Published in final edited form as:

J Med Chem. 2009 October 22; 52(20): 6489–6493. doi:10.1021/jm900908p.

Structure-guided development of selective TbcabB inhibitors

Jeremy P. Mallari^{†,‡}, Anang A. Shelat[‡], Aaron Kosinski[‡], Conor R. Caffrey[§], Michele Connelly[‡], Fangyi Zhu[‡], James H. McKerrow[§], and R. Kiplin Guy^{†,‡,*}

[†]Graduate Program in Chemistry and Chemical Biology, University of California, San Francisco CA 94143-2280

[‡]Department of Chemical Biology and Therapeutics, St. Jude Children's Research Hospital, Memphis TN, 38105

[§]Department of Cellular and Molecular Pharmacology, University of California, San Francisco CA 94143-2280

Abstract

The trypanosomal cathepsin TbcabB is essential for parasite survival and is an attractive therapeutic target. Herein we report the structure-guided development of TbcabB inhibitors with specificity relative to rhodesain and human cathepsins B and L. Inhibitors were tested for enzymatic activity, trypanocidal activity, and general cytotoxicity. These data chemically validate TbcabB as a drug target, and demonstrate that it is possible to potently and selectively inhibit TbcabB relative to trypanosomal and human homologues.

Introduction

The burden of Human African Trypanosomiasis (HAT) on the health and economies of sub-Saharan Africa, coupled with the well-known shortcomings of the currently available drugs, clearly illustrates the need for new drug development.¹ RNAi studies have demonstrated the essentiality of the trypanosomal cathepsin TbcabB, both *in vitro* and *in vivo*.^{2, 3} There is growing evidence that cysteine protease inhibitors hold promise for the discovery of new HAT drugs.^{4–9}

To minimize possible off-target effects, it is generally desirable for potential therapeutics to selectively inhibit the molecular target relative to homologues. We have previously reported that the purine nitrile scaffold produced several potent inhibitors of TbcabB and that all of those compounds were also fairly non-selective, being inhibitors of human cathepsin L.⁵ Furthermore, the lead compound from that series, **4**, showed moderate activity against human cathepsin B. Cathepsin L has been implicated by conditional knockout studies in regulating pro-opiomelanocortin hormone production.^{10, 11} While the role of cathepsin B alone remains unclear,¹⁰ the coordinate action of cathepsin B and cathepsin L has been implicated in maintaining normal brain physiology in adults.¹¹ For these reasons, this suboptimal selectivity profile raised concerns as to the feasibility of developing non-toxic TbcabB inhibitors.

*To whom correspondence should be addressed. Tel.: (901) 495-5714. Fax: (901) 495-5715. kip.guy@stjude.org.

Supporting Information Available

Spectroscopic data and LCMS data for all listed compounds. Experimental details for molecular docking, heatmap, and regression modeling studies. Experimental details for inhibitor synthesis, protease purification, and cellular and biochemical assays. This material is available free of charge via the Internet at <http://pubs.acs.org>.

In addition, the first generation compounds were found to be potent inhibitors of the trypanosomal cathepsin L-like protease rhodesain (*unpublished data*), indicating that rhodesain inhibition may be a determinant of trypanocidal activity. Although TbcabB is known to be essential,^{2, 3, 5, 7} the interplay between TbcabB and rhodesain inhibition in killing the parasite is not well understood.

We sought to address these questions by developing specific TbcabB inhibitors, relative to cathepsin L, cathepsin B, and rhodesain using a structure-guided approach. Molecular docking studies were conducted to explore possible strategies for achieving selectivity and to guide inhibitor design. The hypotheses arising from these studies were tested by making and studying the inhibitory effects of the designed inhibitors. Herein, we report the development of a series of TbcabB inhibitors with greatly enhanced selectivity against rhodesain and human cathepsins B and L. This series includes the most potent TbcabB inhibitors to date, and these compounds were shown to kill *T.brucei* with a high degree of selectivity relative to a panel of four mammalian cell lines (BJ, HEK 293, HEP G2, and Raji).

Inhibitor Design

The previously reported purine nitrile inhibitors of TbcabB also inhibit rhodesain and human cathepsin L.⁵ To explore possible structural explanations for the overlap between the inhibition profiles of these proteases, lead TbcabB inhibitor **4** was modeled as a covalent adduct with cysteine and docked to the previously reported homology model of TbcabB⁵ and crystal structures of human cathepsin L and rhodesain (PDB codes 1mhw and 2p86, respectively) (Figure 1, see Supporting Information for modeling details).

Compound **4** was predicted to make similar interactions with each protease, consistent with the lack of selectivity observed for this inhibitor. In each model, the 3-hydroxypropyl side chain of the ligand projects into solvent on the prime side of the protease binding pocket. The N9 amine forms a hydrogen bond to the carbonyl of either Gly72 (TbcabB), Gly68 (cathepsin L), or Gly66 (rhodesain). Finally, the 3,4-dichlorophenyl ring makes strong Van der Waals contact with the well-defined, hydrophobic S2 pockets of each protease.

Although the inhibitor's predicted binding orientation is similar across the three enzymes, modeling suggests potentially exploitable differences in the S2 and S3 pockets. In TbcabB, residues His179 to Gly188 form a loop oriented towards the prime side of the active site cleft. Consequently, the entrance to the S2 pocket near Asp165 is much wider in comparison to those of cathepsin L and rhodesain. In contrast, the homologous loop region in cathepsin L points away from the prime side, in part because of a disulphide bridge between Cys156 and Cys204. As a result, Met161 truncates the S2 pocket in cathepsin L. At the other side of the active site cleft, Asp73 projects toward solvent and acts to constrict the S3 pocket in TbcabB, while in cathepsin L the orientation of Tyr72 results in a much wider S3 pocket which bridges the S2 site via Leu69. Rhodesain shares structural traits from both the TbcabB and cathepsin models: Leu160 plays a similar role to Met161 in cathepsin L, but Leu67 and Phe61 occlude the S3 pocket just as Asp73 does in TbcabB.

In summary, the S2 pocket of TbcabB is expected to be much larger and more negatively charged than the S2 pockets of rhodesain and cathepsin L, whereas the S3 pocket is most accessible in cathepsin L. It was envisioned that the differences between the proteases' S2 binding sites could be exploited by increasing steric bulk at the 6-amino substituent in order to improve inhibitor potency and selectivity for TbcabB. The first inhibitor series explored this hypothesis by incorporating structurally varied aryl moieties at the 3 position of the 6-amino benzyl ring (Table 1).

Chemistry

Intermediate **2** was synthesized by the general route (Scheme 1) previously described.^{5, 12} Briefly, **1** was reacted with 3-bromobenzylamine in 2-butanol to install the 6-amino substituent. The crude reaction was concentrated *in vacuo* and re-suspended in dimethylformamide (DMF) with K₂CO₃. Alkylation at N9 was accomplished by heating the crude reaction product with 3-bromopropanol. Purification was accomplished by flash chromatography, and overall yield for the two reactions was 60%. Subsequent reaction with sodium cyanide in dimethyl sulfoxide (DMSO) with microwave acceleration afforded intermediate **2**, which was purified by preparative C₁₈ chromatography with a resulting yield of 70%. Installation of the distal aryl ring was performed by Suzuki cross coupling. The desired aryl boronic acid, intermediate **2**, Na₂CO₃, and Pd(PPh₃)₄ were reacted in 1,4 dioxane with microwave acceleration. The target inhibitors **3 (a-ae)** were purified by preparative C₁₈ chromatography with yields of 10% to 60%. $\geq 95\%$ purity of target compounds was confirmed by LCMS on both C₄ and C₁₈ columns.

Determining the activity of purine inhibitors **3(a-ae)**

Inhibitors **3(a-ae)** were tested *in vitro* to determine inhibitory activity against TbcA, rhodesain, cathepsin L, and cathepsin B (Table 2, 3). Compounds were also assayed to determine anti-proliferative activity against the cultured bloodstream form of *T. brucei* (Table 4). Finally, cytotoxicity was evaluated in BJ, HEK 293, Raji, and HEP G2 cell lines to determine a therapeutic window.

The initial inhibitor set focused on introducing sterically and electronically varied aryl groups at the 3' position of the 6-amino benzyl ring (Table 2). This design strategy was quite successful, producing several inhibitors with greater than 5-fold enhanced potency against TbcA relative to parent compound **4** (EC₅₀ = 1.4 μ M). It was found that a wide range of aryl moieties were suitable, including small heterocycles, substituted phenyl groups, and larger fused ring systems. Activity against cathepsin B in this series is almost non-existent, especially among the most potent TbcA inhibitors. In addition, many compounds in this series (**3c**, **3h**, **3o**, **3p**, **3v**) showed modest (2–3 fold) selectivity for TbcA over cathepsin L and rhodesain, a marked improvement over the parent compound, which preferentially inhibits both proteases (approximately 10- and 25-fold, respectively) over TbcA. **3b** was the most selective TbcA inhibitor relative to both cathepsin L (~5-fold) and rhodesain (~25-fold). This inhibitor was one of only two compounds (**3b** and **3n**) to incorporate an ortho-substituted biphenyl group, and it was hypothesized that this substitution pattern constrains the biphenyl moiety in a conformation that gives rise to TbcA selectivity.

To test this hypothesis, a second series of compounds was synthesized to further optimize the ortho-substituent at the distal aryl ring (Table 3). This round of optimization yielded a compound with exquisite selectivity for TbcA relative to rhodesain (**3ab**, >100-fold selective) as well as the compound most selective for TbcA relative to cathepsin L (**3aa**, >10-fold selective). Compounds **3ac–3ae** all displayed good selectivity relative to rhodesain, and modest selectivity relative to cathepsin L. These results indicate that ortho-substitution at the distal phenyl ring is indeed an important determinant of selectivity against both rhodesain and cathepsin L.

To further evaluate the therapeutic potential of this compound series, each inhibitor was screened against a panel of four mammalian cell lines (BJ, Raji, Hep G2, and HEK 293). Previous work demonstrated purine derived nitriles are generally non-toxic in HEP G2 and Raji, and only mildly toxic in HEK 293.⁵ Similar trends were observed in this compound series. Almost every inhibitor displayed EC₅₀ values weaker than 10–20 μ M across all four cell lines. Overall, HEK 293 was the most sensitive cell line while Hep G2 and BJ were the least sensitive. To determine a therapeutic window, a cellular therapeutic index was defined as:

Cellular therapeutic index=(Lowest EC₅₀ from mammalian cell panel)/(EC₅₀ vs. *T.brucei*)

Several inhibitors display a therapeutic index between 25- to 50-fold across four different cell lines, further supporting the proposition that targeting TbcabB is a promising therapeutic strategy. It is interesting to note that potency against TbcabB correlates with therapeutic index, although selectivity for TbcabB relative to cathepsin L does not necessarily improve therapeutic index.

Additional docking studies to rationalize cathepsin L and rhodesain activity

Although good selectivity was achieved in a subset of inhibitors, many compounds in this study were far more promiscuous than expected. It was believed that the smaller S2 pockets in cathepsin L and rhodesain could not accommodate aryl substitution of the 6-amino benzyl ring. In order to gain insight into this unexpected activity, compounds **3a**, **3aa**, **3ab**, **3i**, and **3x** were docked against TbcabB, cathepsin L, and rhodesain. This inhibitor set exhibits a wide range of selectivity: **3a**, **3i** and **3x** are active against all three proteases, **3aa** is TbcabB selective, and **3ab** is inactive against rhodesain..

As expected, the wider S2 pocket in TbcabB easily accommodates the distal aryl ring of the 6-amino side chains. In contrast, most cathepsin L ligands occupy the S3 pocket and most rhodesain compounds spill out of the S2 pocket by packing against Leu160 (Figure 2). The two exceptions are the 4-chlorophenyl compound, **3i**, which adopts a similar pose in both cathepsin L and rhodesain by pointing into a hydrophobic pocket beyond S2, and the 2-chlorophenyl compound, **3aa**, which displaces Phe61 and occupies the S3 pocket in the rhodesain model (Figure S1).

These molecular modeling studies offer no clear explanation for the observed selectivity of the ortho-substituted compounds. The covalent nature of the protein-ligand complex complicates the comparison of binding free energies across different compounds. Nevertheless, it is apparent that the aryl substituted 6-amino benzyl side chain is much more flexible than previously imagined. Additional optimization of cathepsin L selectivity might be achieved by exploiting the different electronic environments of the S2 pocket in TbcabB and the S3 pocket of cathepsin L. It may also be productive to adjust the placement of the distal aryl ring by replacing of the 6-amino benzyl linker. These results offer a promising direction for further development of TbcabB specific inhibitors, and support the idea that TbcabB can be effectively and specifically targeted in a therapeutic context.

Protease inhibition and trypanocidal activity

Although the purine nitrile scaffold is thought to kill *T.brucei* by inhibition of TbcabB, the contribution of rhodesain inhibition to trypanocidal activity remains poorly understood. This inhibitor series provides new chemical tools for understanding trypanocidal activity as a function of TbcabB and rhodesain inhibition. With a set of structurally similar inhibitors that exhibit a range of selectivity profiles, it is now possible to examine the individual importance of these two putative protease targets in determining trypanocidal activity.

The heatmap in Figure 3 summarizes the protease and trypanocidal activity of the compounds in this study. All ortho-substituted analogs partition into the right cluster (**3f–3n**) formed from the first bifurcation of the dendrogram along the x-axis. Two conclusions can be drawn from this analysis. First, it is clear that ortho-substitution is generally favorable for TbcabB, but it is not well tolerated in either rhodesain or cathepsin B. Second, when comparing activity profiles

along the y-axis, the TbcAtB protease assay behaves more similarly to the trypanocidal assay, whereas the rhodesain assay clusters more closely to that of cathepsin L.

To further investigate how trypanocidal activity relates to TbcAtB inhibition, linear regression analysis was performed to model *T. brucei* EC₅₀ as a function of the following independent variables: log-normalized TbcAtB and rhodesain EC₅₀, PAMPA-based (Parallel Artificial Membrane Permeability Assay) percent membrane retention and log-normalized permeability, and two calculated descriptors (SlogP_VSA9 and a_count) which appeared in the regression model describing the first generation series.⁵ An iterative model building process yielded an equation with three terms (standard errors in parenthesis, see Supporting Information for regression details):

$$T. brucei \log_{10}(EC_{50}) = [0.85(0.13)][TbcAtB \log_{10}(EC_{50})] + [0.12(0.06)][rhodesain \log_{10}(EC_{50})] - [0.11(0.06)][\log_{10}(\text{permeability})] + [0.54(0.13)]$$

Increasing potency (decreasing EC₅₀) against TbcAtB and rhodesain, and increasing PAMPA permeability results in greater trypanocidal activity. In this model, the TbcAtB term is dominant whereas permeability and rhodesain inhibition contribute similarly.

The regression model fits the training set data with an adjusted $r^2 = 0.667$ (Figure 4), and demonstrates reasonable predictive accuracy as assessed by cross-validation: the average predictive squared correlation coefficient, q^2 ,¹⁸ is 0.578 and the average root-mean squared error (RMSE) is 0.208 (in $\log_{10}(EC_{50})$ units). The TbcAtB term is highly significant by ANOVA ($p < 10^{-5}$), whereas the rhodesain and permeability covariates are marginally significant ($p = 0.063$ and $p = 0.091$, respectively). Importantly, rhodesain inhibition and PAMPA permeability are not statistically significant by themselves. The addition of both variables together significantly improves model quality relative to TbcAtB alone (Δq^2 is $[0.035 - 0.054]$ and $\Delta RMSE$ is $[-0.012 - -0.008]$ at the 95% confidence interval) to justify inclusion.

It is clear from the plot of experimental versus predicted *T. brucei* EC₅₀ that the refined model does a poor job of predicting the activity of the first generation compounds. However, the TbcAtB term was not the dominant term in the earlier study. Moreover, the two calculated descriptors (SlogP_VSA9 and a_count) are no longer relevant to the current model. It is likely that the first generation series, which had greater chemical diversity at both the N6 and N9 positions, exerted some of its trypanocidal activity via off-target effects, or was more prone to differences in membrane permeability or intracellular localization.

In contrast, data from the current study strongly suggests that TbcAtB inhibition is primarily responsible for the anti-parasitic activity of the aryl substituted 6-aminophenyl purine nitriles. The importance of rhodesain inhibition appears to be minimal. This conclusion is clearly illustrated by the TbcAtB selective ortho-substituted series (compounds **3b** and **3aa–3ae**), which maintains potent trypanocidal activity despite poor potency against rhodesain.

Conclusion

This paper describes the structure-guided optimization of selectivity in a series of purine nitrile inhibitors. By exploiting predicted differences at the S2 pockets of TbcAtB, rhodesain, and cathepsin L, a series of selective inhibitors with enhanced potency for TbcAtB was developed. The activity profiles of these compounds across a panel of proteases strongly support the proposed binding model for TbcAtB, and provide a structural rationale for the further development of selective protease inhibitors. In addition, these results suggest that it is possible to effectively target TbcAtB with selectivity relative to the human cathepsins, reaffirming the potential of TbcAtB as a therapeutic target. Finally, it was found that inhibition of TbcAtB was

a key determinant of trypanocidal activity for this compound series while inhibition of rhodesain was relatively unimportant.

Experimental section

Synthesis

Amines and alkyl bromides were purchased from Sigma Aldrich. Sodium cyanide, potassium carbonate, sodium carbonate, and 2,6-dichloropurine **1** were purchased from Sigma Aldrich. Boronic acids were purchased from Sigma Aldrich, TCI America, Maybridge, Alfa Aesar, and Fluka. All reaction solvents were anhydrous. LCMS data were collected on a Waters system using XTerra C₁₈ and YMC Pro C₄ columns (MeOH:H₂O, 1% HCOOH). Intermediate nitrile **2** was synthesized as previously described.⁵

Synthesis of 3a–3ae—2 (23 mg, 0.06 mmol), Na₂CO₃ (25 mg, 0.24 mmol, 4.0 eq.), Pd (PPh₃)₄ (14 mg, 0.012 mmol, 0.2 eq.), and the appropriate aryl boronic acid (0.12 mmol, 2.0 eq.) were dissolved in anhydrous 1,4-dioxane (0.5 mL, 0.12 M) in a 0.5 mL sealed microwave reaction vial (Biotage). The vial was heated to 150° C for increments of 5–15 minutes in a Biotage Initiator microwave until the reaction was complete as shown by LC/MS, a total of 20–90 minutes. The crude product was purified by reverse-phase HPLC (Waters XTerra preparative C₁₈ column, ACN:H₂O, 1% HCOOH, linear gradient from 40–70% ACN) to afford target inhibitor **3a–3ae** with a resulting yields of 10–70%

Cloning, expression, and purification of TbcatB

TbcatB was cloned and prepared as previously described.⁵

Cellular and biochemical assays

Inhibitors were screened for protease inhibition, trypanocidal activity, and toxicity against mammalian cell lines as previously described.^{5, 7}

Molecular docking studies

The previously reported TbcatB homology model⁵, the crystal structure of human cathepsin L (PDB code 1mhw), and the crystal structure of *T. brucei* rhodesain (PDB code 2p86) served as the starting point for the molecular docking studies. The MOE program (Chemical Computing Group, version 2008.10) was used for visualization, initial protein preparation, and building the modified cysteine residues. Protein models were charged at physiological pH using the PDB2PQR¹³ and PropKa¹⁴ programs. All molecular dynamic simulations were performed using AMBER10 in conjunction with AMBER Tools version 1.0,¹⁵ and visualized in VMD 1.8.6 (<http://www.ks.uiuc.edu/Research/vmd/>).¹⁶ Diagnostics plots and statistics were calculated using custom PERL and R scripts.¹⁷ Default parameters and settings were used unless otherwise noted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Abbreviations

HAT, Human African Trypanosomiasis; DMF, dimethylformamide; DMSO, dimethyl sulfoxide; ACN, acetonitrile; PAMPA, parallel artificial membrane permeability assay.

Acknowledgements

This work was supported by the American Lebanese Syrian Associated Charities (ALSAC) of St. Jude Children's Research Hospital, National Institute of Allergy and Infectious Diseases grant AI35707, Drugs for Neglected Diseases initiative, and the Sandler Family Supporting Foundation. The authors would also like to thank Cindy Choy and David Smithson for their advice and feedback.

References

1. WHO. African trypanosomiasis (sleeping sickness). WHO; 2006. [Online] <http://www.who.int/mediacentre/factsheets/fs259/en/>
2. Abdulla MH, O'Brien T, Mackey ZB, Sajid M, Grab DJ, McKerrow JH. RNA Interference of *Trypanosoma brucei* Cathepsin B and L Affects Disease Progression in a Mouse Model. *PLoS Negl Trop Dis* 2008;2(9):e298. [PubMed: 18820745]
3. Mackey ZB, O'Brien TC, Greenbaum DC, Blank RB, McKerrow JH. A cathepsin B-like protease is required for host protein degradation in *Trypanosoma brucei*. *J Biol Chem* 2004;279(46):48426–48433. [PubMed: 15326171]
4. Scory S, Caffrey CR, Stierhof YD, Ruppel A, Steverding D. *Trypanosoma brucei*: killing of bloodstream forms in vitro and in vivo by the cysteine proteinase inhibitor Z-phe-ala-CHN2. *Exp Parasitol* 1999;91(4):327–333. [PubMed: 10092476]
5. Mallari JP, Shelat AA, O'Brien T, Caffrey CR, Kosinski A, Connelly M, Harbut M, Greenbaum D, McKerrow JH, Guy RK. Development of potent purine-derived nitrile inhibitors of the trypanosomal protease Tbc4B. *J Med Chem* 2008;51(3):545–552. [PubMed: 18173229]
6. Troeberg L, Morty RE, Pike RN, Lonsdale-Eccles JD, Palmer JT, McKerrow JH, Coetzer TH. Cysteine proteinase inhibitors kill cultured bloodstream forms of *Trypanosoma brucei*. *Exp Parasitol* 1999;91(4):349–355. [PubMed: 10092479]
7. Mallari JP, Shelat A, Kosinski A, Caffrey CR, Connelly M, Zhu F, McKerrow JH, Guy RK. Discovery of trypanocidal thiosemicarbazone inhibitors of rhodesain and Tbc4B. *Bioorg Med Chem Lett* 2008;18(9):2883–2885. [PubMed: 18420405]
8. Fujii N, Mallari JP, Hansell EJ, Mackey Z, Doyle P, Zhou YM, Gut J, Rosenthal PJ, McKerrow JH, Guy RK. Discovery of potent thiosemicarbazone inhibitors of rhodesain and cruzain. *Bioorg Med Chem Lett* 2005;15(1):121–123. [PubMed: 15582423]
9. Greenbaum DC, Mackey Z, Hansell E, Doyle P, Gut J, Caffrey CR, Lehrman J, Rosenthal PJ, McKerrow JH, Chibale K. Synthesis and structure-activity relationships of parasitocidal thiosemicarbazone cysteine protease inhibitors against *Plasmodium falciparum*, *Trypanosoma brucei*, and *Trypanosoma cruzi*. *J Med Chem* 2004;47(12):3212–3219. [PubMed: 15163200]
10. Deussing J, Roth W, Saftig P, Peters C, Ploegh HL, Villadangos JA. Cathepsins B and D are dispensable for major histocompatibility complex class II-mediated antigen presentation. *Proc Natl Acad Sci U S A* 1998;95(8):4516–4521. [PubMed: 9539769]
11. Felbor U, Kessler B, Mothes W, Goebel HH, Ploegh HL, Bronson RT, Olsen BR. Neuronal loss and brain atrophy in mice lacking cathepsins B and L. *Proc Natl Acad Sci U S A* 2002;99(12):7883–7888. [PubMed: 12048238]
12. Altmann E, Cowan-Jacob SW, Missbach M. Novel purine nitrile derived inhibitors of the cysteine protease cathepsin K. *J Med Chem* 2004;47(24):5833–5836. [PubMed: 15537340]
13. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* 2004;32(Web Server issue):W665–W667. [PubMed: 15215472]
14. Li H, Robertson AD, Jensen JH. Very fast empirical prediction and rationalization of protein pKa values. *Proteins* 2005;61(4):704–721. [PubMed: 16231289]
15. Case, DADTA.; Cheatham, TE., III; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Crowley, M.; Walker, RC.; Zhang, W.; Merz, KM.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, KF.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, SR.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, DH.; Seetin, MG.; Sagui, C.; Babin, V.; Kollman, PA. AMBER 10.

16. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph* 1996;14(1):27–28. 33–38.
17. Team, RDC. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2005. <http://www.R-project.org>
18. Schüürmann G, Ebert RU, Chen J, Wang B, Kühne R. External validation and prediction employing the predictive squared correlation coefficient test set activity mean vs training set activity mean. *J Chem Inf Model* 2008;48(11):2140–2145. [PubMed: 18954136]

Figure 1. Targeting the S2 pocket to increase TbcA selectivity. Compound 4 (space-filling representation) docked to Connolly surface depictions of (a) TbcA, (b) cathepsin L, and (c) rhodesain. Polar pockets are magenta, hydrophobic pockets are green, and exposed surfaces are red.

Figure 2.

The second generation purine nitriles exploit different binding pockets in (a) TbcA, (b) cathepsin L, and (c) rhodesain. The phenyl substituted compound, **3a**, adopts the most common binding mode observed for each protease (see Supporting Information for additional models).

Figure 3.

Heatmap of EC₅₀ values for the second generation purine nitriles. Compounds are hierarchically clustered by Tanimoto chemical similarity along the x-axis, whereas the assays are hierarchically clustered along the y-axis according to activity profile. Compounds **3y** and **3z** were excluded from analysis because they incorporate an unusually large phenylsulfonyl indole substitution at the Ar position.

Figure 4.

Experimental trypanocidal activity versus trypanocidal activity predicted from the three-term linear model. The red line has unity slope and represents perfect predictions.

Scheme 1. Synthesis of purine-derived nitriles^a

^aConditions: (i) 3-bromobenzylamine, 2-butanol, 60°C, 3 to 12 h; (ii) 3-bromopropanol, K₂CO₃, DMF, 60°C, 10 to 16h; (iii) NaCN, DMSO, μ W, 180°C, 160 W, 5 to 20 min; (iv) Ar-B(OH)₂, Na₂CO₃, Pd(PPh₃)₄, 1,4-dioxane, μ W, 150°C, 300 W, 5 to 40 min.

Table 1
Aryl substituents

a		l		w
b		m		x
c		n		y
d		o		z
e		p		aa
f		q		ab
g		r		ac
h		s		ad
i		t		ae
j		u
k		v

Table 2

Protease inhibition of initial inhibitor series

Compound	Ar	IC ₅₀ vs. Thcat B (μM)	IC ₅₀ vs. Rhodesain (μM)	IC ₅₀ vs. Cat L (μM)	IC ₅₀ vs. Cat B (μM)
3a	phenyl	1.8	0.7	0.8	8.4 ± 0.8
3b	2-CF ₃ -phenyl-	0.28	8	1.5	>15
3c	3-vinyl-phenyl-	0.8	1.5	1.5	>50
3d	3-Cl-phenyl-	0.8	1.2	0.7	>50
3e	3-methoxycarbonylphenyl-	0.8	0.5	0.8	>15
3f	3-acetamidophenyl-	1.6	0.40	0.4	>15
3g	3-dimethylamino-phenyl-	1.2	0.9	1.4	>15
3h	4-vinyl-phenyl-	0.9	1.8	1.1	>50
3i	4-Cl-phenyl-	1.7	1.4	1.9	>15
3j	4-methoxycarbonylphenyl-	1.6	1.2	1.8	>50
3k	4-methoxy-phenyl-	1.2	0.8	1.1	>50
3l	3,4-Cl ₂ -phenyl-	1.2	0.9	1.6	>50
3m	3,5-Cl ₂ -phenyl-	1.2	1.1	0.9	>50
3n	3-bromo-2-isopropoxy-phenyl-	0.9	1.5	1.3	>50
3o	pentafluorophenyl-	0.9	5.2	2.7	>15
3p	2-naphthyl-	0.23	0.7	0.8	>50
3q	furan-2-yl-	1.0	0.31	0.84	0.4
3r	furan-3-yl-	0.46	0.03	0.6	0.5
3s	thiophen-2-yl-	0.6	0.8	0.9	0.2
3t	thiophen-3-yl-	0.37	0.24	0.48	0.1
3u	pyridine-3-yl-	0.6	0.26	1.6	0.3
3v	2,4-dimethoxypyrimidin-5-yl-	2.9	4	6	>15

Compound	Ar	IC ₅₀ vs. Thcat B (μM)	IC ₅₀ vs. Rhodesain (μM)	IC ₅₀ vs. Cat L (μM)	IC ₅₀ vs. Cat B (μM)
3w	indol-5-yl-	0.25 ±	0.5 ±	0.19 ±	13 ±
3x	indol-6-yl-	0.24 ±	0.26 ±	0.46 ±	13 ±
3y	1-phenylsulfonyl-indol-2-yl	>15	9 ±	1.8 ±	>50
3z	1-phenylsulfonyl-indol-3-yl	>15	2.0 ±	1.0 ±	>50

Optimization of ortho-substituted phenyl moiety

Compound	Ar	IC ₅₀ vs. Tbeat B (μM)	IC ₅₀ vs. Rhodesain (μM)	IC ₅₀ vs. Cat L (μM)	IC ₅₀ vs. Cat B (μM)					
3b	2-CF3-phenyl-	0.28	±	0.03	8	±	1.5	±	0.6	>15
3aa	2-Cl-phenyl-	1.3	±	0.1	20	±	15	±	7	>15
3ab	2-Me-phenyl-	0.39	±	0.03	>50	±	1.1	±	0.5	>15
3ac	2, 6 dimethyl phenyl-	0.27	±	0.03	>15	±	1.0	±	0.4	>15
3ad	2- <i>i</i> -pr-phenyl-	0.31	±	0.03	>15	±	0.9	±	0.3	>15
3ae	2-methoxycarbonyl-phenyl-	1.0	±	0.2	>10	±	1.4	±	0.6	>15

Table 4
Trypanocidal activity and potency against trypanosomal proteases

Compound	Ar	IC ₅₀ vs. Tbeat B (μM)	IC ₅₀ vs. Rhodesain (μM)	EC ₅₀ vs. <i>T.brucei</i> (μM)
3a	Phenyl	1.8	0.2	0.7
3aa	2-Cl-phenyl	1.3	0.1	20
3ab	2-Me-phenyl	0.39	0.03	>50
3ac	2, 6 dimethyl phenyl	0.27	0.03	>15
3ad	2- <i>ipr</i> -phenyl	0.31	0.03	>15
3ae	2-methoxycarbonyl-phenyl	1.0	0.2	>10
3b	2-CF ₃ -phenyl	0.28	0.03	8
3c	3-vinyl-phenyl	0.8	0.2	1.5
3d	2-Cl-phenyl	0.8	0.1	1.2
3e	3-methoxycarbonylphenyl	0.8	0.1	0.5
3f	3-acetamidophenyl	1.6	0.2	0.40
3g	3-dimethylamino-phenyl	1.2	0.2	0.9
3h	4-vinyl-phenyl	0.9	0.4	1.8
3i	4-Cl-phenyl	1.7	0.2	1.4
3j	4-methoxycarbonylphenyl	1.6	0.3	1.2
3k	4-methoxy-phenyl	1.2	0.9	0.8
3l	3,4-Cl ₂ -phenyl	1.2	0.2	0.9
3m	3,5-Cl ₂ -phenyl	1.2	0.1	1.1
3n	3-bromo-2-isopropoxy-phenyl	0.9	0.2	1.5
3o	pentafuorophenyl	0.9	0.3	5.2
3p	2-naphthyl	0.23	0.03	0.7
3q	furan-2-yl	1.0	0.3	0.31
3r	furan-3-yl	0.46	0.1	0.03
3s	thiophen-2-yl	0.6	0.3	0.8
3t	thiophen-3-yl	0.37	0.09	0.24
3u	pyridin-3-yl	0.6	0.2	0.26
3v	2,4-dimethoxypyrimidin-5-yl	2.9	0.7	4
3w	indol-5-yl	0.25	0.03	0.5
3x	indol-6-yl	0.24	0.09	0.26
3y	1-phenylsulfonyl-indol-2-yl	>15		9

Compound	Ar	IC ₅₀ vs. Tbeat B (μM)	IC ₅₀ vs. Rhodesain (μM)	EC ₅₀ vs. <i>T.brucei</i> (μM)	
3z	1-phenylsulfonyl-1-indol-3-yl-	>15	2.0 ±	0.4 ±	0.4