Dissecting the interstrand crosslink DNA repair system of *Trypanosoma cruzi*

Monica Zavala Martinez[1;#], Francisco Olmo[2], Martin C. Taylor[2], Fabrice Caudron[1,$] and Shane R Wilkinson[1*]

[1]School of Biological & Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK.
[2]Department of Infection Biology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Corresponding author.
*Dr Shane Wilkinson, School of Biological and Chemical Sciences, Queen Mary University of London, London, E1 4NS UK.
Email: s.r.wilkinson@qmul.ac.uk; Phone +44 (0)20 7882 3057

[#]Current address: Trypanosome Molecular Biology, Institut Pasteur, Université Paris Cité, Paris, France.
[$]Current address: IGMM, Univ Montpellier, CNRS, Route de Mende, 34293 Montpellier, France.

Abstract
DNA interstrand crosslinks (ICLs) are toxic lesions that can block essential biological processes. Here we show *Trypanosoma cruzi*, the causative agent of Chagas disease, is susceptible to ICL-inducing compounds including mechlorethamine and novel nitroreductase-activated prodrugs that have potential in treating this infection. To resolve such lesions, cells co-opt enzymes from "classical" DNA repair pathways that alongside dedicated factors operate in replication-dependent and -independent mechanisms. To assess ICL repair in *T.cruzi*, orthologues of SNM1, MRE11 and CSB were identified and their function assessed. The *T.cruzi* enzymes could complement the mechlorethamine susceptibility phenotype displayed by corresponding yeast and/or *T.brucei* null confirming their role as ICL repair factors while GFP-tagged TcSNM1, TcMRE11 and TcCSB were shown to localise to the nuclei of insect and/or intracellular form parasites. Gene disruption demonstrated that while each activity was non-essential for *T.cruzi* viability, nulls displayed a growth defect in at least one life cycle stage with TcMRE11-deficient trypomastigotes also compromised in mammalian cell infectivity. Phenotyping revealed all nulls were more susceptible to mechlorethamine than controls, a trait complemented by re-expression of the deleted gene. To assess interplay, the gene disruption approach was extended to generate *T.cruzi* deficient in TcSNM1/TcMRE11 or in TcSNM1/TcCSB. Analysis demonstrated these activities functioned across two ICL repair pathways with TcSNM1 and TcMRE11 postulated to operate in a replication-dependent system while TcCSB helps resolve transcription-blocking lesions. By unravelling how *T.cruzi* repairs ICL damage, specific inhibitors targeting repair components could be developed and used to increase the potency of trypanocidal ICL-inducing compounds.

Keywords
Chagas disease; Trypanosomes; DNA interstrand crosslink repair; homologous recombination; Transcription coupled-nucleotide excision repair; metallo-ß-lactamase.

1.0 Introduction

Across Latin America, an estimated 6 million people suffer from Chagas disease (CD), a chronic debilitating illness that has a high (~30%) mortality rate. This infection primarily afflicts poor rural communities, having a major impact on the socio-economic development of regions least able to deal with the associated economic burden, often trapping sufferers and their families in a disease-poverty cycle. The causative agent of CD is *Trypanosoma cruzi*, a protozoan that is spread between mammalian populations through the spread hematophagous feeding behaviour of triatomine insects. When a triatomine takes a blood meal, it defecates on the mammalian host. Infectious, non-dividing metacyclic trypomastigote parasites present in the insect's faeces then enters the host through the bite site or mucus membranes, such as the conjunctiva. In the host, trypomastigotes invade cells near the site of entry, where they transform into intracellular amastigotes. The amastigotes divide by binary fission, differentiate into infectious, non-dividing bloodstream trypomastigotes that upon rupturing of the mammalian cell, are released into the host's circulatory systems. The trypomastigotes can then infect cells at sites throughout the mammalian host thereby propagating the intracellular phase of the life cycle, or can be taken up by a triatomine as it feeds. In the insect's midgut, the trypomastigotes differentiate into replicative epimastigotes that can enter the hindgut where they transform into metacyclic trypomastigotes ready for transmission to the next victim. Through the implementation of vector control strategies and improved housing, disease prevalence at endemic sites has dramatically fallen over the past 25 years such that CD has now been eliminated from Chile, Uruguay and several regions of Argentina and Brazil. However, because this disease is a zoonosis (*T. cruzi* can infect all mammals), it is imperative that these insect control/housing programmes are maintained as CD disease could re-emerge as a problem in areas where it had been thought of as being eradicated. Worryingly, and due to congenital and sexual transmission, blood transfusion and organ transplantation coupled with population migration and autochthonous transmissions cycles, CD has started to emerge as a global public health problem. The Centers for Disease Control and Prevention estimate that up to 300,000 people living in the USA are infected with *T. cruzi* while it has been proposed that up to 4.2% of Latin American born nationals now residing in Europe harbour the parasite, a value that rises to 18.0 % for people born in Bolivia [1, 2].

Current treatment of CD is based on the two nitroheterocylic prodrugs benznidazole and nifurtimox with their anti-parasitic selectivities dependent on their activation mechanism. In a reaction catalysed by a NADH-dependent, FMN-containing mitochondrial type I nitroreductase (NTR) [3], a conserved nitro group linked to an imidazole (in the case of benznidazole) or furan (in the case of nifurtimox) ring undergoes two sequential 2 electron reduction events to form nitroso and then hydroxylamine forms [4, 5]. The latter derivatives are unstable and can undergo further, non-enzymatic processing to generate highly reactive end products via a series of adduct forming intermediates [4-7]. Based on the mutagenic properties the two prodrugs display to NTR competent bacterial and/or trypanosomal cells, DNA appears to constitute a major target [8-12]. Although the precise nature of this interaction remains unclear, it has been reported that both compounds can lead to base mismatch pairing [10] with, in the case of benznidazole, this potentially arising from incorporation of oxidised nucleotides into the parasite genome [13]. Such lesions can then promote DNA damage through formation of point mutations and/or double-strand DNA breaks (DSBs) [8, 11, 14]. As humans lack orthologues of the trypanosomal type I NTR, this parasite activity has been exploited to screen distinct groups of nitroaromatic and quinone-based compounds for anti-*T. cruzi* properties. From this, several nitrobenzyl-based prodrugs that contain functional groups which have

potential to promote DNA interstrand crosslinks (ICLs) were shown to display significant anti-parasitic effects while exhibiting low mammalian cell toxicity [15-17].

Formed when the complementary strands within the DNA double helix become covalently linked by bifunctional alkylating agents, ICLs represent a particularly dangerous lesion that can block essential cellular processes that require DNA strand separation and can lead to chromosomal breakage, rearrangements, and cell death [18-21]. To maintain integrity and functionality, all cells have evolved multiple mechanisms to remove such lesions from their genomes [20, 22]. Although the precise factors involved in this process have yet to be fully elucidated different mechanisms are known to operate at various stages in the cell cycle, employing components from the "classical" DNA repair pathways (nucleotide excision repair (NER), mismatch repair (MMR), translesion synthesis (TLS) and homologous recombination (HR)) in conjunction with ICL repair specific proteins (*e.g.*, PSO2/SNM1, FAN1) [23-28]. In transcriptionally active mammalian cells, ICL formation often results in distortion of the DNA double helix and the stalling of an RNA polymerase complex. These events can be recognised by factors such as by the transcription coupled-nucleotide excision repair (TC-NER) helicase CSB, with these subsequently recruiting other NER enzymes including XPG and XPF-ERCC1, and SNM1A to the lesion site [29-31]. In a process known as "unhooking" and mediated by the nucleolytic activities of XPF-ERCC1 and XPG, the sugar-phosphate backbone on one of the DNA strands is cleaved at sites upstream and downstream of the ICL with nucleases, including SNM1A, subsequently degrading the released sequence up to and beyond the crosslink. The resultant gap is filled by damage tolerant TLS DNA polymerases such as Pol ζ, Pol η, Pol ι, Pol Θ or Pol κ, and the double strand DNA structure restored by DNA ligase [32-34]. Once this strand has been repaired, the ICL is completely removed from the DNA by a second round of NER incision, TLS DNA polymerase and DNA ligase activities. A variant of this mechanism can operate in non-transcriptionally activity regions of a genome but here recognition of the ICL-induced chromatin structure/DNA conformation alterations is mediated by components of the MMR and global genome-nucleotide excision repair (GG-NER) pathways [29, 35-37].

During the S-phase of the mammalian, amphibian and fish cell cycle, the so-called Fanconi Anaemia (FA) pathway predominates to remove ICLs from the nuclear genome [38, 39]. Here, the ICL-mediated collapse of DNA replication forks is recognised by the multi-subunit FA core complex that guides and activates the FA recruitment complex. Once triggered, this recruitment complex directs a range of effector proteins from the "classical" DNA repair pathways that operate to resolve the ICL. In dividing yeast cells, an abridged FA pathway is postulated to operate. Here, stalling of a replisome(s) results in the recruitment of MPH1, MUTSα and ancillary factors to the lesion site [27, 28, 40]. These function to regress the replication fork and direct endonuclease (*e.g.*, XPG and XPF-ERCC1), exonuclease (*e.g.*, EXO1, PSO2), TLS DNA polymerase and DNA ligase activities to remove the ICL from the genome with repair of DSBs that arise from this process occurring via the HR pathway.

Informatic and/or functional studies on *Trypanosoma brucei*, the causative agent of African trypanosomiasis, has shown this parasite expresses several activities that operate in the ICL repair systems of other organisms [41-43]. These function across at least two distinct networks with CSB, EXO1 and SNM1 operating to repair ICLs that block transcription complexes while the HR factors MRE11, RAD51 and BRCA2 resolve lesions that cause stalling of DNA replication forks [43]. Here, we assess the susceptibility of *T. cruzi* to various classes of ICL inducing agent and using a classical genetics-based approach, analyse whether DNA repair enzymes from this parasite's TC-NER (TcCSB) and HR (TcMRE11) pathways also play a role

in resolving ICLs alongside that of the ICL repair specific factor, TcSNM1. As *T. cruzi* and *T. brucei* diverged from each other in the mid-Cretaceous period, around 100 million years before present [44], and that many aspects of their biology are distinct, our comparative analysis revealed that the ICL repair networks expressed by these two trypanosomal species are similar but not identical.

2.0 Material and methods

2.1 Cell culturing.

Trypanosoma cruzi CL-Brener and clone MHOM/BR/78/Sylvio-X10.6 epimastigotes were grown at 27 °C in RPMI-1640 (without glutamate) (Lonza) medium supplemented with 0.5 % (w/v) trypticase, 0.5 % (w/v) HEPES pH 8.0, 30 µM hemin, 10 % (v/v) heat inactivated foetal calf serum (FCS), 2 mM sodium glutamate, 2 mM sodium pyruvate, 250 units ml^{-1} penicillin and 250 µg ml^{-1} streptomycin [45]. Recombinant *T. cruzi* (Table S1) were maintained in this medium containing 10µg ml^{-1} blasticidin, 5 µg ml^{-1} puromycin, 100 µg ml^{-1} G418 and/or 100 µg ml^{-1} hygromycin.

To establish *T. cruzi*-infected mammalian cells, parasites from a stationary phase (~15-day old) epimastigote culture were used to infect L6 rat myoblasts overnight at 37 °C under a humid 5 % (v/v) CO_2 atmosphere. The infected monolayers were washed with RPMI-1640 without glutamate medium to remove any residual non-internalized trypanosomes and cultures were incubated at 37 °C under a humid 5 % (v/v) CO_2 atmosphere in RPM1-1640 medium supplemented with 0.5 % (w/v) HEPES pH 8.0, 2 mM sodium glutamate, 2 mM sodium pyruvate, 250 units ml^{-1} penicillin, 250 µg ml^{-1} streptomycin and 1 % (v/v) heat inactivated FCS, with the medium changed every 3-4 days. After 10 to 14 days, tissue culture derived bloodstream trypomastigotes were observed in the medium. These were collected and used to infect fresh mammalian cell monolayers.

Bloodstream form *Trypanosoma brucei brucei* (Lister 427 strain; clone MITat1.2) and derivatives null for Tb*SNM1* (Tb*snm1Δ*), TbMRE11 (Tb*mre11Δ*) or TbCSB (Tb*csbΔ*)) were maintained at 37 °C under a humid 5 % (v/v) CO_2 atmosphere and in HMI-9 medium (Invitrogen) containing 0.3 % (w/v) sodium bicarbonate, 0.0014 % (v/v) β-mercaptoethanol and 10 % (v/v) FCS [46]. Recombinant *T. brucei* were selected and maintained in this medium supplemented with 2 µg ml^{-1} blasticidin, 2 µg ml^{-1} puromycin and/or 1 µg ml^{-1} phleomycin.

L6 rat myoblasts were grown at 37 °C under a humid 5 % (v/v) CO_2 atmosphere in RPMI-1640 medium containing 0.5 % (w/v) HEPES pH 8.0, 2 mM sodium glutamate, 2 mM sodium pyruvate, 250 units ml^{-1} penicillin, 250 µg ml^{-1} streptomycin and 10 % (v/v) heat inactivated FCS.

S. cerevisiae strains BY4742 (*MATα his3-Δ1 leu2-Δ0 lys2-Δ0 ura3-Δ0*) and a *PSO2*Δ derivative obtained from the Open Biosystems (Thermo Scientific) knock-out collection were maintained in yeast extract-peptone broth containing 2 % (w/v) glucose. Transformed cells were grown in Synthetic Complete Dropout medium lacking leucine (Sigma).

2.2. Chemicals & treatments.

Mechlorethamine (HN2) was obtained from Cambridge Bioscience while CB1954, methyl methanesulfonate (MMS), hydrogen peroxide (H_2O_2) and hydroxyurea (HU) were all sourced from Sigma-Aldrich. UV irradiation was performed with a Stratalinker UV crosslinker (Stratagene). Other DNA damaging agents were obtained from Drug Synthesis and Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis,

National Cancer Institute, USA. Benznidazole, nifurtimox and fexinidazole were supplied by Prof Simon Croft and Prof John Kelly (both London School of Hygiene and Tropical Medicine, UK) while RH1 was donated by Prof Frank Guziec Jr (Southwestern University, USA). The trypanocidal selective compounds blasticidin, puromycin, G418 and hygromycin were purchased from Melford Laboratories Ltd.

2.3. Informatics.

The Tc*snm1* (TcCLB.511003.50; TCSYLVIO_003789), Tc*ccsb* (TcCLB.508675.20; TcCLB.506983.60; TCSYLVIO_010473) and Tc*mre11* (TcCLB.509099.70; TcCLB.509319.60; TCSYLVIO_000429) open reading frames (ORF) and flanking regions were identified following ortholog, textual, InterPro Domain or BLAST searches of the TriTrypDB
. Deduced protein sequences were aligned to the counterparts expressed by other organisms using CLUSTALΩ (www.ebi.ac.uk/Tools/msa/clustalo), domain structures evaluated using HMMR (www.ebi.ac.uk/Tools/hmmer/) and regions of homology highlighted using BoxShade (embnet.vital-it.ch/software/BOX_form.html). Prediction of subcellular localisation was determined using PSORT [47], WoLFPSORT [48], NucPred [49] and Euk-mPLoc 2.0 [50] while potential post-translational modification sites were identified using UbPred [51] and GPS-SUMO [52]. 3D approximations of protein structures were determined using Phyre2 [53], with models visualized and edited using the PyMOL2 software (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.).

2.4. Construction of recombinant parasite lines.

The vectors used to interrupt a target gene in the *T. cruzi* genome were generated as follows. For Tc*snm1*, sequences corresponding to 5′ and 3’ flanking regions were amplified from *T. cruzi* genomic DNA and sequentially cloned either side of a resistance cassette that includes the gene encoding for blasticidin-S deaminase (*bla*) or puromycin N-acetyltransferase (*pac*). The vectors designed to disrupt Tc*ccsb* or Tc*mre11* were made in a similar way except that the amplified fragments were derived from the 5′ and 3’ ends of the coding sequence and these were cloned either side of a resistance cassette that included the gene encoding for *bla*, *pac*, hygromycin B phosphotransferase (*hyg*) or neomycin phosphotransferase (*neo*): The primers and primer combinations used to generate the 5’ and 3’ sequences and the sizes of the amplified fragments are shown in Tables S2 & S3, and in Figure S3. Constructs were linearized (SacI/KpnI for the *hyg*, *neo* and *pac* constructs and SacII/ KpnI for the *bla* vector) then transformed into hydroxyurea-treated *T. cruzi* X10.6 epimastigotes [54] using Tb-BSF buffer (5 mM KCl, 0,15 mM $CaCl_2$, 90 mM Na_2HPO_4, 50 mM HEPES & pH 7.3) [55] and an Amaxa® Nucleofector™ (Lonza AG) set to program X-014 [56]. As the *T. cruzi* genes targeted here are single copy in a diploid genome, two rounds of nucleofection were needed to firstly create heterozygous and then null mutant lines. Integration of the Tc*snm1*, Tc*ccsb* and Tc*mre11*-based gene disruption constructs into the *T. cruzi* genome resulted in deletion of 62, 54 and 39 % of the corresponding ORF, respectively. The regions removed from the genome represent sequences postulated to be required for nuclease (TcSNM1; TcMRE11), DNA binding (TcMRE11) and ATPase (TcCSB) activities.

The vectors used to ectopically express GFP tagged proteins were made as follows: DNA fragments containing *gfp* minus its STOP codon or its START codon were amplified from pTEX-eGFP [57] and the resultant BamHI/XhoI-digested DNAs cloned into the corresponding sites of pTRIX [16] to generate pTRIX-^{GFP}X and pTRIX-X^{GFP}. The Tc*snm1* or Tc*mre11* ORFs, or Tc*ccsb* minus its stop codon were amplified from *T. cruzi* X10.6 genomic DNA and cloned into the HindIII/XhoI sites of pTRIX-^{GFP}X (for Tc*snm1* or Tc*mre11*) or the BamHI/HindIII

sites of pTRIX-X^{GFP} (for Tc*ccsb*) generating 5' GFP tagged versions of Tc*snm1* or Tc*mre11* and a 3' GFP tagged version of Tc*ccsb*. Constructs were linearized with SacI/KpnI and transformed into hydroxyurea-treated *T. cruzi* CL-Brener (or null lines for complementation studies) with recombinant parasites selected using G418 [54].

To construct the *T. brucei* complementation vector, a *gfp*-Tc*snm1* gene fusion amplicon, generated from pTRIX-^{GFP}TcSNM1, was SbfI/AscI-digested and cloned into the corresponding sites of pTubEX [58]. The vector was linearized with NotI/XhoI and fragments transformed into BSF *T.brucei* 221 or Tb*snm1l∆* and recombinant parasites selected using phleomycin.

To construct the yeast complementation vector, full length Tc*snm1* was amplified *T. cruzi* X10.6 genomic DNA, the amplicon XbaI/HindIII-digested, cloned into the corresponding sites of a pYCYlac111 derivative [59] and transformed into *S. cerevisiae* BY4742 and *pso*2Δ.

2.5 Validation of recombinant parasites.

To demonstrate that integration of a gene interruption cassette had occurred at the correct genetic loci and confirm that the *T. cruzi* null mutant line was no longer expressing the targeted gene(s), diagnostic PCRs were performed on parasite genomic DNA or cDNA templates. For gene interruption constructs, the primer combinations used generated amplicons specific for the intact targeted gene or the *hyg*-, *neo*-, *pac*- and/or *bla*-disrupted allele.

To demonstrate that a GFP-based construct was present in the *T. cruzi* or *T. brucei* genomes and confirm that a fusion gene/protein was expressed, PCRs on genomic DNA or cDNA templates, and western blots on protein lysates were performed using material extracted from recombinant parasites. To detect fusion events where a DNA repair gene had been tagged at its 5' end, amplifications were carried out using a GFP-specific forward and gene-specific reverse primer while in cases where the 3' end of the target gene had been tagged, a gene-specific forward and GFP-specific reverse primer were employed. Fragments amplified from cDNA using Phusion® High Fidelity DNA polymerase (New England Biolabs) were sequenced (Source BioScience) with the resultant outputs used to confirm the in-frame fusion. For all analysed DNA's, additional reactions aimed at detecting Tc*tert* were performed to check the integrity of the template and as loading control: The Tc*tert* gene codes for the *T. cruzi* telomerase reverse transcriptase The primer sequences and combinations used in cell line validation are listed in Tables S2 and S3. To detect recombinant protein, western blots were probed using an anti-GFP monoclonal antibody (C163) (Invitrogen).

2.6. Parasite growth studies.

T. cruzi epimastigotes in the logarithmic phase of growth were seeded at 5×10^5 parasites ml^{-1} in parasite growth medium and incubated at 27 °C under a 5 % (v/v) CO_2 atmosphere in a well of a 6-well plate. Each day, the cell density of each culture was measured using a Neubauer haemocytometer. When the number of parasites reached ~2 x 10^7 ml^{-1}, a new culture seeded at 5×10^5 parasites ml^{-1} was set up. This analysis was carried out over a 20-day period. Growth curves were generated using GraphPad Prism (GraphPad Software Inc.). All growth assays were performed in triplicate and each count at each time point expressed as a mean ± standard deviation. Mean generation times were calculated using http://www.doubling-time.com/compute.php.

L6 rat myoblasts seeded at 1×10^4 cells ml^{-1} adhered to a chamber slide were infected with amastigote-free, tissue culture derived bloodstream trypomastigotes at a ratio of 5 trypanosomes per cell and cultures incubated at overnight 37 ° in a humid 5 % (v/v) CO_2

atmosphere. Monolayers were then copiously washed in *T. cruzi* amastigote growth medium to remove non-internalized parasites before replacing with fresh medium. At time intervals, phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.76 mM KH_2PO_4, pH 7.6) (Sigma Aldrich) washed cells were fixed with 2 % (w/v) paraformaldehyde (PFA) (Sigma-Aldrich) in PBS for 10 minutes. Following disassembly of the chamber walls, mammalian and parasite DNAs were stained using Vectashield Mounting Medium containing DAPI (Vectorshield Laboratories) and cells visualized using a DMRA2 upright epifluorescent microscope (Leica) fitted with monochrome Hamamatsu Orca-ER camera in conjunction with Volocity 3D software (Perkin Elmer). Images were processed using ImageJ [60]. At each time point, at least 250 mammalian (infected/non-infected) cells were identified and the % of L6 myoblasts infected with *T. cruzi* calculated. For infected cells, the number of amastigotes per L6 myoblast was also determined. All assays were performed on four independent cultures with the data expressed as a mean % L6 myoblasts infected ± standard deviation or mean number of amastigotes per L6 myoblast ± standard deviation. The statistical significance of any differences was assessed using the Student's *t* test calculator (GraphPad Software Inc.).

2.7. Microscopy and protein analysis.

PBS-washed *T. cruzi* epimastigotes were fixed with 2 % (w/v) paraformaldehyde in PBS, washed in PBS before air drying aliquots containing ~10^5 cells onto printed microscope slides (Hendley Essex Ltd). For *T. cruzi* amastigotes, monolayers of parasite-infected L6 cells grown for 3 to 5 days at 37 °C in a 5 % (v/v) CO_2 atmosphere and in 8-well chamber slides (ThermoFisher Scientific), were washed with phosphate buffered saline (PBS), fixed 2 % (w/v) paraformaldehyde in PBS, and washed once more in PBS before air dried. Parasite and mammalian DNAs were stained using Vectashield Mounting Medium containing DAPI (Vectorshield Laboratories) and cells visualized using a sCMOS camera on a DeltaVision Ultra High-Resolution microscope (GE Healthcare Life Sciences) fitted with an Olympus 60X NA 1.4 objective lens and images captured with the DeltaVision SoftWorx software (GE Healthcare Life Sciences). For TcCSB-GFP expression studies, *T. cruzi* epimastigotes engineered to the recombinant protein were treated with UV light (1200 J m^{-2}) and allowed to recover for 2 hours prior to fixation, DAPI co-staining and microscopic evaluation. All images were processed, and corrected fluorescence intensities determined using ImageJ.

For western blot analysis, protein extracts from ~2 x 10^6 trypanosomes were probed with mouse anti-GFP (Invitrogen), rabbit anti-γH2A [61, 62] and/or rabbit anti-enolase [63] antisera used at 1:5,000, 1:1,000 or 1:10,000 dilution, respectively, followed by a goat anti-Mouse IRDye®680LT or IRDye®800CW goat anti-rabbit IgG antibodies (LI-COR), both diluted at 1:10,000. Detection of the near infrared signal was monitored using an Odyssey® CLx infrared imaging system (LI-COR).

2.8. Susceptibility screens.

T. cruzi epimastigotes or BSF *T. brucei* in the logarithmic phase of growth were seeded at 5 x 10^5 ml^{-1} or 1 × 10^4 ml^{-1}, respectively in 200 µl growth medium containing different concentrations of the compound under study. After incubation for 4 days at 27 °C for *T. cruzi*, or for 3 days at 37 °C for *T. brucei* both under a humid 5 % (v/v) CO_2 atmosphere, resazurin (Sigma Aldrich) was added to each well at a final concentration of 12.5 µg ml^{-1} (or 2.5 µg per well). The plates were incubated overnight at 27 °C for *T. cruzi* or 8 hours at 37 °C for *T. brucei* both under a humid 5 % CO_2 before measuring the fluorescence of each culture using a Gemini Fluorescent Plate reader (Molecular Devices) set at λ_{EX} = 530 nm and λ_{EM} = 585 nm with a filter cut off at 550 nm. A drug/treatment concentration that inhibits cell growth by 50% (EC_{50})

was established using the non-linear regression tool on GraphPad Prism (GraphPad Software Inc.).

For UV treatment, *T. cruzi* epimastigotes in the logarithmic phase of growth were seeded at 2.5 × 10^6 cells ml^{-1} in 200 µl growth medium and dispensed into the wells of 24-well plates. Cells were UV irradiated using Stratalinker® UV crosslinker (Stratagene) and the culture volume adjusted to 1 ml with growth medium. Aliquots (200 µl) were transferred to the wells of a 96-well plate and then incubated at 27 °C in a humid 5 % (v/v) CO_2 atmosphere for 4 days prior to addition of resazurin.

T. cruzi amastigote susceptibility assays were set up and analysed using the previously described amastigote growth assay protocol with treatment initiated once a *T.* cruzi/L6 myoblast infection had been established (~24 post-infection). All assays were performed on four independent cultures with the data expressed as a mean % L6 myoblasts infected ± standard deviation or mean number of amastigotes per L6 myoblast± standard deviation. The statistical significance of any differences was assessed using the Student's *t* test calculator (GraphPad Software Inc.).

2.9 Cell cycle analysis using fluorescence-activated cell sorting.
Exponentially growing *T. cruzi* epimastigotes were treated with 20 mM HU for 24 hours, washed twice with PBS and suspended in fresh growth medium. Cultures were incubated at 27 °C with aliquots (~1 x 10^6 cells) taken every 3 hours over a 48-hour period. Cells were washed once with PBS, fixed with 2 % (w/v) PFA for 20 minutes then suspended in fresh PBS. *T. cruzi* DNA was stained with Hoechst 33342 (ThermoFisher Scientific) with ~10^4 cells sorted using a 2020 Attune NxT Flow Cytometer (ThermoFisher Scientific) and the data visualized using the FlowJo version 10.7.1 (Becton, Dickinson and Company; 2019).

3. Results

3.1 Susceptibility of *T. cruzi* to ICL inducing agents.

Many antimicrobial and anticancer drugs mediate their activities by promoting DNA damage with ICL-inducing compounds being particularly toxic. Previous screening studies against bloodstream form *T. brucei* have shown that such chemicals are effective trypanocidal agents, with those that function as NTR1-activated prodrugs often displaying specificity towards the parasite [15-17, 42, 64]. To assess whether the anti-parasitic properties extend across to other trypanosomes, various structures that promote ICL formation, including those containing nitrogen mustard, aziridinyl, platinum complex, nitrosourea or alkyl sulfonate groupings, were screened against *T. cruzi* epimastigotes. Out of the 21 compounds tested, 7 had no effect on *T. cruzi* growth at concentrations up to 500 µM, including busulfan, the only alkyl sulfonate analysed here, and six of the eight nitrogen mustards (Figure 1A). These were excluded from further studies. For the remainder, the effective concentration that inhibits parasite growth by 50% (EC_{50}) was determined (Figure 1A). This revealed that multiple compounds from all the remaining classes of ICL inducing agents displayed growth inhibition activity against *T. cruzi*, yielding EC_{50} values of ~2 nM up to ~200 µM. As a group, the aziridinyl-containing agents displayed the highest potency towards *T. cruzi* with a subset that contain a NTR1-activated nitro grouping exhibiting the highest anti-parasitic activities: EC_{50} values of 2.4 and ~ 340 nM were recorded for RH1 and CB1954, respectively.

In many eukaryotes, including trypanosomes, nuclear DNA damage is often associated with the phosphorylation of the histone H2A(X) to form γH2A(X) [65, 66]. To assess whether HN2, the archetypal ICL inducing compound, can promote this post translational modification in *T.*

cruzi, parasite protein extracts derived from epimastigotes exposed to this nitrogen mustard (100 μM; overnight) were probed with antisera against the modified histone and compared against untreated controls (Figure 1B): As control, enolase was analysed in parallel. A single 17 kDa band corresponding to γH2A was detected in all lysates. This band was more in intense in extracts derived from treated parasites relative to controls and confirms that HN2 does promote DNA damage in *T. cruzi*.

3.2 Identification of *T. cruzi* ICL repair enzymes.

SNM1/PSO2 plays a key role in the repair of ICL DNA damage in most eukaryotes analysed [67]. Interrogation of the *T. cruzi* genome database identified a single 2148 bp ORF located on chromosome 34 in the CL Brener Non-Esmeraldo-like haplotype with potential to encode for an 80.4 kDa enzyme belonging to this protein family (TcSNM1; TriTryp Gene ID: TcCLB.511003.50). The full-length protein is 42 % identical to the *T. brucei brucei* TREU927 and *T. evansi* strain STIB 805 orthologues (TriTryp Gene ID: Tb927.4.1480 and TevSTIB805.4.1530, respectively) and has 27–32 % identity to leishmanial sequences. In comparison to its *S. cerevisiae*, *A. thaliana* and *H. sapiens* counterparts, TcSNM1 displayed 17, 18 and 24 % sequence identity, respectively. As with other members of the SNM1/PSO2 family, TcSNM1 contains a metallo-ß-lactamase (MBL; PF12706) domain which here is found in its amino terminal region (residues 35 to 233), and a ß-CASP domain (named after its representative members **C**PSF, **A**rtemis, **S**NM1 and **P**SO2; PF10996) (residues 232–533) within which is found the DRMBL (DNA repair metallo-ß-lactamase; PF07522) domain (residues 474 to 533) (Figure 2A). These regions encompass 5 conserved motifs including a characteristic HxHxDH signature (motif 2), all centred on charged residues that coordinate zinc cofactor binding. In addition, the DRMBL domain also plays a role in defining nucleic acid specificity [68]. Analysis of this region revealed TcSNM1 contains a valine residue at the key point in this sequence (V532) indicating this trypanosomal enzyme is involved in DNA processing. Sequence motifs present in other members of the SNM1/PSO2 family, such as the the UBZ and PIP-like box PCNA interaction motifs of yeast PSO2 and human SNM1A, and the telomere or DNA-PKc binding signatures found in the carboxyl extensions of human SNM1B/Apollo or SNM1C/Artemis, are apparently absent from TcSNM1.

Using the Phyre2 package, a 3D approximation of TcSNM1 based on the human DNA cross-link repair 1A crystal structure (5AHR; Protein Data Bank) [69] was generated (Figure 2B): The model had 100% confidence over 323 amino acid (or 45%) coverage to the template. This revealed that in keeping with other MBL/β-CASP proteins, TcSNM1 can be divided into two distinct structural domains [69-71]. One domain forms a characteristic MBL-fold composed of 11 β sheets split across two regions and having a mixed arrangement. These two regions are sandwiched between two sets of α helices giving an αβ/βα organisation with this representing the MBL-fold. The other domain is inserted between the 10^{th} and 11^{th} β sheet of the MBL-fold and contains the β-CASP domain. Alignment of TcSNM1 model with the known structures of its human SNM1A and SNM1B/Apollo counterparts showed that several residues from motifs 2-4 in the parasite enzyme can be spatial positioned to facilitate binding of two zinc ions, with other amino acids, such as those found in motifs 1 and 5, supporting these interactions (Figure 2C): H87, H89 (both from motif 2), H171 (motif 3) and D193 (motif 4) plus two water/hydroxylate ions are postulated to mediate binding to one zinc ion while D193 along with D91, H92 (both from motif 2) and three water/hydroxylate ions may mediate binding to the second zinc atom [69]. Although TcSNM1 and the two human proteins have similar MBL-folds, they differ in the surface charge distribution. On the side of the proteins containing the zinc binding region, TcSNM1 is like SMN1B/Apollo and distinct from SNM1A: SNM1A has a predominant positive charge in this area whereas TcSNM1 and SMN1B/Apollo have

negatively and positively charged sections (Figure 2D) [69] with the zinc binding region (the active site) represents one of the positively charged areas. In contrast, the reverse side of both human proteins are predominantly negative while TcSNM1 is only moderately so.

To assess whether TcSNM1 is a *bona fide* SNM1/PSO2, the full-length gene was amplified and cloned into a yeast expression vector with the recombinant plasmid then introduced into wild type and *pso2* null mutant (*pso2Δ)* strains of *S. cerevisiae*. The susceptibility of these cells towards HN2 was evaluated with the resultant growth profiles allowing EC_{50} values to be determined (Figures 3A and B). *S. cerevisiae* cells lacking *pso2* were ~3-fold more susceptible to the nitrogen mustard than wild type. When TcSNM1 was expressed in either line, EC_{50} values like that noted for the wild type were obtained. This analysis was then extended to *T. brucei* by expressing a version of TcSNM1 tagged at its amino terminal with GFP in wild type and Tb*snm1* null mutant (Tb*snm1Δ)* lines (Figures 3C and D). *T. brucei* lacking SNM1 were ~10-fold more susceptible to HN2 than wild type, a phenotype that could be partly reversed by expression of the recombinant *T. cruzi* orthologue. Together, these experiments clearly show that the *T. cruzi* enzyme can readily complement for the *pso2Δ* and Tb*snm1Δ* mutations in *S. cerevisiae* and *T. brucei*, respectively, and establishes that TcSNM1 is a functional member of the PSO2/SNM1 family.

Other enzymes that often contribute to the repair of ICL lesions include MRE11 and CSB [30, 43, 72, 73]. Using the *T. brucei* sequences [41, 74, 75] as bait, the *T. cruzi* MRE11 and CSB counterparts were identified from the TriTryDB database. This revealed *T. cruzi* possesses a single 2250 bp ORF (TriTryp Gene ID: TcCLB.509319.60) located on chromosome 23 in the CL Brener Non-Esmeraldo-like haplotype with potential to code for an 82 kDa orthologue of MRE11 (designated here as TcMRE11). The full-length protein was 66 % identical to *T. brucei brucei* TREU927 MRE11 (TriTryp Gene ID: Tb927.02.4390) and has 37–42 % identity to its *S. cerevisiae*, *A. thaliana* and *H. sapiens* counterparts. As with other members of this family of nucleases, TcMRE11 contains a calcineurin-like phosphoesterase (PF00149) (residues 14-254) and Mre11 DNA binding (PF04152) (residues 297-470) domains (Figure S1A) with 3D approximation models (Figure S1B) showing the structural organisation of these regions. Comparison of TcMRE11 model with its *Chaetomium thermophilum* and human counterparts showed that several residues in the calcineurin-like phosphoesterase domain of the *T. cruzi* enzyme can be correctly positioned to facilitate binding of two manganese ions (Figure S1C) [76, 77]: In TcMRE11, binding to one manganese ion is postulated to involve the residues D21, H23, D62, H252 and two molecules of water with D60, N130, H218, N250 and water potentially contributing to the binding of the second metal atom.

In the case of CSB, searches identified a single 3345 bp ORF (Gene ID: TcCLB.506983.60) located on chromosome 11 in the *T. cruzi* CL Brener Non-Esmeraldo-like haplotype as having homology to the *T. brucei* orthologue. The full-length, 124 kDa protein (designated here as TcCSB) has 57 % identical to the *T. brucei* TREU927 CSB (TriTryp Gene ID: Tb927.7.4080) and 37–43 % identity to its *S. cerevisiae, A. thaliana* and *H. sapiens* counterparts. Based on its primary sequence, TcCSB contains two domains, with a SNF2 Family N-terminal (PF00176; residues 420 to 751) and helicase conserved C-terminal (PF00271; residues 804 to 916) domains (Figure S2A). These contain sequences that constitute Walker motifs (I to VI) which facilitate ATP binding and hydrolysis. Molecular modelling of TcCSB showed the structural organisation of these domains (Figure S2B) and the relative position of key regions involved in ATP processing (Figures S2A and C).

A similar *T. brucei* complementation strategy to that outlined for TcSNM1 was used to determine if the above TcMRE11 and TcCSB homology-based assignments were correct. GFP tagged version of TcMRE11 or TcCSB were expressed in the appropriate *T. brucei* null and the susceptibility of the rescue lines towards HN2 assessed (Figure 4). From the resultant dose response curves, the TbMRE11 and TbCSB null lines were 3- and 6-fold more susceptible, respectively, to the nitrogen mustard than wild type, phenotypes that could be partly reversed by expression of the recombinant *T. cruzi* orthologue. This confirms that the *T. cruzi* enzymes can readily complement for the corresponding mutation in *T. brucei* and establishes them as *bona fide* MRE11 or CSB family members.

3.3 Nuclear localisation of TcSNM1, TcMRE11 and TcCSB.

The PSORT and WoLFPSORT algorithms predict that TcSNM1 is found in the parasite's nucleus and targeted there via a "four pattern" KRRR nuclear localisation signal (residues 434-437) (Figure 2A). To confirm this, Tc*snm1* without its ATG start codon was amplified and ligated in-frame and downstream of a sequence coding for the enhanced green fluorescence protein (eGFP) lacking its STOP codon in the pTRIX *T. cruzi* expression vector backbone. The resultant construct was linearized, introduced into *T. cruzi* epimastigotes and recombinant parasite clones selected. Following PCR validation using gDNA and cDNA templates, and by Western blot analysis (Figures 5A and B) of the resultant lines, the GFP-TcSNM1 fluorescence pattern in *T. cruzi* epimastigotes and amastigotes was evaluated (Figure 5C). In both parasite stages, a signal coincidental with the larger, faint DAPI-staining compartment was observed in most (~90 %) cells confirming that TcSNM1 was found in the parasite nucleus: For the remaining ~10 % cells no GFP signal was detected while the smaller, intensely DAPI staining region corresponds to the kinetoplast, the mitochondrial genome of this organism. For *T. cruzi* epimastigotes, fluorescence was observed throughout the nucleus while in intracellular amastigotes, a punctate pattern was observed. This apparent difference in TcSNM1 localisation may reflect expression levels in the two parasite stages at the ribosomal array into which the expression vector has integrated into.

Based on the WolfPSORT and Euk-mPLoc 2.0 packages, TcMRE11 and TcCSB are predicted to have a nuclear localisation with PSORT and NucPred indicating that targeting of the latter may be via two "four pattern" nuclear localisation signal (296-KRRR-299 and 780-KRRR-784) (Figure S2A): No clear nuclear localisation signal(s) was identified in TcMRE11. To investigate the localisation, vectors that facilitate expression of GFP tagged version of TcMRE11 or TcCSB were made and electroporated into *T. cruzi* epimastigotes: In the ribosomal integrative expression system used, Tc*mre11* was tagged at its 5' end with *gfp* to generate GFP-TcMRE11 in the parasite while Tc*csb* was tagged at its 3' end to facilitate TcCSB-GFP expression. Following PCR validation using gDNA and cDNA templates (Figures 5A and B), and for TcMRE11, by Western blot analysis (Figure 1B), the pattern of fluorescence in GFP-TcMRE11 or TcCSB-GFP expressing *T. cruzi* epimastigote and amastigote was evaluated. For TcMRE11, a pattern like that observed for TcSNM1 was noted such that in ~90 % of cells a nuclear wide signal was detected in epimastigote parasites with a punctate pattern observed in amastigotes (Figure 5C): For the remaining ~10 % of cells, no GFP signal was detected). In the case of TcCSB-GFP expressing *T. cruzi* epimastigotes, about 0.5 % of cells displayed a fluorescence pattern throughout the nucleus with no signal detected in the remaining parasites (Figure 5D). Treatment of cultures with UV light (1200 J m^{-2}) followed by a 2-hour recuperation period increased the number of cells displaying the nuclear GFP signal to around 10 % indicating that this damaging treatment may affect recruitment or stability of TcCSB at lesion sites (Figure 5C & D). When the localisation studies were extended to

amastigotes, no fluorescence signal was observed in non-UV treated cells with UV treatment resulting in death of the host (and hence parasite) cells.

To examine whether HN2 treatment affects GFP-TcSNM1 or GFP-TcMRE11 expression, recombinant *T. cruzi* epimastigotes were treated with the ICL inducing agent (180 µM) and the fluorescence intensity of single cells followed with time: The low number of TcCSB-GFP expressing cells precluded such an analysis on this line. For TcSNM1 and by 4 hours post-treatment, the GFP signal started to increase above background peaking 10 hours into the experiment (Figure 6A). After this the fluorescence intensity declined and by 24 hours post-treatment were equivalent to pre-treatment cultures. As GFP-TcSNM1 is ectopically expressed in *T. cruzi* one potential mechanism involved in regulating enzyme levels may by through traits present in the protein sequence itself (e.g. post translational modification). Informatic analysis of TcSNM1 using UbPred and GPS-SUMO predicts that it contains several lysine residues which may undergo ubiquitation (K579, K632, K649 & K656) and/or sumolyation (K632) with one or more of these sites potentially playing a role in regulating protein abundance. Alternatively, the HN2 treatment itself may impose a selection for those cells that express sufficient protein for survival, such as by selecting those parasites with greater plasmid copies. For TcMRE11 and by 4 hours post-treatment, the GFP signal started to increase above background and reached a peak 10 hours into the experiment (Figure 6B). In contrast to GFP-TcSNM1 expression, GFP-TcMRE11 fluorescence remained at an elevated level 24 hours post-treatment.

3.4 Functional analysis of TcSNM1 in *T. cruzi*.

A gene disruption approach was used to investigate the function and importance of TcSNM1 in *T. cruzi*. Here, the 5' and 3' regions of the Tc*snm1* gene were cloned either side of a cassette containing blasticidin or puromycin resistance markers. The constructs were linearized, and the fragments transformed into *T. cruzi* X10/6 epimastigotes. As the *T. cruzi* X10/6 nuclear genome is diploid, to remove both allelic copies of the target gene two rounds of DNA transformation were performed to create heterozygous and then null mutant lines, with clonal populations selected in both cases.

To assess that the introduced DNA fragments had integrated into the correct genetic loci, PCRs were performed using gDNA as template: The primers and primer combinations used to generate amplicons specific for the intact gene or a disrupted allele and expected outcomes are shown in Tables S2 and 3 and Figure S3A. Using a primer combination designed to detect intact Tc*snm1*, PCR generated a band of the expected size (0.15 kb) from gDNA extracted wild type cells with no band observed in gDNA isolated from the Tc*snm1Δ* line (Figure S4A). In contrast, primer combinations used to detect the *pac* or *bla* interrupted alleles generated amplicons of the predicted size (both 0.6 kb) only from gDNA purified from the null line: No bands were amplified when using wild type gDNA as template.

To show that the null line no longer expressed Tc*snm1*, PCRs were performed on cDNA templates. When using primer combinations designed to detect a Tc*snm1* specific amplicon, a single band of 0.15 kb was observed in cDNA synthesised from total RNA extracted from wild type parasites with no band detected in material derived from Tc*snm1Δ* line (Figure S4A). To confirm that RNA had been extracted from both cell lines and that cDNA had indeed been made, control reactions amplifying Tc*tert* were conducted in parallel. For all tested samples, an amplicon of the expected size (~0.14 kb) was detected.

As *T. cruzi* epimastigotes null for TcSNM1 (the line was designated as Tc*snm1Δ*) could be readily generated, this nuclease is clearly not essential for the growth of this parasite stage. To phenotype the line, the growth characteristics of *T. cruzi* epimastigotes (Figure 7A) and amastigotes (Figure 7B) along with its ability to infect mammalian cells (Figure 7C) were assessed. In relation to the growth of insect form stage, a slight but statistically significant growth difference was noted with the mutant having a doubling time of around 23 hours as compared to the wild type being around 19 hours. For the other traits tested (amastigote growth and infectivity), TcSNM1 deficient parasites behaved like wild type. When the characterisation was extended to evaluate susceptibility the Tc*snm1Δ* line behaved similarly to wild type when challenged with the DNA damaging agents MMS, H_2O_2, HU, nitroureas (*e.g.,* semustine and carmustine) and platinum-based complexes (*e.g.,* carboplatin and cisplatin), and to the clinically used trypanocidal drugs nifurtimox, benznidazole and fexinidazole (Figures 1A and 7D). In contrast, Tc*snm1Δ* cells were shown to be more susceptible to nitrogen mustard (HN2) and aziridinyl compounds (triethylmelamine, ThioTEPA, mitomycin C, RH1 and CB1954) (Figures 1A and 7E), and UV light than controls (Figure 7D). Intriguingly, this altered phenotype toward HN2 was not observed in amastigotes as the TcSNM1 nulls were as equally susceptible to this ICL inducing agent as controls (Figure 7B).

To show that the HN2 susceptibility phenotype displayed by *T. cruzi* epimastigotes was due to TcSNM1 deficiency, a complementation strategy was used. A vector that facilitates constitutive ectopic expression of eGFP tagged Tc*snm1* was generated and introduced into Tc*snm1Δ* cells. Following PCR validation, the susceptibility of the complemented line toward HN2 was determined and the dose response curves/EC_{50} values compared with those obtained using the wild type and Tc*snm1* null mutants (Figure 7E). As previously noted, *T. cruzi* lacking SNM1 were ~3-fold more susceptible to the nitrogen mustard than wild type with ectopic expression of the tagged nuclease in the mutant partially complementing this phenotype.

3.5 Functional analysis of TcMRE11 and TcCSB in *T. cruzi*.

To assess the role of TcMRE11 or TcCSB in *T. cruzi*, gene disruption vectors containing the 5' and 3' regions from each gene flanking a blasticidin or puromycin resistance cassette were made. These were linearized and sequentially transformed into *T. cruzi* X10/6 epimastigotes to generate heterozygote then null mutant clonal populations: The predicted effect on the *T. cruzi* genome for each gene disruption event is depicted (Figures S3B and C). As for Tc*snm1*, PCR validations were performed on gDNA or cDNA templates with these amplifications confirming that each DNA fragment had integrated into the correct genetic locus and that the null parasites were no longer expressing full length Tc*mre11* (Figure S4B) or Tc*csb* (Figure S4C) gene transcripts.

As *T. cruzi* epimastigotes null for TcMRE11 (line designated as Tc*mre11Δ*) or TcCSB (line designated as Tc*csbΔ*) could be readily generated the encoded activities are not essential for the growth of this parasite stage. Phenotyping of the two lines revealed that the epimastigote (Figure 8A) and amastigote (Figure 8B) forms of both mutants displayed growth defects compared to controls. In the case of cells lacking TcMRE11, insect form parasites had a doubling time ~1.7-fold greater than wild type (Tc*mre11Δ* and wild type have doubling times of around 32 and 19 hours, respectively) while each infected mammalian cell contained about 3-fold fewer null mutant amastigotes relative to controls (on average, ~25 *T. cruzi* wild type amastigotes were counted in each infected mammalian cell with about 8 Tc*mre11Δ* amastigotes noted per infected mammalian cell). Cell cycle analysis of Tc*mre11Δ* epimastigotes that had been HU arrested at the G1/S boundary and then released following treatment withdrawal revealed fewer parasites were in the G2 phase relative to wild type and Tc*snm1Δ* (Figure S5).

As such, TcMRE11 null cells may progress more readily through the G2/M checkpoint such that any DNA damage is not repaired prior to the cell transiting into mitosis. Alternatively, fewer mutant cells may be able to progress through the S phase, perhaps due their death because of critical replication-associated damage, resulting in the detection of fewer G2 cells. For the Tc*csbΔ* line, a slight but statistically significant differences in epimastigote and amastigote growth were noted with the insect form nulls having a doubling time (around 23 hours) ~1.2-fold slower than controls while 40% fewer amastigotes (~15) were observed in each infected mammalian cell compared to controls (~25). To determine if lack of either DNA repair factor affected infectivity, the % of mammalian cells containing amastigotes was evaluated (Figure 8C). This revealed only about 6 % of the L6 myoblasts cells infected with Tc*mre11Δ* contained parasites, ~50 % lower than the value recorded for wild type and Tc*csbΔ* (12 to 13 %). Taken together, this indicates that Tc*csbΔ* trypomastigote form parasites can infect mammalian cells (at least L6 myoblasts) as efficiently as wild type with the intracellular amastigote forms growing slower than controls. In contrast, Tc*mre11Δ* trypomastigotes are compromised in their ability to invade mammalian cells and then divide as amastigotes.

When the susceptibility of epimastigotes lacking TcMRE11 to the DNA damaging agents MMS, H_2O_2, UV light and HU, and to the clinically used trypanocidal drugs nifurtimox, benznidazole and fexinidazole was evaluated, the mutant line behaved like wild type (Figure 8D): Note agents such as bleomycin that promote double strand DNA breaks and are commonly used in MRE11-based phenotype studies have no growth inhibitory effect on *T. cruzi*. When the screens were extended to ICL generating agents, TcMRE11-deficient parasites behaved similarly to the Tc*snm1Δ* line, being up to 5-fold more susceptible to HN2, triethylmelamine, ThioTEPA, mitomycin C, RH1 and CB1954 than controls (Figures 1A & 8E). Again, and in keeping with the Tc*snm1Δ* line, the altered HN2 susceptibility phenotype displayed by epimastigote form parasites was not observed in amastigotes with TcMRE11 nulls being as susceptible to this ICL inducing agent as controls (Figure 8B).

When the susceptibility characteristics of Tc*csbΔ* epimastigotes was examined, trypanosomes lacking TcCSB displayed wild type sensitivities towards HU, H_2O_2, nifurtimox and benznidazole but were more susceptible to UV light, MMS, fexinidazole, HN2 and the aziridinyls, triethylmelamine, ThioTEPA, mitomycin C, RH1 and CB1954 (Figures 1A, 8D & 8F). In contrast to the Tc*snm1Δ* and Tc*mre11* lines, the altered HN2 susceptibility phenotype displayed by Tc*csbΔ* epimastigotes extended to the intracellular, amastigote form parasites (Figure 8B).

To show that the HN2 susceptibility phenotype displayed by the Tc*mre11Δ* or Tc*csbΔ* epimastigotes was due to deficiency of the DNA factor, GFP tagged versions of each enzyme (GFP-TcMRE11 or TcCSB- GFP) were constitutively expressed in the corresponding null line (Figures 8E and F). Analysis of the dose response curves exhibited by the various lines toward the nitrogen mustard demonstrated that *T. cruzi* lacking MRE11 or TbCSB were ~3-fold more susceptible to HN2 than the control with ectopic expression of the tagged enzyme in the appropriate mutant line complementing (completely in the case of TcMRE11 or partially for TcCSB) for the susceptibility phenotype.

3.6 Assessing the interplay between ICL repair enzymes of *T. cruzi*.

To evaluate the functional relationship between components of the *T. cruzi* repair system, the genes encoding for TcMRE11 or TcCSB were disrupted using *hyg*- and *neo*-based vectors in Tc*snm1*-deficient *T. cruzi* epimastigotes to create a series of double null mutants. Following selection, all putative double null clones (Tc*snm1*ΔTc*mre11*Δ or Tc*snm1*ΔTc*csb*Δ) were

validated using the previously described PCR-based strategies with this confirming that the introduced DNA fragments had integrated into the correct site of the *T. cruzi* genome and that the recombinant parasites were no longer expressing either disrupted gene (Figures 6D and E). Despite numerous attempts lines lacking TcMRE11 and TcCSB failed to generate culturable cells suggesting that loss of this combination is synthetically lethal in *T. cruzi*.

Analysis of the double null mutants revealed that parasites lacking TcSNM1 and TcCSB grew at an equivalent rate as the single null lines while cells deficient in TcSNM1 and TcMRE11 exhibited a longer doubling time (approximately 50½ hours) than the *T. cruzi* Tc*mre11*Δ line which in turn had a longer doubling time than wild type (Figures 9A-C). Further analysis of the Tc*snm1*ΔTc*mre11*Δ line revealed it readily loses its ability to regulate cell cycle progression (Figure S5). Phenotypic screens were then conducted to investigate the susceptibility of the *T. cruzi* double null mutants to HN2 and investigate any epistatic/non-epistatic relationships between DNA repair factors (Figures 9D and E). These revealed that cells lacking TcSNM1 and TcMRE11 exhibit dose response sensitivities equivalent to Tc*snm1*Δ or Tc*mre11*Δ parasites with these recombinant lines being more susceptible to the ICL inducing agent than wild type (Figure 9D). The observed pattern indicates that TcSNM1 functions epistatically with TcMRE11. In contrast, *T. cruzi* lacking TbSNM1 and TbCSB were hypersensitive towards HN2 relative to wild type and the Tc*snm1*Δ or Tc*csb*Δ lines showing that these two enzymes function in a non-epistatic fashion and do not operate in the same ICL repair system (Figure 9E).

The double null mutant data shows that *T. cruzi* expresses at least two distinct ICL repair systems with one involving the concerted action of TbSNM1 and TbMRE11 with the other using TbCSB. It is tempting to speculate that given TcMRE11 functions predominantly in the trypanosomal homologous recombination system, the MRE11/SNM1- dependent system helps remove ICLs encountered during DNA replication while TcCSB plays a role transcription coupled ICL repair mechanisms.

4. Discussion

Cells are continually exposed to endogenous metabolites and environmental stimuli that can promote DNA damage. ICLs represent a dangerous lesion type that can accumulate in cell's genome and affect cell cycle progression. As rapidly growing cells are prone to this form of injury, compounds able to promote this damage are of interest in treatment of certain cancers and infectious diseases [20, 78, 79]. Here, we report that ICL-inducing agent display anti-parasitic activity towards *T. cruzi*. The most potent were compounds containing multiple aziridine groupings with structures previously shown to function as type I NTR-activated prodrugs (RH1, CB1954) [16, 64] being the most effective. We also observed that treatment of *T. cruzi* with HN2 caused phosphorylation of histone H2A to generate γH2A. Formation of this epigenetic marker in trypanosomes is indicative of accumulation of DSBs [65, 66] and confirms that in *T. cruzi*, HN2 can promote DNA damage within this parasite's nuclear genome.

Different ICL-inducing agents promote distinct cross linkages that can be distinguished based on the nucleotides which they covalently bind and the distorting effect they have on the DNA sugar-phosphate backbone [80]. These diverse lesions then influence which factors are deployed in their resolution [81, 82] with distorting ICLs detected and repaired throughout the cell cycle while non-helix distorting damage are predominantly encountered during the S phase [82, 83]. To survive such insults, all cells possess multiple DNA replication-independent and -dependent ICL repair mechanisms involving components from the 'classical' DNA repair

pathways alongside dedicated enzymes. Determining the makeup of these mechanisms in each species is complicated as evolutionary diverse organisms often utilise distinct factors at contrasting cell cycle stages to resolve ICLs [38, 84]. Based on mechanisms used by other eukaryotes, MRE11 and CSB from the HR and TC-NER pathways, respectively, alongside the SNM1/PSO2 nucleases have all been shown to contribute to ICL repair. Searches of genome databases identified *T. cruzi* gene orthologues coding for each of these enzymes with their deduced amino acid sequences containing the characteristic domain structure expected for a protein family member. To confirm these identifications, heterologous complementation studies were performed: this involved the ectopic expression of the *T.cruzi* enzyme in *S. cerevisiae* and/or *T. brucei* null for the corresponding activity then assessing the susceptibility phenotype of the resultant lines to HN2. In all cases, the *T. cruzi* proteins were able to partially (TcSNM1, TcCSB) or completely (TcMRE11) revert the susceptibility phenotype displayed by the null toward the nitrogen mustard thereby confirming that each trypanosomal enzyme does play a role in ICL repair. Further, and based on the partial complementation profiles, the mechanisms by which TcSNM1 and TcCSB are recruitment to damage sites and/or interact with other factors may be subtly different to those found in yeast and *T. brucei*. For example, *S. cerevisiae* PSO2 (and human SNM1A) contains ubiquitin binding zinc finger (UBZ) and PIP-like box motifs that may facilitate interaction with monoubiquitinated PCNA, a factor which accumulates in response to ICL-stalled replication forks [85]. However, equivalent sequences could not be discerned in TcSNM1. If such motifs are crucial for the recruitment of SNM1/PSO2 to damage sites then for the *T. cruzi* enzyme such interactions maybe facilitated by a highly divergent, non-conserved UBZ/PIP sequences, through alternative interaction domains, or *via* adapter protein(s).

The endogenous function of TcSNM1, TcMRE11 and TcCSB are non-essential to all *T. cruzi* life cycle stages tested: both copies of each gene could be readily deleted in epimastigote parasites, with each line able to transit through the metacyclic trypomastigote, amastigote and bloodstream trypomastigote forms. Additional phenotyping identified each nulls trait in relation to the growth properties of the two replicative forms, their ability to infect mammalian cells, and susceptibility to DNA damaging agents. Evaluation of Tc*snm1*-deficient *T. cruzi* revealed they behaved similarly to wild type in terms of amastigote growth and their ability to invade host cells. However, epimastigote trypanosomes displayed a slightly longer doubling time than wild type with susceptibility studies demonstrating the importance of SNM1 to the parasite in ICL repair: null cells were more susceptible to ICL-inducing DNA alkylating agents than controls while displaying a wild type-like sensitivity to most other genome damaging treatments. This ICL susceptibility phenotype solely due to loss of TcSNM1 as re-expression of this activity in nulls restored wild type-like sensitivity towards HN2. As with its *T. brucei* orthologue and in contrast to the mammalian and yeast equivalents, TcSNM1 functions only in the repair ICL lesions generated by chemicals that contain multiple nitrogen mustard or azirindinyl grouping [42, 86, 87]. This may be because platinum (e.g., cisplatin, carboplatin) or nitrosourea (e.g., semustine, carmustein *etc*) based compounds exert their trypanocidal effects by inducing forms of DNA damage (*e.g.*, DNA intrastrand crosslinks; monoalkylated bases) that are more readily resolved by other mechanisms, through activating signal transduction pathways, stimulating redox cycling, acting as enzyme inhibitors or alkylating other biological molecules [88-91].

When extended to Tc*csb*, phenotyping revealed bloodstream trypomastigote nulls were as infectious to mammalian cells as wild type although epimastigote and amastigote growth was compromised. These growth defects indicate that during its life cycle, *T.cruzi* is continually exposed to DNA damaging-stimuli that require CSB to resolve. As both the insect and

intracellular replicative forms of the parasite were cultured in the dark, it is unlikely that this damage involved formation of bulky, UV-induced photolesions. Instead, and as noted for the counterparts expressed by other eukaryotes, TcCSB may operate in cellular systems and pathways outside of its 'classical' role in TC-NER [92]. By comparison to its human orthologue [93-96], TcCSB may play a role in guiding transcription-coupled HR repair and base excision repair pathways that operate to resolve transcription-blocking DNA lesions (e.g., oxidised bases and DSBs) generated by reactive oxygen and nitrogen species. *T.cruzi* will be continually exposed to such DNA damaging agents throughout its cell and life cycles with such metabolites generated *via* endogenous reactions as well as through interactions with the environment, including those generated by a host cell as part of its response to intracellular form parasites. Further phenotyping demonstrated that TcCSB-deficient *T.cruzi* were more susceptible to UV irradiation, typical of yeast and mammalian cells that lack this activity [97, 98], and to nitrogen mustard/azirindinyl-containing ICL-inducing agents. This latter susceptibility phenotype could be partially rescued by re-expression of the deleted gene in the null which confirms that TcCSB plays a key role in *T.cruzi*'s ICL repair system.

Evaluation of TcMRE11 null *T. cruzi* demonstrated that the growth of the replicative insect and intracellular forms along with bloodstream trypomastigote infectivity were all significantly compromised. These traits indicate that throughout its life cycle, *T. cruzi* is continually exposed to endogenous and host cell derived metabolites which generate DSBs. The growth defect displayed by the null may stem from failure to detect such breaks at an appropriate checkpoint in the cell cycle with these lesions exerting their deleterious effects at an inappropriate stage of the cycle, potentially when the cell is unable to mount a suitable repair response [99-101]. Regarding its infectivity phenotype, TcMRE11-deficient bloodstream trypomastigotes may have reduced capacity to efficiently resolve any DSBs in their nuclear genome, damage potentially arising from host cell responses triggered during the invasion process [102]. Consequently, some host cells within a population may be effectively "resistant" to this null, at least for a while. Alternatively, null bloodstream trypomastigotes may be able to readily infect mammalian cells but are compromised in their ability to become established in the amastigote form. Intriguingly, upon entering a mammalian cell *T. cruzi* secretes factors that elicit production of reactive oxygen species by the host cell [103]. In TcMRE11 competent parasites, such oxidants act as growth factors [104, 105] but in cells lacking this activity, such metabolites may have a deleterious effect resulting in death and clearance of the trypanosome from the intracellular milieu. Further phenotyping revealed that TcMRE11-deficient *T. cruzi* epimastigotes were more susceptible nitrogen mustard/azirindinyl-containing ICL-inducing agents than controls, a phenotype that could be complemented by re-expression of Tc*mre11* in the null genetic background, again providing additional evidence that this HR factor plays a key role in the *T.cruzi* ICL repair system. Intriguingly, amastigotes lacking TcMRE11 were as sensitive to the archetypal nitrogen mustard, HN2, as controls, a trait also noted for TcSNM1-deficient cells but in contrast with the phenotype displayed by TcCSB null. The difference in the HN2 susceptibility profiles displayed by intracellular form *T.cruzi* null suggest that this parasite form expresses at least two distinct ICL repair mechanisms.

After establishing that TcSNM1, TcMRE11 and TcCSB are involved in *T. cruzi*'s ICL repair system, the interplay between these factors in resolving such damage was assessed. This involved extending the gene disruption studies to generate parasites null for two activities then evaluating each lines sensitivity to HN2. Trypanosomes lacking TcSNM1 and TcCSB or TcSNM1 and TcMRE11 could be readily generated although in the latter situation recombinant cells displayed a major growth defect suggesting these two enzymes may function synergistically. In contrast, and despite multiple attempts, a viable line lacking TcCSB and

TcMRE11 could not be created indicating these two DNA repair enzymes have a synthetically lethal interaction. The susceptibility screens allowed us to conclude that *T. cruzi* expresses at least two ICL repair systems with TcSNM1 and TcMRE11 cooperating in one mechanism to resolve HN2-induced lesions with TcCSB functioning in a second. This interaction pattern is distinct from that reported in *T. brucei* and yeast where the SNM1 orthologue (along with CSB in the case of *T. brucei*) operates to resolve transcription-blocking lesions with MRE11 participating in a DNA replication-dependent pathway [43, 72, 106]. In human cells, the functional relationship between these three enzymes appears to be more complex. For SNM1A, a human SNM1 isoform most like TcSNM1, a direct interaction with CSB has been established [30] indicating its involvement in replication-independent ICL repair mechanisms. However, recruitment of this SNM1A to ICL lesion sites parallels that of HR factors, including MRE11 [107] while mammalian cells depleted of this activity exhibit an increase in replication-associated DSBs [108]. Further, a second human isoform, SNM1B, can readily interact with MRE11 to help resolve ICLs that block replisome progression [109, 110]. Why equivalent factors are deployed in diverse ways to fix ICLs is unclear but could reflect the requirement of certain eukaryotic microorganisms to divide quickly. In rapidly dividing yeast and *T.brucei*, any DNA lesions, especially DSBs, must be speedily resolved if cell cycle progression is to swiftly occur. As HR often play a significant role in DSB repair then any ancillary interaction(s) that slow the rate at which this pathway operates may impinge on the rate of cell division. As such, accessory associations such as SNM1 interacting with MRE11, may not evolve or if established, could be selected against. In slower dividing cells such as mammalian cells and *T. cruzi*, the pressure to rapidly resolve DNA damage and quickly transit through the cell cycle is not so great, allowing such ancillary interactions to evolve.

We have now shown that resolution of ICLs by *T. cruzi* involves SNM1 along with components from the HR (MRE11) and TC-NER (CSB) pathways. Based on pathways reported in other organisms, we postulate that TcSNM1 and TcMRE11 are components of a HR-dependent ICL repair mechanism that resolves replisome-stalling lesions while TcCSB most likely functions in a replication-independent process to detect and guide removal of transcription-blocking crosslinks. Application of comparative and functional genomic approaches will facilitate in the identification of other factors that operate in this parasites ICL repair system and reveal any similarities to those expressed by mammalian host cells. This will also highlight any differences that could be of potential interest in the development of novel chemotherapies targeting CD. It is envisaged that unearthing an inhibitor(s) which selectively targets a *T. cruzi* ICL repair component could render the parasite more susceptible to compounds that promote this form of DNA damage.

Acknowledgement
We thank Prof Richard McCulloch (University of Glasgow) for the anti-*T. brucei* γH2A antisera, Prof Simon Croft (London School of Hygiene and Tropical Medicine) for benznidazole and nifurtimox, Prof John Kelly (London School of Hygiene and Tropical Medicine) for fexinidazole, and Prof Frank Guziec Jr (Southwestern University) for RH1. Monica Zavala Martinez was funded by a CONACyT PhD studentship.

5. References

[1] C. Bern, S.P. Montgomery, An estimate of the burden of Chagas disease in the United States, Clin Infect Dis, 49 (2009) e52-54.
[2] A. Requena-Mendez, E. Aldasoro, E. de Lazzari, E. Sicuri, M. Brown, D.A. Moore, J. Gascon, J. Munoz, Prevalence of Chagas disease in Latin-American migrants living in Europe: a systematic review and meta-analysis, PLoS Negl Trop Dis, 9 (2015) e0003540.

[3] S.R. Wilkinson, M.C. Taylor, D. Horn, J.M. Kelly, I. Cheeseman, A mechanism for cross-resistance to nifurtimox and benznidazole in trypanosomes, Proc Natl Acad Sci U S A, 105 (2008) 5022-5027.
[4] B.S. Hall, C. Bot, S.R. Wilkinson, Nifurtimox activation by trypanosomal type I nitroreductases generates cytotoxic nitrile metabolites, J Biol Chem, 286 (2011) 13088-13095.
[5] B.S. Hall, S.R. Wilkinson, Activation of benznidazole by trypanosomal type I nitroreductases results in glyoxal formation, Antimicrob Agents Chemother, 56 (2012) 115-123.
[6] E.G. Diaz de Toranzo, J.A. Castro, B.M. Franke de Cazzulo, J.J. Cazzulo, Interaction of benznidazole reactive metabolites with nuclear and kinetoplastic DNA, proteins and lipids from Trypanosoma cruzi, Experientia, 44 (1988) 880-881.
[7] A. Trochine, D.J. Creek, P. Faral-Tello, M.P. Barrett, C. Robello, Benznidazole biotransformation and multiple targets in Trypanosoma cruzi revealed by metabolomics, PLoS Negl Trop Dis, 8 (2014) e2844.
[8] S.G. Goijman, A.C. Frasch, A.O. Stoppani, Damage of Trypanosoma cruzi deoxyribonucleic acid by nitroheterocyclic drugs, Biochem Pharmacol, 34 (1985) 1457-1461.
[9] R.C. Ferreira, L.C. Ferreira, Mutagenicity of nifurtimox and benznidazole in the Salmonella/microsome assay, Braz J Med Biol Res, 19 (1986) 19-25.
[10] A. Buschini, L. Ferrarini, S. Franzoni, S. Galati, M. Lazzaretti, F. Mussi, C. Northfleet de Albuquerque, T. Maria Araujo Domingues Zucchi, P. Poli, Genotoxicity revaluation of three commercial nitroheterocyclic drugs: nifurtimox, benznidazole, and metronidazole, J Parasitol Res, 2009 (2009) 463575.
[11] M.C. Campos, J. Phelan, A.F. Francisco, M.C. Taylor, M.D. Lewis, A. Pain, T.G. Clark, J.M. Kelly, Genome-wide mutagenesis and multi-drug resistance in American trypanosomes induced by the front-line drug benznidazole, Sci Rep, 7 (2017) 14407.
[12] E. Rose, J.L. Carvalho, M. Hecht, Mechanisms of DNA repair in Trypanosoma cruzi: What do we know so far?, DNA repair, 91-92 (2020) 102873.
[13] M.A. Rajao, C. Furtado, C.L. Alves, D.G. Passos-Silva, M.B. de Moura, B.L. Schamber-Reis, M. Kunrath-Lima, A.A. Zuma, J.P. Vieira-da-Rocha, J.B. Garcia, I.C. Mendes, S.D. Pena, A.M. Macedo, G.R. Franco, N.C. de Souza-Pinto, M.H. de Medeiros, A.K. Cruz, M.C. Motta, S.M. Teixeira, C.R. Machado, Unveiling benznidazole's mechanism of action through overexpression of DNA repair proteins in Trypanosoma cruzi, Environmental and molecular mutagenesis, 55 (2014) 309-321.
[14] A. Dattani, I. Drammeh, A. Mahmood, M. Rahman, J. Szular, S.R. Wilkinson, Unraveling the antitrypanosomal mechanism of benznidazole and related 2-nitroimidazoles: From prodrug activation to DNA damage, Mol Microbiol, 116 (2021) 674-689.
[15] B.S. Hall, X. Wu, L. Hu, S.R. Wilkinson, Exploiting the drug-activating properties of a novel trypanosomal nitroreductase, Antimicrob Agents Chemother, 54 (2010) 1193-1199.
[16] C. Bot, B.S. Hall, N. Bashir, M.C. Taylor, N.A. Helsby, S.R. Wilkinson, Trypanocidal activity of aziridinyl nitrobenzamide prodrugs, Antimicrob Agents Chemother, 54 (2010) 4246-4252.
[17] L.Q. Hu, X.H. Wu, J.Y. Han, L. Chen, S.O. Vass, P. Browne, B.S. Hall, C. Bot, V. Gobalakrishnapillai, P.F. Searle, R.J. Knox, S.R. Wilkinson, Synthesis and structure-activity relationships of nitrobenzyl phosphoramide mustards as nitroreductase-activated prodrugs, Bioorganic & Medicinal Chemistry Letters, 21 (2011) 3986-3991.
[18] M.L. Dronkert, R. Kanaar, Repair of DNA interstrand cross-links, Mutat Res, 486 (2001) 217-247.
[19] P.J. McHugh, V.J. Spanswick, J.A. Hartley, Repair of DNA interstrand crosslinks: molecular mechanisms and clinical relevance, The lancet oncology, 2 (2001) 483-490.

[20] A.J. Deans, S.C. West, DNA interstrand crosslink repair and cancer, Nature reviews. Cancer, 11 (2011) 467-480.
[21] B. Sengerova, A.T. Wang, P.J. McHugh, Orchestrating the nucleases involved in DNA interstrand cross-link (ICL) repair, Cell cycle, 10 (2011) 3999-4008.
[22] K.F. Grossmann, A.M. Ward, M.E. Matkovic, A.E. Folias, R.E. Moses, S. cerevisiae has three pathways for DNA interstrand crosslink repair, Mutat Res, 487 (2001) 73-83.
[23] L.J. Barber, T.A. Ward, J.A. Hartley, P.J. McHugh, DNA interstrand cross-link repair in the Saccharomyces cerevisiae cell cycle: overlapping roles for PSO2 (SNM1) with MutS factors and EXO1 during S phase, Mol Cell Biol, 25 (2005) 2297-2309.
[24] P. Lehoczky, P.J. McHugh, M. Chovanec, DNA interstrand cross-link repair in Saccharomyces cerevisiae, FEMS Microbiol Rev, 31 (2007) 109-133.
[25] K. Kratz, B. Schopf, S. Kaden, A. Sendoel, R. Eberhard, C. Lademann, E. Cannavo, A.A. Sartori, M.O. Hengartner, J. Jiricny, Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents, Cell, 142 (2010) 77-88.
[26] C. MacKay, A.C. Declais, C. Lundin, A. Agostinho, A.J. Deans, T.J. MacArtney, K. Hofmann, A. Gartner, S.C. West, T. Helleday, D.M. Lilley, J. Rouse, Identification of KIAA1018/FAN1, a DNA repair nuclease recruited to DNA damage by monoubiquitinated FANCD2, Cell, 142 (2010) 65-76.
[27] D.L. Daee, E. Ferrari, S. Longerich, X.F. Zheng, X. Xue, D. Branzei, P. Sung, K. Myung, Rad5-dependent DNA repair functions of the Saccharomyces cerevisiae FANCM protein homolog Mph1, J Biol Chem, 287 (2012) 26563-26575.
[28] T.A. Ward, Z. Dudasova, S. Sarkar, M.R. Bhide, D. Vlasakova, M. Chovanec, P.J. McHugh, Components of a Fanconi-like pathway control Pso2-independent DNA interstrand crosslink repair in yeast, PLoS genetics, 8 (2012) e1002884.
[29] H. Zheng, X. Wang, A.J. Warren, R.J. Legerski, R.S. Nairn, J.W. Hamilton, L. Li, Nucleotide excision repair- and polymerase eta-mediated error-prone removal of mitomycin C interstrand cross-links, Mol Cell Biol, 23 (2003) 754-761.
[30] T. Iyama, S.Y. Lee, B.R. Berquist, O. Gileadi, V.A. Bohr, M.M. Seidman, P.J. McHugh, D.M. Wilson, 3rd, CSB interacts with SNM1A and promotes DNA interstrand crosslink processing, Nucleic Acids Res, 43 (2015) 247-258.
[31] S. Hashimoto, H. Anai, K. Hanada, Mechanisms of interstrand DNA crosslink repair and human disorders, Genes Environ, 38 (2016) 9.
[32] S. Sarkar, A.A. Davies, H.D. Ulrich, P.J. McHugh, DNA interstrand crosslink repair during G1 involves nucleotide excision repair and DNA polymerase zeta, EMBO J, 25 (2006) 1285-1294.
[33] T.V. Ho, A. Guainazzi, S.B. Derkunt, M. Enoiu, O.D. Scharer, Structure-dependent bypass of DNA interstrand crosslinks by translesion synthesis polymerases, Nucleic Acids Res, 39 (2011) 7455-7464.
[34] S. Sharma, C.E. Canman, REV1 and DNA polymerase zeta in DNA interstrand crosslink repair, Environmental and molecular mutagenesis, 53 (2012) 725-740.
[35] J. Zhao, A. Jain, R.R. Iyer, P.L. Modrich, K.M. Vasquez, Mismatch repair and nucleotide excision repair proteins cooperate in the recognition of DNA interstrand crosslinks, Nucleic Acids Res, 37 (2009) 4420-4429.
[36] B.S. Thoma, M. Wakasugi, J. Christensen, M.C. Reddy, K.M. Vasquez, Human XPC-hHR23B interacts with XPA-RPA in the recognition of triplex-directed psoralen DNA interstrand crosslinks, Nucleic Acids Res, 33 (2005) 2993-3001.
[37] K.M. Vasquez, J. Christensen, L. Li, R.A. Finch, P.M. Glazer, Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions, Proc Natl Acad Sci U S A, 99 (2002) 5848-5853.

[38] H. Dong, D.W. Nebert, E.A. Bruford, D.C. Thompson, H. Joenje, V. Vasiliou, Update of the human and mouse Fanconi anemia genes, Hum Genomics, 9 (2015) 32.
[39] R. Ceccaldi, P. Sarangi, A.D. D'Andrea, The Fanconi anaemia pathway: new players and new functions, Nat Rev Mol Cell Biol, 17 (2016) 337-349.
[40] P.J. McHugh, T.A. Ward, M. Chovanec, A prototypical Fanconi anemia pathway in lower eukaryotes?, Cell cycle, 11 (2012) 3739-3744.
[41] C.R. Machado, J.P. Vieira-da-Rocha, I.C. Mendes, M.A. Rajao, L. Marcello, M. Bitar, M.G. Drummond, P. Grynberg, D.A. Oliveira, C. Marques, B. Van Houten, R. McCulloch, Nucleotide excision repair in Trypanosoma brucei: specialization of transcription-coupled repair due to multigenic transcription, Mol Microbiol, 92 (2014) 756-776.
[42] J.A. Sullivan, J.L. Tong, M. Wong, A. Kumar, H. Sarkar, S. Ali, I. Hussein, I. Zaman, E.L. Meredith, N.A. Helsby, L. Hu, S.R. Wilkinson, Unravelling the role of SNM1 in the DNA repair system of Trypanosoma brucei, Mol Microbiol, 96 (2015) 827-838.
[43] A. Dattani, S.R. Wilkinson, Deciphering the interstrand crosslink DNA repair network expressed by Trypanosoma brucei, DNA repair, 78 (2019) 154-166.
[44] J. Stevens, H. Noyes, W. Gibson, The evolution of trypanosomes infecting humans and primates, Mem Inst Oswaldo Cruz, 93 (1998) 669-676.
[45] G. Kendall, A.F. Wilderspin, F. Ashall, M.A. Miles, J.M. Kelly, Trypanosoma cruzi glycosomal glyceraldehyde-3-phosphate dehydrogenase does not conform to the 'hotspot' topogenic signal model, EMBO J, 9 (1990) 2751-2758.
[46] H. Hirumi, K. Hirumi, Continuous cultivation of Trypanosoma brucei blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers, J Parasitol, 75 (1989) 985-989.
[47] K. Nakai, P. Horton, PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization, Trends Biochem Sci, 24 (1999) 34-36.
[48] P. Horton, K.J. Park, T. Obayashi, N. Fujita, H. Harada, C.J. Adams-Collier, K. Nakai, WoLF PSORT: protein localization predictor, Nucleic Acids Res, 35 (2007) W585-587.
[49] M. Brameier, A. Krings, R.M. MacCallum, NucPred--predicting nuclear localization of proteins, Bioinformatics, 23 (2007) 1159-1160.
[50] K.C. Chou, H.B. Shen, A new method for predicting the subcellular localization of eukaryotic proteins with both single and multiple sites: Euk-mPLoc 2.0, PLoS One, 5 (2010) e9931.
[51] P. Radivojac, V. Vacic, C. Haynes, R.R. Cocklin, A. Mohan, J.W. Heyen, M.G. Goebl, L.M. Iakoucheva, Identification, analysis, and prediction of protein ubiquitination sites, Proteins, 78 (2010) 365-380.
[52] Q. Zhao, Y. Xie, Y. Zheng, S. Jiang, W. Liu, W. Mu, Z. Liu, Y. Zhao, Y. Xue, J. Ren, GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs, Nucleic Acids Res, 42 (2014) W325-330.
[53] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis, Nat Protoc, 10 (2015) 845-858.
[54] F. Olmo, F.C. Costa, G.S. Mann, M.C. Taylor, J.M. Kelly, Optimising genetic transformation of Trypanosoma cruzi using hydroxyurea-induced cell-cycle synchronisation, Mol Biochem Parasitol, 226 (2018) 34-36.
[55] G. Schumann Burkard, P. Jutzi, I. Roditi, Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters, Mol Biochem Parasitol, 175 (2011) 91-94.
[56] L. Pacheco-Lugo, Y. Diaz-Olmos, J. Saenz-Garcia, C.M. Probst, W.D. DaRocha, Effective gene delivery to Trypanosoma cruzi epimastigotes through nucleofection, Parasitol Int, 66 (2017) 236-239.

[57] S.R. Wilkinson, D.J. Meyer, M.C. Taylor, E.V. Bromley, M.A. Miles, J.M. Kelly, The Trypanosoma cruzi enzyme TcGPXI is a glycosomal peroxidase and can be linked to trypanothione reduction by glutathione or tryparedoxin, J Biol Chem, 277 (2002) 17062-17071.
[58] M.C. Taylor, H. Kaur, B. Blessington, J.M. Kelly, S.R. Wilkinson, Validation of spermidine synthase as a drug target in African trypanosomes, Biochem J, 409 (2008) 563-569.
[59] T.V. Novoselova, R.S. Rose, H.M. Marks, J.A. Sullivan, SUMOylation regulates the homologous to E6-AP carboxyl terminus (HECT) ubiquitin ligase Rsp5p, J Biol Chem, 288 (2013) 10308-10317.
[60] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis, Nat Methods, 9 (2012) 671-675.
[61] J.A. Stortz, T.D. Serafim, S. Alsford, J. Wilkes, F. Fernandez-Cortes, G. Hamilton, E. Briggs, L. Lemgruber, D. Horn, J.C. Mottram, R. McCulloch, Genome-wide and protein kinase-focused RNAi screens reveal conserved and novel damage response pathways in Trypanosoma brucei, PLoS Pathog, 13 (2017) e1006477.
[62] E. Briggs, K. Crouch, L. Lemgruber, G. Hamilton, C. Lapsley, R. McCulloch, Trypanosoma brucei ribonuclease H2A is an essential R-loop processing enzyme whose loss causes DNA damage during transcription initiation and antigenic variation, Nucleic Acids Res, 47 (2019) 9180-9197.
[63] V. Hannaert, M.A. Albert, D.J. Rigden, M.T. da Silva Giotto, O. Thiemann, R.C. Garratt, J. Van Roy, F.R. Opperdoes, P.A. Michels, Kinetic characterization, structure modelling studies and crystallization of Trypanosoma brucei enolase, Eur J Biochem, 270 (2003) 3205-3213.
[64] E.L. Meredith, A. Kumar, A. Konno, J. Szular, S. Alsford, K. Seifert, D. Horn, S.R. Wilkinson, Distinct activation mechanisms trigger the trypanocidal activity of DNA damaging prodrugs, Mol Microbiol, 106 (2017) 207-222.
[65] O. Fernandez-Capetillo, A. Lee, M. Nussenzweig, A. Nussenzweig, H2AX: the histone guardian of the genome, DNA repair, 3 (2004) 959-967.
[66] L. Glover, D. Horn, Trypanosomal histone gammaH2A and the DNA damage response, Mol Biochem Parasitol, 183 (2012) 78-83.
[67] E. Cattell, B. Sengerova, P.J. McHugh, The SNM1/Pso2 family of ICL repair nucleases: from yeast to man, Environmental and molecular mutagenesis, 51 (2010) 635-645.
[68] I. Callebaut, D. Moshous, J.P. Mornon, J.P. de Villartay, Metallo-beta-lactamase fold within nucleic acids processing enzymes: the beta-CASP family, Nucleic Acids Res, 30 (2002) 3592-3601.
[69] C.K. Allerston, S.Y. Lee, J.A. Newman, C.J. Schofield, P.J. McHugh, O. Gileadi, The structures of the SNM1A and SNM1B/Apollo nuclease domains reveal a potential basis for their distinct DNA processing activities, Nucleic Acids Res, 43 (2015) 11047-11060.
[70] C.R. Mandel, S. Kaneko, H. Zhang, D. Gebauer, V. Vethantham, J.L. Manley, L. Tong, Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease, Nature, 444 (2006) 953-956.
[71] A. Dorleans, I. Li de la Sierra-Gallay, J. Piton, L. Zig, L. Gilet, H. Putzer, C. Condon, Molecular basis for the recognition and cleavage of RNA by the bifunctional 5'-3' exo/endoribonuclease RNase J, Structure, 19 (2011) 1252-1261.
[72] A.F. Lam, B.O. Krogh, L.S. Symington, Unique and overlapping functions of the Exo1, Mre11 and Pso2 nucleases in DNA repair, DNA repair, 7 (2008) 655-662.
[73] M. Enoiu, J. Jiricny, O.D. Scharer, Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway involving transcription-coupled repair and translesion synthesis, Nucleic Acids Res, 40 (2012) 8953-8964.

[74] N.P. Robinson, R. McCulloch, C. Conway, A. Browitt, J.D. Barry, Inactivation of Mre11 does not affect VSG gene duplication mediated by homologous recombination in Trypanosoma brucei, J Biol Chem, 277 (2002) 26185-26193.
[75] K.S. Tan, S.T. Leal, G.A. Cross, Trypanosoma brucei MRE11 is non-essential but influences growth, homologous recombination and DNA double-strand break repair, Mol Biochem Parasitol, 125 (2002) 11-21.
[76] F.U. Seifert, K. Lammens, K.P. Hopfner, Structure of the catalytic domain of Mre11 from Chaetomium thermophilum, Acta Crystallogr F Struct Biol Commun, 71 (2015) 752-757.
[77] Y.B. Park, J. Chae, Y.C. Kim, Y. Cho, Crystal structure of human Mre11: understanding tumorigenic mutations, Structure, 19 (2011) 1591-1602.
[78] Y. Chen, L. Hu, Design of anticancer prodrugs for reductive activation, Med Res Rev, 29 (2009) 29-64.
[79] S.R. Wilkinson, C. Bot, J.M. Kelly, B.S. Hall, Trypanocidal activity of nitroaromatic prodrugs: current treatments and future perspectives, Curr Top Med Chem, 11 (2011) 2072-2084.
[80] H.B. Rycenga, D.T. Long, The evolving role of DNA inter-strand crosslinks in chemotherapy, Curr Opin Pharmacol, 41 (2018) 20-26.
[81] V. Beljanski, L.G. Marzilli, P.W. Doetsch, DNA damage-processing pathways involved in the eukaryotic cellular response to anticancer DNA cross-linking drugs, Mol Pharmacol, 65 (2004) 1496-1506.
[82] M.B. Smeaton, E.M. Hlavin, T. McGregor Mason, A.M. Noronha, C.J. Wilds, P.S. Miller, Distortion-dependent unhooking of interstrand cross-links in mammalian cell extracts, Biochemistry, 47 (2008) 9920-9930.
[83] J. Zhang, J.M. Dewar, M. Budzowska, A. Motnenko, M.A. Cohn, J.C. Walter, DNA interstrand cross-link repair requires replication-fork convergence, Nat Struct Mol Biol, 22 (2015) 242-247.
[84] M. McVey, Strategies for DNA interstrand crosslink repair: insights from worms, flies, frogs, and slime molds, Environmental and molecular mutagenesis, 51 (2010) 646-658.
[85] K. Yang, G.L. Moldovan, A.D. D'Andrea, RAD18-dependent recruitment of SNM1A to DNA repair complexes by a ubiquitin-binding zinc finger, J Biol Chem, 285 (2010) 19085-19091.
[86] M.L. Dronkert, J. de Wit, M. Boeve, M.L. Vasconcelos, H. van Steeg, T.L. Tan, J.H. Hoeijmakers, R. Kanaar, Disruption of mouse SNM1 causes increased sensitivity to the DNA interstrand cross-linking agent mitomycin C, Mol Cell Biol, 20 (2000) 4553-4561.
[87] D. Bonatto, L.F. Revers, M. Brendel, J.A. Henriques, The eukaryotic Pso2/Snm1/Artemis proteins and their function as genomic and cellular caretakers, Braz J Med Biol Res, 38 (2005) 321-334.
[88] K.U. Schallreuter, F.K. Gleason, J.M. Wood, The mechanism of action of the nitrosourea anti-tumor drugs on thioredoxin reductase, glutathione reductase and ribonucleotide reductase, Biochim Biophys Acta, 1054 (1990) 14-20.
[89] G. Pagano, P. Manini, D. Bagchi, Oxidative stress-related mechanisms are associated with xenobiotics exerting excess toxicity to Fanconi anemia cells, Environ Health Perspect, 111 (2003) 1699-1703.
[90] Z.H. Siddik, Cisplatin: mode of cytotoxic action and molecular basis of resistance, Oncogene, 22 (2003) 7265-7279.
[91] C.A. Rabik, M.E. Dolan, Molecular mechanisms of resistance and toxicity associated with platinating agents, Cancer treatment reviews, 33 (2007) 9-23.
[92] V. Tiwari, B.A. Baptiste, M.N. Okur, V.A. Bohr, Current and emerging roles of Cockayne syndrome group B (CSB) protein, Nucleic Acids Res, 49 (2021) 2418-2434.

[93] T. Stevnsner, M. Muftuoglu, M.D. Aamann, V.A. Bohr, The role of Cockayne Syndrome group B (CSB) protein in base excision repair and aging, Mech Ageing Dev, 129 (2008) 441-448.
[94] H. Menoni, J.H. Hoeijmakers, W. Vermeulen, Nucleotide excision repair-initiating proteins bind to oxidative DNA lesions in vivo, J Cell Biol, 199 (2012) 1037-1046.
[95] N.L. Batenburg, E.L. Thompson, E.A. Hendrickson, X.D. Zhu, Cockayne syndrome group B protein regulates DNA double-strand break repair and checkpoint activation, EMBO J, 34 (2015) 1399-1416.
[96] L. Wei, A.S. Levine, L. Lan, Transcription-coupled homologous recombination after oxidative damage, DNA repair, 44 (2016) 76-80.
[97] C. Troelstra, A. van Gool, J. de Wit, W. Vermeulen, D. Bootsma, J.H. Hoeijmakers, ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes, Cell, 71 (1992) 939-953.
[98] A.J. van Gool, R. Verhage, S.M. Swagemakers, P. van de Putte, J. Brouwer, C. Troelstra, D. Bootsma, J.H. Hoeijmakers, RAD26, the functional S. cerevisiae homolog of the Cockayne syndrome B gene ERCC6, EMBO J, 13 (1994) 5361-5369.
[99] C.T. Carson, R.A. Schwartz, T.H. Stracker, C.E. Lilley, D.V. Lee, M.D. Weitzman, The Mre11 complex is required for ATM activation and the G2/M checkpoint, EMBO J, 22 (2003) 6610-6620.
[100] E. Olson, C.J. Nievera, E. Liu, A.Y. Lee, L. Chen, X. Wu, The Mre11 complex mediates the S-phase checkpoint through an interaction with replication protein A, Mol Cell Biol, 27 (2007) 6053-6067.
[101] K.M. Garner, A. Eastman, Variations in Mre11/Rad50/Nbs1 status and DNA damage-induced S-phase arrest in the cell lines of the NCI60 panel, BMC Cancer, 11 (2011) 206:201-213.
[102] G.R. Goes, P.S. Rocha, A.R. Diniz, P.H. Aguiar, C.R. Machado, L.Q. Vieira, Trypanosoma cruzi Needs a Signal Provided by Reactive Oxygen Species to Infect Macrophages, PLoS Negl Trop Dis, 10 (2016) e0004555.
[103] G.P. Dos Santos, F.M. Abukawa, N. Souza-Melo, L.M. Alcantara, P. Bittencourt-Cunha, C.B. Moraes, B.K. Jha, B.S. McGwire, N.S. Moretti, S. Schenkman, Cyclophilin 19 secreted in the host cell cytosol by Trypanosoma cruzi promotes ROS production required for parasite growth, Cell Microbiol, 23 (2021) e13295.
[104] C.N. Paiva, D.F. Feijo, F.F. Dutra, V.C. Carneiro, G.B. Freitas, L.S. Alves, J. Mesquita, G.B. Fortes, R.T. Figueiredo, H.S. Souza, M.R. Fantappie, J. Lannes-Vieira, M.T. Bozza, Oxidative stress fuels Trypanosoma cruzi infection in mice, J Clin Invest, 122 (2012) 2531-2542.
[105] P.P. Dias, R.F. Capila, N.F. do Couto, D. Estrada, F.R. Gadelha, R. Radi, L. Piacenza, L.O. Andrade, Cardiomyocyte oxidants production may signal to T. cruzi intracellular development, PLoS Negl Trop Dis, 11 (2017) e0005852.
[106] F.M. Munari, L.F. Revers, J.M. Cardone, B.F. Immich, D.J. Moura, T.N. Guecheva, D. Bonatto, J.P. Laurino, J. Saffi, M. Brendel, J.A. Henriques, Sak1 kinase interacts with Pso2 nuclease in response to DNA damage induced by interstrand crosslink-inducing agents in Saccharomyces cerevisiae, Journal of photochemistry and photobiology. B, Biology, 130 (2014) 241-253.
[107] C.T. Richie, C. Peterson, T. Lu, W.N. Hittelman, P.B. Carpenter, R.J. Legerski, hSnm1 colocalizes and physically associates with 53BP1 before and after DNA damage, Mol Cell Biol, 22 (2002) 8635-8647.
[108] A.T. Wang, B. Sengerova, E. Cattell, T. Inagawa, J.M. Hartley, K. Kiakos, N.A. Burgess-Brown, L.P. Swift, J.H. Enzlin, C.J. Schofield, O. Gileadi, J.A. Hartley, P.J. McHugh, Human

SNM1A and XPF-ERCC1 collaborate to initiate DNA interstrand cross-link repair, Genes Dev, 25 (2011) 1859-1870.
[109] J.B. Bae, S.S. Mukhopadhyay, L. Liu, N. Zhang, J. Tan, S. Akhter, X. Liu, X. Shen, L. Li, R.J. Legerski, Snm1B/Apollo mediates replication fork collapse and S Phase checkpoint activation in response to DNA interstrand cross-links, Oncogene, 27 (2008) 5045-5056.
[110] J.M. Mason, I. Das, M. Arlt, N. Patel, S. Kraftson, T.W. Glover, J.M. Sekiguchi, The SNM1B/APOLLO DNA nuclease functions in resolution of replication stress and maintenance of common fragile site stability, Hum Mol Genet, 22 (2013) 4901-4913.

Figures

A

compound	*T. cruzi* (EC_{50})			
	wild type	*Tcsnm1Δ*	*Tcmre11Δ*	*TccsbΔ*
Nitrogen mustards				
melphalan, chlorambucil, trofosamide, cyclophosphamide, bendmustine, ifosfamide	>500	nd	nd	nd
mechlorethamine	114.8 ± 1.6	38.9 ± 1.7 (3.6)*	28.3 ± 1.6 (4.1)*	21.5 ± 0.6 (5.3)*
prednimustine	216.5 ± 21.9	nd	nd	nd
Aziridines				
triethylmelamine	4.3 ± 0.1	2.2 ± 0.1 (2.0)*	2.3 ± 0.1 (1.9)*	2.9 ± 0.2 (1.5)*
ThioTEPA	41.0 ± 1.0	12.2 ± 0.6 (3.4)*	10.6 ± 1.6 (3.9)*	34.3 ± 1.9 (1.2)
mitomycin C	3.9 ± 0.3	1.5 ± 0.1 (2.6)*	0.9 ± 0.03 (4.3)*	2.1 ± 0.2 (1.9)*
RH1 (in nM)	2.4 ± 0.05	0.6 ± 0.04 (4)*	0.5 ± 0.03 (4.8)*	0.2 ± 0.01 (12)*
CB1954 (in nM)	341.5 ± 35.5	72.1 ± 2.3 (4.7)*	91.4 ± 3.1 (3.7)*	41.1 ± 4.9 (8.3)*
Platinum-based				
carboplatin	75.3 ± 4.0	64.1 ± 3.6 (1.2)	nd	nd
cisplatin	152.1 ± 7.4	149.8 ± 5.3 (1.0)	nd	nd
Nitrosoureas				
semustine	218.0 ± 6.8	257.5 ± 5.9 (0.8)	nd	nd
carmustine	34.9 ± 0.9	30.8 ± 1.4 (1.1)	nd	nd
lomustine	39.1 ± 1.4	nd	nd	nd
nimustine	66.5 ± 12.3	nd	nd	nd
streptozotocin	54.4 ± 27.4	nd	nd	nd
Alkyl sulfonate				
busulfan	>500	nd	nd	nd

B

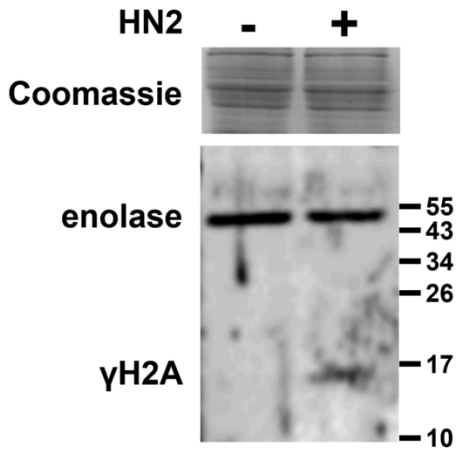


Figure 1: Effect of ICL inducing agents on *T. cruzi* epimastigotes.

A. Dose response curves of *T. cruzi* X10.6 (wild type) epimastigotes and the *T. cruzi* Tc*snm1Δ*, Tc*mre11Δ* and Tc*csbΔ* null lines to various ICL inducing agents was determined from which EC_{50} values (given in µM except where noted) were extrapolated. All values are means ± standard deviation from four independent growth inhibition assays. The fold difference in susceptibility between wild type and the null lines are given in parentheses. The asterisk indicates significant differences in susceptibility (P<0.0001) between wild type and genetically modified cells, as assessed by Student's *t* test (GraphPad Software). nd is not determined.

B. Mechlorethamine (HN2) promotes DNA damage in *T. cruzi*. Cell lysates from *T. cruzi* wild type cultured overnight in the presence (+) of 100 µM mechlorethamine were analysed by western blot using anti-*T. brucei* γH2A antiserum: non-treated (-) cells were analysed in parallel. An anti-*T. brucei* enolase antiserum and a Coomassie stained gel were used as loading controls. Molecular weight markers are in kilodaltons.

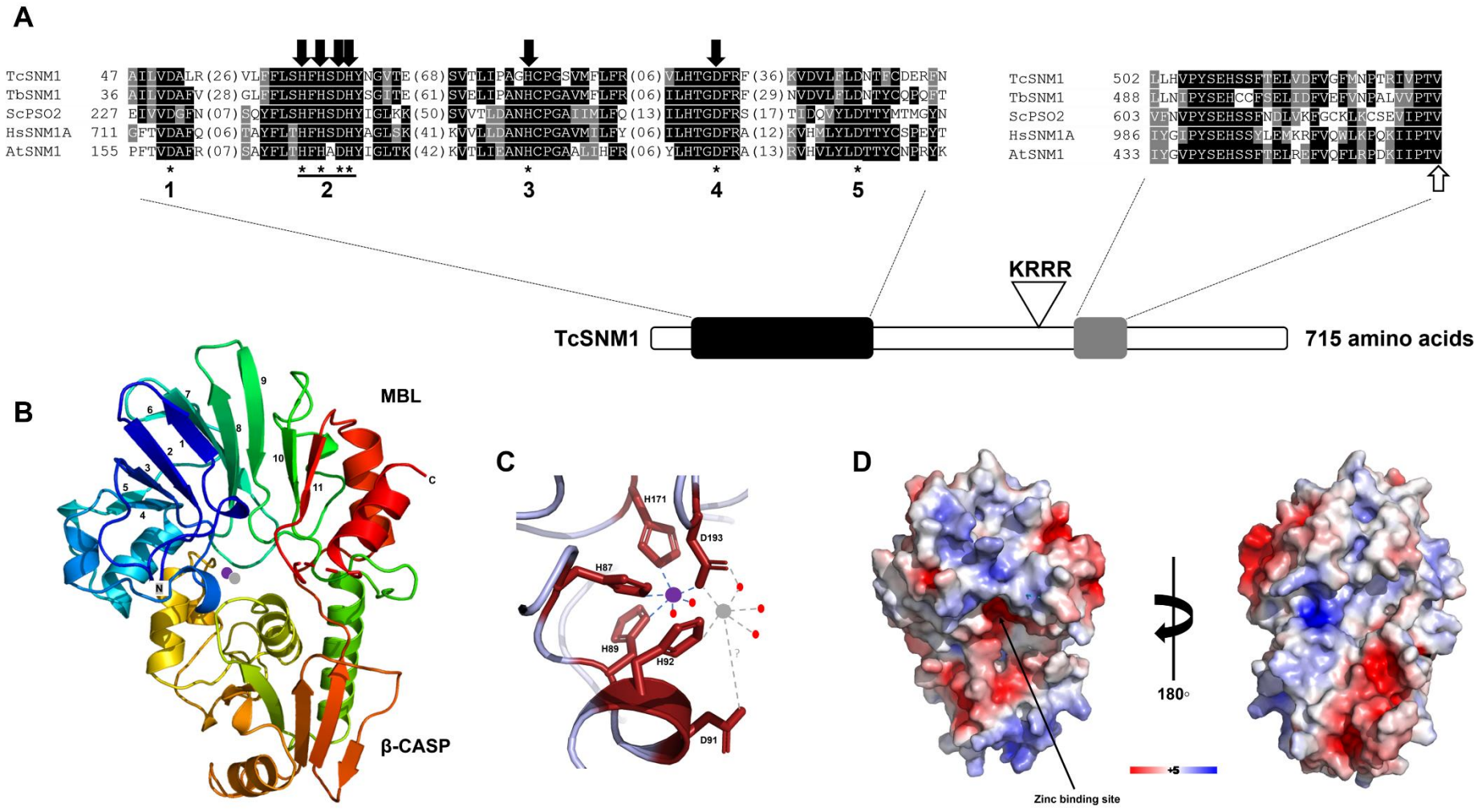
Figure 2. Informatic analysis of TcSNM1

A. The sequence corresponding to the Lactamase_B_2 Beta-lactamase superfamily (MBL; PF12706; black box) and DNA repair metallo-beta-lactamase (DRMBL; PF07522; grey box) domains of TcSNM1 were aligned with other members of the SNM1/PSO2 family of nucleases. Residues that are identical in at least three of the five sequences and conserved substitutions are highlighted in black and grey, respectively. Amino acids postulated to help coordinate zinc co-factor binding are denoted by a down arrow with the upward pointing arrow highlighting a residue (V532) that distinguishes DNA from RNA processing metallo-β-lactamases. The KRRR motif corresponds to a putative nuclear 'pattern 4' targeting signal while motifs 1 to 5 represent regions typically found in members of the SNM1/PSO2 family. The sequences aligned are *T. cruzi* TcSNM1 (GenBank RNC43493), *T. brucei* TbSNM1 (XP_844298), *Saccharomyces cerevisiae* ScPSO2 (NP_013857), *Homo sapiens* HsSNM1A (NP_001258745) and *Arabidopsis thaliana* AtSNM1 (NP_189302).

B. A 3D approximation model of TcSNM1 (tritrypdb.org: TCSYLVIO_003789) based on the human SNM1A (PDB: 5AHR; [66]) crystal structure was constructed using Phyre2 and visualized using PyMOL: The model has 100% confidence over 323 amino acid (or 45%) coverage to the template. The amino (N; in blue) and carboxyl (C; in red) termini are indicated. The numbers (1 to 11) highlight the β-sheets present in the MBL domain (top) with the region containing the DRMBL domain (bottom) found as an insert between β-sheets 10 and 11 of the MBL domain. The zinc ions are shown as: 1. a purple sphere, with this representing the metal ion mapped onto the 5AHR sequence, and 2. a grey sphere, with this corresponding to the putative site where a second metal ion is believed to bind.

C. The assignment of residues involved in zinc binding was made following comparison against the human SNM1 and SNM1B (PDB: 5AHO) sequences, respectively [66]. The residues from motifs 2 (H87; H89), 3 (H171) and 4 (D193) involved in binding (blue dotted line) to the 'mapped' zinc ion (purple sphere) and those from motifs 2 (D91; H92) and 4 (D193) proposed to bind (grey dotted line) the 2nd putative zinc ion (grey sphere) are shown. The red spheres correspond to activated water/hydroxylate ions [66].

D. The electrostatic surface potential of TcSNM1 shown in two orientations. Regions in red are more electronegative with blue representing more electropositive sites. The zinc binding site is indicated.


A
TcSNM1 47 AILVDALR(26)VLFFLSHFHSDHYNGVTE(68)SVTLIPAGHCPGSVMFLFR(06)VLHTGDFRF(36)KVDVLFLDNTFCDERFN
TbSNM1 36 AILVDAFV(28)GLFFLSHFHSDHYSGITE(61)SVELIPANHCPGAVMFLFR(06)ILHTGDFRF(29)NVDVLFLDNTYCQPQFT
ScPSO2 227 EIVVDGFN(07)SQYFLSHFHSDHYIGLKK(50)SVVTLLDANHCPGAIIMLFQ(13)ILHTGDFRS(17)TIDQVYLDTTYMTMGYN
HsSNM1A 711 CFTVDAFQ(06)TAYFLTHFHSDHYAGLSK(41)KVVLLLDANHCPGAVMILFY(06)ILHTGDFRA(12)KVHMLYLDTTYCSPEYT
AtSNM1 155 PFTVDAFR(07)SAYFLTHFHADHYIGLTK(42)KVTLIFANHCPGAALIHFR(06)YLHTGDFRA(13)RVHVLYLDTTYCNPRYK
1 2 3 4 5
TcSNM1 502 LLHVPYSEHSSFTELVDFVGFMNPTRIVPTV
TbSNM1 488 LLNIPYSEHCCFSELIDFVPFVNPALVVPTV
ScPSO2 603 VFNVPYSEHSSFNDLVKFGCKLKCSEVIPTV
HsSNM1A 986 IYGIPYSEHSSYLEMKRFVQWLKPQKIIPTV
AtSNM1 433 IYGVPYSEHSSFTELREFVQFLRPDKIIPTV
KRRR
TcSNM1
715 amino acids
B
MBL
β-CASP
N
C
C
H171
D193
H87
H89
H92
D91
D
180°
+5
Zinc binding site

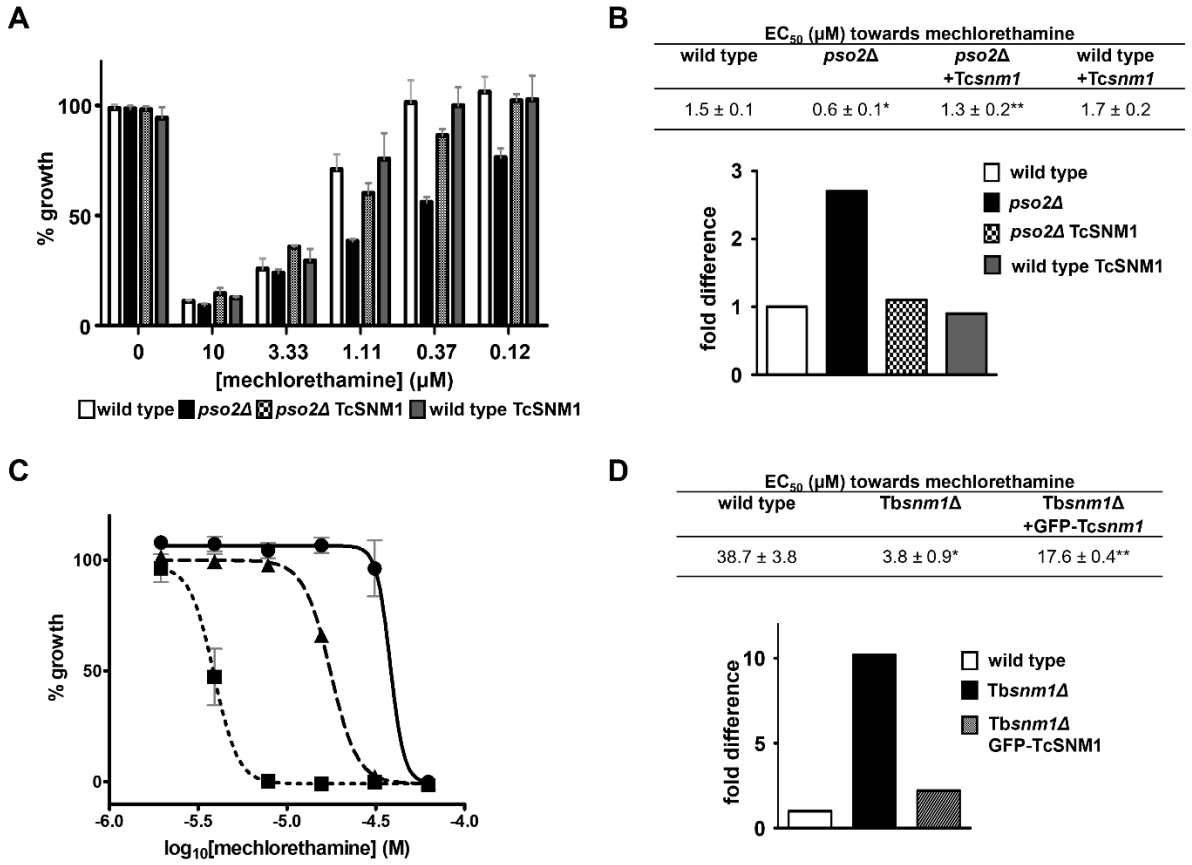


Figure 3. TcSNM1 complements the yeast *pso2*Δ and *T. brucei snm1*Δ susceptibility phenotype.

A. Susceptibility of wild type (black bar), *pso2*Δ (white bar), *pso2*Δ expressing Tc*snm1* (hatched bar) and wild type expressing Tc*snm1* (grey bar) yeast strains towards different concentrations of mechlorethamine.

B. The susceptibility of the yeast strains toward mechlorethamine, as judged by their EC_{50} values, were determined, and expressed as a fold difference relative to wild type (in parenthesis). All data are mean values ± standard deviations from experiments performed in quadruplicate. Statistically significance differences (P<0.0001) between *wild type and the *pso2Δ* strain, and **between the *pso2Δ* and the *pso2Δ* ectopically expressing TcSNM1 lines were assessed by the Student's *t*-test.

C. Dose response curves of *T. brucei* wild type (●), Tb*snm1Δ* (■) and Tb*snm1Δ* expressing GFP-TcSNM1 (▲) towards mechlorethamine.

D. The susceptibility of the *T. brucei* lines, as judged by their EC_{50} values, were compared, and expressed as a fold difference relative to wild type. All data are mean values ± standard deviations from experiments performed in quadruplicate. Statistically significance differences (P<0.0001) between *wild type and the Tb*snm1Δ* line, and **between the Tb*snm1Δ* and the Tb*snm1Δ* ectopically expressing GFP-TcSNM1 lines were assessed by the Student's *t*-test.

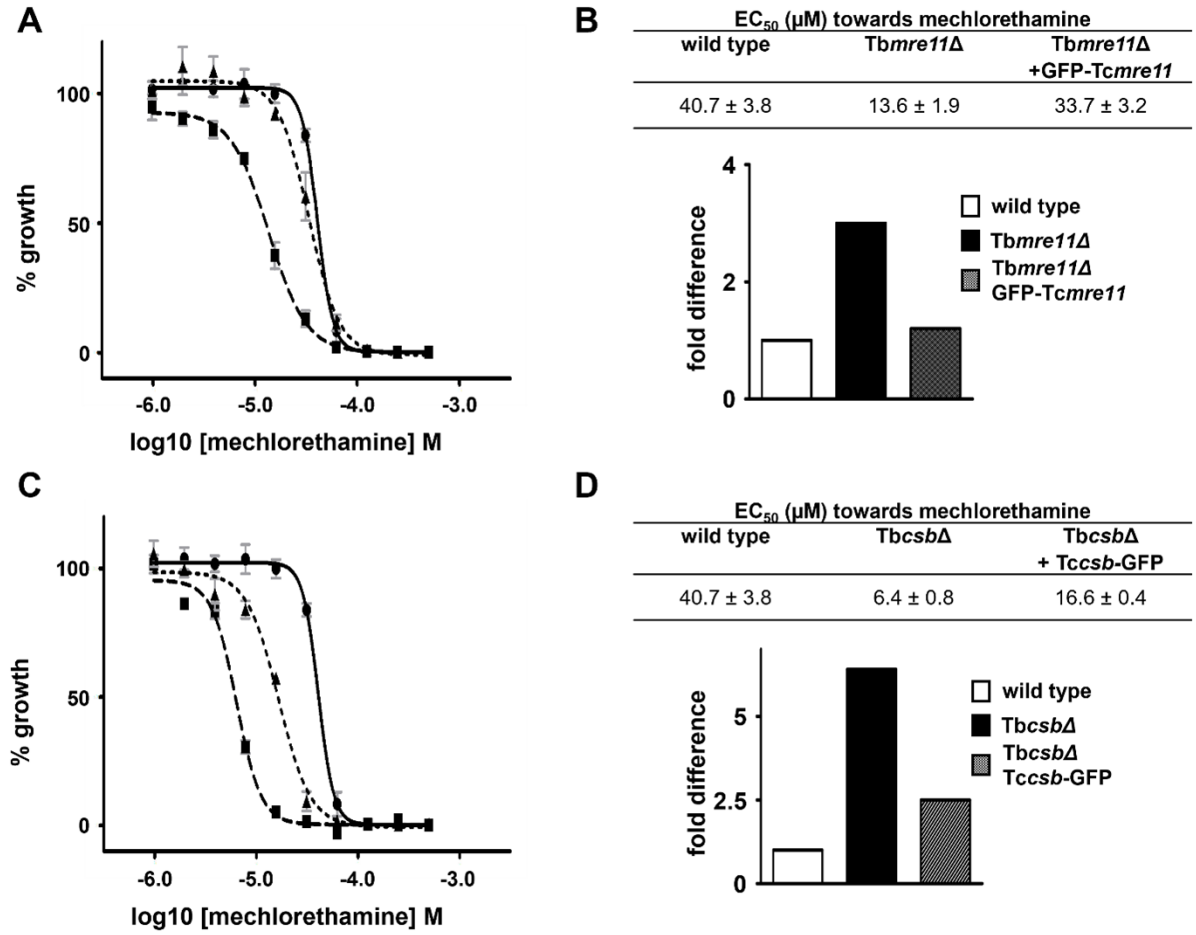
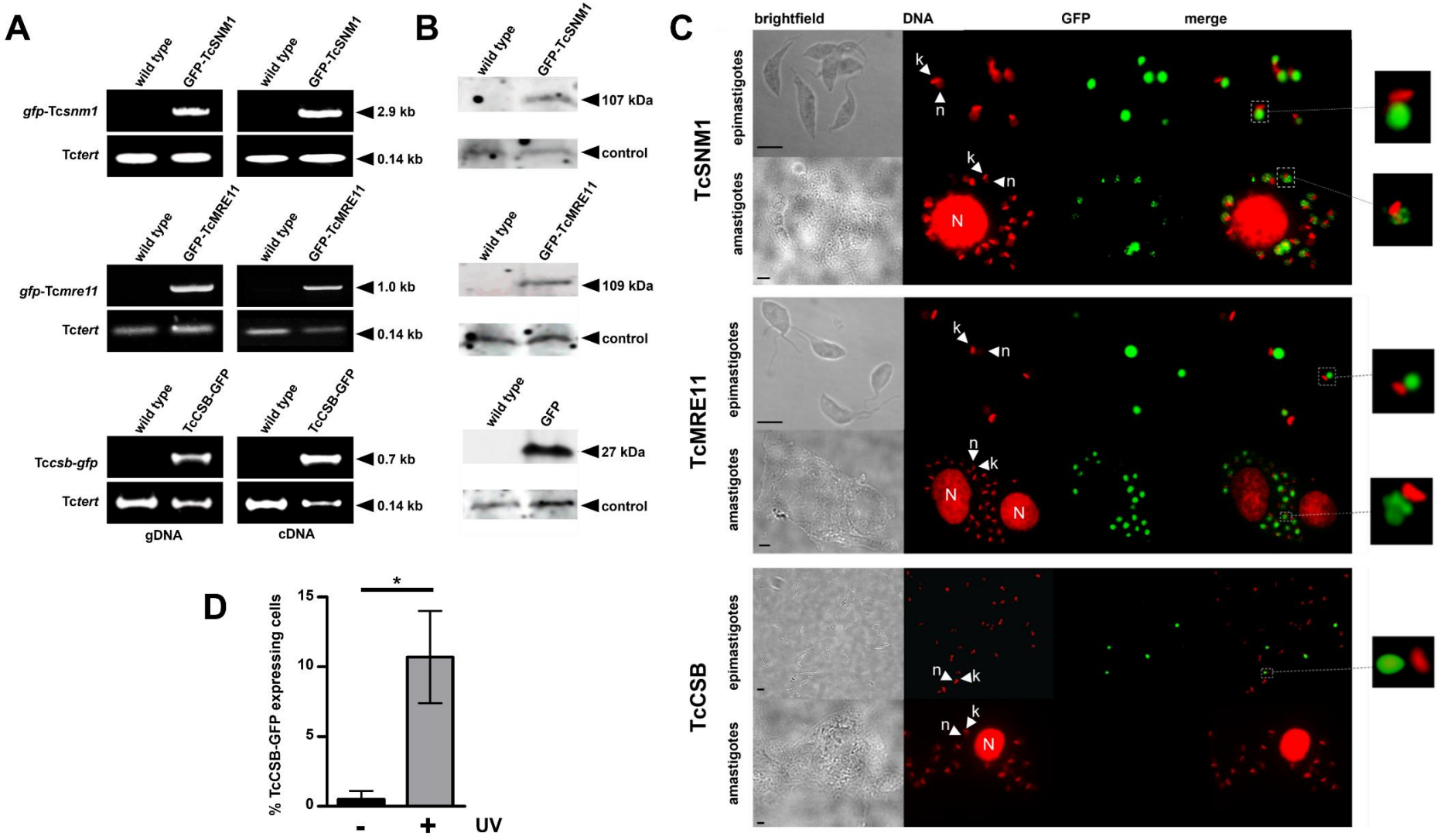


Figure 4. TcMRE11 and TcCSB complement the mechlorethamine susceptibility phenotype displayed by the corresponding *T. brucei* null mutant.

A & C. Dose response curves of *T. brucei* wild type (●), Tb*mre11Δ* (■) and Tb*mre11Δ* expressing GFP-TcMRE11 (▲) (A) or *T. brucei* wild type (●), Tb*csbΔ* (■) and Tb*csb Δ* expressing TcCSB-GFP(▲) (C) towards mechlorethamine.

B & D. The susceptibility of the *T. brucei* lines, as judged by their EC_{50} values, were compared, and expressed as a fold difference relative to wild type. All data are mean values ± standard deviations from experiments performed in quadruplicate. Statistically significance differences (P<0.0001) between *wild type and the Tb*mre11Δ* or Tb*csbΔ* lines, and **between Tb*mre11Δ* or Tb*csbΔ* and their corresponding complemented lines were assessed by the Student's *t*-test.

Figure 5. Localisation of TcSNM1, TcMRE11 and TcCSB in *T. cruzi*.
A. Amplicons (in kb) corresponding to *gfp*-Tc*snm1* (upper panel) *gfp*-Tc*mre11* (middle panel) or *Tccsb-gfp* (lower panel) were generated from template genomic DNA (gDNA) or cDNAs derived from total RNA extracted from the *T. cruzi* lines indicated. The integrity of DNA samples was evaluated by amplification of a 0.14 kb control fragment, Tc*tert*. The primer sequences and combinations used for each amplification are listed in Tables S2 and S3, respectively.
B. Expression of GFP-TcSNM1 (upper panel) or GFP-TcMRE11 (middle panel) was examined by probing blots containing cell lysates from *T. cruzi* wild type and the tagged GFP expressing line using an anti-GFP antibody. Protein from 2×10^7 cells was loaded in each track and an unidentified cross-reactive epitope (lower panel) was used as loading control.
C. *T. cruzi* epimastigotes and amastigotes ectopically expressing GFP-TcSNM1 (upper panels), GFP-TcMRE11 (middle panels) or TcCSB-GFP (lower panels) were fixed, co-stained with DAPI (DNA; red) and examined by fluorescence microscope. The pattern of colocalization (merged), including a close up of the nuclei from each cell type, are shown. In parasites expressing the fusion protein, the GFP signal was coincidental with the nucleus (n; large, faint DAPI spot); the smaller, bright DAPI spot corresponds to the kinetoplast (k), the trypanosome mitochondrial genome with N representing the host cell nucleus. No TcCSB-GFP signal was observed in *T. cruzi* amastigotes while the number of epimastigotes expressing this fusion was enhanced from ~2.5 % to ~20 % following UV treatment: The *T. cruzi* epimastigote TcCSB-GFP image shown is following UV treatment. Bar, 5 μm.
D. Fixed, UV-treated (+) *T. cruzi* epimastigotes engineered to express TcCSB-GFP were co-stained with DAPI and examined by fluorescence microscope. For each experiment, the fluorescence pattern displayed by ~100 cells was assessed and the % of cells presenting with a GFP signal determined: this signal was nuclear in all cases. Untreated cells (-) were analysed in parallel. Each data point represents the mean % TcCSB-GFP expressing cells ± standard deviation from four independent experiments. *Statistically significance difference (P<0.001) between untreated and UV-treated cells was assessed by the Student's *t*-test. In total, 43 of the 402 (or 10.7 %) UV-treated trypanosomes exhibited a TcCSB-GFP signal while in untreated controls, only 2 out of 400 (or 0.5 %) parasites displayed this fluorescence pattern.


A
wild type
GFP-TcSNM1
gfp-Tcsnm1
Tctert
2.9 kb
0.14 kb
wild type
GFP-TcMRE11
gfp-Tcmre11
Tctert
1.0 kb
0.14 kb
wild type
TcCSB-GFP
Tccsb-gfp
Tctert
0.7 kb
0.14 kb
gDNA
cDNA
B
wild type
GFP-TcSNM1
107 kDa
control
wild type
GFP-TcMRE11
109 kDa
control
wild type
GFP
27 kDa
control
C
brightfield
DNA
GFP
merge
TcSNM1
epimastigotes
amastigotes
TcMRE11
epimastigotes
amastigotes
TcCSB
epimastigotes
amastigotes
k
n
N
D
% TcCSB-GFP expressing cells
15
10
5
0
*
-
+
UV

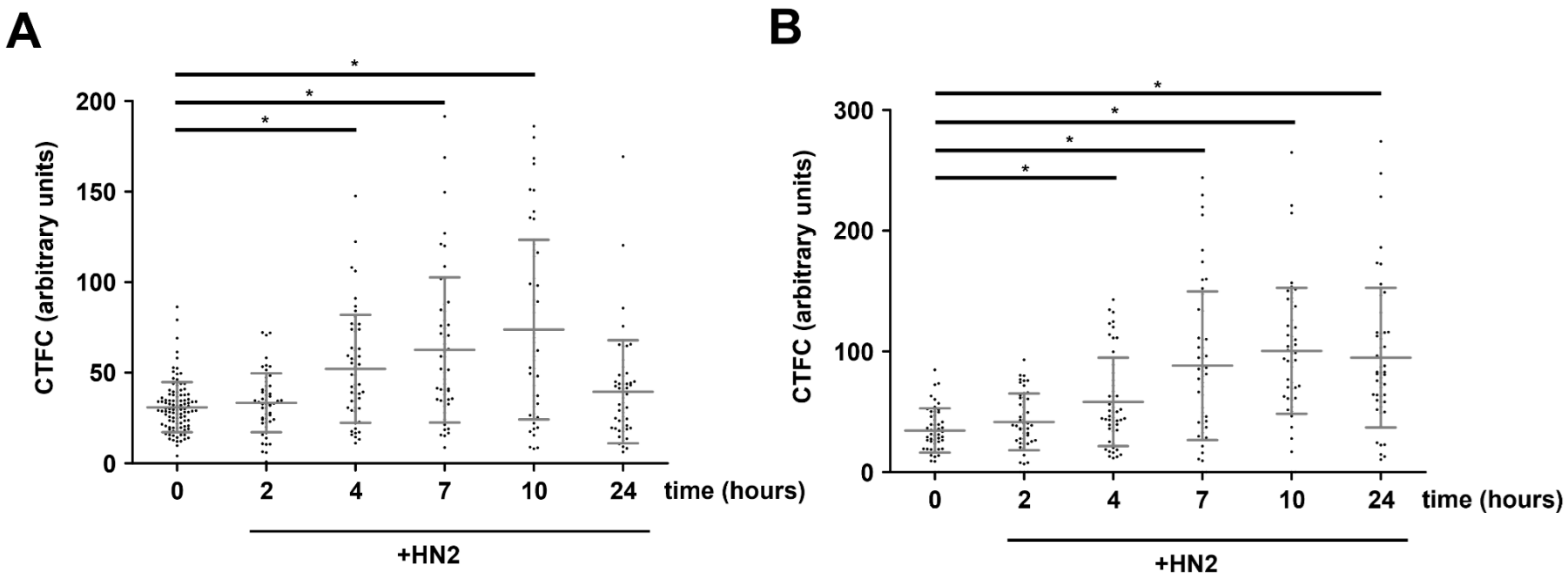


Figure 6. Effect of mechlorethamine on TcSNM1 and TcMRE11 expression in *T. cruzi.*

A & B. The corrected total cell fluorescence (CTFC) of individual *T. cruzi* cells expressing GFP tagged TcSNM1 (**A**) or TcMRE11 (**B**) was determined at time intervals following addition of mechlorethamine (HN2) to cultures. Each data point represents the fluorescence of an individual nuclei (n = 50 to 110), with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. Statistically significance difference between each pair was assessed by the Student's *t*-test; *corresponds to $P<0.0001$.

Figure 7. Characterisation of *T. cruzi* SNM1 null line.

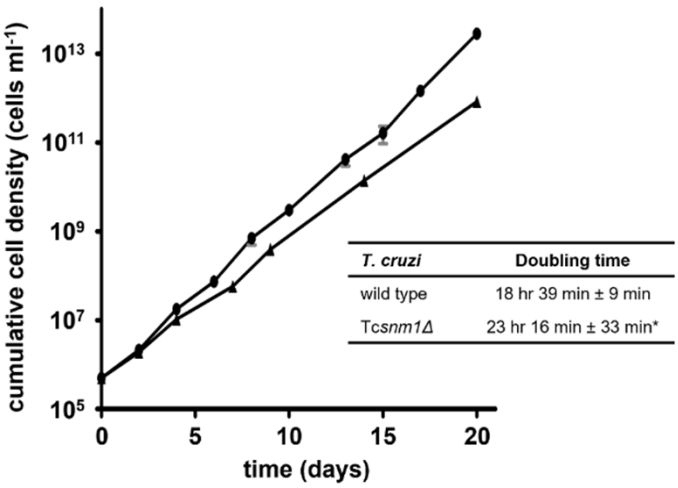
A. The cumulative cell density of *T. cruzi* Tc*snm1*Δ (▲) was followed for 20 days and compared against wild type (●) cultures grown in parallel. Each data point represents the mean cell density ± standard deviation from three independent cultures. Doubling times (insert) were calculated from the cumulative cell density growth curves using the online tool https://doubling-time.com/compute.php. *A statistically significance difference (P<0.0001) as assessed by the Student's *t*-test in doubling time was noted between the null mutant line and wild type.

B & C. L6 myoblasts were infected overnight with *T. cruzi* bloodstream form trypomastigotes. After washing to removed non-internalised parasites, infections were incubated for a further 4 days. Cells were DAPI stained and the number of amastigotes per infected L6 myoblasts (**B**) and number of infected L6 myoblasts (**C**) determined. For **B**, amastigotes were cultured in the presence (+) of mechlorethamine (7.5 μM) and compared against untreated controls (-). The data shown in **B** represents a measure of amastigote growth while **C** represents a measure of parasite infectivity. Each data point represents the mean cell density ± standard deviation from four independent cultures.

D. Susceptibility as judged by EC_{50} values (in μM except where noted) of *T. cruzi* X10.6 (wild type) epimastigotes and the *T. cruzi* Tc*snm1Δ* null line to DNA damaging and trypanocidal agents were determined. All values are means ± standard deviation from four independent growth inhibition assays. The fold difference in susceptibility between wild type and the null line is given in parentheses. The asterisk indicates significant differences in susceptibility (P<0.0001) between wild type and genetically modified cells, as assessed by Student's *t* test (GraphPad Software).

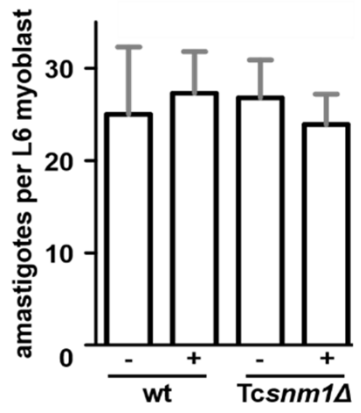
E. Dose response curves and EC_{50} values (in μM) (insert) of *T. cruzi* Tc*snm1Δ* (■) and its complemented counterpart (▲) towards mechlorethamine was determined and compared against wild type cells (●). All data points are mean values ± standard deviations from experiments performed in quadruplicate. The susceptibility of the *T. cruzi* null and complement lines against mechlorethamine, as judged by their EC_{50} values, were compared, and expressed as a fold difference relative to wild type (right hand panels). Statistically significance differences (P<0.0001) between *wild type and the null mutant, and **between the null mutant and the complemented line were assessed by the Student's *t*-test.

A

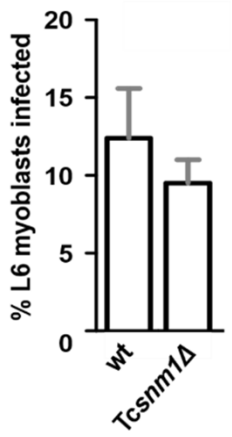


T. cruzi	Doubling time
wild type	18 hr 39 min ± 9 min
Tc*snm1Δ*	23 hr 16 min ± 33 min*

B



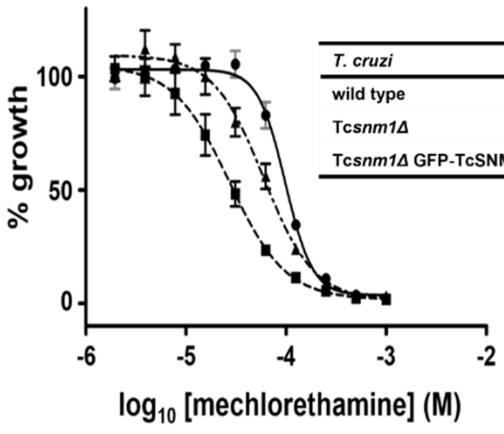
C



D

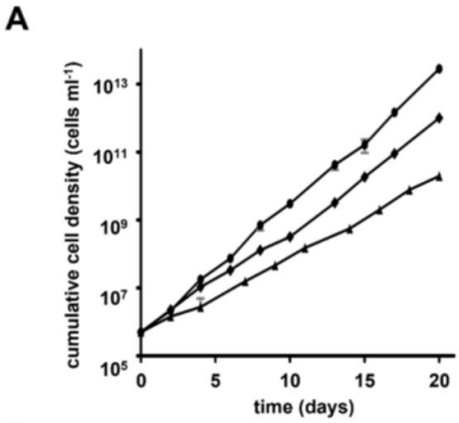
compound	*T. cruzi* (EC_{50})	
	wild type	*Tcsnm1Δ*
MMS	96.5 ± 2.0	107.6 ± 4.6 (0.9)
Hydroxyurea	227.0 ± 4.8	224.5 ± 8.1 (1)
UV irradiation (in J m^{-2})	30.0 ± 1.6	19.0 ± 1.0 (1.6)*
H_2O_2	132.3 ± 0.9	139.2 ± 1.3 (1)
benznidazole	23.2 ± 1.0	24.6 ± 1.4 (0.9)
fexinidazole	6.0 ± 0.1	5.4 ± 0.1 (1.1)
nifurtimox	5.7 ± 0.4	5.2 ± 0.3 (1.1)

E

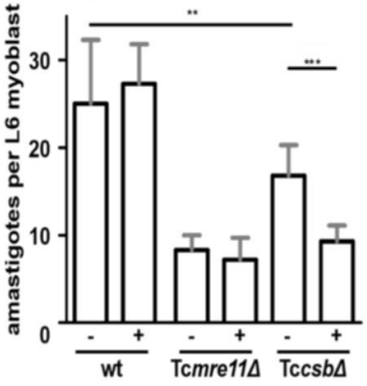


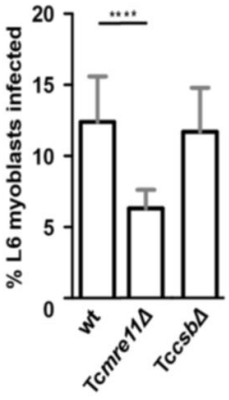
T. cruzi	EC_{50} (μM)
wild type	89.9 ± 7.9
Tc*snm1Δ*	38.6 ± 2.8*
Tc*snm1Δ* GFP-TcSNM1	80.7 ± 4.3**

Figure 8. Functional analysis of TcMRE11 and TcCSB in *T. cruzi*.
A. The cumulative cell density of *T. cruzi* Tc*mre11*Δ (▲) and Tc*csb*Δ (♦) lines were followed for 20 days and compared against wild type *T. cruzi* (●) cultures grown in parallel. Each data point represents the mean cell density ± standard deviation from three independent cultures. Doubling times (insert) were calculated from the cumulative cell density growth curves using the online tool https://doubling-time.com/compute.php. *Statistically significance differences (P<0.0001) as assessed by the Student's *t*-test in doubling times was between each null mutant line and the wild type
B & C. L6 myoblasts were infected with *T. cruzi* bloodstream form trypomastigotes overnight cultures. Infections were incubated for a further 4 days. Cells were DAPI stained and the number of amastigotes per infected L6 myoblasts (**B**) and number of infected L6 myoblasts (**C**) counted. For **B**, amastigotes were cultured in the presence (+) of mechlorethamine (HN2; 7.5 μM) and compared against untreated controls (-). The data shown in **B** represents a measure of amastigote growth while **C** represents a measure of parasite infectivity. Each data point represents the mean cell density ± standard deviation from four independent cultures. Statistically significance differences between pairs were assessed by the Student's *t*-test; *, **, *** and **** corresponds to P = 0.0043, 0.0029, 0.0089 and 0.0123, respectively.
D. Susceptibility as judged by EC_{50} values (in μM except where noted) of *T. cruzi* X10.6 (wild type) epimastigotes and the *T. cruzi* Tc*mre11Δ or* Tc*csbΔ* null lines to DNA damaging and trypanocidal agents were determined. All values are means ± standard deviation from four independent growth inhibition assays. The fold difference in susceptibility between wild type and a null line is given in parentheses. The asterisk indicates significant differences in susceptibility (P<0.0001) between wild type and genetically modified cells, as assessed by Student's *t* test (GraphPad Software).
E & F. Dose response curves and EC_{50} values (in μM) (inserts) of *T. cruzi* Tc*snm1Δ* (■) and its complemented counterpart (▲) (**E**) and *T. cruzi* Tc*csbΔ* (■) and its complemented counterpart (▲) (**F**) towards mechlorethamine were determined and compared against wild type cells (●). All data points are mean values ± standard deviations from experiments performed in quadruplicate. The susceptibility of the *T. cruzi* null and complement lines against mechlorethamine, as judged by their EC_{50} values, were compared, and expressed as a fold difference relative to wild type (right hand panels). Statistically significance differences (P<0.0001) between *wild type and the null mutant, and **between the null mutant and the complemented line were assessed by the Student's *t*-test.

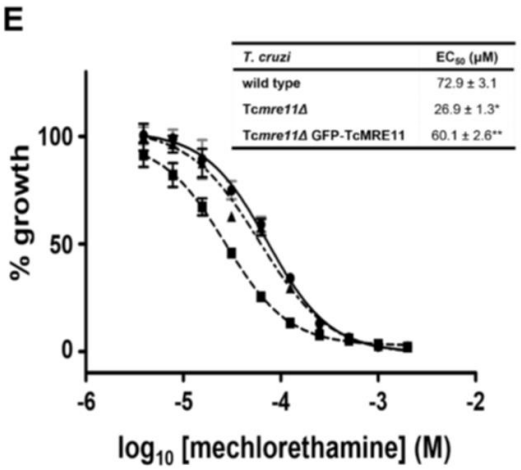


T. cruzi	Doubling time
wild type	18 hr 39 min ± 9 min
Tc*mre11Δ*	31 hr 9 min ± 21 min*
Tc*csbΔ*	22 hr 58 min ± 21 min*

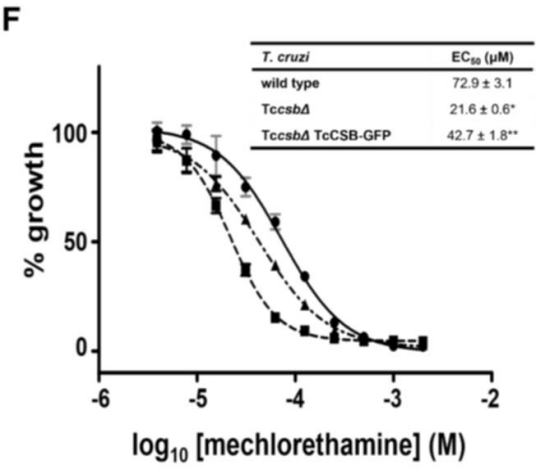




compound	*T. cruzi* (EC_{50})		
	wild type	*Tcmre11Δ*	*TccsbΔ*
MMS	96.5 ± 2.0	107.6 ± 3.0 (0.9)	33.8 ± 1.3 (2.9)*
hydroxyurea	227.0 ± 4.8	240.3 ± 9.4 (0.9)	199.1 ± 2.1 (1.1)
UV irradiation (in J m^{-2})	30.0 ± 1.6	31.6 ± 1.4 (0.9)	4.1 ± 0.8 (7.3)*
H_2O_2	132.3 ± 0.9	138.8 ± 4.4 (1)	142.3 ± 3.9 (0.9)
benznidazole	23.2 ± 1.0	25.5 ± 0.7 (0.9)	25.9 ± 0.9 (0.9)
fexinidazole	6.0 ± 0.1	6.6 ± 0.5 (0.9)	3.5 ± 0.2 (1.7)*
nifurtimox	5.7 ± 0.4	5.4 ± 0.3 (1.1)	7.0 ± 1.0 (0.8)



T. cruzi	EC_{50} (μM)
wild type	72.9 ± 3.1
Tc*mre11Δ*	26.9 ± 1.3*
Tc*mre11Δ* GFP-TcMRE11	60.1 ± 2.6**



T. cruzi	EC_{50} (μM)
wild type	72.9 ± 3.1
Tc*csbΔ*	21.6 ± 0.6*
Tc*csbΔ* TcCSB-GFP	42.7 ± 1.8**

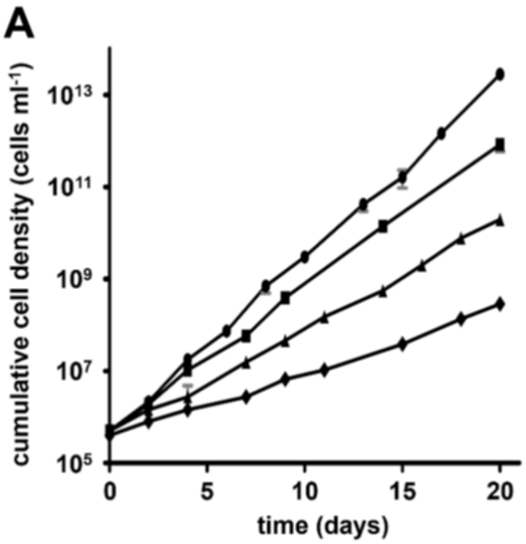
Figure 9. Characterisation of *T. cruzi* DNA repair null mutants.

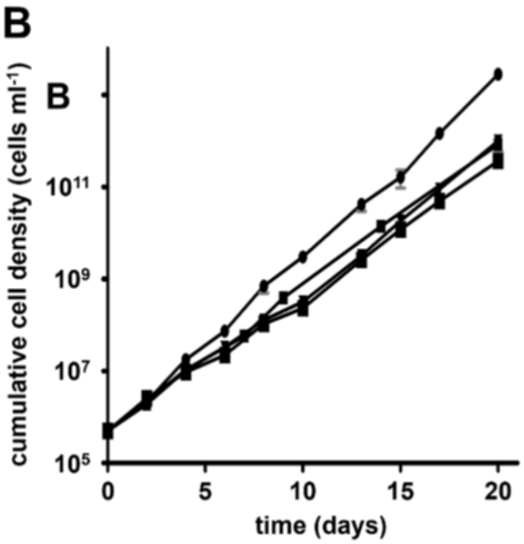
A. The cumulative cell density of *T. cruzi* Tc*snm1*Δ (■), Tc*mre11*Δ (▲) and Tc*snm1*ΔTc*mre11*Δ (♦) lines was followed for 20 days and compared against wild type *T. cruzi* (●) cultures grown in parallel. Each data point represents the mean cell density ± standard deviation from three independent cultures.

B. The cumulative cell density of *T. cruzi* Tc*snm1*Δ (■), Tc*csb*Δ (▲) and Tc*snm1*ΔTc*csb*Δ (♦) lines was followed for 20 days and compared against wild type *T. cruzi* (●) cultures grown in parallel. Each data point represents the mean cell density ± standard deviation from three independent cultures.

C. Doubling times were calculated from the cumulative cell density growth curves using the online tool https://doubling-time.com/compute.php. All data are mean values ± standard deviations from experiments performed in triplicate. The figures in parenthesis represents the fold difference in doubling time relative to wild type.

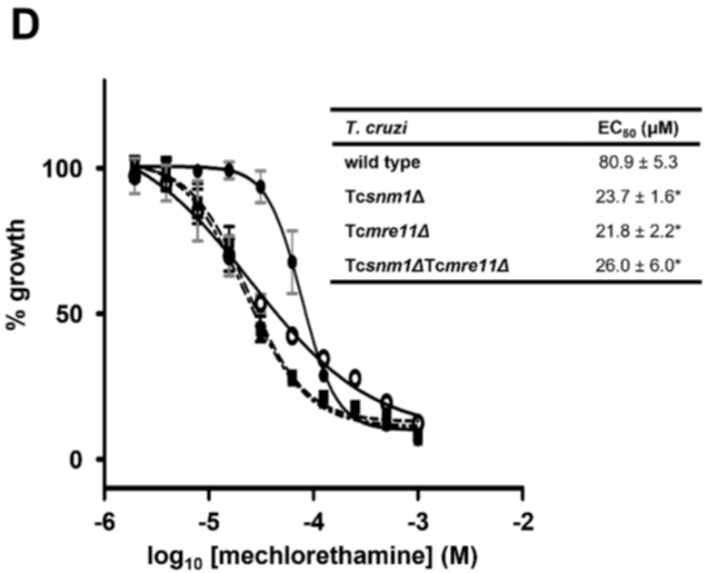
D & E. Dose response curves and EC_{50} values (in μM) (inserts) of *T. cruzi* Tc*snm1Δ*, Tc*mre11Δ* and Tc*snm1Δ*Tc*mre11Δ* (**D**), and *T. cruzi* Tc*snm1Δ*, Tc*csbΔ* and Tc*snm1Δ*Tc*csbΔ* (**E**) mutant lines towards mechlorethamine was determined and compared against wild type cells. All data points are mean values ± standard deviations from experiments performed in quadruplicate. The susceptibility of the *T. cruzi* null lines against mechlorethamine, as judged by their EC_{50} values, were compared and expressed as a fold difference relative to wild type. Statistically significance differences (P<0.0001) between *wild type and the single/double null mutant lines, and **between the single null and double null mutant lines were assessed by the Student's *t*-test.



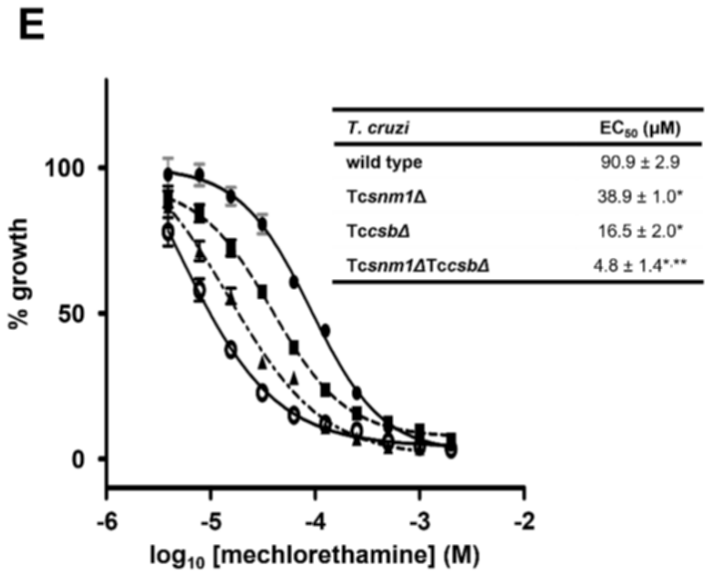


C

T. cruzi line	Doubling time ± SD (fold difference)
wild type	18 hr 39 min ± 9 min
Tc*snm1Δ*	23 hr 16 min ± 33 min (1.2)
Tc*mre11Δ*	31 hr 9 min ± 21 min (1.7)
Tc*snm1Δ*Tc*mre11Δ*	50 hr 36 min ± 46 min (2.7)
Tc*csbΔ*	22 hr 58 min ± 21 min (1.2)
Tc*snm1Δ*Tc*csbΔ*	24 hr 47 min ± 50 (1.3)



T. cruzi	EC_{50} (μM)
wild type	80.9 ± 5.3
Tc*snm1Δ*	23.7 ± 1.6*
Tc*mre11Δ*	21.8 ± 2.2*
Tc*snm1Δ*Tc*mre11Δ*	26.0 ± 6.0*



T. cruzi	EC_{50} (μM)
wild type	90.9 ± 2.9
Tc*snm1Δ*	38.9 ± 1.0*
Tc*csbΔ*	16.5 ± 2.0*
Tc*snm1Δ*Tc*csbΔ*	4.8 ± 1.4*,**

Supplementary material

Dissecting the interstrand crosslink DNA repair system of *Trypanosoma cruzi*

Monica Zavala Martinez, Francisco Olmo, Martin C. Taylor, Fabrice Caudron and Shane R Wilkinson.

Supplementary figures, legends and tables

Table S1. *T. cruzi* lines used in this project.

All null lines are based on the *T. cruzi* clone MHOM/BR/78/Sylvio-X10.6 (or X10.6) while parasites used in protein localisation are based on *T. cruzi* CL-Brener. *ble*, *bla*, *hyg*, *neo* and *pac* correspond to the genes that encode for bleomycin phosphotransferase, blasticidin-S-deaminase, hygromycin B phosphotransferase, neomycin phosphotransferase and puromycin N-acetyltransferase, respectively.

Cell line	**genotype**
X10.6	
X10.6 Tc*snm1*Δ	Tc*snm1*::*bla*/*pac*
X10.6 Tc*mre11*Δ	Tc*mre11*::*bla*/*pac*
X10.6 Tc*csb*Δ	Tc*csb*::*bla*/*pac*
X10.6 Tc*snm1*Δ Tc*mre11*Δ	Tc*snm1*::*bla*/*pac*; Tc*mre11*::*hyg*/*neo*
X10.6 Tc*snm1*Δ Tc*csb*Δ	Tc*snm1*::*bla*/*pac*; Tc*csb*::*hyg*/*neo*
X10.6 Tc*snm1*Δ *gfp*-Tc*snm1*	Tc*snm1*::*bla*/*pac*; TRIX-*gfp*-Tc*snm1 (neo)*
X10.6 Tc*mre11*Δ *gfp*-Tc*mre11*	Tc*mre11*::*bla*/*pac*; TRIX-*gfp*-Tc*mre11 (neo)*
X10.6 Tc*csb*Δ Tc*csb-gfp*	Tc*csb*::*bla*/*pac*; TRIX-Tc*csb-gfp (neo)*
CL-Brener	
CL-Brener *gfp*-Tc*snm1*	TRIX- *gfp*-Tc*snm1 (neo)*
CL-Brener *gfp*-Tc*mre11*	TRIX-*gfp*-Tc*mre11 (neo)*
CL-Brener Tc*csb-gfp*	TRIX-Tc*csb-gfp (neo)*

Table S2. Primers used in this study.

The bases in uppercase correspond to regions of homology present in the target DNA, lowercase italics correspond to restriction sites incorporated into the primers to facilitate cloning while lowercase non-italics represents a nucleotide clamp that aid in the binding of the named restriction enzymes.

Primer name[1]	Primer sequence
Tcsnm1-KO1 (P3)	ggg*gagctcccgcgg*ATGACTATAGAGGCCAAAAAG
Tcsnm1-KO2	ggg*tctaga*CTGGCGTAGATGGTTCCGTGG
Tcsnm1-KO3	aaa*gggccc*CTGCGTCAGCGTTACTCCAAT
Tcsnm1-KO4	ggggg*tacc*TCAGACATCTATAATAACACA
Tcsnm1-q3 (P1)	AACTTTCCCTCTCGTGCAGAT
Tcsnm1-q6 (P2)	AATGACCGCCACAGAAATCGC
Tcsnm1-NE1	aa*aaagctt*ACTATAGAGGCCAAAAGGTGCGGGGG
Tcsnm1-NE2	aaa*ctcgag*TCAGACATCTATAATAACACAGTCAT
Tcsnm1-3-CL	aaa*gtcgacggcgcgcc*TCAGACATCTATAATAACACA
TcSNM1-X10-pTrcHisR	aaa*aagcttggcgcgcc*TCAGACATCTATAATAACACA
Tcmre11-KO1 (P8)	aaa*gagctcccgcgg*ATGGCGAGTGCGGGGACTGCC
Tcmre11-KO2	ggg*tctaga*CCAAAGGCGTCCTTTTACCGG
Tcmre11-KO3	aaa*gggccc*TGCAGAGGGTATTTTGGTGCC
Tcmre11-KO4 (P10)	aaa*ggtacc*TCAACTTTGACCTGCCCACTT
Tcmre11-q3 (P6)	GCAACAAATGGTTCCCACACC
Tcmre11-q4 (P7)	ACTACGCAGTGGAAATGGCGT
Tcmre11-NE1	aaa*aagctt*GCAAGTGCGGGGAATGCCGCGTCGTC
Tcmre11-NE2	aaa*ctcgag*TCAACTTTGACCTGCCCACTTTGAAA
Tcmre11-3-CL	aaa*gtcgacggcgcgcc*TCAACTTTGACCTGCCCACTT
Tccsb-KO1 (P15)	ggg*gagctcccgcgg*CGGTTCTCCAGTTTCGTCTGG
Tccsb-KO2	ggg*tctaga*GCAACTTCCTCTAACCGCCG
Tccsb-KO3	aaa*gggccc*CATGGCGCATTGGTCAAACAC
Tccsb-KO4	aaa*ggtacc*TCAATCTCCATGTTCACCTTT
Tccsb-q3 (P13)	TCCGGCACTGTGAAGTTGGAA
Tccsb-q4 (P14)	AGCCACCGCCAGAATTCATGA
Tccsb-NE1	aaa*ggatcc*ATGGCCGATGATCTTTCGCAGTTGGG
Tccsb-NE2	aaa*aagctt*ATCTCCATTTTCACCTTTTACTTCTC
Tccsb-XGFP_F	GGCAACGACGATGACGAACAA
Tccsb-5-CL	aaa*gagctccctgcagg*ATGGCCGATGATCTTTCGCAG
eGFP-BamHI	ggg*ggatcc*ATGGTGAGCAAGGGCGAGGAGCTGTT
eGFP-XhoI-HindIII	aaa*ctcgagaagctt*GTACAGCTCGTCCATGCCGA
eGFP-BamHI-HindIII	ggg*ggatccaagctt*GTGAGCAAGGGCGAGGAGAG
eGFP-XhoI	aaa*ctcgag*TTACTTGTACAGCTCGTCCATGCCGA
GFP-5’	ATGGTGAGCAAGGGCGAGGAG
GFP-5’-MF	GAAGGGCATCGACTTCAAGGA
GFP-3'-MR	CTCCTTGAAGTCGATGCCCTT
GFP-5-SacI-SbfI	aaa*gagctccctgcagg*ATGGTGAGCAAGGGCGAGGAG
GFP-3-SalI-SgsI	aaa*gtcgacggcgcgcc*TTACTTGTACAGCTCGTCCAT
Tctert-F	CACCTGGAAAAGACGCCGTCC
Tctert-R	TTCAGCAACCACCGCGCGTTG
NEO-2 (P11)	ggg*atcgat*TCAGAAGAACTCGTCAAGAAG
HYG-1 (P12)	ggg*aagctt*ATGAAAAAGCCTGAACTCACC
HYG-2 (P17)	ggg*atcgat*CTATTCCTTTGCCCTCGGACG
PAC-1 (P16)	CGGGCTTGCGGGTCATGCACCA
PAC-2 (P9)	TGGTGCATGACCCGCAAGCCC
PAC-3 (P5)	CACCGTGGGCTTGTACTCGGT
BLA-1 (P4)	GGATTCTTCTTGAGACAAAGGCTTGGC

[1]P1 to 17 correspond to the primer naming system used in Figure S3.

Table S3. Primers: Function and combinations.

Gene	Function	Primer combination[1]
Tc*snm1*	**Gene disruption vector:**	
	5' targeting fragment	Tcsnm1-KO1/Tcsnm1-KO2
	3' targeting fragment	Tcsnm1-KO3Tcsnm1-KO4
	Null mutant validation (gDNA):	
	Tc*snm1* allele	Tcsnm1-q3/Tcsnm1-q6 (P1/P2)
	bla-interrupted Tc*snm1* allele	Tcsnm1-KO1/BLA-1 (P3/P4)
	pac-interrupted Tc*snm1* allele	Tcsnm1-KO1/PAC-3 (P3/P5)
	Null mutant validation (cDNA):	
	Tc*snm1* mRNA	Tcsnm1-q3/Tcsnm1-q6 (P1/P2)
	GFP expression vector	
	gfp-Tc*snm1*	Tcsnm1-NE1/Tcsnm1-NE2
	validating GFP line (gDNA; cDNA)	
	gfp-Tc*snm1* allele & mRNA	GFP-5'/Tcsnm1-NE2
	Yeast complementation vector	
	Tc*snm1*	Tcsnm1-KO1/TcSNM1-X10-pTrcHisR
	***T. brucei* complementation vector**	
	gfp-Tc*snm1*	GFP-5-SacI-SbfI/Tcsnm1-3-CL
Tc*mre11*	**Gene disruption vector:**	
	5' targeting fragment	Tcmre11-KO1/Tcmre11-KO2
	3' targeting fragment	Tcmre11-KO3/Tcmre11-KO4
	Null mutant validation (gDNA):	
	Tc*mre11* allele	Tcmre11-q3/Tcmre11-q4 (P6/P7)
	bla-interrupted Tc*mre11*allele	Tcmre11-KO1/BLA-1 (P8/P4)
	pac-interrupted Tc*mre11*allele	PAC-2/Tcmre11-KO4 (P9/P10)
	neo-interrupted Tc*mre11* allele	Tcmre11-KO1/NEO-2 (P8/P11)
	hyg-interrupted Tc*mre11* allele	HYG-1/Tcmre11-KO4 (P12/P10)
	Null mutant validation (cDNA):	
	Tc*mre11* mRNA	Tcmre11-q3/Tcmre11-q4 (P6/P7)
	GFP expression vector	
	gfp-Tc*mre11*	Tcmre11-NE1/Tcmre11-NE2
	validating GFP line (gDNA; cDNA)	
	gfp-Tc*mre11* allele & mRNA	GFP-5'-MF/Tcmre11-KO2
	***T. brucei* complementation vector**	
	gfp-Tc*mre11*	GFP-5-SacI-SbfI/Tcmre11-3-CL
Tc*csb*	**Gene disruption vector:**	
	5' targeting fragment	Tccsb-KO1/Tccsb-KO2
	3' targeting fragment	Tccsb-KO3/Tccsb-KO4
	Null mutant validation (gDNA):	
	Tc*csb* allele	Tbcsb-q3/Tbcsb-q4 (P13/P14)
	bla-interrupted Tc*csb* allele	Tccsb-KO1/BLA-1 (P15/P4)
	pac-interrupted Tc*csb* allele	Tccsb-KO1/PAC-1 (P15/P16)
	neo-interrupted Tc*csb* allele	Tccsb-KO1/NEO-2 (P15/P11)
	hyg-interrupted Tc*csb* allele	Tccsb-KO1/HYG-2 (P15/P17)
	Null mutant validation (cDNA):	
	Tc*csb* mRNA	Tbcsb-q3/Tbcsb-q4 (P13/P14)
	GFP expression vector	
	Tc*csb-gfp*	Tccsb1-NE1/Tccsb1-NE2
	validating GFP line (gDNA; cDNA)	
	Tc*csb-gfp* allele & mRNA	Tccsb-XGFP_F/GFP-3'-MR
	***T. brucei* complementation vector**	
	Tc*csb-gfp*	GFP-3-SalI-SgsI/Tccsb-5-CL
egfp	***T. cruzi* GFP expression vectors**	
	5' tagging vector	eGFP-BamHI/eGFP-XhoI-HindIII
	3' tagging vector	eGFP-BamHI/HindIII-eGFP-XhoI
Tc*tert*	**Null mutant/ GFP line validation**	
	Tc*tert* allele & mRNA	Tctert-F/Tctert-R

[1]P1 to 17 correspond to the primer naming system used in Figure S3.

Figure S1.

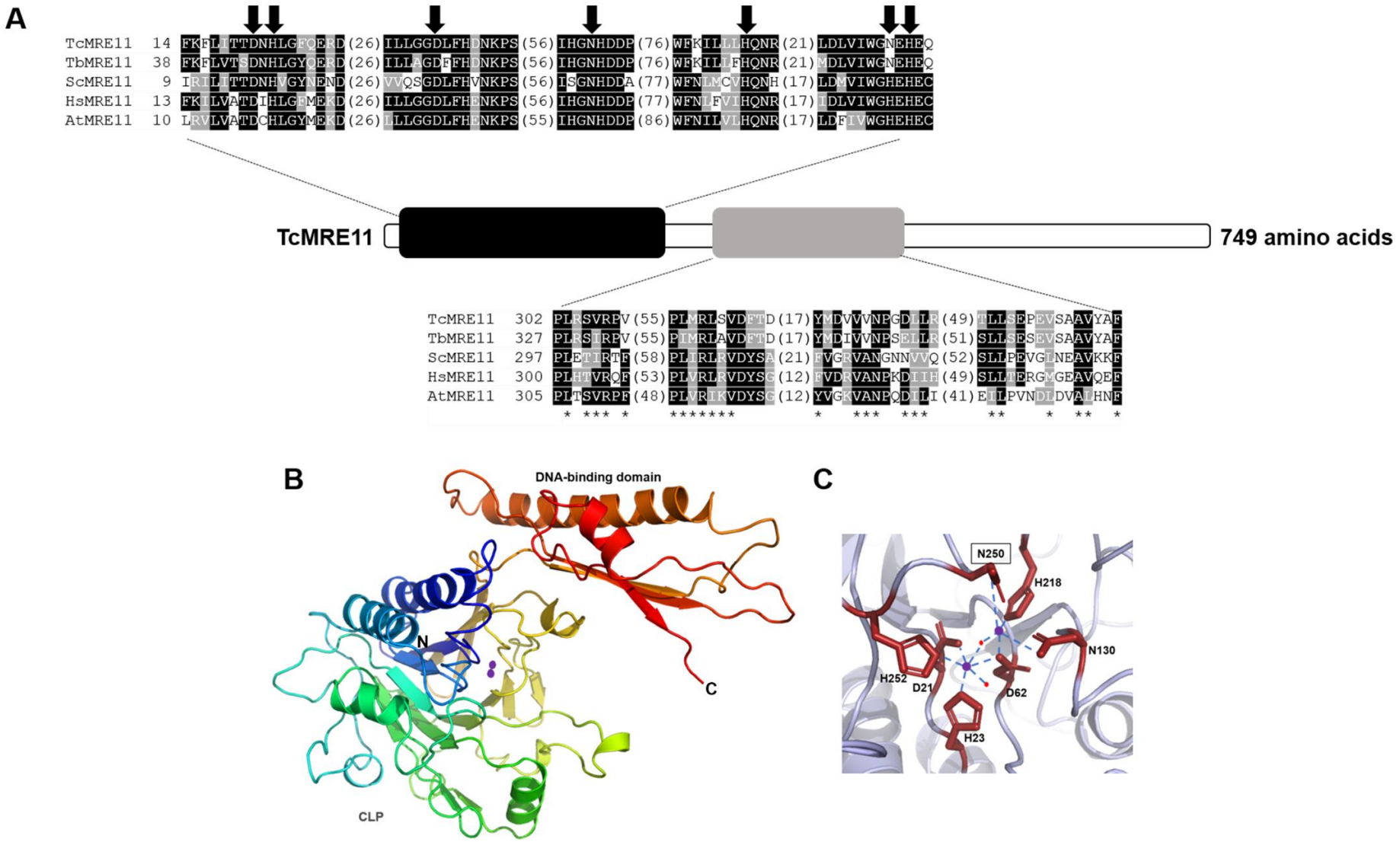

A
TcMRE11 14 FKFLITTDNHLGFQERD(26)ILLGGDLFHDNKPS(56)IHGNHDDE(76)WFKILLLHQNR(21)LDLVIWGNEHEQ
TbMRE11 38 FKFLVTSDNHLGYQERD(26)ILLAGDFFHDNKPS(56)IHGNHDDE(76)WFKILLFHQNR(21)MDLVIWGNEHEQ
ScMRE11 9 IRILITTDNHVGYNEND(26)VVQSGDLFHVNKPS(56)ISGNHDDA(77)WFNLMCVHQNH(17)LDMVIWGHEHEC
HsMRE11 13 FKILVATDIHLGFMEKD(26)ILLGGDLFHENKPS(56)IHGNHDDE(77)WFNLFVIHQNR(17)IDLVIWGHEHEC
AtMRE11 10 LRVLVATDCHLGYMEKD(26)LLLGGDLFHENKPS(55)IHGNHDDE(86)WFNILVLHQNR(17)LDFIVWGHEHEC
TcMRE11
749 amino acids
TcMRE11 302 PLRSVRPV(55)PLMRLSVDFTD(17)YMDVVVNEGDLLR(49)TLLSEPEVSAAVYAF
TbMRE11 327 PLRSIRPV(55)PIMRLAVDFTD(17)YMDIVVNESELLR(51)SLLSESEVSAAVYAF
ScMRE11 297 PLETIRTF(58)PLIRLRVDYSA(21)FVGRVANGNNVVQ(52)SLLPEVGLNEAVKKF
HsMRE11 300 PLHTVRQF(53)PLVRLRVDYSG(12)FVDRVANEKDIIH(49)SLLTERGMGEAVQEF
AtMRE11 305 PLTSVRPF(48)PLVRIKVDYSG(12)YVGKVANEQDILI(41)EILPVNDLDVALHNF
B
DNA-binding domain
N
C
CLP
C
N250
H218
N130
H252
D21
D62
H23

Figure S1. Informatic analysis of TcMRE11

A. The sequence corresponding to the calcineurin-like phosphoesterase (PF00149; residues 14-254; black box) and Mre11 DNA-binding (PF04152; residues 297-470; grey box) domains of TcMRE11 (GenBank XP_816077) were aligned with the equivalent regions from other members of the MRE11 family of nucleases: *T. brucei* TbMRE11 (CAD20051), *Saccharomyces cerevisiae* ScMRE11 (P32829), *Homo sapiens* HsMRE11 (NP_005582) and *Arabidopsis thaliana* AtMRE11 (NP_200237). Residues that are identical in at least three of the five sequences and conserved substitutions are highlighted in black and grey, respectively. Amino acids in the calcineurin-like phosphoesterase domain postulated to coordinate metal (magnesium) co-factor binding are denoted with a down arrow while residues highlighted with an asterisk in the DNA-binding domain representing amino acids that constitute the HMM logo.

B. A 3D approximation of TcMRE11 (tritrypdb.org: TCSYLVIO_000429) based on the *Chaetomium thermophilum* MRE11 (PDB: 4YKE; [1]) crystal structure was constructed using Phyre2 [2] and visualized using PyMOL: The model had 100 % confidence over 45 % coverage to the template. The amino (N; in blue) and carboxyl (C; in red) termini are indicated. The position of the calcineurin-like phosphoesterase (CLP; bottom) and Mre11 DNA-binding presumed (DNA-binding; top) domains are shown. The putative binding sites of two manganese ions (as purples spheres) are shown.

C. The assignment of residues involved in manganese binding was made following comparison against MRE11 sequences from the *C. thermophilum* and *Homo sapiens* (PDB: 3T1I) [1, 3]. Binding to one manganese ion is mediated by D21, H23, D62, H252 and two molecules of water (red sphere) with D60, N130, H218, N250 and water contributing to the binding of the second metal atom. The boxed residue (N250) represents a difference in the metal binding site noted between trypanosomal and fungal/human sequences.

Figure S2

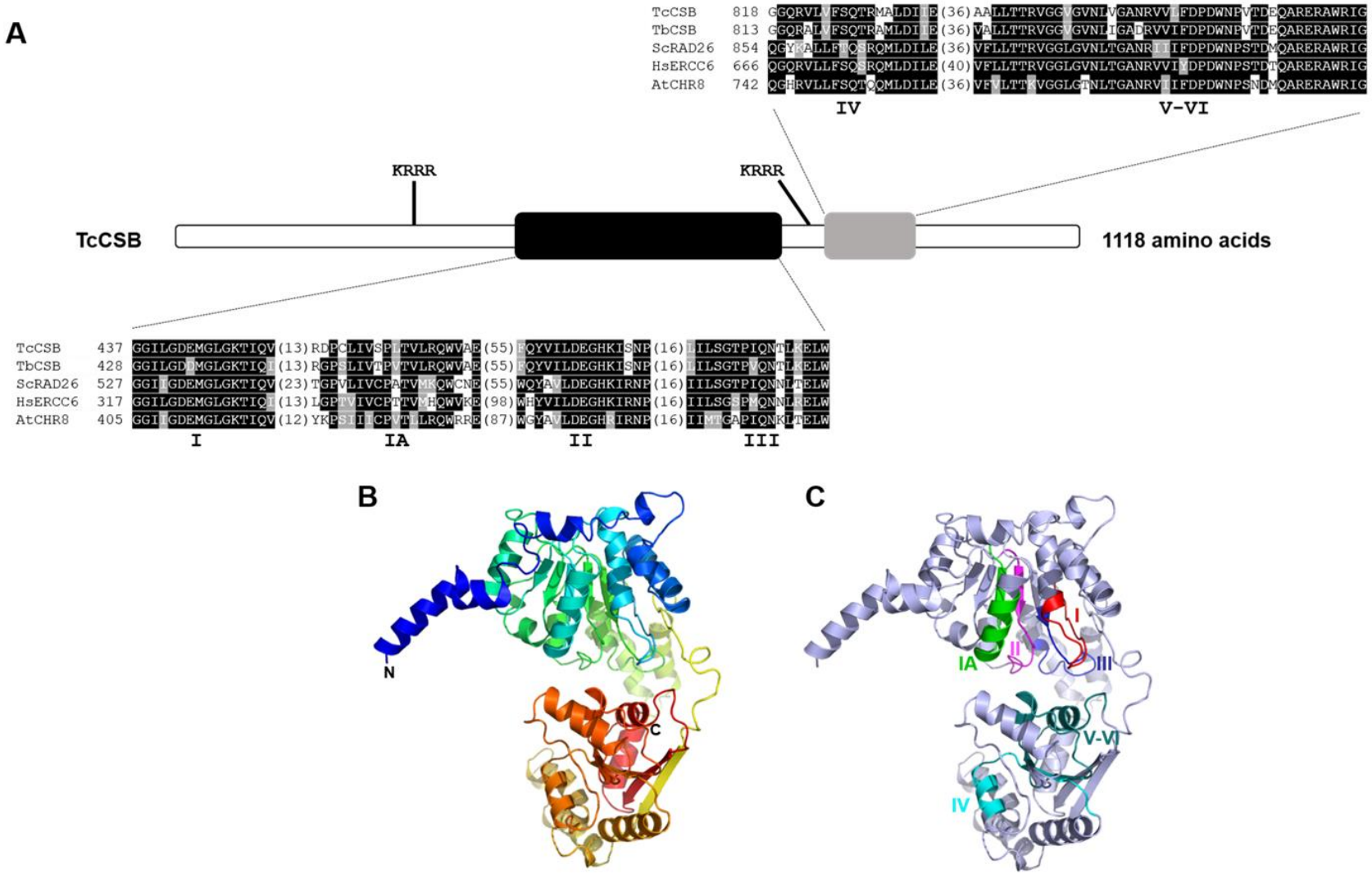

A
TcCSB 818 GGQRVLVFSQTRMALDIIE(36)AALLTTRVGGVGVNLVGANRVVIFDPDWNPVTDEQARERAWRIG
TbCSB 813 GGQRALVFSQTRAMLDIIE(36)VALLTTRVGGVGVNLIGADRVVIFDPDWNPVTDEQARERAWRIG
ScRAD26 854 QGYKALLFTQSRQMLDILE(36)VFLLTTRVGGLGVNLTGANRIIIFDPDWNPSTDMQARERAWRIG
HsERCC6 666 QGQRVLLFSQSRQMLDILE(40)VFLLTTRVGGLGVNLTGANRVVIYDPDWNPSTDTQARERAWRIG
AtCHR8 742 QGHRVLLFSQTQQMLDILE(36)VFVLTTKVGGLGTNLTGANRVIIFDPDWNPSNDMQARERAWRIG
IV
V-VI
KRRR
KRRR
TcCSB
1118 amino acids
TcCSB 437 GGILGDEMGLGKTIQV(13)RDECLIVSPLTVLRQWVAE(55)FQYVILDEGHKISNE(16)HILSGTPIQNTLKELW
TbCSB 428 GGILGDDMGLGKTIQI(13)RGPSLIVTPVTVLRQWVAE(55)FQYVILDEGHKISNE(16)HILSGTPVQNTLKELW
ScRAD26 527 GGIIGDEMGLGKTIQV(23)TGFVLIVCPATVMKQWCNE(55)WQYAVLDEGHKIRNP(16)IILSGTPIQNNLTELW
HsERCC6 317 GGILGDEMGLGKTIQI(13)LGPTVIVCPTTVMHQWVKE(98)WHYVILDEGHKIRNP(16)IILSGSPMQNNLRELW
AtCHR8 405 GGIIGDEMGLGKTIQV(12)YKPSIIICPVTLLRQWRRE(87)WGYAVLDEGHRIRNP(16)IIMTGAPIQNKLTELW
I
IA
II
III
B
N
C
C
I
IA
II
III
V-VI
IV

Figure S2. Informatic analysis of TcCSB.

A. The sequence corresponding to the SNF2-related (PF00176; residues 420-751; black box) and Helicase conserved C-terminal (PF00271; residues 804-916; grey box) domains of TcCSB (GenBank XP_816077) were aligned with the equivalent regions from other members of the CSB family of helicase: *T. brucei* TbCSB (XP_846031), *Saccharomyces cerevisiae* ScRAD26 (P40352), *Homo sapiens* HsERCC6 (NP_000115) and *Arabidopsis thaliana* AtCHRB (NP_001318246). Residues that are identical in at least three of the five sequences and conserved substitutions are highlighted in black and grey, respectively. The central ATPase motifs that span the two domains are labelled I to VI while the two KRRR motifs correspond to a putative nuclear 'pattern 4' targeting signals.

B. A 3D approximation model of TcCSB (tritrypdb.org: TCSYLVIO_010473) based on the ScRAD26 (PDB: 5VVR; [4]) crystal structure was constructed using Phyre2 [2] and visualized using PyMOL: The model had 100 % confidence over 45 % coverage to the template. The amino (N; in blue) and carboxyl (C; in red) termini are indicated.

C. The regions corresponding to the central ATPase motifs I to VI are highlighted on the 3D approximation model.

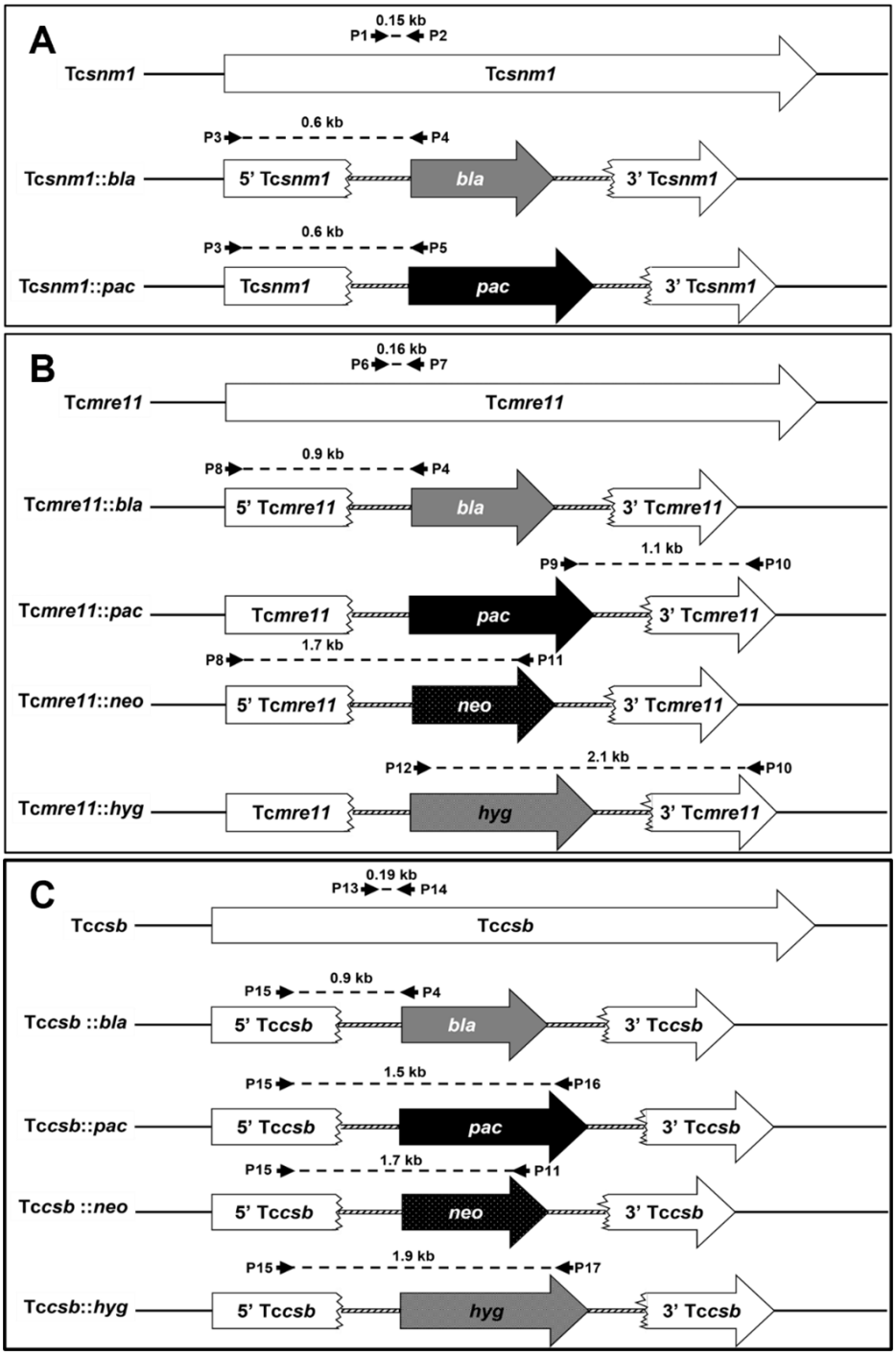


Figure S3. Effect of gene disruption on the *T. cruzi* genome.
Schematic representation of the Tc*snm1* (**A**), *Tcmre11* (**B**) or Tc*csb* (**C**) alleles and the effects of their disruption with DNA fragments containing sequences encoding for blasticidin-S-deaminase (*bla*) or puromycin N-acetyltransferase (*pac*), plus the *T. brucei* tubulin intergenic elements required for processing their mRNAs (hatched boxes). P1 to P3 correspond to the primers used to generate Tc*snm1* specific amplicons from genomic DNA templates, P6 to P8 and P10 to the primers used to generate Tc*mre11* specific amplicons and P11 to P13 to the primers used to generate Tc*csb* specific amplicons. P4, P5, P9 and P14 correspond to *bla* and *pac* specific primers. The size of the predicted amplicons, denoted by a dashed line, is given. The primer sequences and combinations are listed in Supplementary Tables 2 and 3, respectively.

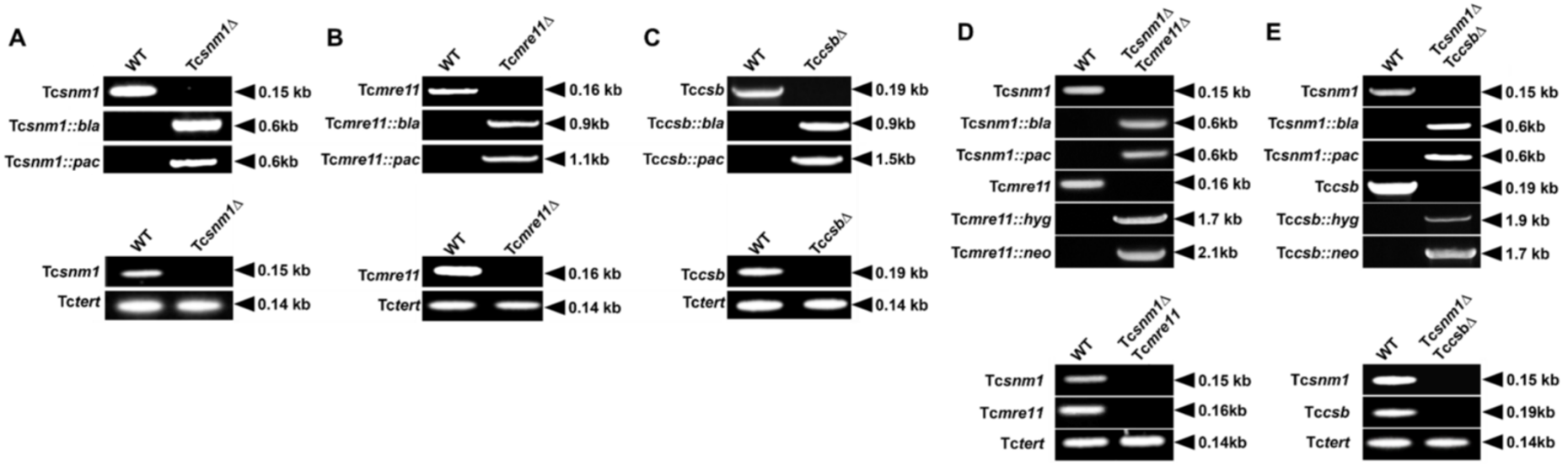


Figure S4. Validation of *T. cruzi* null mutant lines.

A-E (upper panels). Amplicons corresponding to intact Tc*snm1* (**A, D & E**), *Tcmre11* (**B & D**) or Tc*ccsb* (**C & E**) and their *pac*-, *bla*, *neo* or *hyg*--disrupted counterparts were generated from template genomic DNAs extracted from *T. cruzi* wild type and null mutant lines**.**

A-E (lower panels). Amplicons corresponding to intact Tc*snm1* (**A, D & E**), *Tcmre11* (**B & D**) or Tc*ccsb* (**C & E**) were generated from template cDNAs derived from total RNA extracted from the various *T. brucei* lines indicated. The integrity of cDNAs (and hence RNAs) was evaluated by amplification of a 140 bp control fragment, Tc*tert*:

The primer sequences and combinations used for each amplification are listed in Tables S2 and S3, respectively.

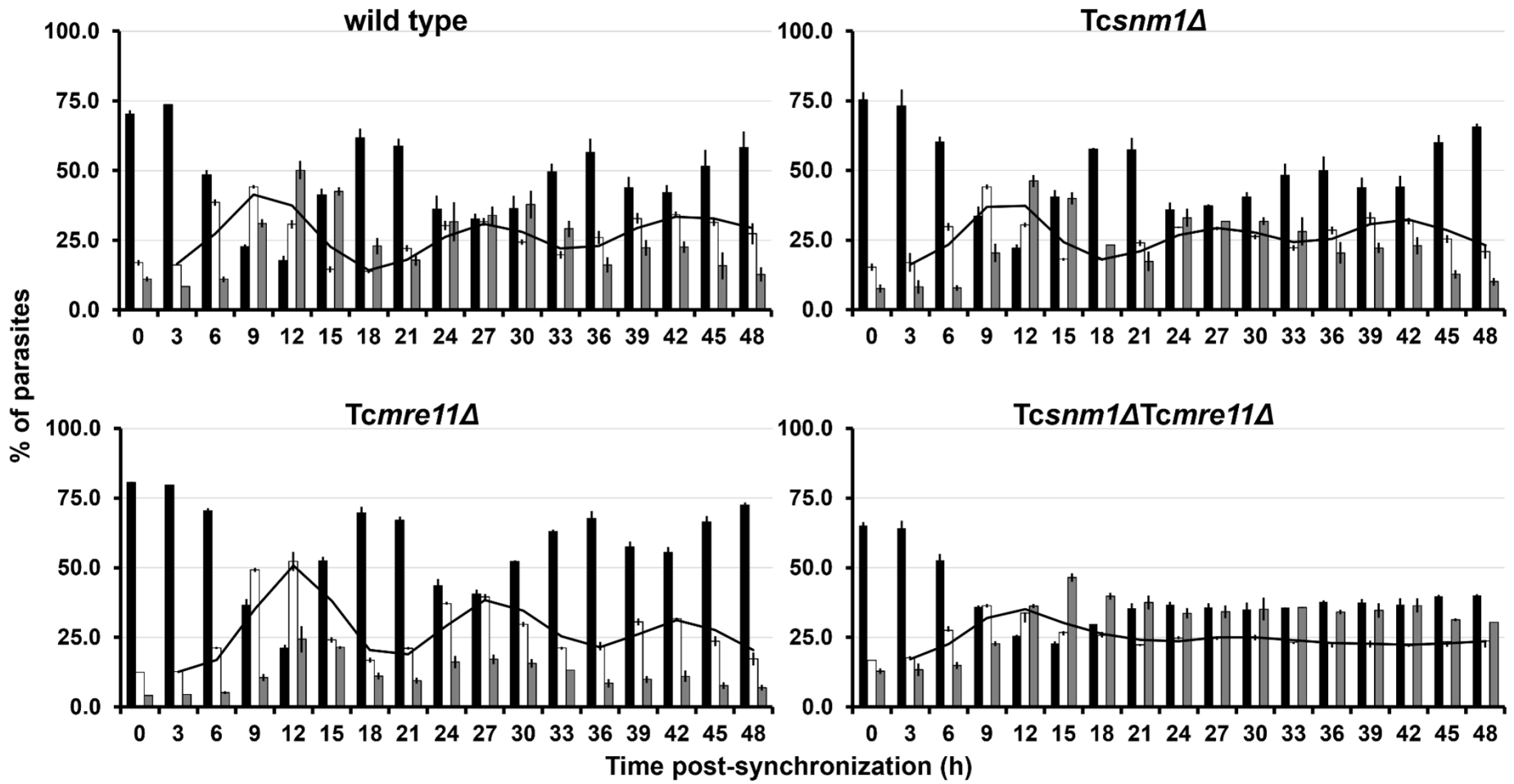


Figure S5. Cell cycle analysis of *T. cruzi* mutants.
Exponentially growing *T. cruzi* were arrested at the G1/S boundary by HU treatment. Following HU removal, progression through the cell cycle for approximately 10,000 trypanosomes was determined every 3 hours over a 48-hour period using flow cytometry. The black, white, and grey bars represent the G1, S and G2 phase of the cell cycle, respectively, while the solid line shows progression of the S phase of each cell line at the timepoints. Each data point represents the mean % number of parasites at each cell cycle stage ± standard deviation from three readings.

References

[1] F. U. Seifert, K. Lammens and K. P. Hopfner, Structure of the catalytic domain of Mre11 from Chaetomium thermophilum, Acta Crystallogr F Struct Biol Commun, 71 (2015) 752-757.
[2] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis, Nat Protoc, 10 (2015) 845-858.
[3] Y.B. Park, J. Chae, Y.C. Kim, Y. Cho, Crystal structure of human Mre11: understanding tumorigenic mutations, Structure, 19 (2011) 1591-1602.
[4] J. Xu, I. Lahiri, W. Wang, A. Wier, M.A. Cianfrocco, J. Chong, A.A. Hare, P.B. Dervan, F. DiMaio, A.E. Leschziner, D. Wang, Structural basis for the initiation of eukaryotic transcription-coupled DNA repair, Nature, 551 (2017) 653-657.