Electronic Supplementary Material (ESI) for RSC Medicinal Chemistry.

Collaborative virtual screening to elaborate an imidazo[1,2-a]pyridine hit series for visceral leishmaniasis

Supporting Information

Contents:

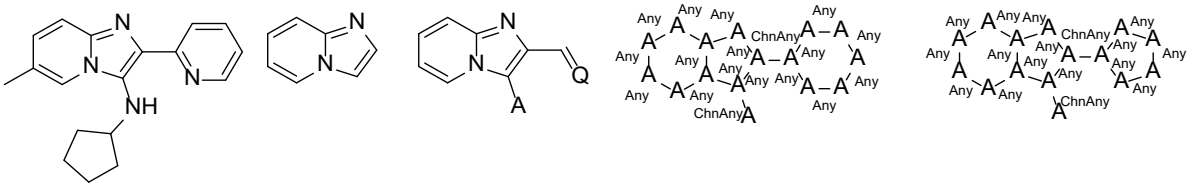
Computational Methods

The following computational methods were employed for virtual ligand based screening

Company A: All similarity searches were done using Pipeline Pilot Protocols (Pipeline Pilot 2018, Dassault Systems Biovia Corp). Molecular fingerprints for the Company A library and seed molecules were generated using ECFP4 fingerprints. Subsequently, Tanimoto similarity coefficients were computed and Company A library molecules with Tanimoto similarity threshold of >0.3 were selected as close analogs. The threshold was adjusted in the range of 0.3 - 0.7 depending on the number of close analogs for each seed - For example, if nothing came up at 0.7, then decreasing similarity incrementally by 0.05 increments down to 0.3 until enough hits were identified. Other times a combination of Tanimoto similarity and

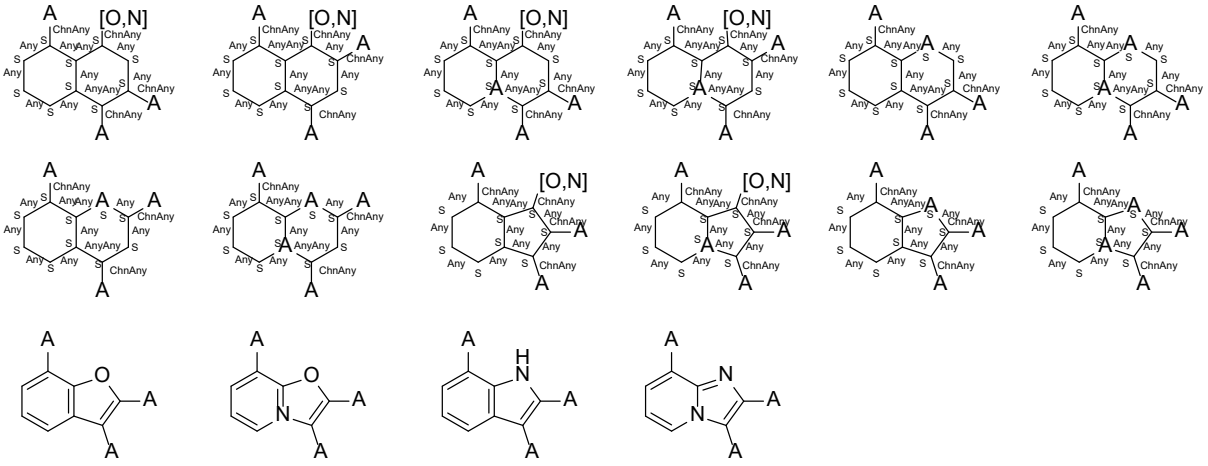
substructure search was used. Available compounds were sourced and plated for testing. Sourcing compounds was based on availability in sufficient amounts for plating

Company B: *For 1st virtual screen Fig 1A* Company B created a set of queries shown below. It was drawn in ChemDraw.



4108 compounds were identified in Company B library by substructure search based on these queries. Company B calculated their similarity value against the seed structure using Morgan fingerprint approach ECFP4 or FCFP4 and then picked up all compounds met criteria in which *Tanimoto similarity* is more than 0.5. In addition Company B picked up some unique compounds by visual inspection despite these compounds having *Taniomto similarity* is below 0.5. 96 compounds were submitted to DNDi to remove overlapped compounds with those of other consortium members.

For 2nd virtual screen Fig 1B the same process was followed using the set of queries shown below.



2168 compounds were identified in our library by the same way described previously. Finally 96 compounds were cherry-picked by visual inspection and then submitted to DNDi

Company C- *For 1st virtual screen Fig 1A* Similarity search against the Company C collection using ECFP4 and an in-house implementation of Daylight fingerprint with Tanimoto index cutoff set to 0.60 and 0.70, respectively, led to 284 hits. After applying compound quality and IP filters, the list was reduced to 180 unencumbered structures. *For 2nd virtual screen Fig 1B* Similarity search against the Company C collection using ECFP4 and an in house implementation of Daylight fingerprint with Tanimoto index cutoff set to 0.70, led to 764 hits. After applying compound quality criteria and IP filters, the list was reduced to 722 unencumbered structures. The final list of 139 compounds was obtained by selecting cluster representatives and focusing the selection to non-publicly known compounds aided by a proprietary in-house database.

Company D - Several hundreds of diverse compounds were gathered by three kinds of search, Tanimoto similarity search by Daylight fingerprint and ChemAxon fingerprint and automated substructure search from structure **1** or **2**. Second, the diverse compounds were ranked by alignment score to the pharmacophore model of the series compound structures, then the top one hundred compounds were supplied.

Company E - 150 compounds were selected by Tanimoto similarity calculation using the FCFP4 fingerprint. The selection of compounds consisted of 85 compounds with a similarity of 0.40 or more and other 65 compounds selected by clustering compounds with a similarity of 0.35 or more and less than 0.40.

***Leishmania donovani* assay**:

Cell preparation: THP-1 cells (ATCC, Manassas, VA, USA) were cultured at 37 °C in 5% CO_2 in a humidified incubator. THP-1 cells were washed with RPMI media (Welgene Inc., Gyeongsangbuk-do, Republic of Korea) supplemented with 10% (v/v) fetal bovine serum (FBS) (Gibco, Waltham, MA, USA), 1% (v/v) Penicillin-Streptomycin (10,000 U/mL) (Gibco) and plated directly to the 384-well tissue culture microplate (Greiner Bio-One, Kremsmünster, Austria) at 1.0 x 104 cells per well with 50ng/mL of phorbol 12-myristate 13-acetate (PMA) (Sigma-Aldrich, St. Louis, MO, USA). Wellmate Microplate Dispenser

(American Laboratory Trading, Inc., CT, USA) was used to dispense the cells 30 µl/well. It took 48 hours for THP-1 to be fully differentiated and ready for the infection.

Parasite preparation and infection: *Leishmania donovani* was seeded 4 days prior to the infection at $1x10^6$ parasites/ml in M199 medium (Gibco) supplemented with 10% (v/v) FBS (Gibco). On the infection day (48 hours after the cells plating), the parasite density was quantified and the parasites centrifuged (pelleted) and resuspended at 1.0×10^8 parasites/ml in DPBS (Welgene). To purify metacyclic forms, parasites were incubated with 50 µg/mL Lectin (Sigma-Aldrich) for 30 min at 28 ºC shaking incubator. Prior to centrifuge, parasite and Lectin were mixed gently. Then parasites were washed twice at 40 g for 5 min with DPBS. Parasites were then washed with RPMI and quantified again for infection. Wellmate Microplate Dispenser was used to dispense 20 µl/well (2.0×10^5 parasites/well) at the multiplicity of infection (MOI) values of 20:1 parasite to host cell.

Reference compounds, nuclei staining and fluorescence microscopy: The anti-leishmanial reference drugs used were amphotericin B (Sigma-Aldrich), miltefosine (Merck, Darmstadt, Germany). The reference drugs and tested compounds were added by CyBi®-Well multichannel pipettor (Analytik Jena AG, Jena, Germany) after 24 hours infection and incubated at 37 ºC and 5% CO_2 for 4 days. The cells and parasites were then fixed with 4% paraformaldehyde (CureBio, Seoul, Republic of Korea) and stained with 5µM DRAQ5 (Biostatus Ltd, Leicestershire, UK) in DPBS. Fluorescent images (4 images per well) were acquired from each single well using an Operetta automated confocal microscope (PerkinElmer, Inc., Waltham, MA, USA) at 635 mm excitation filter, 20X lens magnification enabling the determination of the infection ratio of the parasites by image analysis. An algorithm was developed with Columbus™ software (PerkinElmer) to identify and individualize the macrophages by setting an intensity threshold to discriminate background (extracellular space) from foreground (macrophage cells area). The macrophage nuclei and internalized parasites were identified and counted if localized in the foreground previously selected. With this technique, extracellular parasites were not considered in the calculations, and the infection ratio was determined by the number of infected cells divided

by the total number of cells after normalization based on the controls. The average infection ratios from the positive and negative controls were normalized to 0% and 100% infection, and the infection ratio read from each compound activity was proportionally distributed within this range. Z-factor was used for protocol validation and active compound selection acceptance. The values produced by the algorithm were further analyzed using XLfit (IDBS, Guildford, UK) in Microsoft Excel for half-maximal effective concentration (EC_{50}) value determinations.

***Trypanosoma cruzi* assay**:

Cell and parasite preparation U2OS (human bone osteosarcoma host cells, ATCC, Manassas, VA, USA) cells were cultured in DMEM-High glucose media (Welgene Inc., Gyeongsangbuk-do, Republic of Korea) supplemented with 10% (v/v) fetal bovine serum(FBS) (Gibco, Waltham, MA, USA), 1% (v/v) Penicillin-Streptomycin (10,000 U/mL) (Gibco) at 37°C, 5% CO_2. *T. cruzi* parasites used for the infection were Tissue Culture Trypomastigotes (TCT). To generate this stage of parasites, metacyclic trypomastigotes were obtained from a late stage epimastigote culture. The metacyclic trypomastigotes were used to infect the monkey kidney cell line LLC-MK2 (ATCC). Parasites 7 days after infection were collected from the supernatant of the LLC-MK2 infected culture and re-infected new culture of LLC-MK2 in DMEM-Low glucose media (Welgene Inc.) supplemented with 2% (v/v) FBS (Gibco), 1% (v/v) Penicillin-Streptomycin (10,000 U/mL) (Gibco).

Assay for bioactive compounds against *intracellular T. cruzi* The compounds were prepared in 2-fold serial dilution 10-points from 10mM (200X stock) with 100% DMSO (v/v) (Sigma-Aldrich, St. Louis, MO, USA) into 384-well polypropylene microplates (Greiner Bio-One, Kremsmünster, Austria). For the controls, Benznidazole (Carbosynth Ltd, Berkshire, UK) and Posaconazole (Carbosynth Ltd) were used as a positive controls and 0.5% DMSO (v/v) as a negative controls. The 0.3 µL of compounds or reference compounds were dispensed in the 10 µL of DPBS (Welgene) containing wells of 384-well tissue culture microplate (Greiner) with CyBi®-Well multichannel pipettor (Analytik Jena AG, Jena, Germany). U2OS cells and parasites were mixed in DMEM-Low glucose media (Welgene Inc.) supplemented with 2%

(v/v) FBS (Gibco), 1% (v/v) Penicillin-Streptomycin (10,000 U/mL) (Gibco) at the multiplicity of infection (MOI) values of 12.5:1 parasite to host cell. Prepared cells and parasites mixtures were dispensed into the assay plate at 50µL/well and incubated for 72 hours at 37°C, 5% CO_2. After incubation was done, the cells and parasites were stained using 5µM DRAQ5 (Biostatus Ltd, Leicestershire, UK) in 4% paraformaldehyde (CureBio, Seoul, Republic of Korea). Cell images were acquired with the Operetta automated confocal microscope (PerkinElmer, Inc., Waltham, MA, USA) at 635 mm excitation filter, 20X lens magnification. Images were captured 4 images per well covering 45% of the well and analyzed with Columbus™ software (PerkinElmer). The number of cells and intracellular parasites were counted by the DRAQ5 stained nuclei of both host cells and parasites. The large host cell nuclei were first detected and the cytoplasm of the host cells was then identified by a cytoplasm detection script. The intracellular parasites were defined as spots within the host cell cytoplasm using spot analysis and host cells containing more than 3 spots were considered to be infected. The infection ratio was determined by the number of infected cells divided by total number of cells and normalized values based on the positive and negative controls to 0% and 100% infection. To estimate EC_{50} and CC_{50} values, data were fitted to sigmoid dose-response one-site-fit model 205 (4 Parameter Logistic Model) of XLfit (IDBS, Guildford, UK) in Microsoft Excel.

Mouse / Human Liver Microsomes Stability Assay

Mouse and human liver microsomes were purchased from Xenotech or Corning and stored in a freezer (lower than -60°C) before use. β-nicotinamide adenine dinucleotide phosphate reduced form, tetrasodium salt, was purchased from Chem-impex international, Cat.No.00616. Control compounds: Testosterone, diclofenac and propranolol. Preparation of Working Solutions was as follows:

- Stock Solution: 10 mM test compound in DMSO.
- Intermediate solution: 100 µM test or control compounds in 100% methanol (Concentration of Organic Solvent: 1% (v/v) DMSO and 99% (v/v) methanol)

- Working solution： 10 μM test or control compound in 100 mM potassium phosphate buffer (Concentration of Organic Solvent: 0.1% (v/v) DMSO and 9.9% (v/v) methanol)

A total of eight 96-well sample plates were prepared for incubation, labelled as T0, T5, T10, T20, T30, T60, Blank60 and NCF60. The first 6 plates corresponded to time points 0, 5, 10, 20, 30 and 60 min. The Blank60 plate referred to the no-test-compound treatment and sampled only at 60 min after incubation. The NCF60 plate referred to the treatment of not adding the NADPH (replaced by the potassium phosphate buffer) and sampled only at 60 min after incubation. Each well on the T0, T5, T10, T20, T30, T60 and NCF60 plates were added with 10.0 μL of the Test or Control Compound Working Solution and 80.0 μL of the Microsome Working Solution containing 0.5mg/mL animal or human microsomes protein, while microsomes were added to the Blank60 plate without the test or control compounds. These plates were then pre-warmed in a 37°C water bath for approximately 10 min. To start the reaction, 10.0 μL of the NADPH Solution was added to each well on all plates except that 10.0 μL of 100 mM potassium phosphate buffer was added in the NCF60 plate. Consequently, for the wells containing the test or control compounds, the final concentration was 1 μM for test compounds, testosterone, diclofenac and Propranolol , 0.4 mg/mL for animal or human liver microsomes, 0.01% (v/v) for DMSO and 0.99% (v/v) for methanol. After incubation at 37°C, the reaction was stopped at appropriate time points, i.e., 5, 10, 20, 30 and 60 min, by adding 300 μL of the stop solution in 100% acetonitrile containing internal standards into each well of test and control compounds. T0 plate preparation: The T0 plate was ended by adding 300 μL of the stop solution in 100% acetonitrile containing internal standards firstly followed by adding 10.0 μL of the NADPH Working Solution. All plates were shaken and centrifuged at 4000 rpm for 20 min. Then 100 μL of supernatant was taken from each well and diluted with 300 μL pure water before being analyzed by LC/MS/MS.

Kinetic solubility assay

The kinetic solubility assay employed was the shake flask method followed by HPLC analysis. Using a 10 mM stock solution of each compound in 100% DMSO, dilutions were prepared to

a theoretical concentration of 200 µM in PB buffer (2% DMSO final). After incubation for 24 hr and filtering, the supernatant is analyzed by HPLC equipped with UV.

The following working solutions were required:

- pH 2.0 Buffer 1.95g $NaH_2PO_4 \cdot 2H_2O$, dissolved with about 240mL water in a flask, mixed well, pH adjusted to 2.0 with HCl (add 800µL HCl, test pH, then add HCl incrementally to pH 2.0) , transferred to a 250mL volumetric flask, filled with water to the level
- pH 7.4 Buffer 0.39g $NaH_2PO_4 \cdot 2H_2O$, 1.4025g Na_2HPO_4, dissolved with about 240mL water in a flask, mixed well, pH adjusted to 7.4 with 10M NaOH , transferred to a 250mL volumetric flask, filled with water to the level.
- Standard solution preparation: 50% ACN was mixed with 50% of buffer solution to get the diluted solution. A 10 mM stock solution (10 µL/compound) was mixed with a diluted solution (490 µL/compound) to obtain UV standard solutions of 200 µM. The 200µM standard solution was diluted by a factor of 10 and 200 fold to obtain UV standard solutions of 20µM and 1µM. The UV standard solutions of 1, 20 and 200 µM were used as standard samples for kinetic solubility tests.

Method: the test compound was dissolved in DMSO to make the stock solution of 10 mM. At least 200 µL each of stock solution was needed. Amiodarone hydrochloride, Carbamazepine and Chloramphenicol solutions were used as controls. Test compounds and controls (10 mM in DMSO, 10 µL/well) were added into the pH 2.0 buffer (490 µL/well) and pH 7.4 buffer (490 µL/well) respectively and placed in a 96-well plate. The final concentrations of test compounds and DMSO were 200µM and 2%, respectively. The theoretical maximum concentration is 200µM. The solubility solutions were Incubated on a shaker at 500rpm for 24 h at room temperature. 200µL of each solubility solution were transferred into a new MultiScreen fliter plate (with polycarbonate membrane), filtered by Millipore vacuum manifold and the filtrate collected as test samples. Test compound concentrations were determined by HPLC UV using the following process: three UV standard solutions were injected into HPLC from low to high concentration, followed by testing of K.S. filtrate. Testing samples were injected in duplicate. The UV-chromatogram was integrated. The calibration equation was simulated and the kinetic solubility calculated.

Mouse Pharmacokinetic study of compound 23

The mouse pharmacokinetic study was performed with strict adeherence to IACUC guideleines (IACUC study number N20161028).

Following at least 3 days of acclimatization 6 female Balb-C mice of 7-9 weeks of age (sourced from LingChang) were split into two groups of 3 (Group 1: IV; group 2 p.o. administration). Animals were fasted overnight prior to administration of test article, with food returned 4 hours post-dose.

Vehicle preparation and administration was completed as follows:

IV dosing group : the required amount of **23** was weighed into a suitable vial. 20% final volume of NMP was added into the vial and vortexed to get a clear solution. 40% final volume of PEG400 was added into the vial and vortexed to get a clear solution. 40% final volume of NaCl 0.9% was added into the vial and vortexed to get a clear solution. The dose formulation was administered via tail vein to animals following the SOPs of the animal facility. IV dose was filtered prior to administration.

PO dosing: the required amount of **23** was weighed into a suitable vial. 100% final volume of PEG400 was added into the vial and vortexed or stirred to get a clear solution. The dose formulation was administered via oral gavage per facility SOPs.

The dose volume of each animal was determined by the body weight collected on the morning of the dosing day. In both instances stirring was maintained during dosing.

Approximate 30µL of blood was collected from submandibular vein or saphenous vein for each mouse. All blood samples were transferred into 1.5-mL plastic tubes containing K2-EDTA (0.5M, 2µl) as an anti-coagulant and placed on ice until processed for plasma. Blood samples were processed for plasma by centrifugation at approximately 4°C (3000g, 15min). All plasma samples were transferred into pre-labelled polypropylene micro centrifuge tubes, quick-frozen over dry ice and stored at -70°C until LC/MS/MS analysis. Plasma concentration versus time data was analyzed by non-compartmental approaches using the Phoenix WinNonlin 6.3 software program. CL, VD_{ss}, C_0, C_{max}, T_{max}, $T_{½}$, $AUC_{(0-t)}$, $AUC_{(0-inf)}$, $MRT_{(0-t)}$, $MRT_{(0-inf)}$, %F and graphs of plasma concentration versus time profile were reported.