Structure-Based and Molecular Modeling Studies for the Discovery of Cyclic Imides as Reversible Cruzain Inhibitors With Potent Anti-Trypanosoma cruzi Activity

Rafael A. A. Ferreira 1†, Ivani Pauli 2†, Thiago S. Sampaio 1, Mariana L. de Souza 2, Leonardo L. G. Ferreira 2, Luma G. Magalhães 2, Celso de O. Rezende Jr. 1, Rafaela S. Ferreira 2, Renata Krogh 2, Luiz C. Dias 1* and Adriano D. Andricopulo 2*

1 Instituto de Química, Universidade Estadual de Campinas, Campinas, Brazil, 2 Laboratório de Química Medicinal e Computacional, Centro de Pesquisa e Inovação em Biodiversidade e Fármacos, Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, Brazil, 3 Departamento de Bioquímica e Imunologia, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

Chagas disease causes ~10,000 deaths each year, mainly in Latin America, where it is endemic. The currently available chemotherapeutic agents are ineffective in the chronic stage of the disease, and the lack of pharmaceutical innovation for Chagas disease highlights the urgent need for the development of new drugs. The enzyme cruzain, the main cysteine protease of Trypanosoma cruzi, has been explored as a validated molecular target for drug discovery. Herein, the design, molecular modeling studies, synthesis, and biological evaluation of cyclic imides as cruzain inhibitors are described. Starting with a micromolar-range cruzain inhibitor (3a, IC$_{50}$ = 2.2 µM), this molecular optimization strategy resulted in the nanomolar-range inhibitor 10j (IC$_{50}$ = 0.6 µM), which is highly active against T. cruzi intracellular amastigotes (IC$_{50}$ = 1.0 µM). Moreover, most compounds were selective toward T. cruzi over human fibroblasts, which were used as host cells, and are less toxic to hepatic cells than the marketed drug benznidazole. This study enabled the discovery of novel chemical diversity and established robust structure-activity relationships to guide the design of optimized cruzain inhibitors as new trypanocidal agents.

Keywords: Chagas disease, Trypanosoma cruzi, cruzain, SAR, medicinal chemistry, synthesis, inhibitors, molecular docking

INTRODUCTION

Caused by the protozoan Trypanosoma cruzi and endemic in 21 countries in Latin America, Chagas disease kills ~10,000 people each year. This neglected tropical disease has reached non-endemic regions, affecting 8 million people worldwide and putting another 25 million at risk of infection. The USA shows the greatest burden among non-endemic countries, with ~300,000 people estimated to be infected with T. cruzi (Pérez-Molina and Molina, 2018). In addition, Chagas disease is a major cause of infectious cardiomyopathy worldwide, contributing substantially to the global burden of cardiovascular disease (Bern, 2015; Cucunubá et al., 2016). Mortality and a

1https://www.who.int/chagas/en/
reduction in productivity of the affected populations significantly impact the economies of the endemic regions. These economic and social burdens can be translated in numbers that estimate losses of more than US $7.2 billion per year and ~243,600 disability-adjusted life years (DALYs) due to Chagas disease (GBD DALYs and HALE Collaborators, 2016; Arnal et al., 2019).

Even more than a century after the discovery of Chagas disease by Brazilian physician Carlos Chagas in 1909, current chemotherapy for this condition relies on two drugs only—benznidazole (BZ) and nifurtimox (Dias et al., 2014). These nitroheterocyclic compounds, which were identified between the mid-1960s and 1970s, are effective only when administered during the acute stage of the disease, a limitation that leaves millions of chronic chagasic patients without appropriate treatment (Molina et al., 2015). Moreover, benznidazole and nifurtimox cause severe adverse effects in up to 40% of patients, leading to poor adherence to treatment (Pérez-Molina and Molina, 2018). These drawbacks highlight the urgent need for the development of effective and safe drugs for the therapy of Chagas disease (Olivera et al., 2015; Ferreira and Andricopulo, 2019).

The enzyme cruzain (EC 3.4.22.51) is the main cysteine protease of *T. cruzi* and has been explored as a validated molecular target in Chagas disease drug discovery (McKerrow, 1999; Jose Cazzulo et al., 2001). It is expressed throughout the life cycle of *T. cruzi* and is involved in critical biological processes such as the interaction with host cells, parasite reproduction and evasion from the host immunologic system (Engel et al., 1998; Ferreira and Andricopulo, 2017). Cruzain has been validated as a molecular target for Chagas disease drug discovery based on genetic studies of *T. cruzi* and the ability of cruzain inhibitors to decrease parasite burden in vivo (Doyle et al., 2011; Ndao et al., 2014). These studies have recently supported the design and identification of several classes of cruzain inhibitors, including vinyl sulfones, triazoles, pyrimidines, thiosemicarbazones, chalcones, nitroalkenes, and benzimidazoles (Rogers et al., 2012; Ferreira et al., 2014; Avelar et al., 2015; Espíndola et al., 2015; Neitz et al., 2015; Latorre et al., 2016). The vinyl sulfone K777, which is a covalent cruzain inhibitor, showed promising results in preclinical efficacy tests; however, toxicity-related drawbacks prevented the compound from progressing into advanced clinical development (Ndao et al., 2014). The poor safety profile of K777 was associated with the irreversible mode of action of the compound. Following the failure of K777, the pursuit of novel cruzain inhibitors has recently focused on the design of reversible ligands. These investigations, along with the available structural data of cruzain bound with small molecule ligands, have been key to promoting the discovery of novel classes of inhibitors with improved safety profiles. Moreover, these data have enabled the integration of experimental and computational approaches into robust structure-based drug design (SBDD) campaigns that have been key to identifying novel chemical diversity to be explored in Chagas disease drug discovery.

**MATERIALS AND METHODS**

**Molecular Docking**

The three-dimensional structures of the designed cruzain inhibitors were constructed using the standard geometric parameters of SYBYL-X 2.1 (Certaara, Princeton, NJ). Each compound was energetically minimized using the Tripos force field (Clark et al., 1989) and Powell conjugate gradient algorithm (Powell, 1977) with a convergence criterion of 0.05 kcal/molÅ and Gasteiger-Hückel charges (Gasteiger and Marsili, 1980). The designed imide derivatives were docked into the cruzain catalytic site using GOLD 5.3 (Cambridge Crystallographic Data Centre, Cambridge, UK) (Jones et al., 1997). The X-ray structure of cruzain (PDB 3KKU, 1.28 Å) (Ferreira et al., 2010) was prepared by removing the water molecules and adding hydrogen atoms. The active site residues Cys25 and His162 were maintained as negatively charged and protonated, respectively. A sphere with a 10 Å radius centered on the sulfur atom of Cys25 was settled as the binding site. Compounds were docked by applying the GoldScore scoring function with a search efficiency of 200%. Visual analysis of the molecular docking-derived binding conformations was carried out with PyMOL 1.3 (Schrödinger, New York, NY) (Lill and Danielson, 2011).

**Pro-Cruzain Expression, Activation, and Purification**

Cruzain was expressed and purified using a modified version of a previously published protocol (Ferreira et al., 2014). *Escherichia coli* (M15) was preinoculated in Luria Bertani (LB) medium with ampicillin (100 µg/mL) and kanamycin (50 µg/mL) and kept overnight (37°C, 200 rpm). The preinoculum was diluted 10-fold in fresh LB medium (1 L) supplemented with 0.5 M NaCl, 0.2% glucose, 1 mM betaine, 0.5 M sorbitol, 100 µg/mL ampicillin and 50 µg/mL kanamycin, and incubated (37°C, 200 rpm). When the optical density (OD<sub>600</sub>) reached 0.9, the culture was incubated at 47°C for 20 min to induce the expression of chaperones. Next, cruzain expression was induced by adding 0.2 mM isopropyl β-D-thiogalactopyranoside (IPTG), and the culture was kept overnight at 20°C. The cells were harvested by centrifugation (5,000 rpm, 30 min, 4°C) and suspended in 50 mL of lysis buffer (300 mM NaCl, 50 mM Tris-HCl, 1.6 mg/mL lysozyme, pH 8.0). The cells were then lysed by sonication (12 cycles of 30 s), and the suspension was centrifuged (9,000 rpm, 30 min, 4°C). The supernatant was collected, cruzain was precipitated by the addition of 35% ammonium sulfate (2 h), and this suspension was centrifuged (9,000 rpm, 30 min, 4°C). The precipitate was resuspended in lysis buffer and dialyzed to remove the ammonium sulfate. The soluble fraction of cruzain was purified by metal affinity chromatography using a Ni-NTA resin (Qiagen, Hilden, Germany). Contaminant proteins were removed using a washing buffer (300 mM NaCl, 50 mM Tris-HCl, 10 mM imidazole, pH 8.0). Next, cruzain was eluted with an increasing gradient of imidazole (25, 50, 75, 100, and 250 mM). The fractions containing cruzain were dialyzed in 0.1 M acetate buffer (1.5 L, pH 5.5) and concentrated to 0.5 mg/mL for subsequent activation.
Pro-cruzain was incubated in activation buffer (100 mM sodium acetate, 10 mM EDTA, 5 mM DTT, and 1 M NaCl, pH 5.5) at 37°C (Mott et al., 2010). The activation process was monitored at 30 min intervals by the cleavage of the substrate Z-Phe-Arg-AMC and confirmed by SDS-PAGE. Next, the mature enzyme was diluted 20-fold in binding buffer (20 mM sodium phosphate, 150 mM NaCl, pH 7.2) and incubated overnight with thiopropyl Sepharose 6B resin (GE Healthcare Life Sciences, Pittsburgh, PA) at 4°C. Cruzain was eluted using binding buffer supplemented with 20 mM DTT. Fractions containing the enzyme were stored at −80°C in 0.1 M sodium acetate (pH 5.5).

**Cruzain Inhibition Assays**

The catalytic activity of cruzain was monitored by cleavage of the fluorogenic substrate Z-Phe-Arg-aminomethyl coumarin (Z-Phe-Arg-AMC) as previously described (Ferreira et al., 2014). The enzyme kinetics assays were performed in 0.1 M sodium acetate buffer (pH 5.5) with 5 mM dithiothreitol (DTT) and 0.01% Triton X-100. The final concentrations of cruzain and substrate were 20 μM and 50 μM, respectively, except in the tests for the mechanism of inhibition, in which different substrate concentrations were used. The enzyme reaction was monitored for 5 min at 30°C in 96-well black flat bottom plates, and the activity was calculated based on the initial rates compared with a control (DMSO). Wavelengths of 355 nm for excitation and 460 nm for emission were used. IC$_{50}$ values were independently determined by determining rate measurements for at least six inhibitor concentrations, each evaluated in triplicate. To determine the mechanism of inhibition, eight substrate concentrations and four inhibitor concentrations were used, each in triplicate. The mechanism of inhibition was determined by Lineweaver-Burk plots. SigmaPlot 10.0 (Systat Software Inc., Erkrath, Germany) was used to determine the IC$_{50}$ values.

**T. cruzi Intracellular Amastigote Assays**

*In vitro* assays against *T. cruzi* were performed as described previously (Ferreira et al., 2014). The *T. cruzi* Tulahuen strain, which expresses the *E. coli* β-galactosidase gene lacZ (Buckner et al., 1996), was provided by Frederick S. Buckner (University of Washington, Seattle, WA). Stock solutions of the synthesized compounds were prepared in 100% DMSO. Epimastigotes were grown in liver infusion tryptone (LIT) supplemented with 10% fetal calf serum (FCS), penicillin and streptomycin (28°C). Metacyclogenesis from epimastigotes to trypomastigotes was induced by incubation of the epimastigotes in Grace’s insect medium (Sigma-Aldrich, St. Louis, MO) supplemented with 10% FCS (28°C). HFF-1 human fibroblasts were seeded at 2 × 10$^4$ cells/well (80 μL) in 96-well tissue culture plates in RPMI 1640 without phenol red supplemented with 10% FCS and incubated overnight (37°C, 5% CO$_2$). Next, trypomastigotes were added at 1.0 × 10$^4$/well (20 μL), and the plates were incubated (37°C, 5% CO$_2$). After 24 h, 3-fold serial dilutions (50 μL) of the synthesized compounds were added at concentrations ranging from 0.1 to 100 μM, and the plates were incubated (37°C, 5% CO$_2$). Each compound concentration was evaluated in triplicate. All plates included BZ (Sigma-Aldrich) as a positive control and untreated wells (100% growth) as a negative control. After 120 h, 50 μL of chlorophenol red β-D-galactopyranoside (CPRG, Sigma-Aldrich) and IGEPAL CA-630 (Sigma-Aldrich) (0.1%) was added to each well. The absorbance was measured at 570 nm in an automated microplate reader, and the data were transferred to SigmaPlot 10.0 (Systat Software Inc.) for IC$_{50}$ value calculation.

**Cytotoxicity in HFF-1 and HepG2 Cell Lines**

The cytotoxicities of the compounds against HFF-1 fibroblasts and HepG2 hepatocytes were evaluated using the MTS tetrazolium assay (Promega, Madison, WI) (Barltrop et al., 1991). HFF-1 fibroblasts were seeded at 2 × 10$^3$/well in 96-well culture plates in RPMI 1640 without phenol red supplemented with 10% FCS and incubated overnight (37°C, 5% CO$_2$). HepG2 hepatocytes were seeded at 6 × 10$^3$/well in 96-well culture plates in DMEM (CultiLab, Campinas SP) supplemented with 10% FCS and incubated overnight (37°C, 5% CO$_2$). The compounds were added in 3-fold serial dilutions, and the plates were incubated at 37°C with 5% CO$_2$. Each compound concentration was evaluated in triplicate. All plates included doxorubicin (Sigma-Aldrich, St. Louis, MO) as a positive control and untreated wells (100% growth) as a negative control. After 72 h, 20 μL of MTS was added, and the plates were incubated for 4 h. The absorbance was measured at 490 nm in an automated microplate reader, and the data were transferred to SigmaPlot 10.0 (Systat Software Inc., Erkrath, Germany) for IC$_{50}$ value calculation. The percent of nonviable cells was determined and compared to the negative control wells (100% growth).

**Chemistry**

All reactions were performed under an argon atmosphere with dry solvents and magnetic stirring unless stated otherwise. Dichloromethane (DCM) and triethylamine (Et$_3$N) were distilled from CaH$_2$. Tetrahydrofuran (THF) was distilled from sodium/benzophenone. Dimethyl formamide (DMF) was purchased from Sigma-Aldrich (anhydrous) and used without further purification. Yields refer to homogeneous materials obtained after purification of the reaction products by flash column chromatography using silica gel (200–400 mesh) or recrystallization. Analytical thin-layer chromatography was performed on silica-gel 60 and GF (5–40 μm thickness) plates, and visualization was accomplished using UV light, basic potassium permanganate staining or ninhydrin solution followed by heating. $^1$H and proton-decoupled $^{13}$C NMR spectra were acquired in CDCl$_3$, CD$_2$OD, or d$_6$-DMSO at 250 MHz ($^1$H) and 62.5 MHz ($^{13}$C) (Bruker DPX250), at 400 MHz ($^1$H) and 100 MHz ($^{13}$C) (Bruker Avance 400), at 500 MHz ($^1$H) and 125 MHz ($^{13}$C) (Bruker Avance 500), or at 600 MHz ($^1$H) and 150 MHz ($^{13}$C) (Bruker Avance 600). Chemical shifts (δ) are reported in ppm using residual undeuterated solvent as an internal standard (CDCl$_3$ at 7.26 ppm, CD$_2$OD at 3.31 ppm, d$_6$-DMSO at 2.50 ppm, and TMS at 0.00 ppm for $^1$H NMR spectra and CDCl$_3$ at 77.16 ppm, CD$_2$OD at 49.0 ppm, d$_6$-DMSO at 39.52 ppm for $^{13}$C NMR spectra). Multiplicity data are reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, br s = broad singlet, dd = doublet of doublets, dt = doublet of triplets, dddd = doublet of doublet of doublets and m = multiplet. The
multiplicity is followed by the coupling constant(s) in Hz and integration. High-resolution mass spectrometry (HRMS) was measured using electrospray ionization (ESI) (Waters Xevo Q-TOF, Thermo LTQ-FT Ultra, or Thermo Q-Exactive) or electron ionization (EI) (GCT Premier, Waters).

**General Procedures for the Preparation of Alcohols 2a-m**

Method A: The aniline derivative containing the appropriate substituents was dissolved in dry dichloromethane (DCM) under an argon atmosphere with magnetic stirring. The solution was cooled to 0°C followed by the slow addition of acetoxyacetyl chloride (1.1 equivalents). The mixture was stirred for 5 min before the addition of zinc mesh 20 (0.1 equivalents) and removal of the ice bath, allowing the reaction to reach room temperature. The reaction mixture was monitored by thin layer chromatography (TLC), and after 30 min, total conversion was observed. Half of the initial volume of DCM was added to dilute the reaction mixture, and the zinc was removed by filtration. The supernatant was washed with an aqueous sodium bicarbonate solution. The organic phase was dried over magnesium sulfate (MgSO₄) and concentrated to dryness under vacuum. The product was purified by column chromatography on silica gel using a mixture of ethyl acetate/hexanes (50%) as the eluent. The obtained ester was dissolved in methanol at room temperature followed by the addition of solid potassium carbonate (K₂CO₃). The reaction mixture was stirred for 1 h and then quenched by the addition of ethyl acetate. The mixture was washed with aqueous saturated ammonium chloride (NH₄Cl) solution and brine. The resulting organic phase was dried over MgSO₄ and concentrated under reduced pressure. Purification was performed by column chromatography on silica gel using a mixture of ethyl acetate/hexanes (50%) as the eluent.

Method B: The aniline derivative containing the appropriate substituents was dissolved in dry DCM under an argon atmosphere with magnetic stirring. To the resulting solution, dry triethylamine (1.1 equivalents) and acetoxyacetyl chloride (1.1 equivalents) were added and stirring was maintained for 2 h to monitor the conversion by TLC. The mixture was then washed with saturated aqueous NH₄Cl solution and brine solution. The organic phase was dried over MgSO₄ and concentrated under reduced pressure. Purification was performed by column chromatography on silica gel using a mixture of ethyl acetate/hexanes (50%) as the eluent.

**General Procedures for Production of Carboxylic Acids 1 and 9a-n**

Method C: The appropriate cyclic anhydride and amino acid (1.1 equivalents) were dissolved in glacial acetic acid under an argon atmosphere. The mixture was magnetically stirred overnight followed by 4 h of reflux periodically monitored by TLC. After consumption of the cyclic anhydride, the reaction pot was cooled to room temperature. Acetic acid was removed in a rotary evaporator, and ice-cold water was added to the resulting slurry to generate a precipitate. Concentrated hydrochloric acid (HCl) was added, and the solid was collected by filtration, followed by high vacuum drying.

Method D: The appropriate cyclic anhydride and amino acid (1.1 equivalents) were added to a reaction flask containing toluene. To this suspension, TEA (0.1 equivalents) was added and a Dean–Stark apparatus was coupled to the system. The mixture refluxed for 4 h and cooled to room temperature. Toluene was removed in a rotary evaporator, and the residue was dissolved in ethyl acetate and washed twice with 1 M HCl. The organic phase was extracted with saturated aqueous sodium bicarbonate solution. The aqueous phase was acidified with HCl solution to pH 5.0 and extracted with ethyl acetate. The combined organic phases were dried over MgSO₄, concentrated under vacuum and used in the following step without further purification.

Method E: Leucine was dissolved in a 1:1 mixture of hydrobromic acid and water in an ice bath. Sodium nitrite (6.5 M aqueous solution) was added dropwise. The ice bath was removed, and the mixture stirred at room temperature for 2.5 h. The acidic residuals were removed under reduced pressure, and the remaining solution was extracted with diethyl ether 3 times. The combined organic phases were dried over MgSO₄ and concentrated under vacuum.

The residue was dissolved in DCM under an argon atmosphere. The solution was cooled in an ice bath and benzyl alcohol (1.1 equivalents), 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) (1.1 equivalents) and 4-dimethylaminopyridine (DMAP) (0.1 equivalents) were added. The ice bath was removed, and the mixture stirred for 5 h. Then, the reaction mixture was washed with saturated aqueous NH₄Cl solution and brine. The organic phase was dried over MgSO₄ and concentrated under vacuum. The product was purified by column chromatography on silica gel using a mixture of ethyl acetate/hexanes (50%) as the eluent.

A solution of the pure ester was prepared in acetonitrile under an argon atmosphere, and the appropriate cyclic amine was added (2.0 equivalents). To the resulting solution, cesium carbonate (1.1 equivalents) was added followed by stirring for 2 h. The solids were removed by filtration, and the solvent was removed under reduced pressure followed by redissolution in DCM and washing with water. The organic phase was dried over MgSO₄ and concentrated to dryness under vacuum. The product was purified by column chromatography on silica gel using a gradient from 0 to 4% methanol in DCM as the eluent.

The ester was dissolved in methanol, and 10% Pd/C was added (0.1 equivalents). The mixture stirred at room temperature under a hydrogen atmosphere (1 atm) for 30 min. The catalyst was removed by filtration over Celite, and the solvent removed under reduced pressure to generate the pure carboxylic acid.

**General Procedure for the Preparation of Compounds 3a-m and 10a-n**

Method F: A solution of carboxylic acid in dry DCM was prepared, and the alcohol was added (1.1 equivalents). The solution was cooled in an ice bath followed by the addition of EDC (1.2 equivalents) and DMAP (0.1 equivalents). The
solution was allowed to warm to room temperature and stirred for 4 h while monitoring periodically by TLC. The reaction mixture was washed with saturated aqueous NH₄Cl solution and brine. The organic phase was dried over MgSO₄ and concentrated under reduced pressure. The product was purified by column chromatography on silica gel using a mixture of ethyl acetate/hexanes (50%) as the eluent.

2-((3-chloro-4-methoxyphenyl)amino)-2-oxoethyl (2S)-2-(1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-isindol-2-yl)-4-methylpentanoate (3a)

Methods A and C followed by F

Very viscous liquid, 90% ¹H NMR (500 MHz, CDCl₃) δ 8.26 (s, 1H); 7.83 (d, J = 2.5 Hz, 1H); 7.68 (dd, J = 9.0, 2.5 Hz, 1H); 6.90 (d, J = 9.0 Hz, 1H); 5.90 (m, 2H); 5.04 (d, J = 15.6 Hz, 1H); 4.88 (dd, J = 11.0, 4.4 Hz, 1H); 4.40 (d, J = 15.6 Hz, 1H); 3.89 (s, 3H); 3.23 (m, 2H); 2.65 (m, 2H); 2.27 (m, 2H); 2.08 (m, 1H); 1.94 (m, 1H); 1.41 (m, 1H); 0.93 (d, J = 6.6 Hz, 6H). ¹³C NMR (125 MHz, CDCl₃) δ 180.4; 180.3; 168.1; 164.7; 151.9; 131.1; 127.8; 127.7; 122.3; 122.2; 119.4; 112.1; 63.4; 56.3; 51.0; 39.3; 39.2; 37.2; 24.7; 23.5; 23.4; 20.9; 20.9.

HRMS (ESI-Orbitrap): [M+H]⁺ 463.1617; found 463.1636.

2-((4-hydroxyphenyl)amino)-2-oxoethyl (2S)-2-(1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-isindol-2-yl)-4-methylpentanoate (3b)

Methods B and C followed by F

Very viscous liquid, 88% ¹H NMR (500 MHz, DMSO) δ 9.78 (s, 1H); 7.46 (d, J = 9.0 Hz, 1H); 6.89 (dd, J = 9.0 Hz, 1H); 5.85 (m, 1H); 4.79 (dd, J = 11.1, 4.4 Hz, 1H); 4.63 (m, 2H); 3.72 (s, 3H); 3.24 (m, 2H); 2.39 (m, 2H); 1.99 (m, 1H); 1.77 (m, 1H); 1.34 (m, 1H); 0.84 (m, 6H). ¹³C NMR (62.5 MHz, DMSO) δ 179.6; 179.5; 164.3; 155.5; 131.3; 127.4; 120.4; 113.9; 63.4; 55.2; 50.2; 36.0; 24.1; 23.2; 23.0; 20.7.

HRMS (ESI-Orbitrap): [M+H]⁺ 429.1826; found 429.1828.

2-((3-chloro-4-methoxyphenyl)amino)-2-oxoethyl (2S)-2-(1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-isindol-2-yl)-4-methylpentanoate (3c)

Methods B and C followed by F

Very viscous liquid, 91% ¹H NMR (250 MHz, DMSO) δ 10.15 (m, 1H); 7.77 (m, 1H); 7.44 (m, 1H); 7.33 (m, 1H); 7.11 (m, 1H); 5.84 (m, 2H); 5.03 (m, 1H); 4.67 (m, 2H); 3.59 (m, 2H); 3.23 (m, 1H); 2.43 (m, 3H); 2.19 (m, 2H); 1.97 (m, 1H); 1.75 (m, 1H); 1.34 (m, 1H); 0.83 (m, 6H). ¹³C NMR (62.5 MHz, DMSO) δ 179.6; 179.5; 179.4; 173.3; 173.2; 172.4; 169.2; 168.6; 165.7; 165.2; 139.8; 139.7; 133.2; 130.5; 127.5; 127.4; 120.5; 124.9; 123.4; 123.3; 118.8; 117.7; 63.4; 62.7; 52.3; 51.6; 51.5; 50.3; 50.2; 38.4; 36.1; 36.0; 25.3; 24.2; 24.1; 23.2; 23.1; 20.3; 20.7.

HRMS (ESI-Orbitrap): [M+Na]⁺ 445.1362.
Very viscous liquid, 84% $^1\text{H}$ NMR (250 MHz, DMSO) $\delta$ 10.05 (s, 1H); 7.57 (d, $J = 8.8$ Hz, 1H); 6.98 (d, $J = 8.8$ Hz, 1H); 5.80 (m, 2H); 4.88 (m, 1H); 4.66 (m, 1H); 3.24 (m, 1H); 2.40 (m, 2H); 2.19 (m, 2H); 1.98 (m, 1H); 1.77 (m, 1H); 1.35 (m, 1H); 0.85 (m, 6H).

2-(naphthalen-2-ylamino)-2-oxoethyl
(2S)-2-(1,3-dioxo-1,3a,4,7a-hexahydro-2H-isooindol-2-yl)-4-methylpentanoate (3i)

Methods B and C followed by F

Very viscous liquid, 68% $^1\text{H}$ NMR (250 MHz, DMSO) $\delta$ 8.43 (s, 2H); 7.80 (m, 4H); 7.45 (m, 2H); 5.91 (m, 2H); 5.13 (d, $J = 15.5$ Hz, 1H); 4.92 (dd, $J = 10.8$, 4.5 Hz, 1H); 4.47 (d, $J = 15.5$, 1H); 3.23 (m, 2H); 2.68 (m, 2H); 2.22 (m, 3H); 1.97 (m, 1H); 0.95 (d, $J = 6.5$, 6H).

HRMs (ESI-Orbitrap): [M+Na]$^+$ Calculated for C$_{26}$H$_{28}$O$_5$N$_3$Na 471.18959; found 471.18888.

2-((4-fluoro-3-nitrophenyl)amino)-2-oxoethyl
(2S)-2-(1,3-dioxo-1,3a,4,7a-hexahydro-2H-isooindol-2-yl)-4-methylpentanoate (3j)

Methods A and C followed by F

Very viscous liquid, 98% $^1\text{H}$ NMR (500 MHz, CDCl$_3$) $\delta$ 10.51 (s, 1H); 8.57 (d, $J = 9.15$, 1H); 7.66 (d, $J = 2.9$, 1H); 7.24 (dd, $J = 9.3$, 3.0 Hz, 1H); 5.92 (m, 2H); 4.95 (d, $J = 11.3$, 4.4 Hz, 1H); 4.77 (m, 2H); 3.87 (s, 3H); 3.31 (m, 1H); 3.18 (m, 1H); 2.62 (m, 2H); 2.26 (m, 2H); 2.16 (m, 1H); 1.97 (m, 1H); 1.39 (m, 1H); 0.93 (d, $J = 6.6$, 4.5 Hz, 6H); $^{13}$C NMR (125 MHz, CDCl$_3$) $\delta$ 179.5; 179.43; 168.0; 165.0; 155.5; 137.7; 127.7; 126.6; 126.4; 124.0; 122.8; 108.8; 64.0; 55.8; 50.9; 39.2; 38.9; 36.7; 24.6; 23.6; 23.4; 23.0; 20.77.

HRMs (ESI-Orbitrap): [M+Na]$^+$ Calculated for C$_{24}$H$_{25}$O$_{10}$N$_3$Na 474.19045; found 474.19131.

2-((3-chloro-4-methoxyphenyl)amino)-2-oxoethyl
(2S)-2-(1,3-dioxo-1,3a,4,7a-hexahydro-2H-isooindol-2-yl)-4-methylpentanoate (3k)

Methods A and C followed by F

Very viscous liquid, 92% $^1\text{H}$ NMR (250 MHz, DMSO) $\delta$ 10.04 (s, 1H); 7.92 (m, 4H); 7.69 (d, $J = 2.4$, 1H); 7.37 (dd, $J = 8.9$, 2.4 Hz, 1H); 7.10 (d, $J = 9.0$, 1H); 5.04 (dd, $J = 11.2$, 4.4 Hz, 1H); 4.72 (m, 2H); 3.81 (s, 3H); 2.24 (m, 1H); 1.93 (dd, $J = 14.0$, 9.9, 4.5 Hz, 1H); 1.49 (m, 6H); $^{13}$C NMR (62.5 MHz, DMSO) $\delta$ 169.1; 167.2; 164.6; 150.7; 135.0; 131.9; 131.0; 123.5; 121.0; 120.6; 119.2; 112.9; 63.4; 56.1; 49.9; 36.6; 24.5; 22.9; 20.8.

HRMs (ESI-Orbitrap): [M+H]$^+$ Calculated for C$_{28}$H$_{24}$O$_5$NCl 459.1323; found 459.1355.

2-((3-chloro-4-methoxyphenyl)amino)-2-oxoethyl
(2S)-2-(1,3-dioxooctahydro-2H-isooindol-2-yl)-4-methylpentanoate (10a)

Methods A and C followed by F

Very viscous liquid, 91% $^1\text{H}$ NMR (500 MHz, CDCl$_3$) $\delta$ 8.32 (s, 1H); 7.83 (d, $J = 2.5$, 1H); 7.68 (d, $J = 8.9$, 2.6 Hz, 1H); 6.90 (d, $J = 8.9$, 1H); 5.08 (d, $J = 15.6$, 1H); 4.90 (dd, $J = 10.9$, 4.3 Hz, 1H); 4.45 (d, $J = 15.6$, 1H); 3.89 (s, 3H); 2.12 (dd, $J = 14.5$, 10.7, 4.3 Hz, 1H); 1.98 (m, 2H); 1.85 (m, 2H); 1.77 (m, 1H); 1.47 (m, 1H); 1.36 (m, 1H); 0.96 (m, 6H); $^{13}$C NMR (100 MHz, CDCl$_3$) $\delta$ 180.3; 179.4; 168.4; 164.7; 152.0; 131.1; 122.3; 119.5; 112.1; 63.5; 56.3; 50.5; 40.0; 39.8; 37.4; 25.0; 24.0; 23.3; 21.6; 21.1.

HRMs (ESI-Orbitrap): [M+Na]$^+$ Calculated for C$_{23}$H$_{20}$O$_5$NaClNa 487.19118; found 487.16025.

2-((3-chloro-4-methoxyphenyl)amino)-2-oxoethyl
(2S)-2-(2,5-dioxo-2,5-dihydro-1H-pyrroli-1-yl)-4-methylpentanoate (10c)

Methods A and D followed by F
Very viscous liquid, 46% $^1$H NMR (400 MHz, CDCl$_3$) δ 8.12 (sl, 1H); 7.76 (d, $J = 2.6$ Hz, 1H); 7.62 (dd, $J = 8.9$, 2.6 Hz, 1H); 6.90 (d, $J = 8.9$ Hz, 1H); 6.84 (s, 2H); 5.04 (d, $J = 15.5$ Hz, 1H); 4.91 (dd, $J = 11.4$, 4.4 Hz, 1H); 4.47 (d, $J = 15.5$ Hz, 1H); 3.89 (s, 3H); 2.18 (ddd, $J = 14.3$, 11.2, 4.3 Hz, 1H); 1.98 (ddd, $J = 14.3$, 10.0, 4.3 Hz, 1H); 1.47 (m, 1H); 0.96 (dd, $J = 6.6$, 1.2 Hz, 6H).

$^{13}$C NMR (100 MHz, CDCl$_3$) δ 170.4; 168.4; 152.1; 134.6; 130.9; 122.5; 122.4; 63.6; 56.3; 50.4; 38.0; 24.9; 23.0; 21.0.

2-((3-chloro-4-methoxyphenyl)amino)-2-oxoethyl (S)-2-(2,5-dioxopyrrolidin-1-yl)-4-methylpentanoate (10d)

Methods A and D followed by F

Very viscous liquid, 72% $^1$H NMR (400 MHz, CDCl$_3$) δ 8.30 (sl, 1H); 7.79 (d, $J = 2.6$ Hz, 1H); 7.62 (dd, $J = 8.9$, 2.7 Hz, 1H); 6.89 (d, $J = 8.9$ Hz, 1H); 5.01 (d, $J = 15.4$ Hz, 1H); 4.94 (dd, $J = 10.5$, 4.8 Hz, 1H); 4.49 (d, $J = 15.5$ Hz, 1H); 3.88 (s, 3H); 2.86 (m, 4H); 2.05 (m, 2H); 1.45 (m, 1H); 0.96 (dd, $J = 6.7$, 4.6, 6.9 Hz, 1H).

$^{13}$C NMR (100 MHz, CDCl$_3$) δ 177.2; 168.2; 164.7; 152.0; 131.0; 122.3; 119.6; 121.2; 63.4; 56.3; 50.9; 37.3; 28.2; 24.9; 22.9; 21.1.

HRMS (ESI-Orbitrap): [M+Na]$^+$ Calculated for C$_{19}$H$_{23}$O$_5$N$_2$ClNa 433.11423; found 433.11424.

2-(((3-chloro-4-methoxyphenyl)amino)-2-oxoethyl (S)-4-methyl-2-(3-methyl-2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)pentan-10(1H)-one (10e)

Methods A and D followed by F

Very viscous liquid, 81% $^1$H NMR (500 MHz, CDCl$_3$) δ 8.29 (sl, 1H); 7.84 (d, $J = 2.2$ Hz, 1H); 7.69 (dd, $J = 8.8$, 2.4 Hz, 1H); 7.24 (m, 3H); 7.13 (d, $J = 7.1$ Hz, 2H); 6.92 (d, $J = 9.0$ Hz, 1H); 5.70 (m, 2H); 5.12 (m, 2H); 4.43 (d, $J = 15.6$ Hz, 1H); 3.90 (s, 3H); 3.56 (dd, $J = 14.3$, 11.5 Hz, 1H); 3.04 (m, 1H); 2.93 (m, 1H); 2.48 (ddd, $J = 15.7$, 6.05, 2.6 Hz, 1H); 2.29 (ddd, $J = 15.8$, 5.9, 3.0 Hz, 1H); 2.17 (m, 2H).

$^{13}$C NMR (125 MHz, CDCl$_3$) δ 180.2; 179.8; 167.3; 164.6; 152.0; 135.3; 131.0; 129.0; 128.6; 127.4; 121.7; 121.7; 122.3; 122.8; 119.5; 112.1; 63.4; 56.3; 53.2; 39.0; 38.9; 34.3; 23.11; 23.07.

HRMS (ESI-Orbitrap): [M+Na]$^+$ Calculated for C$_{26}$H$_{23}$O$_5$N$_2$ClNa 519.12988; found 519.12918.

$[\alpha]^D_{20} = -24.0$ (1, MeOH).

HPLC flow 3.5 mL/min, polarity hexane:ethanol acetate (60:40).

Rt 16.76 min, area 97.10%; Rt 20.09 min, area 2.90%.

2-((3-chloro-4-methoxyphenyl)amino)-2-oxoethyl (2S)-2-((1,3-dioxo-1,3,3a,4,7,7a-hexahydro-2H-isindol-2-yl)-3-phenylpropanoatoe (10h)

Methods A and D followed by F

Very viscous liquid, 87% $^1$H NMR (250 MHz, CDCl$_3$) δ 8.26 (sl, 1H); 7.83 (d, $J = 2.7$ Hz, 1H); 7.71 (dd, $J = 8.9$, 2.6 Hz, 1H); 7.22 (m, 5H); 6.92 (d, $J = 8.8$ Hz, 1H); 5.70 (m, 2H); 5.12 (m, 2H); 4.42 (d, $J = 15.5$ Hz, 1H); 3.90 (s, 3H); 3.56 (m, 1H); 3.36 (dd, $J = 14.3$, 11.3 Hz, 1H); 2.99 (m, 2H); 2.48 (ddd, $J = 15.8$, 6.0, 2.9 Hz, 1H); 2.22 (m, 3H).

HRMS (ESI-Orbitrap): [M+N$^+$]+ Calculated for C$_{26}$H$_{23}$O$_5$N$_2$ClNa 519.12988; found 519.12965.

$[\alpha]^D_{20} = 0$ (1, MeOH).

HPLC flow 3.5 mL/min, polarity hexane:ethanol acetate (60:40).

Rt 17.16 min, area 50.19%; Rt 20.49 min, area 49.81%.
132.8; 132.5; 131.1; 128.5; 128.1; 127.7; 127.4; 127.2; 127.0; 126.6; 126.3; 125.9; 122.5; 122.4; 119.6; 112.2; 63.5; 56.4; 53.2; 39.0; 38.9; 34.6; 23.1.

HRMS (ESI-Orbitrap): [M+Na]+ Calculated for C30H27O4N2ClNa 569.14553; found 569.14428.

2-((3-chloro-4-methoxyphenyl)amino)-2-oxoethyl (S)-2-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)-3-phenylpropanoate (10j)

Methods A and D followed by F

Very viscous liquid, 23% 1H NMR (500 MHz, CDCl3) δ 8.17 (sl, 1H); 7.57 (d, J = 2.7 Hz, 1H); 7.42 (dd, J = 8.9, 2.5 Hz, 1H); 6.90 (d, J = 8.8 Hz, 1H); 4.83 (d, J = 15.6 Hz, 1H); 4.63 (d, J = 15.6 Hz, 1H); 3.82 (s, 3H); 3.37 (dd, J = 8.8, 5.7 Hz, 1H); 2.59 (m, 4H); 1.60 (m, 7H); 1.44 (m, 2H); 0.94 (dd, J = 9.1, 6.3 Hz, 6H), 13C NMR (125 MHz, CDCl3) δ 170.3; 165.3; 152.4; 135.5; 134.3; 130.9; 114.1; 74.4; 68.3; 66.9; 64.5; 64.4; 57.2; 47.8; 44.5; 44.0; 38.2; 36.4; 26.1; 24.8; 23.5; 23.8; 23.0; 20.9.

HRMS (ESI-Orbitrap): [M+Na]+ Calculated for C23H21O6N3NaFNa 450.19994; found 450.19913.

2-((3-chloro-4-methoxyphenyl)amino)-2-oxoethyl (S)-4-methyl-2-(pyrrolidin-1-yl)pentanoate hydrochloride (10o)

Amorphous solid, 96% 1H NMR (500 MHz, CDCl3) δ 7.67 (m, 1H); 7.38 (m, 1H); 7.02 (m, 1H); 4.83 (m, 2H); 3.83 (m, 2H); 2.67 (d, J = 7.8 Hz, 1H); 2.67 (m, 4H); 1.65 (m, 3H); 0.96 (dd, J = 8.9, 6.1 Hz, 6H), 13C NMR (125 MHz, CDCl3) δ 172.3; 170.5; 167.3; 167.2; 156.2; 153.9; 132.7; 132.9; 123.4; 121.4; 114.1; 74.4; 68.3; 66.9; 64.5; 64.4; 57.2; 47.8; 44.5; 44.0; 38.2; 26.4; 26.1; 23.4; 23.1; 22.4; 22.2.

HRMS (ESI-Orbitrap): [M+H]+ Calculated for C19H23O8N2Cl 398.18346; found 398.18384.

General Procedure for the Preparation of Compounds 3k' and 3m'

The esters, obtained through methods A and D or E followed by F, were dissolved in dry DCM, and a large excess of HCl (4 M solution in dioxane) was added. The resulting solution was stirred at room temperature for 4 h. The solvent mixture was removed under reduced pressure, and the residue was dissolved in DCM followed by crystallization from hexane.

2-((4-(methylamino)phenyl)amino)-2-oxoethyl (1,3-dioxo-1,3,3a,4,7a-hexahydro-2H-isindol-2-yl)-4-methylpentanoate (3k')

Very viscous liquid, 75% 1H NMR (500 MHz, CDCl3) δ 8.04 (sl, 1H); 7.53 (d, J = 8.8 Hz, 2H); 6.60 (d, J = 8.8 Hz, 2H); 5.89 (m, 2H); 5.02 (d, J = 15.3 Hz, 1H); 4.87 (dd, J = 11.1, 4.3 Hz, 1H); 4.40 (d, J = 15.4 Hz, 1H); 3.72 (sl, 1H); 3.20 (m, 1H); 2.83 (s, 3H); 2.64 (m, 2H); 2.25 (m, 2H); 2.11 (m, 1H); 1.94 (m, 1H); 1.40 (m, 1H); 0.92 (d, J = 6.6 Hz, 6H), 13C NMR (125 MHz, CDCl3) δ 180.3; 180.1; 168.0; 164.3; 146.6; 127.7; 121.7; 112.5; 63.5; 51.1; 39.2; 37.1; 31.0; 24.8; 23.5; 23.8; 23.0; 20.9.

HRMS (ESI-Orbitrap): [M+Na]+ Calculated for C23H21O7N3NaFNa 450.19994; found 450.19913.

2-((3-chloro-4-methoxyphenyl)amino)-2-oxoethyl (S)-4-methyl-2-(piperidin-1-yl)pentanoate (10l)

Methods A and E followed by F

Very viscous liquid, 23% 1H NMR (500 MHz, CDCl3) δ 7.96 (sl, 1H); 7.56 (d, J = 2.6 Hz, 1H); 7.38 (dd, J = 8.8, 2.6 Hz, 1H); 6.89 (d, J = 8.9 Hz, 1H); 4.76 (d, J = 15.3 Hz, 1H); 4.68 (d, J = 15.3 Hz, 1H); 3.88 (s, 3H); 3.68 (m, 4H); 3.40 (t, J = 7.3 Hz, 1H); 2.67 (m, 4H); 1.65 (m, 3H); 0.96 (dd, J = 8.9, 6.1 Hz, 6H), 13C NMR (125 MHz, CDCl3) δ 170.8; 165.0; 152.4; 130.1; 122.8; 122.6; 120.0; 112.2; 67.2; 65.8; 62.6; 56.3; 49.8; 36.8; 25.0; 22.6; 22.4.

HRMS (ESI-Orbitrap): [M+Na]+ Calculated for C19H23O8N2ClNa 419.17081; found 419.17050.
in vacuum. The product was purified by column chromatography on silica gel using a mixture of ethyl acetate/hexanes (50%) as the eluent.

**2-((4-aminophenyl)amino)-2-oxoethyl 2-(1,3-dioxo-1,3a,4,7,7a-hexahydro-2H-isindol-2-yl)-4-methylpentanoate (3l)**

Pale amorphous solid, 92% \(^1\)H NMR (500 MHz, CDCl\(_3\)) \(\delta\) 8.06 (sl, 1H); 7.51 (d, \(J = 8.7 \text{ Hz}, 2H\)); 6.67 (d, \(J = 8.7 \text{ Hz}, 2H\)); 5.89 (m, 2H); 5.03 (d, \(J = 15.4 \text{ Hz}, 1H\)) 4.87 (dd, \(J = 11.0, 4.3 \text{ Hz}, 1H\)); 4.40 (d, \(J = 15.4 \text{ Hz}, 1H\)); 3.65 (sl, 2H); 3.20 (m, 2H); 2.64 (m, 2H); 2.26 (m, 2H); 2.10 (m, 1H); 1.94 (m, 1H); 1.40 (m, 1H); 0.92 (d, \(J = 6.6 \text{ Hz}, 6H\)). \(^{13}\)C NMR (125 MHz, CDCl\(_3\)) \(\delta\) 180.09; 179.8; 166.8; 150.8; 132.1; 127.6; 122.0; 120.4; 120.2; 112.7; 56.2; 52.4; 36.2; 24.3; 23.4; 23.3; 23.1; 20.6.

**Procedure for the Preparation of 2-((chloro-4-methoxyphenyl)amino)ethyl 2-(1,3-dioxo-1,3a,4,7,7a-hexahydro-2H-isindol-2-yl)-4-methylpentanoate (7)**

A solution of compound 1 in dry DCM was prepared, and alcohol \(6\) was added (1.1 equivalents). The solution was cooled in an ice bath followed by the addition of EDC (1.2 equivalents) and DMAP (0.1 equivalents). The solution was allowed to warm to room temperature and stirred for 4 h while monitoring periodically by TLC. The reaction mixture was washed with saturated aqueous NH\(_4\)Cl solution and brine. The organic phase was dried over MgSO\(_4\) and concentrated under reduced pressure. The product was purified by column chromatography on silica gel using a mixture of ethyl acetate/hexanes (50%) as the eluent.

Pale oil, 62% \(^1\)H NMR (500 MHz, CDCl\(_3\)) \(\delta\) 6.82 (d, \(J = 8.7 \text{ Hz}, 1H\)); 6.70 (d, \(J = 2.8 \text{ Hz}, 1H\)); 6.51 (dd, \(J = 8.8, 2.8 \text{ Hz}, 1H\)); 5.89 (m, 1H); 5.82 (m, 1H); 4.76 (dd, \(J = 11.3, 4.3 \text{ Hz}, 1H\)); 4.43 (ddd, \(J = 11.1, 7.1, 3.9 \text{ Hz}, 1H\)); 4.18 (m, 1H); 3.82 (m, 3H); 3.30 (m, 2H); 3.10 (m, 2H); 2.60 (m, 2H); 2.23 (m, 2H); 2.09 (m, 1H); 1.84 (m, 1H); 1.35 (m, 1H); 0.89 (dd, \(J = 8.8, 6.5 \text{ Hz}, 6H\)). \(^{13}\)C NMR (125 MHz, CDCl\(_3\)) \(\delta\) 179.64; 179.61; 169.1; 147.6; 142.5; 127.7; 127.6; 123.4; 115.2; 112.1; 64.0; 57.0; 51.2; 43.1; 39.0; 38.9; 36.7; 24.8; 23.4; 23.2; 21.0. HRMS (ESI-Orbitrap) m/z [M + Na] calculated for C\(_{25}\)H\(_{32}\)O\(_3\)N\(_2\)ClNa 471.16572; found 471.16522.

**RESULTS AND DISCUSSION**

In this work, we report the discovery of a series of reversible cruzain inhibitors showing promising trypanocidal activity and low toxicity. The imide derivative 3a (Figure 1A), a reversible cruzain inhibitor that was previously identified from a virtual high-throughput screening (HTS) combined approach (Ferreira et al., 2010), was taken as the initial hit for molecular optimization. The SBDD strategy relied on the molecular docking-predicted binding mode of 3a within the catalytic site of cruzain (Figure 1B). The enzyme-inhibitor interactions rely mainly on a hydrogen bonding network. The secondary amide oxygen fills the S3 subsite and interacts with the main chain nitrogen of Gly66. The secondary amide nitrogen in turn forms a hydrogen bond with the side chain nitrogen of Gln19. One of the imide carbonyl oxygens forms...
Compound 3a was used as the initial hit for the design of imide derivatives as novel cruzain inhibitors with anti-T. cruzi activity. Molecular docking-predicted binding conformation of compound 3a within the active site of cruzain. Cruzain (PDB 3KKU, 1.28 Å) is depicted in cartoon and surface representations. Binding site residues (carbon in gray) and compound 3a (carbon in green) are shown as sticks. Hydrogen bonds are shown as dashed lines. Cruzain subsites are labeled as S1, S1’, S2, and S3.

Synthesis of Imide Derivatives

The synthesis of the initial hit (3a) and analogs with modifications to the aromatic moiety is outlined in Scheme 1. The aromatic ring was manipulated by introducing electron withdrawing and donating groups and hydrogen bond donors and acceptors. Carboxylic acid 1 was coupled under Steglich-like conditions (Pande et al., 2014) with alcohols 2a-m furnishing the intended esters 3a-m. Ester 3k' was obtained through deprotection of 3k under acidic conditions, and compounds 3l' and 3m' were generated from the reduction of 3l and 3m, respectively.

Considering the low stability of the ester group in acidic and basic media, we synthesized a more stable analog with an amide group, as illustrated in Scheme 2. Initially, carboxylic acid intermediate 1 was transformed into an acyl chloride via reaction with oxalyl chloride. The acyl chloride was further used in a coupling reaction with glycine, producing intermediate 4. Intermediate 4 was converted into an acyl chloride, using the same conditions previously mentioned, and coupled with 3-chloro-4-methoxyaniline to generate the target compound 5 (Scheme 2A). Compound 7 (Scheme 2B) was synthesized from 3-chloro-4-methoxyaniline with ethyl carbonate catalyzed by molecular sieves (Kinage et al., 2011) to produce 6, followed by coupling with 1 using Steglich conditions.

Further analogs of compound 3a were designed to explore modifications to the imide nucleus and the hydrophobic isobutyl region (R1). The synthesis was performed according to Scheme 3.
to evaluate the effects of changing the stereochemical and electronic features on the biological activity of the target compounds. Intermediates 9a-c and 9g were prepared according to a previously described procedure (Faghihi, 2008). A mixture of readily available L-amino acids and the corresponding cyclic anhydride in acetic acid stirred overnight, followed by 4 h of reflux to obtain carboxylic acids 9a-c and 9g. The synthesis of compounds 9d-f and 9h-j was carried out by refluxing in toluene with triethylamine as the catalyst. The stereocenter of the carboxylic acids was observed to have the absolute configuration of R (Pande et al., 2014). The amino acid side chains gave rise to hydrophobic fragments. Diazotization of L-leucine (Badiola et al., 2014) followed by bromoacid formation and subsequent nucleophilic substitution with the appropriate amines generated the amine nucleus in intermediates 9k-n, presumably with an inverted configuration of the stereogenic center (S) due to inversion during the nucleophilic substitution reaction. Intermediates 9a-n were coupled under Steglich conditions (Neises and Steglich, 1978) to produce the intended esters 10a-n. Compound 10o was produced by deprotection of the N-Boc analog (10n) under acidic conditions.
Identification of Imide Derivatives as Cruzain Inhibitors

The activities of inhibitor 3a (IC\textsubscript{50} = 2.2 \textmu M) and its analogs having modifications to the imide moiety against cruzain are shown in Figure 2. The incorporation of an aromatic ring into the imide ring to form an isoindoline-1,3-dione core in compound 10a reduced the percent inhibition from 88 to 39% at an inhibitor concentration of 100 \textmu M. In fact, compound 10a assumes a completely different binding conformation, with the imide projecting into the S2 subsite. This causes the loss of key interactions with the enzyme as predicted by the molecular docking simulations (Figure S1A). The coupling of a saturated ring to the imide resulted in a 2-fold decrease in the potency of compound 10b (IC\textsubscript{50} = 4.2 \textmu M) over 3a. The addition of unsaturation to the unsubstituted imide yielded compound 10c (IC\textsubscript{50} = 2.3 \textmu M), which was equipotent with respect to inhibitor 3a. In fact, compound 10c preserves the enzyme-inhibitor interactions observed for 3a, however, the imide carbonyl oxygen is predicted to interact with the side chain nitrogen of Trp184 instead of Gln19. The predicted binding modes of analogs 10b and 10c are shown in Figures S1B,C, respectively). Withdrawing the unsaturation of the imide produced compound 10d (IC\textsubscript{50} = 12.0 \textmu M), which was 6-fold less active than 3a.

Following the positive result observed for compound 10c, we designed compounds 10e and 10f, which have the imide nucleus incorporating an unsaturation but with two different substitution patterns. Compound 10e, with a methyl substituent, was equipotent (IC\textsubscript{50} = 2.5 \textmu M) relative to 10c, whereas compound 10f, featuring two chlorine atoms as substituents, resulted in a 6-fold increase in potency (IC\textsubscript{50} = 0.60 \textmu M). Replacing the imide core with an amine (10k, 10l, 10m, 10n, and 10o) led to a significant drop in the activity of the compounds, highlighting the essentiality of the carbonyl groups. In fact, Figure 1B shows a hydrogen bond between one of the carbonyl groups and the side chain nitrogen of Gln19 at the S1 subsite. Replacing the imide group with other ring systems, such as pyrrolidine, piperidine, and morpholine, abrogated the biological activity of this series. The predicted binding conformations of 10k and 10l (pyrrolidine and piperidine derivatives, respectively) show that the lack of the imide carbonyl groups changes the binding pattern of these compounds compared to that observed for the active analogs. As shown in Figure S2, the isobutyl and cyclic amines do not interact with...
the S1’ and S1 subsites as expected, and become exposed to the solvent.

The molecular optimization strategy also involved modifications to the hydrophobic (isobutyl) fragment of compound 3a (Figure 3A). Removing the isobutyl group to obtain a methylene as the linker between the ester and the imide rendered the resulting compound inactive (10g). Bulky hydrophobic groups proved to be essential for interaction with
the S1’ subsite. Replacing the isobutyl with a methylene led to a completely different binding pattern from that observed for the active compounds. The docking algorithm was unable to place the imide into S1 and the aromatic ring into S2, which rendered compound 10g inactive (Figure S3A). Replacement of the isobutyl by the bulkier and planar benzyl group increased the potency by 2-fold (10h, IC$_{50}$ = 1.4 µM) compared to compound 3a. This activity improvement can be reasoned to be a result of the better complementarity of the benzyl group with the S1’ subsite of the enzyme (Figure 3B). Replacing the benzyl with the bulkier methylene naphthyl moiety led to compound 10i, which was significantly less active, suggesting the ideal volume of a benzyl ring to occupy the S1’ subsite. In fact, the naphthyl group of analog 10i projects into the S2 subsite, which can be the driving force for the percent inhibition value (58%) determined for this compound since S2 usually accommodates bulky hydrophobic groups (Figure S3B). However, this modification caused the imide to lose its interaction with S1, which is a key driving force for the activity of this series.

To probe the relevance of the ester and amide groups for the activity of cruzain inhibitors, compounds 5 and 7 were prepared (Figure 4A). Replacing the ester with an amide, resulting in compound 5, caused a reduction in the percent inhibition from 88 to 38%. Converting the amide carbonyl
into a methylene decreased the biological activity even further, as observed for compound 7 (11% inhibition). As illustrated in Figure 1B, the amide oxygen of compound 3a is predicted to form a hydrogen bond with the main chain nitrogen of Gly66 and the ester carbonyl interacts with Gln19. The loss of these interactions may be one of the reasons for the significant activity drop observed for compound 7. In fact, modifying the ester and amide groups significantly changed the binding mode of both compounds 5 and 7 (Figures 4B,C, respectively) compared to that of the active analogs. Compound 5, which had the two carbonyl groups preserved, forms a hydrogen bond with Gly66 at S3, which can be reasoned as the cause of its higher percent inhibition value (38%) compared to that of analog 7 (11%). Notwithstanding, the imide of both compounds is exposed to the solvent and no interaction with S1’ is observed, which are critical detrimental features of analogs 5 and 7.

The leverage of the 3-chloro-4-methoxyphenyl moiety on cruzain activity was assessed by modifying the substitution pattern of the aromatic ring (Figure 5). Removing either chlorine or methoxyphenyl led to the less active compounds 3b (IC₅₀ = 13.9 µM) and 3c (IC₅₀ = 40.1 µM), respectively. Compound 3b retains the key polar interactions with the enzyme, however, the lack of the chlorine prevents an optimal interaction with the S2 pocket, which is known to be essential for potent cruzain inhibition as reported previously (Ferreira et al., 2014) (Figure S4A). The effect of removing the methoxy group from the S2-interacting aromatic ring proved to be more detrimental to activity than the removal of the chlorine. In fact, compound 3c was predicted to lose most of the key polar interactions with the enzyme (Figure S4B), which resulted in a poorly active compound (IC₅₀ = 40.1 µM). The unsubstituted phenyl derivative (3d) was 8-fold less active (IC₅₀ = 16.7 µM) than the parent 3-chloro-4-methoxyphenyl compound (3a). Additional...
manipulations, such as introducing hydrogen bond donors and acceptors and replacing the chlorine or changing its position, led to poorly active or inactive compounds (3e-h, 3k, 3k', 3l, 3l', 3j, 3m, and 3m'). Figures S4 and S5 show that a suboptimal interaction between the aromatic ring and the S2 subsite lead to the loss of most, if not all, polar contacts that proved to be important for activity. The exception was compound 3l, with naphthyl replacing the phenyl as the S2-interacting aromatic system (Figure S5D). This is consistent with the binding mode of these compounds, which shows that the aromatic moiety interacts with the S2 subsite (Figure 1B). The S2 subsite, which is mostly composed of hydrophobic amino acids, can accommodate bulky groups, as shown in our previous work on benzimidazole derivatives (Ferreira et al., 2014). Such groups can therefore promote an optimum interaction with the S2 subsite, leading to more potent cruzain inhibitors.

Since all compounds described so far were obtained from enantiopure amino acids, we next evaluated how the absolute configuration influences the activity against cruzain. To this end, the racemic form of compound 10h was prepared (10hrac). The synthesis of 10hrac was achieved by the same procedure presented in Scheme 3, using DL-phenylalanine.
instead of L-phenylalanine. Chiral column HPLC and optical rotation analyses confirmed that all reactions did not lead to racemization when L-phenylalanine was used as the starting material. These analyses also confirmed that the racemic compound was obtained when DL-phenylalanine was used as the starting material. The racemate (10h_{rac}) exhibited an IC_{50} value of 1.16 µM against cruzain, while compound 10h had an IC_{50} of 1.4 µM. The very close IC_{50} values obtained for both the racemate and the enantiopure compound demonstrated that the absolute configuration of the stereocenter is not relevant to the activity of this series of compounds.

In the next molecular optimization step, we combined the best fragments from all five regions of initial hit 3a. This strategy produced the most active compound (10j, IC_{50} = 0.6 µM) among the synthesized analogs, which was 4-fold more active than compound 3a (Figure 6A). Compound 10j conserves the same binding mode of the other active compounds in this series, forming the key intermolecular interactions with the S1, S1’, S2, and S3 subsites. The S1’ and S2 subsites are optimally filled with the benzyl and 3-chloro-4-methoxyphenyl rings, respectively. A hydrogen bond is formed between the amide oxygen and Gly66, and the imide carbonyl interacts with Gln19 (Figure 6B).

The aim of this study was to identify new imide derivatives as competitive cruzain inhibitors. Considering the competitive nature of hit compound 3a, the synthesized analogs were expected to follow the same behavior. Hence, we conducted further studies to establish the mechanism of inhibition of compounds 10h and 3i. The competitive mechanism of these compounds was corroborated by the Lineweaver–Burk plots shown in Figure 7. As expected for competitive inhibitors, the maximum velocity (1/V_{max}, intersections with the y-axis) remained unchanged with increasing inhibitor concentrations [I], while the apparent Michaelis–Menten constant (K_{app} = 1/K_{M}, intersections with the x-axis) increased with escalating [I].

**Discovery of Novel Trypanocidal Agents**

The biological activity of 22 compounds was evaluated against *T. cruzi* intracellular amastigotes along with BZ (Tables 1–3). Several compounds showed trypanocidal activity similar to or superior to that of BZ (IC_{50} = 3.0 µM). The data in Table 1 show that the hit compound 3a, despite being a potent cruzain...
indicate that these cyclic amines exert their trypanocidal activity by modulating a molecular target other than cruzain.

32.8 ± 8.0

Table 2 shows the data for compounds modified on the hydrophobic fragment and the amide group. Replacing the isobutyl with a benzyl group resulted in compound 10h, which, although active against cruzain, was inactive against the parasite. Notwithstanding, the benzyl group in compound 10j combined with the maleimide fragment generated the most active
compound against both *T. cruzi* (IC\(_{50} = 1.0 \mu M\)) and cruzain (IC\(_{50} = 0.6 \mu M\)). Compound 7, in which the amide carbonyl was removed, displayed the opposite behavior to that of compound 10h, i.e., 7 was active against *T. cruzi* but inactive on cruzain. As shown in Table 3, except for compounds 3k and 3l, modifications to the aromatic ring were unfavorable for activity against both *T. cruzi* and cruzain.

Figure 8 illustrates an SAR scheme for the synthesized imide derivatives. Following our initial approach, compound 3a was divided into five fragments, and the most relevant SARs were identified. In short, the imide function, although required for cruzain inhibition, is not essential for activity against *T. cruzi*. Benzyl is the ideal hydrophobic fragment for activity for cruzain and is tolerable regarding the activity against *T. cruzi*. Replacement of the 3-chloro-4-methoxyphenyl fragment is, in general, unfavorable for activity against *T. cruzi* and cruzain. Removal of the ester group is unfavorable for activity against cruzain, while the replacement of the secondary amide proved to be tolerable regarding the trypanocidal activity. The gathered data showed a lack of a direct correlation between the phenotypic and target-based results for some compounds; this was expected to some degree given the complexity of the intracellular environment and issues related to transport across membranes. For other compounds, however, the target-based and phenotypic data are clearly correlated. Compounds 10c, 10e, 10f, and 10g, for instance, besides being active against the enzyme, proved to be potent trypanocidal agents.

**In vitro Toxicity and Selectivity**
The cytotoxicity of the synthesized analogs was assessed using HFF-1 human fibroblasts and HepG2 human hepatocytes (Table 4). The selectivity index (SI) was calculated as the ratio between the CC\(_{50}\) values for the human cells and the IC\(_{50}\) values for *T. cruzi*. In general, the designed compounds demonstrated no significant toxicity against HFF-1 and HepG2 cells. Three compounds showed SI values for HFF-1 fibroblasts comparable or superior to that of BZ (SI >33): 10m (SI >55), 7 (SI >66) and 3k (SI >25). It is worth noting that these compounds were active against *T. cruzi*. Another aspect worth mentioning is that the most cytotoxic compounds (10c, 10e, and 10j) are Michael acceptors, which is a possible reason for their cytotoxicity. With respect to cytotoxicity to HepG2 cells, most compounds produced higher SI values than that of BZ, particularly those that are highly active against *T. cruzi*: 10c (SI = 36.6), 10f (SI >32), 10m (SI >35), 10o (SI = 30.6), 10j (SI = 67), and 7 (SI = 42).

**CONCLUSION**
The SBDD strategy applied herein, comprising synthetic organic chemistry, molecular docking, enzyme kinetics and phenotypic assays, resulted in the discovery of potent, reversible and non-peptidic cruzain inhibitors with remarkable trypanocidal activity. The success of this experimental-computational molecular optimization approach can be illustrated by compound 10j, the most potent cruzain inhibitor (IC\(_{50} = 0.6 \mu M\)), which is significantly more active than initial hit 3a (IC\(_{50} = 2.2 \mu M\)). One of the most potent cruzain inhibitors reported to date, compound 10j represents a new chemical class among the known inhibitors of this enzyme. Furthermore, compound 10j shows trypanocidal activity (IC\(_{50} = 1.0 \mu M\)) that is 3-fold higher than that of the clinically used drug BZ (IC\(_{50} = 3.0 \mu M\)). Other promising compounds are 10c, 10f, 10m, 10n, 10o, 10j, 7 and 3k, which showed trypanocidal activity comparable to that of BZ and SI values (HFF-1/*T. cruzi*) >10.
The target-based results enabled the identification of relevant SARs, which allowed the uncovering of pivotal structural aspects that drive the enzyme-inhibitor molecular recognition and the activity of the investigated compounds. Moreover, these findings provide substantial insights into the design of reversible cruzain inhibitors, which can be useful to surmount the drawbacks associated with irreversible ligands. An important aspect to remark is the in vitro toxicity profile of some compounds, which indicates that they may be safer than the drugs currently available for the treatment of Chagas disease. In a context characterized by a lack of therapeutic innovation and serious safety and efficacy issues, the cruzain inhibitors described herein can be explored as novel chemical matter in forthcoming Chagas disease drug discovery campaigns.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS

RAF: conceptualization, writing, molecular design, and organic synthesis. IP: conceptualization, writing, molecular design, and enzyme kinetics. TS: organic synthesis. CR: organic synthesis. MS: conceptualization, molecular design, and enzyme kinetics. LM and RK: in vitro experiments. LF: writing, molecular design, and in vitro experiments. LD: conceptualization, supervision, organic synthesis, and writing. AA: conceptualization, supervision, molecular design, and writing.

FUNDING

The National Council for Scientific and Technological Development (CNPq), Brazil, the Coordination for the Improvement of Higher Education Personnel (CAPES), Brazil, and the São Paulo Research Foundation (FAPESP), Brazil, grants 2013/07600-3 and 2011/13789-6, for financial support and postdoctoral fellowship (grant 158926/2014-5).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article including NMR spectra (Figures S6–S70) can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2019.00798/full#supplementary-material

REFERENCES

Arnal, A., Waleckx, E., Rico-Chávez, O., Herrera, C., Dumonteil, E. (2019). Estimating the current burden of Chagas disease in Mexico: A systematic review and meta-analysis of epidemiological surveys from 2006 to 2017. PLoS Negl. Trop. Dis. 13:e0006859. doi: 10.1371/journal.pntd.0006859

Avelar, L. A., Camilo, C. D., de Albuquerque, S., Fernandes, W. B., Gonzalez, C., Kenny, P. W., et al. (2015). Molecular design, synthesis and trypanocidal activity of dipeptidyl nitriles as cruzain inhibitors. PLoS Negl. Trop. Dis. 7:e0003916. doi: 10.1371/journal.pntd.0003916

Badiola, E., Fiser, B., Gómez-Bengoa, E., Mielgo, A., Olazola, I., Urruzuno, I., et al. (2014). Enantioselective construction of tetrasubstituted stereogenic carbons through Bronsted base catalyzed michael reactions: α-hydroxy enones as
Ferreira, R. S., Simeonov, A., Jadhav, A., Eidam, O., Mott, B. T., Keiser, M. J., Ferreira, L. L. G., and Andricopulo, A. D. (2019). Drugs and vaccines in GBD DALYs and HALE Collaborators. (2016). Global, regional, and national life expectancy (HALE), 1990-2015: a systematic analysis for the Global Burden of Disease Study. Lancet 388, 1603–1658. doi: 10.1016/S0140-6736(16)31473-X

Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol. 267, 727–748. doi: 10.1006/jmbi.1996.0897

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Ferreira, Pauli, Sampaio, de Souza, Ferreira, Magalhães, Rezende, Ferreira, Krogh, Dias and Andricopulo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.