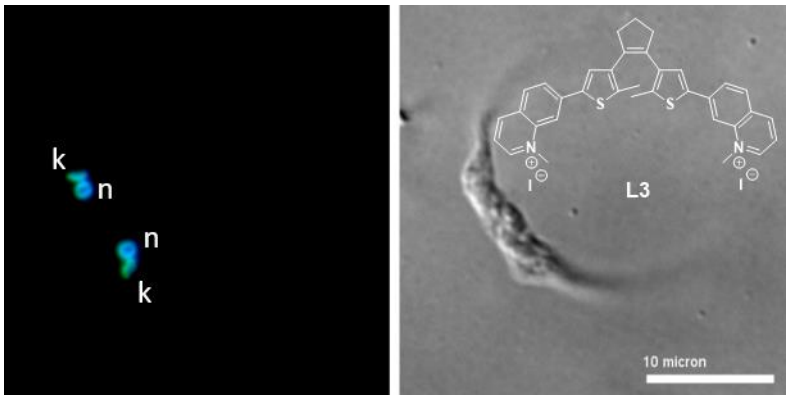
DNA G-Quadruplexes in the genome of Trypanosoma cruzi as potential therapeutic targets for Chagas disease: dithienylethene ligands as effective antiparasitic agents

Manuel Pérez-Soto#, Javier Ramos-Soriano#, Pablo Peñalver#, Efres Belmonte-Reche, Michael O'Hagan, Anne Cucchiarini, Jean-Louis Mergny, María del Carmen Galán*, Manuel Carlos López López*, María del Carmen Thomas*, Juan Carlos Morales*

Brief title:

DNA G-quadruplexes as therapeutic targets for Chagas disease

Graphical Abstract



DNA G-Quadruplexes in the genome of *Trypanosoma cruzi* as potential therapeutic targets for Chagas disease: dithienylethene ligands as effective antiparasitic agents

Manuel Pérez-Soto[1,#], Javier Ramos-Soriano[2,#], Pablo Peñalver[1,#], Efres Belmonte-Reche[3], Michael O'Hagan[2], Anne Cucchiarini[4], Jean-Louis Mergny[4], María del Carmen Galán[2]*, Manuel Carlos López López[1,]*, María del Carmen Thomas[1,]*, Juan Carlos Morales[1,]*

[1] *Departamento de Biología Molecular, Instituto de Parasitología y Biomedicina López Neyra, CSIC, PTS Granada, Avenida del Conocimiento, 17, Armilla, 18016 Granada, Spain.*

[2] *School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom.*

[3] *GENYO, Centre for Genomics and Oncological Research: Pfizer/University of Granada / Andalusian Regional Government, PTS Granada, Av. de la Ilustración, 114, 18016 Granada, Spain.*

[4] *Laboratoire d'optique et Biosciences, Ecole Polytechnique, Inserm U1182, CNRS UMR7645, Institut Polytechnique de Paris, Palaiseau, France.*

[#] authors that contributed equally to this work

* corresponding authors

M.C.Galan@bristol.ac.uk

mclopez@ipb.csic.es

mcthomas@ipb.csic.es

jcmorales@ipb.csic.es

Abstract: Chagas disease is caused by the parasite *Trypanosoma cruzi* and affects over 7 million people worldwide. The two actual treatments, Benznidazole (Bzn) and Nifurtimox, cause serious side effects due to their high toxicity leading to treatment abandonment by the patients. In this work, we propose DNA G-quadruplexes (G4) as potential therapeutic targets for this infectious disease. We have found 174 putative quadruplex forming sequences per 100,000 nucleotides in the genome of *T. cruzi* and confirmed G4 formation of three frequent motifs. We synthesized a family of 14 quadruplex ligands based in the dithienylethene (DTE) scaffold and demonstrated their binding to these identified G4 sequences. Several DTE derivatives exhibited micromolar activity against epimastigotes of four different strains of *T. cruzi*, in the same concentration range as Bzn. Compounds **3** and **4** presented remarkable activity against trypomastigotes, the active form in blood, of *T. cruzi* SOL strain (IC_{50} = 1.5-3.3 μM, SI =25-40.9), being around 40 times more active than Bzn and displaying much better selectivity indexes.

Introduction

Chagas disease (ChD) is one of the Neglected tropical diseases declared by the World Health Organization and is caused by the parasite *Trypanosoma cruzi*. ChD affects over 7 million people, mostly located in Latin America, but it is spreading to other continents due to human migration.[1] It has different transmission paths: through the triatomine bug (vector-borne), orally (food-borne), mother-to-child (congenital) transmission, through blood contact, organ transplantation and laboratory accidents. Still today, ChD has a high mortality rate, causing 12,000 deaths annually worldwide. The disease courses with an initial acute phase characterized by high parasitemia rates in spite of which is in most cases asymptomatic or presents non-specific symptoms. This phase lasts approximately 1-2 months and is followed by a chronic phase that lasts a lifetime. 30-40% of chronic patients suffer cardiac, digestive and neurological disorders which compromise the lives of the patients.

Two drugs are approved for the treatment of ChD, the first-line drug Benznidazole (Bzn) and the second-line drug Nifurtimox. They are effective in the acute phase of the disease and in

congenital transmission. The administration is also recommended in the chronic phase, although the cure rate importantly decreases. Both of them cause serious side effects due to their high toxicity what leads to frequent treatment abandonment by the patients.[2] Therefore, new drug candidates with greater efficacy, tolerability, and safety are urgently needed to enlarge ChD treatment options. Most of new drugs being investigated keep exploring previous scaffolds used in other antiparasitic or antifungal drugs such as nitroaryl,[3] 1,3-thiazole[4,5] or 1,2,3-triazole rings.[6] Similarly, classical targets such as cruzain,[7,8] pteridine reductase,[9] the proteasome[10,11] or DNA minor grooves[12] have yet to lead to new potential drugs entering clinical trials. New targets are needed that may lead to the finding of more effective and less toxic drugs than those currently available.

G-quadruplex (G4) are secondary DNA/RNA structures that form in guanine-rich areas of the genome. The central motif is a four-guanine tetrad stabilized through Hoogsteen hydrogen bonds. These G-tetrads stack on top of each other helped by interactions with cations such as Na^+ or K^+ that are sandwiched between them. Depending on the DNA sequence, cations, loop connectivity and the *syn* or *anti* guanine configuration along the oligonucleotide folding, G4 will display a variety of topologies. Certain key regions of the genome such as promoters, enhancers, insulators, origins of replications and telomeres are enriched with G4 sequences. In fact, G4s play important roles in cellular and genetic processes such as oncogene promoter regulation,[13,14] cellular transcriptome shaping,[15,16] DNA replication,[17] DNA damage and genome instability,[18,19] as well as RNA transcription, translation, and processing.[20,21] Consequently, they have been proposed as new therapeutic targets in cancer and neurodegenerative diseases.[22–24]

Putative quadruplex sequences (PQS) have been identified not only in humans, but also in the genomes of other mammals,[25] yeast,[26] prokaryotes, including bacteria[27–29] and viruses[30–32] using bioinformatics tools. We previously reported the presence of PQS in the genomes of parasites such as *Trypanosoma brucei*, *Leishmania major* and *Plasmodium falciparum*, and proposed them as new therapeutic targets in parasitic infectious diseases.[33] Similarly, other research groups have explored the potential presence of quadruplexes in the genome of

Plasmodium falciparum.[34,35] We found that sugar conjugated naphthalene diimides previously shown to be G4 ligands[36] also bind G4s found exclusively in *T. brucei* and present very notable antiparasitic activity.[33,37,38] Eventually, other quadruplex ligands such as quarfloxin, CX-5461[39] and compounds based on stiff stilbene,[40] naphthalene diimide,[41] perylene diimide,[42] phenantroline,[43,44] quinoxaline[45] and quinoline scaffolds[46] have been reported with antitrypanosomal and antiplasmodial activity.[47]

In this work, we studied the presence of PQS in the genome of *Trypanosoma cruzi* and carried out biophysical studies confirming that some of these frequent PQS actually form stable quadruplex structures. We found that quadruplex ligands based on a DTEstructure bind these G4 structures from *T. cruzi* and present relevant cytotoxicity against their epimastigote and their trypomastigote forms together with displaying a notable therapeutic window. Finally, localization of compound **3** inside *T. cruzi* parasites was also investigated.

Results and discussion

Putative G-Quadruplex Sequences (PQS) Search in T. cruzi genomes

Nine *T. cruzi* genomes (CL Brener, Esmeraldo, DM28(1), DM28(2), Ycl6, Ycl4, Y, S23b, 231) were analysed with G4-iM Grinder, an R-based algorithm that locates quantifies and qualifies Putative Quadruplex Sequences (PQSs) in any genome.[48] We found 164,686 PQSs in the genome of Ycl6 strain of *T. cruzi* which was used as a reference, corresponding to a density of 174 PQSs per 100,000 nucleotides. 21,403 of these PQSs had also a *high* probability of forming G4s *in vitro* according to G4-iM Grinder, which equals a density of 23 PQS with *high* G4-probability per 100,000 nucleotides.

The densities of PQSs in the genomes of other *T. cruzi* strains analyzed fell within the range observed in the reference Ycl6 strain, spanning from 137 to 188 PQSs per 100,000 nucleotides, with PQS densities exhibiting *high* G4-probability ranging from 21 to 33 per 100,000

nucleotides. For comparative analysis, we investigated PQS densities in several other parasites, both related and unrelated, as well as in the human genome (Figure 1).

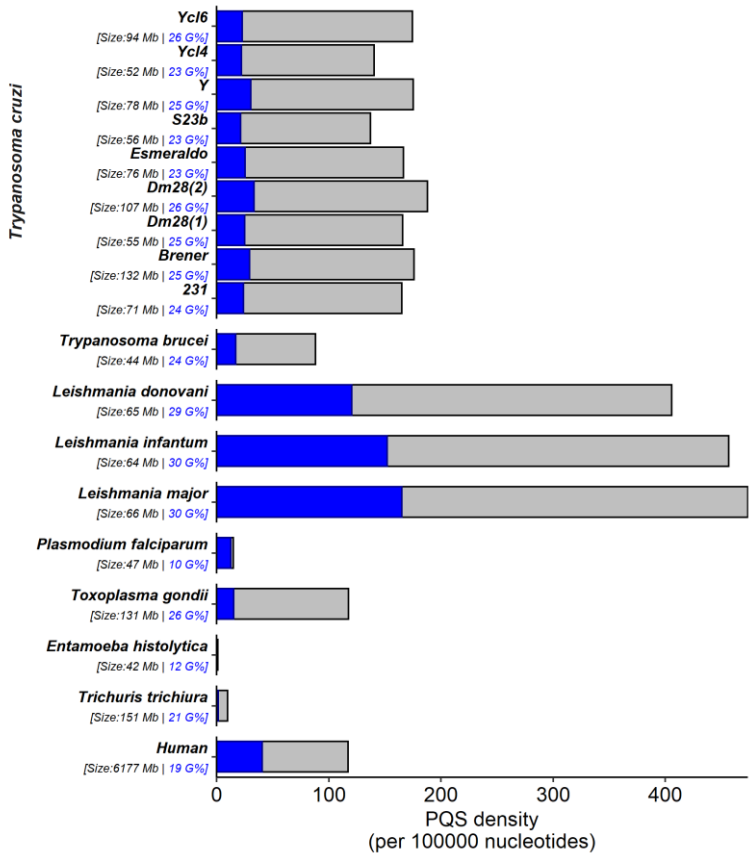


Figure 1. *Trypanosoma cruzi*'s PQS density results compared to parasites and the human genomes. Bars in grey are the PQSs' genomic densities independently of the score. In dark blue, the density of results that score at least 40 (with a high probability of formatting G4 *in vitro*). Density is expressed as the number of PQS per 100,000 nucleotides.

Trypanosoma cruzi's PQS densities were found to be higher than those found in almost all other parasites examined, except *Leishmania*, and even higher than that found in the human genome. The density of sequences with *high* G4-probability in *T. cruzi* were also higher than those found in other parasites (again except *Leishmania*), yet slightly smaller when compared to humans (30 *vs* 40 in humans). It is noteworthy that the genomic content of GC and the density of PQSs did not exhibit a clear correlation. For instance, despite the lower GC content in the human genome compared to other organisms, it displayed a higher abundance of high-scoring PQSs.

Conversely, parasites like *T. trichiuria* (whipworm) and *T. gondii* (the causative agent of toxoplasmosis) exhibited similar or higher GC content than both *T. cruzi* and the human genomes but displayed a scarcity of overall *high* G4-probable PQSs. On the opposite side, *P. falciparum* manifested elevated levels of *high* G4-probable PQSs in the range of those found for the *T. cruzi* genomes analyzed, despite having an exceptionally AT-rich genome. These observations highlight the significance of GC organization rather than solely GC content for PQS distribution, and overall indicate the PQS-richness and quality of *Trypanosoma cruzi*'s genome due to an enhanced GC organization and richness. Furthermore, PQS with *in vitro* confirmed G4s already in literature were identified in the range of the thousands (Table S1). For the reference Ycl6 strain, 5,657 PQSs were identified with these sequences that are already confirmed to form G4 *in vitro*, which corresponds to a density of 7.6 PQSs per 100,000 nucleotides.

Almost half of these PQSs with confirmed G4s were related to the telomeric sequences (22Ag – with the motif GGGTTA), which were identified in all of *T. cruzi's* genomes examined. In the reference Ycl6 strain, these PQSs were located at the beginning and/or end of some chromosomes (3, 4, 5, 8, 10, 17, 18, 23, 24, 26, 31 and 32). Here, the PQSs repeat one after another forming long GGGTTA clusters, organizing in a fashion similar to what is observed in the human genome. Hence, these can probably give rise to a higher-order quadruplex structure, composed of several individual 22Ag G4's as monomers of the bigger supra-structure. However, these clusters are not linked to any genes described in the annotation files (as most likely they are part of the telomeres). 1,616 PQSs were identified with the known 2G_L1.NAR G4 (a two-tetrad G4 in literature, **GG**T**GG**T**GG**T**GG**) located in all the chromosomes of the reference strain. These PQSs (entries 6 to 11, Table S2) are here hosted by dispersed note family protein 1 (DGF-1) and Mucin-associated surface protein (MASP) genes, amongst many other genes. Other known G4s were also identified with high frequency of appearance, such as the PAX9 and T3069 or T30923 G4s (with the motifs GGGA and GGGT, respectively) although these were not associated to any known genomic annotations.

Next, we identified PQSs with *high* G4-probability across all *T. cruzi* genomes, which have not been experimentally confirmed to form G4 *in vitro* and/or have not been yet listed in the G4-library of G4-iM Grinder. This analysis revealed tens of thousands of such PQSs (Table S3) with the reference Ycl6 strain alone harboring 16,847 PQSs exhibiting *high* G4-probability.

The genomic features of the reference strain Ycl6 chromosome's were then matched with the locations of the PQSs with known G4 and PQSs with *high* G4-probability. Of the 14,336 coding genes within the 40 chromosomes (excluding contigs), 3,150 genes presented at least one PQS with a known G4 sequence or a PQS with *high* probability of forming a G4 (28 % of all genes; Figure 2, A). The proportion of results were higher for pseudogenes (39 % of all pseudogenes). The DGF-1 group presented over 400 PQSs within its gene, whilst others such as the retrotransposon hot stop, trans-sialidase and surface protease genes were similarly enriched in PQSs that may notably influence their biological roles (Figure 2, B). PQSs identified in the reference genome with *high* G4-probability and with high frequency of appearance were examined in detail to determine their relationships with these specific genes.

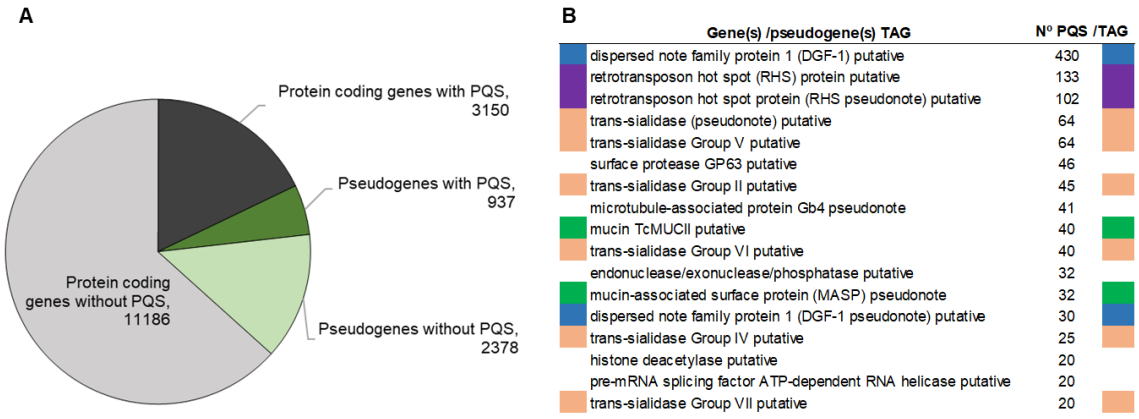


Figure 2. A. Genes/Pseudogenes per PQSs presence. B. Highest PQSs counts per Gene / Pseudogene TAGs. The color code denotes related TAGs including in blue DGF-1, in purple RHS, in light orange trans-sialidases and in green MASP TAGs.

Next, we focused on the sequence GGGCGGACGGTGGTGGTGGACGGCG (named DCr) located in the Dispersed gene protein family-1 gene and repeated 110 times. We selected

this PQS for further biophysical studies to confirm if it is capable of G4-formation *in vitro*. We also selected GGGGACGGGAATGGGGGTGCATGAGGGG (named RCr and repeated 18 times), identified in the Retrotransposon hot spot (RHS), GGGAGGGACGGATGGGCAGAAACGGG (named TCr and repeated 8 times), located in the Trans-Sialidase family gene and GTGGAGGGGGAGGGTCATGGGG (named MCr and repeated 28 times), identified in the Mucin-related family tags for further biophysical evaluation.

G4 Formation by T. cruzi selected PQS

We selected the four PQS mentioned in the previous section due to their novelty and high frequency number in the genome of *T. cruzi* (DCr, RCr, TCr and MCr). Firstly, we carried out a fluorescence assay with the four sequences in the presence of thioflavin T (ThT) or *N*-methyl mesoporphyrin IX (NMM). These fluorescent probes notably increase their fluorescence in the presence of G-quadruplex structures.[49–52] We used 22CTA (antiparallel topology), bcl2 (parallel topology) and ckit87up (hybrid topology) as quadruplex controls. We observed that all sequences increased the fluorescence of ThT (Figure 3A) and NMM (Figure S1) although DCr to a much lower extent.

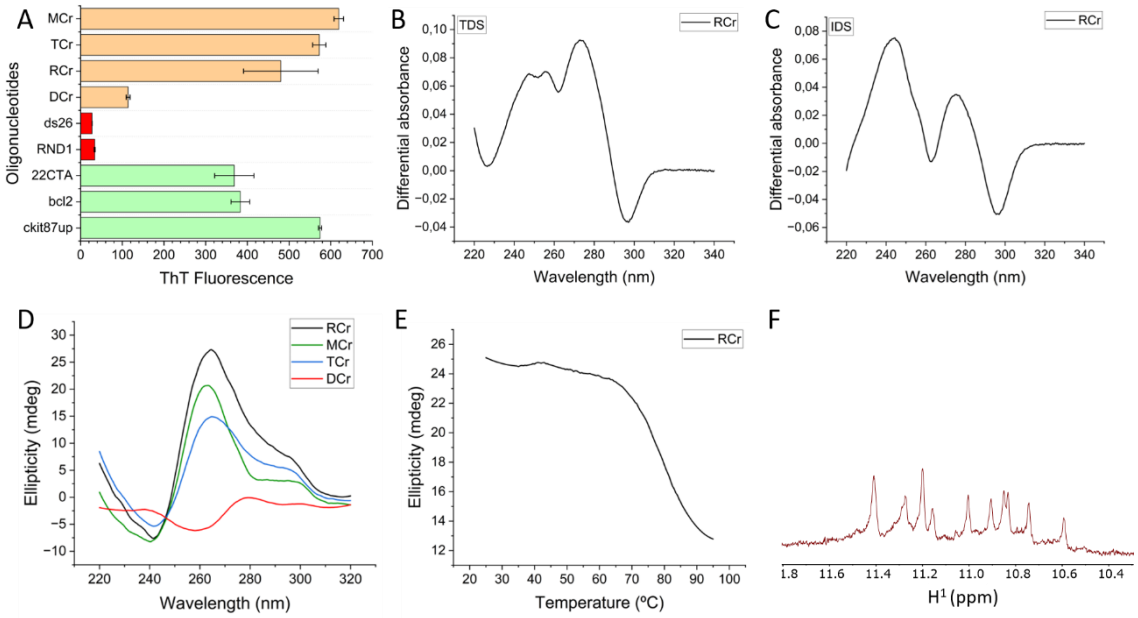


Figure 3. Characterization of *T. cruzi* sequences: DCr, RCr, TCr and MCr (A) Thioflavin T fluorescence assay, positive and negative controls are shown in green and red, respectively; (B) isothermal difference spectra at 25 °C (IDS) of sequence RCr; (C) thermal difference spectra of sequence RCr (TDS); (D) CD spectra in the presence of 100 mN K^+ concentration of all four *T. cruzi* sequences; (E) CD-monitored thermal denaturation experiment for RCr sequence; (F) NMR spectra of imino protons of the sequence RCr in 10 % D_2O and 90% of buffer 20 mM potassium phosphate pH=6.9, 70 mM KCl.

Secondly, we investigated UV thermal difference spectroscopy (TDS) for the four sequences (Figure 3B and S2) by subtracting the spectra recorded at 95 and 25 °C. We observed a typical signature of G-quadruplex folding with major positive peaks at 247, 257, and 274 nm, together with a negative peak at 296 nm, for the sequences RCr, TCr and MCr, but not for DCr that showed only a maximum at 270 nm. Next, isothermal difference spectra experiments (IDS) were carried out for the four sequences by taking the difference between the absorbance spectra of folded and unfolded oligonucleotides. These spectra were recorded with and without potassium cation (100 mM KCl) at 25 °C (Figure 3C and S3). Again, the spectra of RCr, TCr and MCr, but not DCr showed a characteristic signature of a G4 structure, with major positive peaks at 244 and 274 nm and a negative peak at 296 nm.

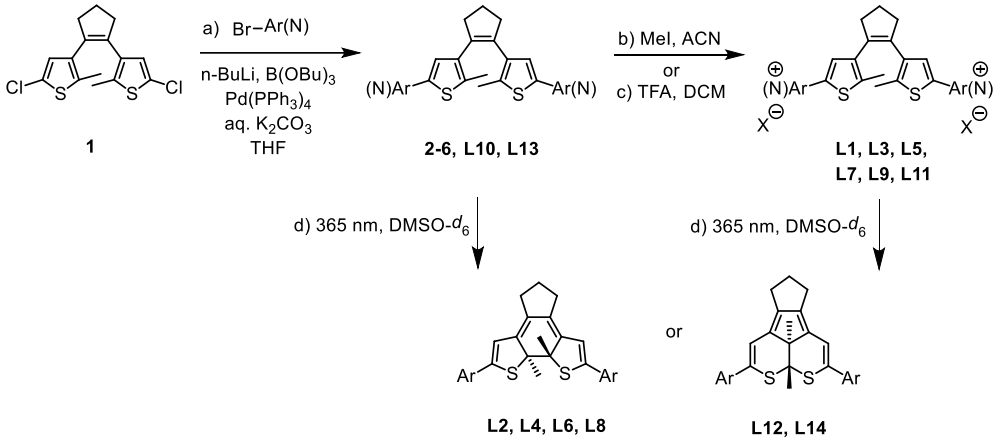
Thirdly, we performed circular dichroism (CD) spectroscopy for all the sequences in the presence of standard K^+ concentration to allow potential G-quadruplex formation. RCr, TCr and MCr showed positive signals at 260 and 295 nm and a negative signal at 240 nm, suggesting a predominantly parallel G-quadruplex conformation (Figure 3D). Moreover, these three sequences showed a typical CD melting curve with ΔT_m's between 55 and 85 °C (Figure S4). Surprisingly, the CD signal of RCr sequence did not reach zero values of ellipticity even at the highest temperature registered (95 °C) suggesting a notably stable G-quadruplex structure (Figure 3E). In fact, the CD signal displayed for the RCr sequence indicated the presence of a quadruplex structure, not only at 100 mM K^+ concentration, but also at 10 mM and even 1 mM K^+ concentration (Figure S5). However, the CD spectra of DCr sequence did not show a typical G4 signature and apparently, no secondary structure is present.

Finally, proton NMR spectra were recorded for the four sequences at 25 °C in aqueous solution. Spectra for RCr, TCr and MCr displayed the characteristic imino protons in the 10−12 ppm range and none at higher field (Figure 3F and S6), suggesting the exclusive folding into G4 structures. However, whereas the spectrum of RCr showed clear separated signals for each of the 12 imino protons, the spectra displayed for TCr and MCr were a complex mixture of signals, indicating the likely coexistence of more than one conformation. Interestingly, RCr kept showing the imino signals even at 80 °C, indicating that the secondary structure is extremely stable even at very high temperatures (Figure S7). Again, DCr did not display the imino signals typical of G4 structures.

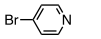
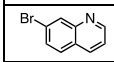
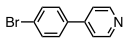
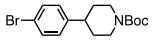
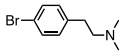
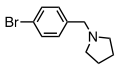
Synthesis of dithienylethene derivatives (L1-L14)

Recently, we reported the preparation of 1-methylpyridin-1-ium dithienylethene (DTE) derivative **L1** as a reversible photoswitchable G4 ligand.[53] Irradiation with 365 nm visible light converted compound **L1** into the closed version **L2** (Table 1, entry 1). Preliminary results obtained on the antiparasitic activity of **L1** and **L2** against *T. cruzi* Y strain were in the range of that found for first-line drug Bzn. These results encouraged us to prepare a small family of DTE derivatives with modified side chains and amino groups. To that end, we synthesized new derivatives following a straightforward strategy as previously used to access **L1** (Table 1, see ESI for full experimental details). In brief, Suzuki coupling of 1,2-bis(5-chloro-2-methylthiophen-3-yl)cyclopent-1-ene (**1**) with the corresponding bromoaryl derivatives afforded compounds **2-6**, **L10** and **L13** in 58-80% yield. Alkylation with iodomethane (step b) or Boc deprotection in TFA/DCM (step c) provided compounds **L1**, **L3**, **L5**, **L7**, **L9** and **L11** as the corresponding salts in excellent yields (70-99%). With the open forms of the ligands in hand, 10 mM solution of each sample were irradiated using a 365 nm UV lamp to obtain the corresponding closed form of G4 ligand candidates. The conversion of these photoreactions was determined accurately by NMR spectroscopy at different times (see ESI). For ligands **L3**, **L5** and **L7**, closed ligand forms **L4**, **L6** and **L8**, respectively, were successfully obtained (Figures S8-S10). However, after several hours of irradiation, **L8** did not afford the corresponding closed form ligand, whereas irradiation of **L11** and **L13** led to the formation of annulated ring systems **L12** and **L14**, respectively (Figures S11-S12). We attributed the formation of the annulated byproducts for these specific substrates to photochemical side reactions, which upon excitation of the ring-closed isomer with UV light undergo a formal 1,2-dyotropic rearrangement as previously reported in the literature for analogous systems e.g. perhydrocyclopentene and perfluorocyclopentene derivatives.[54–56]

Table 1. General procedure to obtain DTE G4 ligand candidates **L1-L14**



1 | **2-6, L10, L13** | **L1, L3, L5, L7, L9, L11** | **L2, L4, L6, L8** | **L12, L14**

Br-Ar(N)	Step a (%)	Step b (%, X = I)	Step c (X = TFA)	Step d (open/closed/byproduct)
	**2** (26%)[a]	**L1** (78%)[a]	-	**L2** (0/100/0)[a]
	**3** (80%)	**L3** (70%)	-	**L4** (13/87/0)
	**4** (60%)	**L5** (quant.)	-	**L6** (8/92/0)
	**5** (58%)	-	**L7** (99%)	**L8** (4/96/0)
	**6** (78%)	**L9** (98%)	-	-
	**L10** (79%)	**L11** (99%)	-	**L12** (0/3/97)
	**L13** (72%)	-	-	**L14** (1/3/96)

[a] Reported in ref. 53

In vitro antiparasitic activity and cytotoxicity of the dithienylethene derivatives L1-L14

The antiparasitic activity of the DTE derivatives was evaluated against epimastigotes of *T. cruzi* Y, SOL, BM Lopez and DA strains (Table 2). Cytotoxicity was determined in MRC-5 human lung fibroblast cells and this cell line was also used to determine the selectivity indexes (SI) against the parasites (SI = IC_{50} MRC-5 / IC_{50} *T. cruzi*). We also included three classical quadruplex ligands, pyridostatin (PDS), TMPyP4 (TM), Braco-19 (Braco) and the reference drug Benznidazole (Bzn) as controls.

Compounds **L2-L5**, **L10**, **L11** and **L13** resulted in IC_{50} values in the micromolar range for *T. cruzi* Y strain and were better than that of Bzn. Interestingly, the N-pyrrolidine-4-methylphenyl- **L10** and the N-dimethyl-4-phenyl- dithienylethene **L13** showed IC_{50} values in the

low micromolar range for all examined *T. cruzi* strains that were better than those displayed by Bzn. The three classical G-quadruplex ligands tested, PDS, TM and Braco showed poor antiparasitic activity against all *T. cruzi* strains, with IC_{50} values over 47.5 μM.

Regarding the selectivity index with respect to MRC-5 cells, we found that **L13** and **L10** SI values were only between 1.2 and 1.6- fold whereas for derivatives **L3** and **L4** SI were between 2 and 4- fold, similar to those observed for Bzn (Table 2). The therapeutic window observed for **L3** and **L4** encouraged us to measure their antiparasitic activity against trypomastigotes, the infectious form of *T. cruzi* present in the bloodstream of patients suffering ChD (Table 3). Remarkably, compounds **L3** and **L4** showed in IC_{50} values in the low micromolar range (1.5 and 3.3 μM, respectively) with an SI value of 40.9- and 25- fold, respectively, a much better antiparasitic activity than that measured for Bzn (IC_{50} value of 57.1 μM and SI >1).

Table 2. Antiparasitic activity of DTE derivative **L1-L14** against epimastigotes of four strains of *T. cruzi* (Y, SOL, BM Lopez and DA) and cytotoxicity values for MRC5 control cell line. Selectivity Index (SI) values are also shown. SI are IC_{50} MRC5/ IC_{50} *T. cruzi*. Classical G4 ligands, pyridostatin (PDS), TMPyP4 (TM) and Braco-19 (Braco). Benznidazole (Bzn) is a control drug. Best IC_{50} and SI values are highlighted in bold.

	IC_{50}					**Selective Index (SI)**			
	MRC5 (M)	Y	SOL	BM Lopez	DA	M/Y	M/SOL	M/BM	M/DA
L1	2.4 ± 0.2	23.1 ± 21.2	> 27.4	> 28.6	> 28.2	<1	<1	<1	<1
L2	3.3 ± 0.01	**13.9 ± 11.2**	> 18.3	> 25.6	30.2 ± 13.6	<1	<1	<1	<1
L3	61.4 ± 10.1	**19.5 ± 3.2**	20.8 ± 4.5	30.1 ± 5.1	21.2 ± 0.6	**3.1**	**2.9**	**2.0**	**2.9**
L4	82.5 ± 15.1	**20.9 ± 2.2**	22.3 ± 6.1	40.7 ± 17.5	20.6 ± 4.1	**3.9**	**3.7**	**2.0**	**4**
L5	2.2 ± 0.1	**14.0 ± 10.3**	37.15 ± 4.6	25.25 ± 0.9	28.4 ± 13.8	<1	<1	<1	<1
L6	4.9 ± 2.3	42.8 ± 27.3	> 50	60.1 ± 14.6	36.6 ± 6.2	<1	<1	<1	<1
L7	8.0 ± 2.2	29.5 ± 20.7	30.8 ± 7.1	29.8 ± 17.7	13.9 ± 15.6	<1	<1	<1	<1
L8	4.3 ± 0.4	> 50	> 50	> 50	32.0 ± 13.5	<1	<1	<1	<1
L9	20.5 ± 1.4	22.8 ± 1.5	20.4 ± 0.0	22.7 ± 4.1	14.4 ± 0.4	<1	1	<1	**1.4**
L10	2.6 ± 0.07	**3.6 ± 1.9**	**3.8 ± 1.3**	**4.0 ± 0.6**	**2.7 ± 1.0**	<1	<1	<1	<1
L11	2.7 ± 0.2	**14.2 ± 6.6**	21.15 ± 4.7	19.1 ± 0.4	6.8 ± 2.4	<1	<1	<1	<1
L12	6.1 ± 3.2	31.1 ± 13.7	45.1 ± 1.0	> 48.4	13.4 ± 1.3	<1	<1	<1	<1
L13	7.2 ± 0.04	**6.6 ± 4.2**	**9.75 ± 2.9**	**4.6 ± 0.8**	**6.1 ± 1.3**	**1.1**	<1	**1.6**	**1.2**
L14	22.5 ± 1.4	> 34	> 50	> 50	40.1 ± 12.2	<1	<1	<1	<1
PDS	5.4 ± 0.1	> 100	> 100	> 100	> 100	<1	<1	<1	<1
TM	> 25	> 100	81.2 ± 6.2	70.6 ± 2.1	69.0 ± 1.2	<1	<1	<1	<1
Braco	8.3 ± 3.0	> 100	> 100	49.4 ± 3.2	47.5 ± 2.2	<1	<1	<1	<1
Bzn	47.2 ± 1.5	20.3 ± 2.7	11.2 ± 2.8	6.3 ± 4.3	n.d.	**2.3**	**4.2**	**7.5**	n.d.

Table 3. Antiparasitic activity of compounds **L3** and **L4** against trypomastigotes of *T. cruzi* SOL and cytotoxicity values for MRC5 control cell line. Selectivity Index (SI) values are also shown. SI are IC_{50} MRC5/ IC_{50} *T. cruzi*. Benznidazole (Bzn) is a control drug.

	IC_{50}		Selective Index (SI)
	MRC5	*T. cruzi SOL*	MRC5/*T.cruzi*
L3	61.4 ± 10.1	1.5 ± 0.1	40.9
L4	82.5 ± 15.1	3.3 ± 1.1	25
Bzn	47.2 ± 1.5	57.1 ± 3.2	< 1

Binding of compounds L3 and L4 to DNA quadruplex sequences

A potential mode of action for the antichagasic activity of this family of compounds is through their binding to DNA G-quadruplexes of *T. cruzi*. Since we have confirmed that the RCr, TCr and MCr sequences of the genome of *T. cruzi* fold into G4 structures, we proceeded to examine the potential binding of **L3** and **L4** to them. Firstly, we conducted FRET thermal melting assays of **L3** and **L4** against fluorophore-labelled RCr, TCr and MCr G4s from *T. cruzi*, Human telomere (F21T) and MYC as control G4s and FdxT as a DNA control duplex. In this experiment, we monitor the ligand-induced stabilization of the quadruplex structure by the change in apparent melting temperature (ΔT_m) of the folded species (see Figure 4A, 4B and S13). Both compounds **L3** and **L4** showed clear binding to all the quadruplex structures examined with ΔT_m values of 17-20 °C, except for binding to MCr quadruplex that was around 14 °C. A small difference in G4 stabilization was observed between **L3** and **L4**, and both clearly established their selectivity for G4 DNA over duplex DNA.

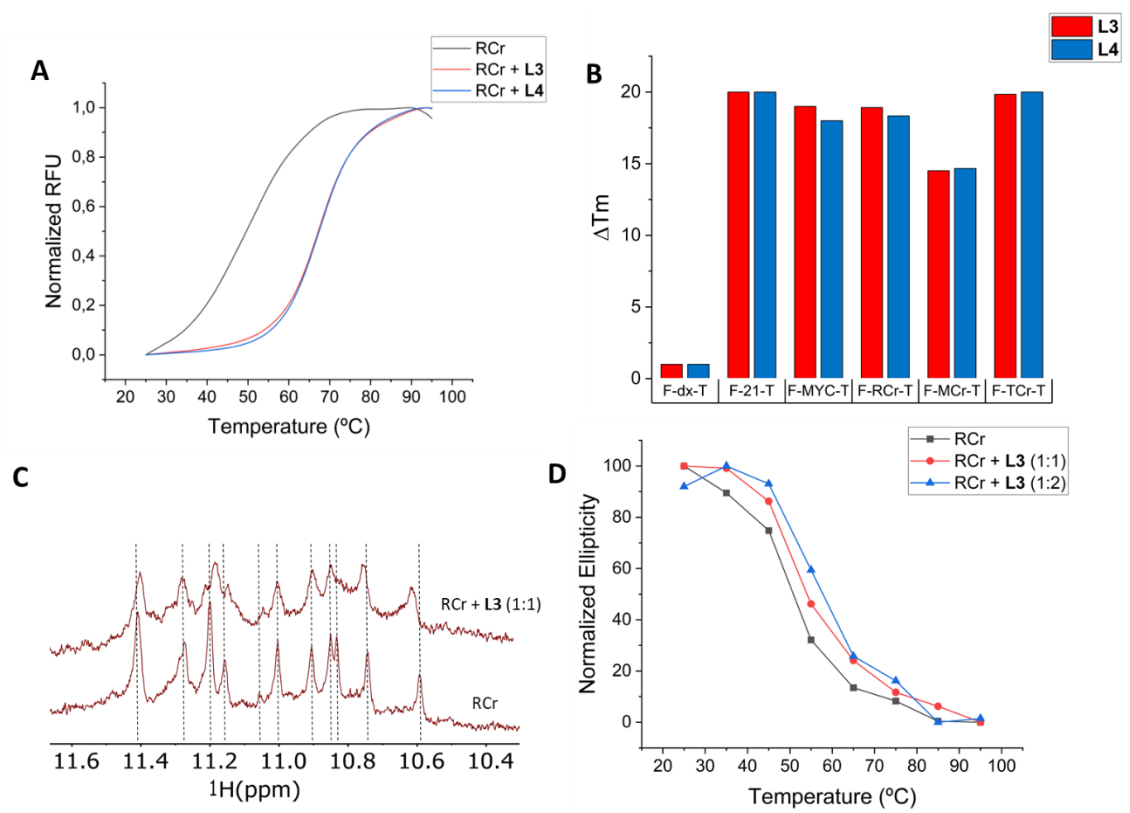


Figure 4. A) FRET melting assay. FRET melting curves of RCr (fluorescently labeled) quadruplex (0.2 μM) in the absence and in the presence of the tested compounds **L3** and **L4** (10 μM) in 10 mM lithium cacodylate, pH 7.2, containing 10 mM KCl + 90 mM LiCl. (B) FRET melting assay. Thermal stabilization induced by the tested compounds **L3** and **L4** (5 μM) on the different DNA quadruplexes (F21T, FMycT, F-RCr-T, F-MCr-T and F-TCr-T) and on a duplex DNA (F-dx-T) at 0.2 μM concentration in 10 mM lithium cacodylate, pH 7.2, containing 10 mM KCl + 90 mM LiCl. The errors in ΔT_m were a maximum of ± 1 °C, except for RCr + **L3** which was ± 1.42 °C; (C) imino region of the ^{1}H-NMR spectra of RCr in aqueous buffer solution (10% D_2O/ 90% aqueous buffer 20 mM potassium phosphate pH =6.9, 70 mM KCl) in the absence and presence of compound **L3**. (D) CD melting curve of RCr quadruplex (3 μM) in the absence and presence of compound **L3** in in 10 mM lithium cacodylate, pH 7.2, 1 mM KCl + 99 mM LiCl.

Secondly, we carried out further FRET melting experiments under competing conditions of increasing concentrations of double stranded DNA (ds26, 0, 1, 3 and 10 μM, see Figure S9). The results showed that ΔT_m values barely changed when **L3** and **L4** are bound to F21T G4 in the presence of duplex DNA, confirming their selectivity for quadruplex over duplex DNA. Similar

results were also observed for the ligands when binding to RCr quadruplex from *T. cruzi* (see also Figure S14).

Thirdly, ^{1}H-NMR studies were undertaken to investigate the binding of **L3** and **L4** to RCr, TCr and MCr G4s from *T. cruzi*. A slight broadening of all imino proton signals was observed when **L3** was added to an aqueous solution of the RCr quadruplex (Figure 4C). This result together with δ shift for 8 out of the 11 signals, confirmed binding of compound **L3**. One of the imino protons (12 are expected from the three potential tetrads in the quadruplex) is not observed since it must be overlapped with other imino protons in the spectra. A similar behavior was observed when **L4** was added to the RCr quadruplex, also supporting binding (Figure S15). In the case of addition of **L3** or **L4** to the TCr or MCr quadruplexes (Figure S15) the imino region of the ^{1}H-NMR spectra showed some signal broadening but it was quite complex to interpret.

Fourthly, CD melting experiments were carried out on RCr and MCr sequences in the absence and presence of compounds **L3** and **L4** to test for binding. We observed an increase in ΔTm for both compounds when added into a 1:2 ratio to RCr quadruplex (6.98 and 3.81 °C, respectively, see Figure 4D and S16, and Table S5). These results indicate that both derivatives bind this G4 found in the genome of *T. cruzi* with compound **L3** showing a stronger binding. Similar results were found in the case of MCr in the presence of our ligands (Figure S17).

Finally, we conducted UV experiments to further confirm binding of **L3** and **L4** to RCr and MCr G4 sequences and their selectivity for quadruplexes over duplex DNA. In both cases, the maximum observed in the UV spectra of **L3** and **L4** decreased considerably in absorption and moved to higher wavelength upon addition of the corresponding G4 sequences (Figure S18). In contrast, addition of duplex DNA showed a slight decrease in the absorbance of the maximum in the UV spectra and a small shift to lower wavelength (Figure S19), confirming both ligands exhibit G4 selectivity over duplex DNA.

Localization of compound L3 in Trypanosoma cruzi strains

We investigated the localization of compound **L3** in the *T. cruzi* BM López strain by fluorescence microscopy using its intrinsic fluorescence properties. In the case of compound **L4**, its fluorescence intensity was too low to allow this type of experiment. After 2h of incubation at 37 °C, **L3** was detected inside the parasites with no clear location at this time point (Figure 5). At 4h, **L3** can be localized mainly in the kinetoplast and partially in the nucleus. Then, at a longer incubation time such as 24h the compound can be observed clearly found within the nucleus and kinetoplast what could suggest that this type of G4-ligands could reach and possibly target DNA G-quadruplex structures found in the parasite's genome.

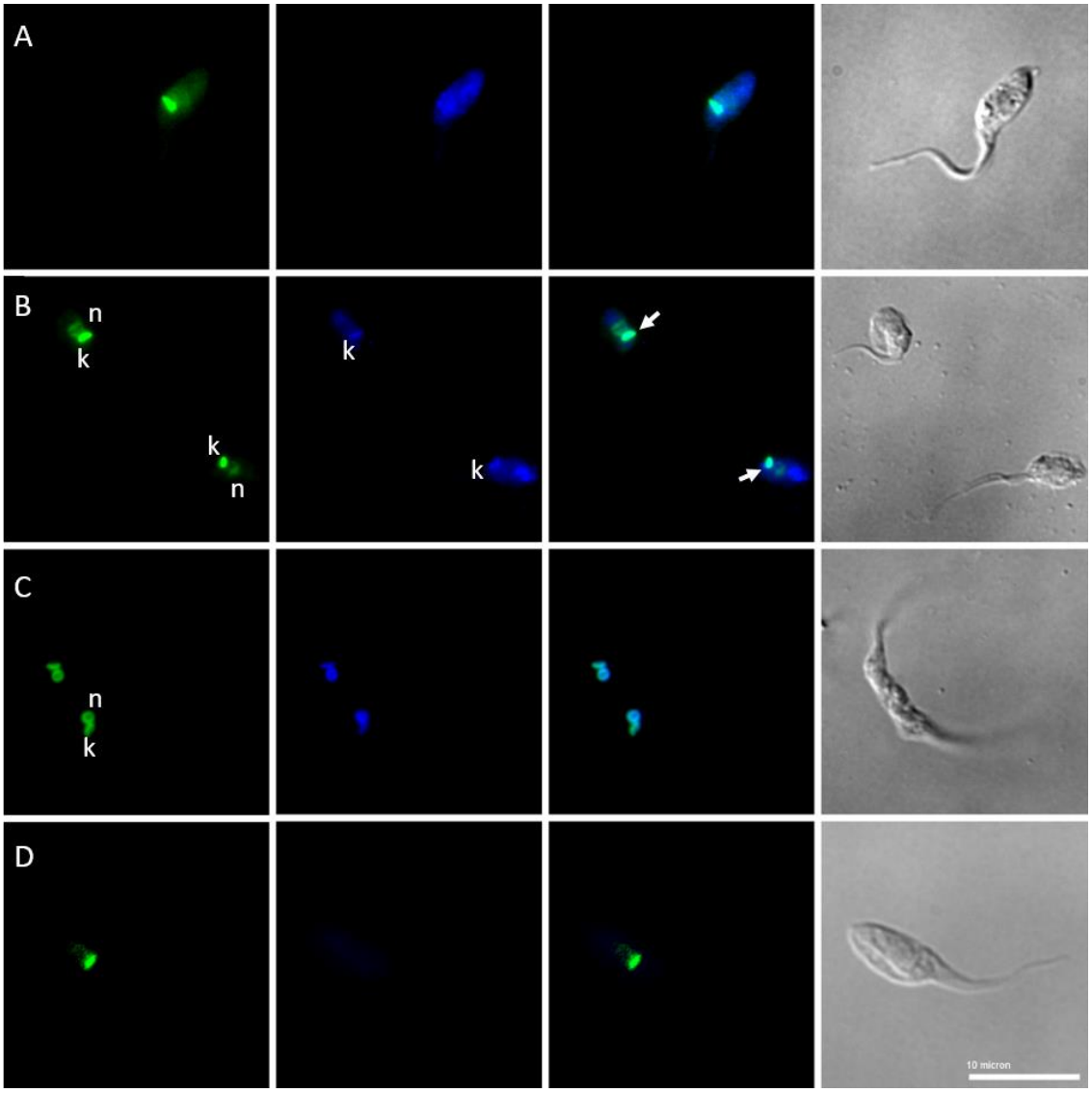


Figure 5. Confocal images of *T. cruzi* BM López parasites after incubation at 20 μM concentration with compound **3** for 2h (A), 4h (B) and 24h (C) and the negative control (D). From left to right nuclei staining, emission of compound **L3**, overlay and clear field, respectively. n= nucleus; k= kinetoplasm.

Conclusions

In summary, we have found a higher number of PQSs in the genome of *T. cruzi* than those in other parasites except Leishmania. At the same time, those PQSs with probability of forming G4s were higher than those found in other parasites (again except *Leishmania*), yet slightly smaller when compared to humans. Among those PQS sequences that appear with high frequency in the

genome of *T. cruzi*, we have proved with different biophysical techniques that three of them (RCr, TCr and MCr) form quadruplex structures. These results prompted us to propose G4s as potential therapeutic targets against *T. cruzi* infections causing ChD as it happened recently with other parasitic, bacterial and viral infections.

After screening several families of G4 ligands, two dithienylethene derivatives **L1** and **L2** showed antiparasitic activity against *T. cruzi* Y strain similar to that of Bzn, a drug currently used to treat patients suffering ChD. Then, we identified **L3** and **L4** from a small library of new DTE derivatives synthesized as compounds with comparable efficacy and therapeutic window to Bzn against epimastigotes of different strains of *T. cruzi*. Remarkably, **L3** and **L4** showed low micromolar activity against trypomastigotes of *T. cruzi* SOL being 38- and 17-fold more active than Bzn and with a much wider therapeutic window.

Next, we studied binding of **L3** and **L4** to quadruplex sequences using FRET, CD, UV and NMR experiments. We confirmed that **L3** and **L4** were capable of binding to different quadruplex sequences such as the G4 telomere and MYC, but also to those found in the genome of *T. cruzi* such as RCr, MCr and TCr. Fluorescence microscopy studies indicated that **L3** entered rapidly into *T. cruzi* parasites in short times, and finally localize into the nucleus and kinetoplast after 24h where they could bind to their potential G4 targets. Accordingly, we propose that G4s in the genome of *T. cruzi* could be potential therapeutic targets, and **L3** and **L4** seem to assess their antiparasitic activity using this mode of action.

Acknowledgments

JCM thanks the funding by the Spanish Ministerio de Ciencia, Innovación y Universidades (PID2021-127109OB-I00). EBR received funding support from a Maria Zambrano fellowship, financed by NextGenerationEU/European Union. MCG thanks ERC-COG: 648239, EPSRC Funds EP/R043361/1, TS/R014329/1 and GCRF EP/T020288/1 for financial support and JRS thanks MSCA fellowship (project 843720-BioNanoProbes). A.C. and J.L.M. thank the generous

benefactors of the Fondation X for their support. MCL and MCT were supported by grant PID2019-109090RB-I00 /AEI/10.13039/501100011033 from the Programa Estatal I+D+I, Spanish Ministry of Science, Innovation and Universities (MICINN). We thank A. López-Barajas from IPBLN-CSIC for the relevant technical assistance in the biological studies.

Conflict of Interest

The authors declare no conflict of interest

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords

Trypanosome cruzi, Chagas disease, DNA G-quadruplex, therapeutic target, ligand, drug

References

[1] "World Health Organization, Chagas disease," can be found under https://www.who.int/n%0Aews-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis), **2024**.

[2] C. B. Scarim, D. H. Jornada, R. C. Chelucci, L. de Almeida, J. L. dos Santos, M. C. Chung, *Eur. J. Med. Chem.* **2018**, *155*, 824–838.

[3] B. I. Pelizaro, J. C. Z. Batista, G. B. Portapilla, A. R. das Neves, F. Silva, D. B. Carvalho, C. Y. K. Shiguemoto, L. R. Pessatto, E. J. Paredes-Gamero, I. A. Cardoso, P. H. Luccas, M. C. Nonato, N. P. Lopes, F. Galvão, K. M. P. Oliveira, N. S. Cassemiro, D. B. Silva, E. M. Piranda, C. C. P. Arruda, S. de Albuquerque, A. C. M. Baroni, *J. Med. Chem.* **2024**, DOI 10.1021/ACS.JMEDCHEM.3C01745.

[4] M. Haroon, M. C. H. de Barros Dias, A. C. da S. Santos, V. R. A. Pereira, L. A. Barros Freitas, R. B. Balbinot, V. Kaplum, C. V. Nakamura, L. C. Alves, F. A. Brayner, A. C. L. Leite, T. Akhtar, *RSC Adv.* **2021**, *11*, 2487–2500.

[5] M. Cox Holanda de Barros Dias, M. Souza Barbalho, G. Bezerra de Oliveira Filho, M. Veríssimo de Oliveira Cardoso, A. C. Lima Leite, A. C. da Silva Santos, A. C. Cristovão Silva, M. C. Accioly Brelaz de Castro, D. Maria Nascimento Moura, L. F. Gomes Rebello Ferreira, M. Zaldini Hernandes, R. de Freitas e Silva, V. Rêgo Alves Pereira, *Eur. J. Med. Chem.* **2023**, *257*, 115508.

[6] R. C. F. M. Reis, E. G. dos Santos, M. D. Benedetti, A. C. C. Reis, G. C. Brandão, G. N. da Silva, L. A. Diniz, R. S. Ferreira, I. S. Caldas, S. F. P. Braga, T. B. de Souza, *Eur. J. Med. Chem.* **2023**, *258*, 115622.

[7] A. C. L. Leite, R. S. de Lima, D. R. de M. Moreira, M. V. de O. Cardoso, A. C. Gouveia de Brito, L. M. Farias dos Santos, M. Z. Hernandes, A. C. Kiperstok, R. S. de Lima, M. B. P. Soares, *Bioorg. Med. Chem.* **2006**, *14*, 3749–3757.

[8] A. C. L. Leite, D. R. de M. Moreira, M. V. de O. Cardoso, M. Z. Hernandes, V. R. Alves Pereira, R. O. Silva, A. C. Kiperstok, M. da S. Lima, M. B. P. Soares, *ChemMedChem* **2007**, *2*, 1339–1345.

[9] P. Linciano, C. Pozzi, G. Tassone, G. Landi, S. Mangani, M. Santucci, R. Luciani, S. Ferrari, N. Santarem, L. Tagliazucchi, A. Cordeiro-da-Silva, M. Tonelli, D. Tondi, L. Bertarini, S. Gul, G. Witt, C. B. Moraes, L. Costantino, M. P. Costi, *Eur. J. Med. Chem.* **2024**, *264*, 115946.

[10] M. L. Silva, M. K. de Santiago-Silva, M. Fabris, G. P. Camargo, M. de Lima Ferreira Bispo, *Curr. Drug Targets* **2023**, *24*, 781–789.

[11] M. G. Thomas, K. McGonagle, P. Rowland, D. A. Robinson, P. G. Dodd, I. Camino-Díaz, L. Campbell, J. Cantizani, P. Castañeda, D. Conn, P. D. Craggs, D. Edwards, L. Ferguson, A. Fosberry, L. Frame, P. Goswami, X. Hu, J. Korczynska, L. MacLean, J. Martin, N. Mutter, M. Osuna-Cabello, C. Paterson, I. Peña, E. G. Pinto, C. Pont, J. Riley, Y. Shishikura, F. R. C. Simeons, L. Stojanovski, J. Thomas, K. Wrobel, R. J. Young, F. Zmuda, F. Zuccotto, K. D. Read, I. H. Gilbert, M. Marco, T. J. Miles, P. Manzano, M.

De Rycker, *J. Med. Chem.* **2023**, *66*, 10413–10431.

[12] J. J. Nué-Martinez, D. Cisneros, M. del V. Moreno-Blázquez, C. Fonseca-Berzal, J. I. Manzano, D. Kraeutler, M. A. Ungogo, M. A. Aloraini, H. A. A. Elati, A. Ibáñez-Escribano, L. Lagartera, T. Herraiz, F. Gamarro, H. P. de Koning, A. Gómez-Barrio, C. Dardonville, *J. Med. Chem.* **2023**, *66*, 13452–13480.

[13] T. A. Brooks, L. H. Hurley, *Genes Cancer* **2010**, *1*, 641–649.

[14] X. Cui, H. Chen, Q. Zhang, M. Xu, G. Yuan, J. Zhou, *Sci. Rep.* **2019**, *9*, 3966.

[15] J. Spiegel, S. M. Cuesta, S. Adhikari, R. Hänsel-Hertsch, D. Tannahill, S. Balasubramanian, *Genome Biol.* **2021**, *22*, 117.

[16] S. Lago, M. Nadai, F. M. Cernilogar, M. Kazerani, H. Domíniguez Moreno, G. Schotta, S. N. Richter, *Nat. Commun.* **2021**, *12*, 3885.

[17] P. Prorok, M. Artufel, A. Aze, P. Coulombe, I. Peiffer, L. Lacroix, A. Guédin, J.-L. Mergny, J. Damaschke, A. Schepers, C. Cayrou, M.-P. Teulade-Fichou, B. Ballester, M. Méchali, *Nat. Commun.* **2019**, *10*, 3274.

[18] A. V Pavlova, E. A. Kubareva, M. V Monakhova, M. I. Zvereva, N. G. Dolinnaya, *Biomolecules* **2021**, *11*, DOI 10.3390/biom11091284.

[19] A. De Magis, S. G. Manzo, M. Russo, J. Marinello, R. Morigi, O. Sordet, G. Capranico, *Proc. Natl. Acad. Sci.* **2019**, *116*, 816–825.

[20] A. B. Sahakyan, P. Murat, C. Mayer, S. Balasubramanian, *Nat. Struct. Mol. Biol.* **2017**, *24*, 243–247.

[21] P. Murat, G. Marsico, B. Herdy, A. Ghanbarian, G. Portella, S. Balasubramanian, *Genome Biol.* **2018**, *19*, 229.

[22] S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat Rev Drug Discov* **2011**, *10*, 261–275.

[23] J. Figueiredo, J.-L. Mergny, C. Cruz, *Life Sci.* **2024**, *340*, 122481.

[24] J. Xu, H. Huang, X. Zhou, *JACS Au* **2021**, *1*, 2146–2161.

[25] A. Verma, K. Halder, R. Halder, V. K. Yadav, P. Rawal, R. K. Thakur, F. Mohd, A. Sharma, S. Chowdhury, *J Med Chem* **2008**, *51*, 5641–5649.

[26] S. G. Hershman, Q. Chen, J. Y. Lee, M. L. Kozak, P. Yue, L. S. Wang, F. B. Johnson, *Nucleic Acids Res* **2008**, *36*, 144–156.

[27] P. Rawal, V. B. Kummarasetti, J. Ravindran, N. Kumar, K. Halder, R. Sharma, M. Mukerji, S. K. Das, S. Chowdhury, *Genome Res* **2006**, *16*, 644–655.

[28] M. Wieland, J. S. Hartig, *Nat Protoc* **2009**, *4*, 1632–1640.

[29] R. Cebrián, E. Belmonte-Reche, V. Pirota, A. de Jong, J. C. Morales, M. Freccero, F. Doria, O. P. Kuipers, *J. Med. Chem.* **2022**, *65*, 4752–4766.

[30] R. Perrone, M. Nadai, J. A. Poe, I. Frasson, M. Palumbo, G. Palu, T. E. Smithgall, S. N. Richter, *PLoS One* **2013**, *8*, e73121.

[31] S. Amrane, A. Kerkour, A. Bedrat, B. Vialet, M. L. Andreola, J. L. Mergny, *J Am Chem Soc* **2014**, *136*, 5249–5252.

[32] S. Artusi, M. Nadai, R. Perrone, M. A. Biasolo, G. Palu, L. Flamand, A. Calistri, S. N. Richter, *Antivir. Res* **2015**, *118*, 123–131.

[33] E. Belmonte-Reche, M. Martinez-Garcia, A. Guedin, M. Zuffo, M. Arevalo-Ruiz, F. Doria, J. Campos-Salinas, M. Maynadier, J. J. Lopez-Rubio, M. Freccero, J. L. Mergny, J. M. Perez-Victoria, J. C. Morales, *J Med Chem* **2018**, *61*, 1231–1240.

[34] N. Smargiasso, V. Gabelica, C. Damblon, F. Rosu, E. De Pauw, M. P. Teulade-Fichou, J. A. Rowe, A. Claessens, *BMC Genomics* **2009**, *10*, 362.

[35] L. M. Harris, K. R. Monsell, F. Noulin, T. Famodimu M, N. Smargiasso, C. Damblon, P. Horrocks, C. J. Merrick, *Antimicrob. Agents Chemother.* **2018**, *62*, e01828-17.

[36] M. Arevalo-Ruiz, F. Doria, E. Belmonte-Reche, A. De Rache, J. Campos-Salinas, R. Lucas, E. Falomir, M. Carda, J. M. Perez-Victoria, J. L. Mergny, M. Freccero, J. C. Morales, *Chem. Eur. J.* **2017**, *23*, 2157–2164.

[37] M. Zuffo, A. Stucchi, J. Campos-Salinas, M. Cabello-Donayre, M. Martínez-García, E. Belmonte-Reche, J. M. Pérez-Victoria, J. L. Mergny, M. Freccero, J. C. Morales, F. Doria, *Eur. J. Med. Chem.* **2019**, *163*, 54–66.

[38] E. Belmonte-Reche, A. Benassi, P. Peñalver, A. Cucchiarini, A. Guédin, J. L. Mergny, F. Rosu, V. Gabelica, M. Freccero, F. Doria, J. C. Morales, *Eur. J. Med. Chem.* **2022**, *232*, 114183.

[39] H. M. Craven, G. Nettesheim, P. Cicuta, A. M. Blagborough, C. J. Merrick, *Int. J. Parasitol. Drugs Drug Resist.* **2023**, *23*, 106–119.

[40] M. P. O'Hagan, P. Peñalver, R. S. L. Gibson, J. C. Morales, M. C. Galan, *Chem. – A Eur. J.* **2020**, *26*, 6224–6233.

[41] M. Pérez-Soto, P. Peñalver, S. T. G. Street, D. Weenink, M. P. O'Hagan, J. Ramos-Soriano, Y. J. Jiang, G. J. Hollingworth, M. C. Galan, J. C. Morales, *Bioorg. Med. Chem.* **2022**, *71*, 116946.

[42] S. T. G. Street, P. Peñalver, M. P. O. Hagan, G. J. Hollingworth, J. C. Morales, M. C. Galan, *Chem. - A Eur. J.* **2021**, *2*, 7712–7721.

[43] J. Guillon, A. Cohen, R. N. Das, C. Boudot, N. M. Gueddouda, S. Moreau, L. Ronga, S. Savrimoutou, L. Basmaciyan, C. Tisnerat, S. Mestanier, S. Rubio, S. Amaziane, A. Dassonville-Klimpt, N. Azas, B. Courtioux, J.-L. Mergny, C. Mullié, P. Sonnet, *Chem. Biol. Drug Des.* **2018**, *91*, 974–995.

[44] J. Guillon, A. Cohen, C. Boudot, S. Monic, S. Savrimoutou, S. Moreau, S. Albenque-Rubio, C. Lafon-Schmaltz, A. Dassonville-Klimpt, J.-L. Mergny, L. Ronga, M. Bernabeu de Maria, J. Lamarche, C. D. Lago, E. Largy, V. Gabelica, S. Moukha, P. Dozolme, P. Agnamey, N. Azas, C. Mullié, B. Courtioux, P. Sonnet, *Pathogens* **2022**, *11*, DOI 10.3390/pathogens11111339.

[45] J. Guillon, A. Cohen, N. M. Gueddouda, R. N. Das, S. Moreau, L. Ronga, S. Savrimoutou, L. Basmaciyan, A. Monnier, M. Monget, S. Rubio, T. Garnerin, N. Azas,

J. L. Mergny, C. Mullie, P. Sonnet, *J Enzym. Inhib Med Chem* **2017**, *32*, 547–563.

[46] J. Guillon, A. Cohen, C. Boudot, A. Valle, V. Milano, R. N. Das, A. Guédin, S. Moreau, L. Ronga, S. Savrimoutou, M. Demourgues, E. Reviriego, S. Rubio, S. Ferriez, P. Agnamey, C. Pauc, S. Moukha, P. Dozolme, S. Da Nascimento, P. Laumaillé, A. Bouchut, N. Azas, J.-L. Mergny, C. Mullié, P. Sonnet, B. Courtioux, *J. Enzyme Inhib. Med. Chem.* **2020**, *35*, 432–459.

[47] L. Monti, M. Di Antonio, *ChemBioChem* **2023**, *24*, e202300265.

[48] E. Belmonte-Reche, J. C. Morales, *NAR Genomics Bioinforma.* **2020**, *2*, lqz005.

[49] J. Mohanty, N. Barooah, V. Dhamodharan, S. Harikrishna, P. I. Pradeepkumar, A. C. Bhasikuttan, *J. Am. Chem. Soc.* **2013**, *135*, 367–376.

[50] A. Renaud de la Faverie, A. Guedin, A. Bedrat, L. A. Yatsunyk, J. L. Mergny, *Nucleic Acids Res* **2014**, *42*, e65.

[51] S. Xu, Q. Li, J. Xiang, Q. Yang, H. Sun, A. Guan, L. Wang, Y. Liu, L. Yu, Y. Shi, H. Chen, Y. Tang, *Sci. Rep.* **2016**, *6*, 24793.

[52] A. Yett, L. Y. Lin, D. Beseiso, J. Miao, L. A. Yatsunyk, *J. Porphyr. Phthalocyanines* **2019**, *23*, 1195–1215.

[53] M. P. O'Hagan, J. Ramos-Soriano, S. Haldar, S. Sheikh, J. C. Morales, A. J. Mulholland, M. C. Galan, *Chem. Commun.* **2020**, *56*, 5186–5189.

[54] M. Herder, B. M. Schmidt, L. Grubert, M. Pätzel, J. Schwarz, S. Hecht, *J. Am. Chem. Soc.* **2015**, *137*, 2738–2747.

[55] T. Hirose, Y. Inoue, J. Hasegawa, K. Higashiguchi, K. Matsuda, *J. Phys. Chem. A* **2014**, *118*, 1084–1093.

[56] E. C. Harvey, J. Areephong, A. A. Cafolla, C. Long, W. R. Browne, B. L. Feringa, M. T. Pryce, *Organometallics* **2014**, *33*, 447–456.