chemother 46:2038-40.
42. Allmann S, Morand P, Ebikeme C, Gales L, Biran M, Hubert J, Brennand A, Mazet M, Franconi JM, Michels PA, Portais JC, Boshart M, Bringaud F. 2013. Cytosolic NADPH homeostasis in glucose-starved procyclic Trypanosoma brucei relies on malic enzyme and the pentose phosphate pathway fed by gluconeogenic flux. J Biol Chem 288:18494-505.
43. Bustamante JM, Sanchez-Valdez F, Padilla AM, White B, Wang W, Tarleton RL. 2020. A modified drug regimen clears active and dormant trypanosomes in mouse models of Chagas disease. Sci Transl Med 12.

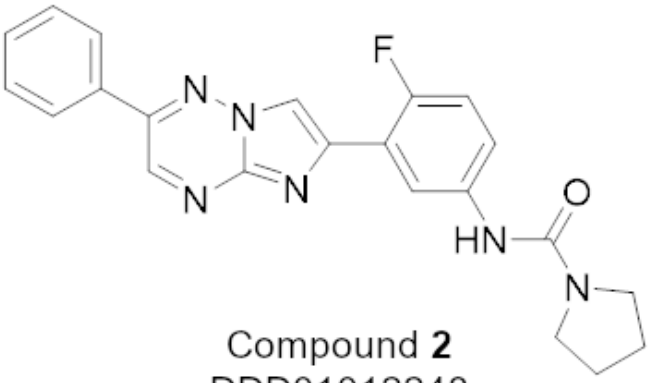
903



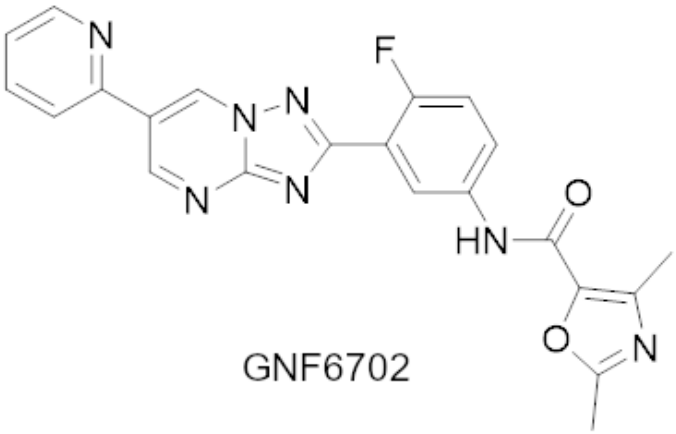
Compound **1**
TCMDC-143194



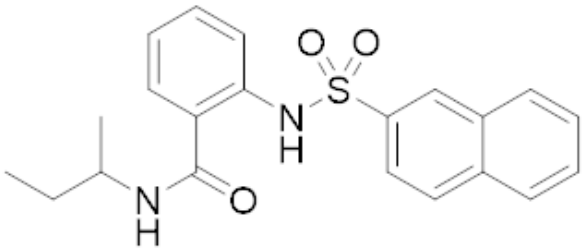
Bortezomib



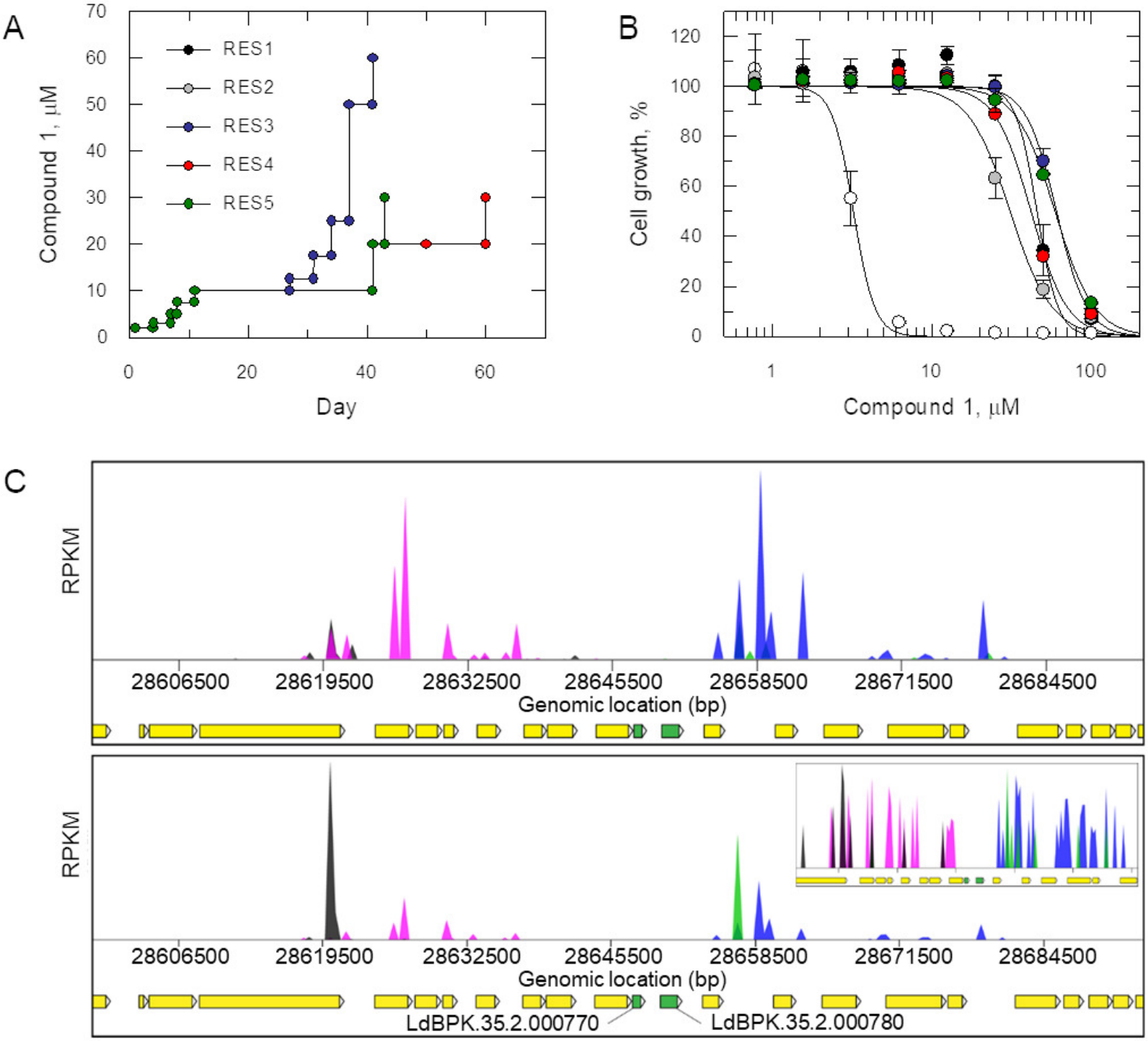
Compound **2**
DDD01012248

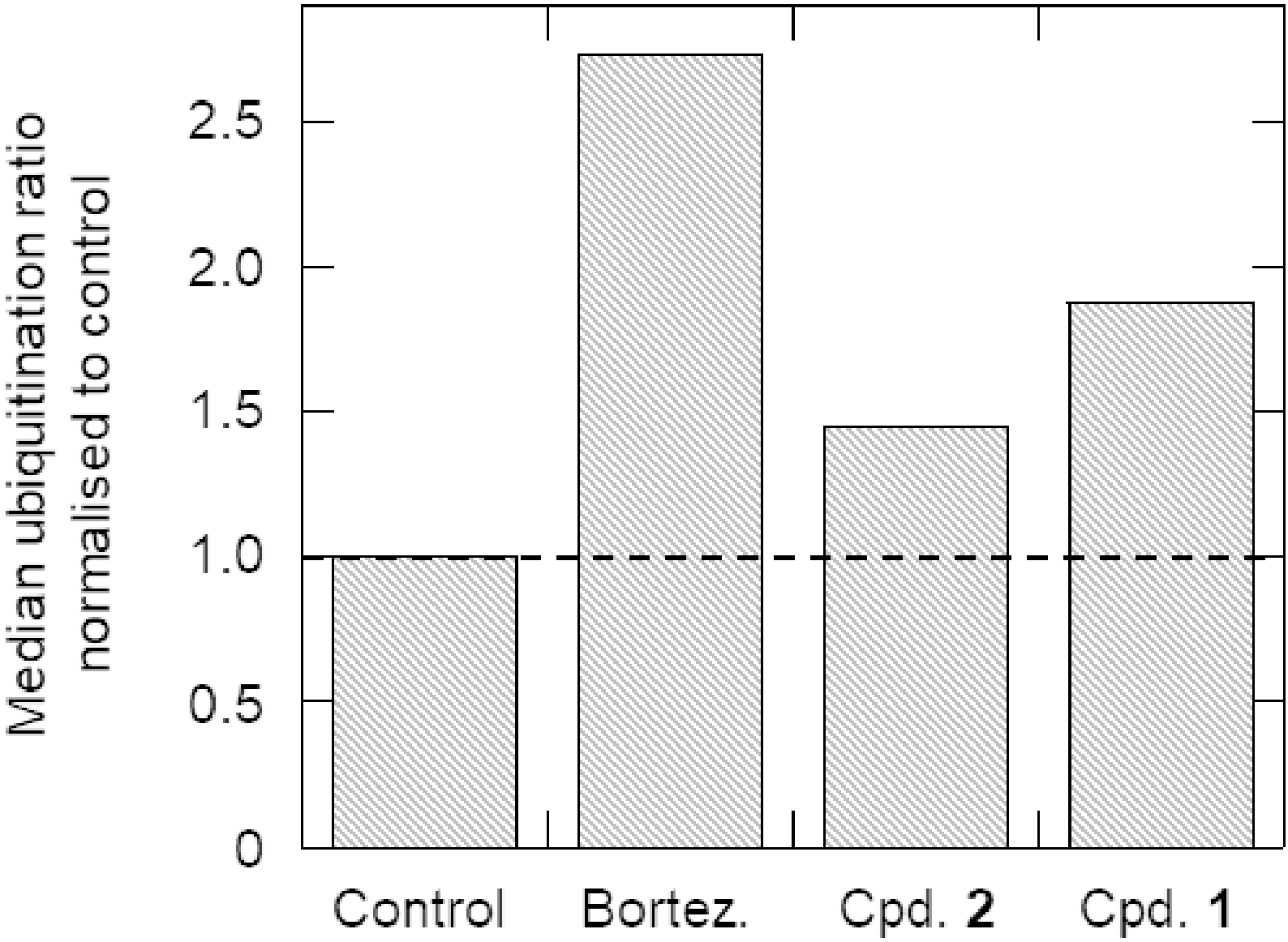


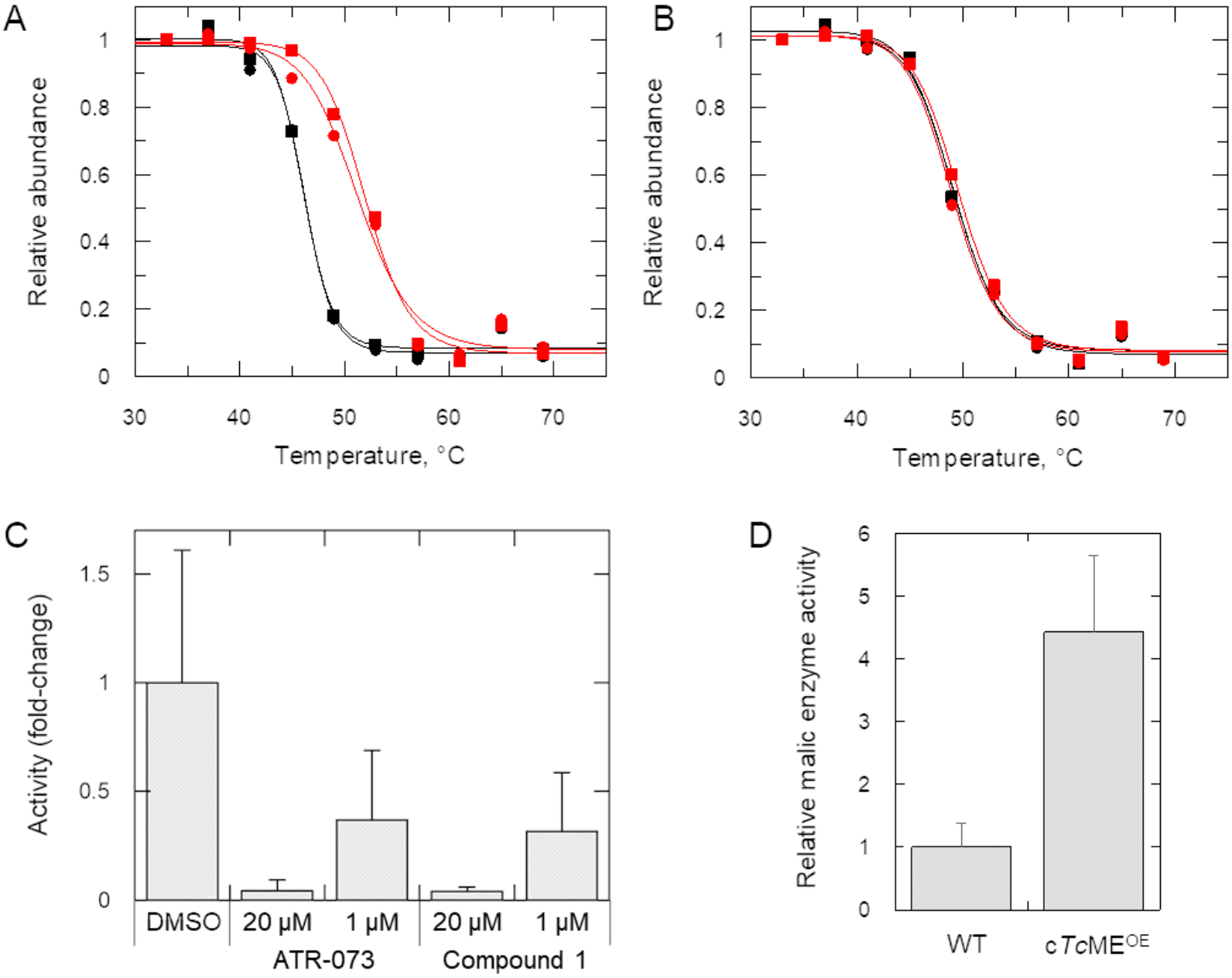
GNF6702

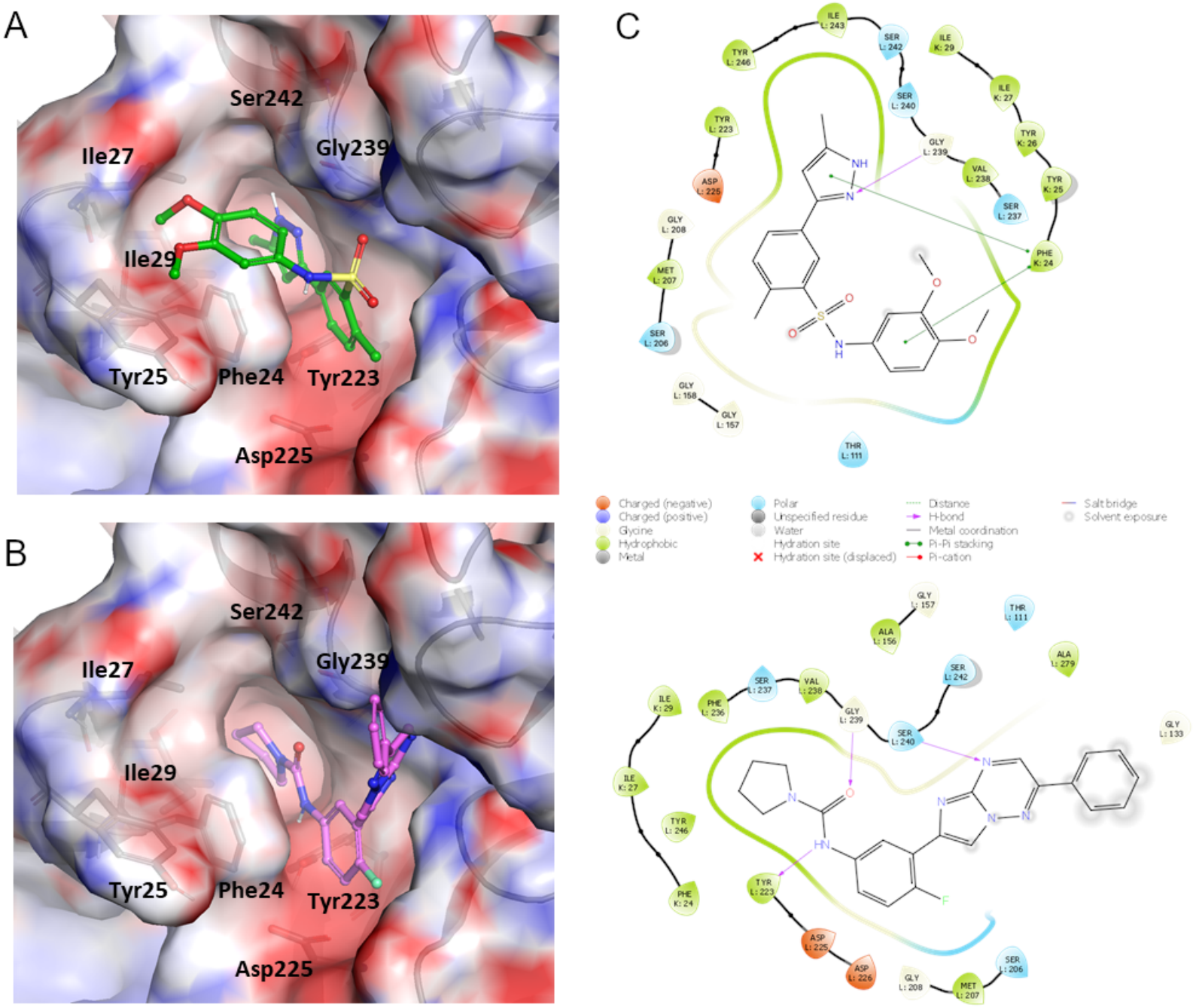


ATR3-073


A
Compound 1, μM
Day
RES1
RES2
RES3
RES4
RES5
B
Cell growth, %
Compound 1, μM
C
RPKM
Genomic location (bp)
28606500
28619500
28632500
28645500
28658500
28671500
28684500
LdBPK.35.2.000770
LdBPK.35.2.000780


Median ubiquitination ratio
normalised to control
2.5
2.0
1.5
1.0
0.5
0
Control
Bortez.
Cpd. 2
Cpd. 1


A
Relative abundance
Temperature, °C
B
Relative abundance
Temperature, °C
C
Activity (fold-change)
DMSO
20 μM
1 μM
20 μM
1 μM
ATR-073
Compound 1
D
Relative malic enzyme activity
WT
cTcME^OE


A
Ser242
Ile27
Gly239
Ile29
Tyr25
Phe24
Tyr223
Asp225
B
Ser242
Ile27
Gly239
Ile29
Tyr25
Phe24
Tyr223
Asp225
C
Charged (negative)
Charged (positive)
Glycine
Hydrophobic
Metal
Polar
Unspecified residue
Water
Hydration site
Hydration site (displaced)
Distance
H-bond
Metal coordination
Pi-Pi stacking
Pi-cation
Salt bridge
Solvent exposure