Journal of Medicinal Chemistry
Brief Article

J. Med. Chem. **2009,** *52,* 6489–6493 **6489**
DOI: 10.1021/jm900908p

Structure-Guided Development of Selective TbcatB 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, California 94143-2280,* [‡]*Department of Chemical Biology and Therapeutics, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, and* [§]*Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94143-2280*

Received June 19, 2009

The trypanosomal cathepsin TbcatB is essential for parasite survival and is an attractive therapeutic target. Herein we report the structure-guided development of TbcatB 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 TbcatB as a drug target and demonstrate that it is possible to potently and selectively inhibit TbcatB relative to trypanosomal and human homologues.

Introduction

The burden of human African trypanosomiasis (HAT[a]) 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 TbcatB, both in vitro and in vivo.[2,3]

We have previously reported that the purine nitrile scaffold produced several potent inhibitors of TbcatB and that all of those compounds were also fairly nonselective, being inhibitors of human cathepsin L.[4] 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.[5,6] While the role of cathepsin B alone remains unclear,[5] the coordinate action of cathepsin B and cathepsin L has been implicated in maintaining normal brain physiology in adults.[6] For these reasons, this suboptimal selectivity profile raised concerns as to the feasibility of developing nontoxic TbcatB inhibitors.

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 TbcatB is known to be essential,[2,3,4] the interplay between TbcatB and rhodesain inhibition in killing the parasite is not well understood.

We sought to address these questions by developing specific TbcatB inhibitors, relative to cathepsin L, cathepsin B, and rhodesain using a structure-guided approach. Herein, we report the development of a series of TbcatB inhibitors with greatly enhanced selectivity against rhodesain and human cathepsins B and L. This series includes the most potent TbcatB inhibitors to date, and these compounds were shown to kill *Trypanosoma brucei* with a high degree of selectivity relative to a panel of four mammalian cell lines (BJ, HEK 293, HEP G2, and Raji).

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

[a] Abbreviations: HAT, human African trypanosomiasis; DMF, dimethylformamide; DMSO, dimethyl sulfoxide; ACN, acetonitrile; PAMPA, parallel artificial membrane permeability assay.

Inhibitor Design

The previously reported purine nitrile inhibitors of TbcatB also inhibit rhodesain and human cathepsin L.[4] To explore possible structural explanations for the overlap between the inhibition profiles of these proteases, lead TbcatB inhibitor **4** was modeled as a covalent adduct with cysteine and docked to the previously reported homology model of TbcatB[4] 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 (TbcatB), 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.

In TbcatB, residues His179 to Gly188 form a loop oriented toward 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 disulfide 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 TbcatB, while in cathepsin L, the orientation of Tyr72 results in a much wider S3 pocket that bridges the S2 site via Leu69. Rhodesain shares structural traits from both the TbcatB 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 TbcatB.

In summary, the S2 pocket of TbcatB 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 TbcatB.

Published on Web 09/21/2009

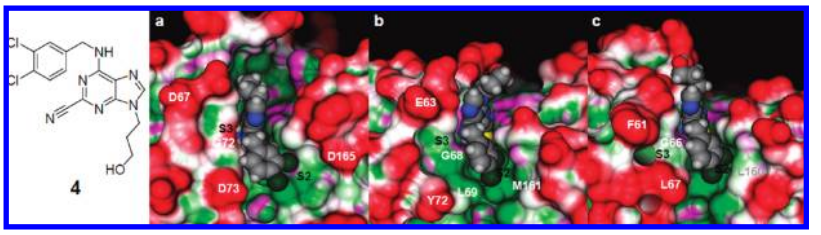
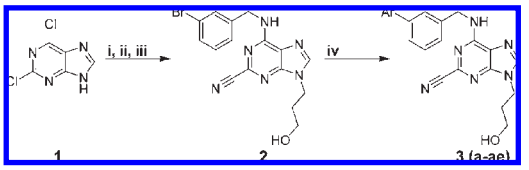


Figure 1. Targeting the S2 pocket to increase TbcatB selectivity. Compound **4** (space-filling representation) docked to Connolly surface depictions of (a) TbcatB, (b) cathepsin L, and (c) rhodesain. Polar pockets are magenta, hydrophobic pockets are green, and exposed surfaces are red.

Scheme 1. Synthesis of Purine-Derived Nitriles[a]



[a] Conditions: (i) 3-Bromobenzylamine, 2-butanol, 60 °C, 3 to 12 h; (ii) 3-bromopropanol, K_2CO_3, DMF, 60 °C, 10 to 16 h; (iii) NaCN, DMSO, μW, 180 °C, 160 W, 5−20 min; (iv) Ar−B(OH)$_2$, Na_2CO_3, Pd(PPh$_3$)$_4$, 1,4-dioxane, μW, 150 °C, 300 W, 5−40 min.

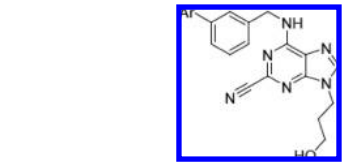
Chemistry

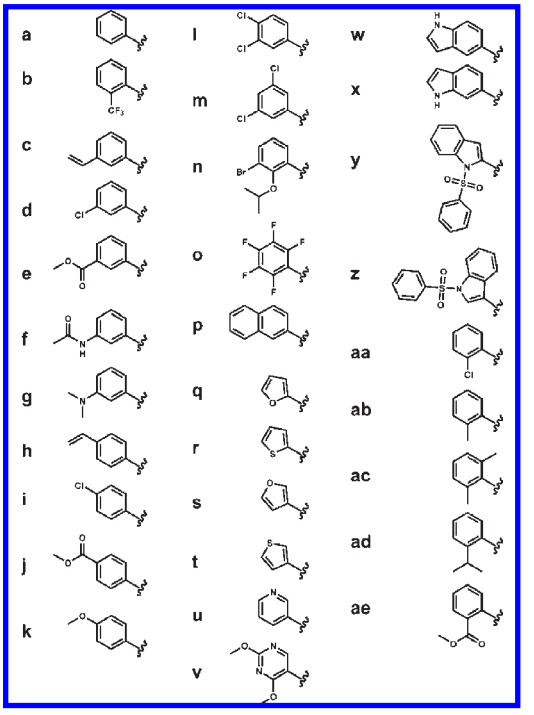
Intermediate **2** was synthesized by the general route (Scheme 1) previously described.[4,7] Briefly, **1** was reacted with 3-bromobenzylamine in 2-butanol to install the 6-amino substituent. The crude reaction was concentrated in vacuo and resuspended in dimethylformamide (DMF) with K_2CO_3. Alkylation 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 C_{18} 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**, Na_2CO_3, and Pd(PPh$_3$)$_4$ were reacted in 1,4 dioxane with microwave acceleration. The target inhibitors **3(a−ae)** were purified by preparative C_{18} chromatography with yields of 10−60%. Purity of target compounds of ≥95% was confirmed by LCMS on both C_4 and C_{18} columns.

Determining the Activity of Purine Inhibitors 3(a−ae). Inhibitors **3(a−ae)** were tested in vitro to determine inhibitory activity against TbcatB, rhodesain, cathepsin L, and cathepsin B (Table 2, 3). Compounds were also assayed to determine antiproliferative 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 TbcatB relative to parent compound **4** (EC_{50} = 1.4 μM). 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 nonexistent, especially among the most potent TbcatB inhibitors. In addition, many compounds in this series (**3c**, **3h**, **3o**, **3p**, **3v**) showed modest (2−3 fold) selectivity for TbcatB over cathepsin L and rhodesain, a marked improvement over the parent compound,

Table 1. Aryl Substitutents



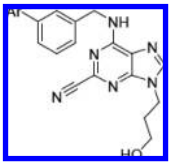


which preferentially inhibits both proteases (approximately 10-and 25-fold, respectively) over TbcatB. Compound **3b** was the most selective TbcatB inhibitor relative to both cathepsin L (∼5-fold) and rhodesain (∼25-fold). This inhibitor was one of only two compounds (**3b** and **3n**) to incorporate an ortho-substituted biphenyl group, and it was hypothesized that this substitution pattern constrains the biphenyl moiety in a conformation that gives rise to TbcatB 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 TbcatB relative to rhodesain (**3ab**, > 100-fold selective) as well as the compound most selective for TbcatB 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.

Each inhibitor was screened against a panel of four mammalian cell lines (BJ, Raji, Hep G2, and HEK 293). Almost every inhibitor displayed EC_{50} values weaker than 10−20 μM

Table 2. Protease Inhibition of Initial Inhibitor Series



compd	Ar	IC_{50} vs ID cat B (μM)	IC_{50} vs rhodesain (μM)	IC_{50} vs cat L (μM)	IC_{50} vs cat B (μM)
3a	phenyl	1.8 ± 0.2	0.7 ± 0.2	0.8 ± 0.3	8.4 ± 0.8
3b	2-CF_3-phenyl-	0.28 ± 0.03	8 ± 3	1.5 ± 0.6	> 15
3c	3-vinyl-phenyl-	0.8 ± 0.2	1.5 ± 0.1	1.5 ± 0.2	> 50
3d	3-Cl-phenyl-	0.8 ± 0.1	1.2 ± 0.4	0.7 ± 0.2	> 50
3e	3-methoxycarbonyl-phenyl-	0.8 ± 0.1	0.5 ± 0.2	0.8 ± 0.2	> 15
3f	3-acetamidophenyl-	1.6 ± 0.2	0.40 ± 0.07	0.4 ± 0.1	> 15
3g	3-dimethylamino-phenyl-	1.2 ± 0.2	0.9 ± 0.3	1.4 ± 0.4	> 15
3h	4-vinyl-phenyl-	0.9 ± 0.4	1.8 ± 0.2	1.1 ± 0.1	> 50
3i	4-Cl-phenyl-	1.7 ± 0.2	1.4 ± 0.1	1.9 ± 0.7	> 15
3j	4-methoxycarbonyl-phenyl-	1.6 ± 0.3	1.2 ± 0.3	1.8 ± 0.3	> 50
3k	4-methoxy-phenyl-	1.2 ± 0.9	0.8 ± 0.1	1.1 ± 0.3	> 50
3l	3,4-Cl_2-phenyl-	1.2 ± 0.2	0.9 ± 0.2	1.6 ± 0.3	> 50
3m	3,5-Cl_2-phenyl-	1.2 ± 0.1	1.1 ± 0.2	0.9 ± 0.2	> 50
3n	3-bromo-2-isopropoxy-phenyl-	0.9 ± 0.2	1.5 ± 0.3	1.3 ± 0.2	> 50
3o	pentaflourophenyl-	0.9 ± 0.3	5.2 ± 0.9	2.7 ± 0.5	> 15
3p	2-napthyl-	0.23 ± 0.03	0.7 ± 0.1	0.8 ± 0.1	> 50
3q	furan-2-yl-	1.0 ± 0.3	0.31 ± 0.07	0.84 ± 0.09	4.7 ± 0.4
3r	furan-3-yl-	0.46 ± 0.1	0.03 ± 0.02	0.6 ± 0.1	8.6 ± 0.5
3s	thiophen-2-yl-	0.6 ± 0.3	0.8 ± 0.3	0.9 ± 0.1	2.5 ± 0.2
3t	thiophen-3-yl-	0.37 ± 0.09	0.24 ± 0.04	0.48 ± 0.07	4.0 ± 0.1
3u	pyridin-3-yl-	0.6 ± 0.2	0.26 ± 0.07	1.6 ± 0.3	3.9 ± 0.3
3v	2,4-dimethoxypyrimidin-5-yl-	2.9 ± 0.7	4 ± 2	6 ± 1	> 15
3w	indol-5-yl-	0.25 ± 0.03	0.5 ± 0.2	0.19 ± 0.04	13 ± 3
3x	indol-6-yl-	0.24 ± 0.09	0.26 ± 0.05	0.46 ± 0.09	13 ± 2
3y	1-phenylsulfonyl-indol-2-yl	> 15	9 ± 3	1.8 ± 0.8	> 50
3z	1-phenylsulfonyl-indol-3-yl	> 15	2.0 ± 0.4	1.0 ± 0.2	> 50

Table 3. Optimization of Ortho-Substituted Phenyl Moiety

compd	Ar	IC_{50} vs TbcatB (μM)	IC_{50} vs rhodesain (μM)	IC_{50} vs Cat L (μM)	IC_{50}vs Cat B (μM)
3b	2-CF_3-phenyl-	0.28 ± 0.03	8 ± 3	1.5 ± 0.6	> 15
3aa	2-Cl-phenyl-	1.3 ± 0.1	20 ± 15	16 ± 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-ipr-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

across all four cell lines. Overall, HEK 293 was the most sensitive cell line, while Hep G2 and BJ were the least sensitive.

Several inhibitors display a therapeutic index between 25- and 50-fold across four different cell lines, further supporting the proposition that targeting TbcatB is a promising therapeutic strategy. It is interesting to note that potency against TbcatB correlates with therapeutic index, although selectivity for TbcatB relative to cathepsin L does not necessarily improve the 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. To gain insight into this unexpected activity, compounds **3a**, **3aa**, **3ab**, **3i**, and **3x** were docked against TbcatB, 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 TbcatB selective, and **3ab** is inactive against rhodesain.

As expected, the wider S2 pocket in TbcatB 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, Supporting Information).

These molecular modeling studies offer no clear explanation for the observed selectivity of the ortho-substituted compounds. Nevertheless, it is apparent that the aryl

Table 4. Trypanocidal Activity and Potency against Trypanosomal Proteases

compd	Ar	IC_{50} vs Tbcat B (μM)	IC_{50} vs rhodesain (μM)	EC_{50} vs *T. brucei* (μM)
3	phenyl	1.8 ± 0.2	0.7 ± 0.2	1.4 ± 0.2
3aa	2-Cl-phenyl-	1.3 ± 0.1	20 ± 15	4.6 ± 0.3
3ab	2-Me-phenyl-	0.39 ± 0.03	> 50	1.05 ± 0.09
3ac	2, 6 dimethyl phenyl-	0.27 ± 0.03	> 15	0.63 ± 0.08
3ad	2-ipr-phenyl-	0.31 ± 0.03	> 15	0.63 ± 0.09
3ae	2-methoxycarbonyl-phenyl-	10 ± 0.2	> 10	0.7 ± 0.1
3b	2-CF_3-phenyl-	0.28 ± 0.03	8 ± 3	0.46 ± 0.08
3c	3-vinyl-phenyl-	0.8 ± 0.2	1.5 ± 0.1	2.6 ± 0.3
3d	3-Cl-phenyl-	0.8 ± 0.1	1.2 ± 0.4	3.7 ± 0.7
3e	3-methoxycarbonyl-phenyl-	0.8 ± 0.1	0.5 ± 0.2	2.7 ± 0.4
3f	3-acetamidophenyl-	1.6 ± 0.2	0.40 ± 0.07	4.1 ± 0.5
3g	3-dimethylamino-phenyl-	1.2 ± 0.2	0.9 ± 0.3	18 ± 0.2
3h	4-vinyl-phenyl-	0.9 ± 0.4	1.8 ± 0.2	2.2 ± 0.2
3i	4-Cl-phenyl-	1.7 ± 0.2	1.4 ± 0.1	1.9 ± 0.4
3j	4-methoxycarbonyl-phenyl-	1.6 ± 0.3	1.2 ± 0.3	4.0 ± 0.6
3k	4-methoxy-phenyl-	1.2 ± 0.9	0.8 ± 0.1	2.9 ± 0.3
3l	3,4-Cl_2-phenyl-	1.2 ± 0.2	0.9 ± 0.2	2.3 ± 0.2
3m	3,5-Cl_2-phenyl-	1.2 ± 0.1	1.1 ± 0.2	3.3 ± 0.4
3n	3-bromo-2-isopropoxy-phenyl-	0.9 ± 0.2	1.5 ± 0.3	1.1 ± 0.1
3o	pentaflourophenyl-	0.9 ± 0.3	5.2 ± 0.9	4.7 ± 0.5
3p	2-napthyl-	0.23 ± 0.03	0.7 ± 0.1	1.1 ± 0.2
3q	furan-2-yl-	1.0 ± 0.3	31 ± 0.07	1.6 ± 0.3
3r	furan-3-yl-	0.46 ± 0.1	0.03 ± 0.02	0.56 ± 0.08
3s	thiophen-2-yl-	0.6 ± 0.3	0.8 ± 0.3	1.0 ± 0.1
3t	thiophen-3-yl-	0.37 ± 0.09	0.24 ± 0.04	0.43 ± 0.06
3u	pyridin-3-yl-	0.6 ± 0.2	0.26 ± 0.07	1.4 ± 0.3
3v	2,4-dimethoxy-pyrimidin-5-yl-	2.9 ± 0.7	4 ± 2	5.9 ± 0.8
3w	indol-5-yl-	0.25 ± 0.03	0.5 ± 0.2	0.38 ± 0.06
3x	indol-6-yl-	0.24 ± 0.09	0.26 ± 0.05	0.8 ± 0.1
3y	1 -phenylsulfonyl-indol-2-yl	> 15	9 ± 3	42 ± 0.05
3z	1 -phenylsulfonyl-indol-3-yl	> 15	2.0 ± 0.4	2.0 ± 0.4

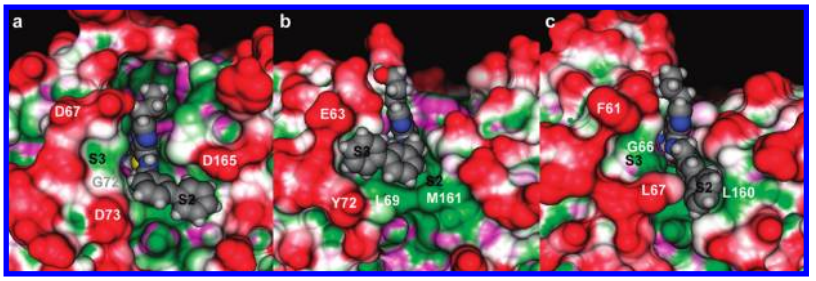


Figure 2. Second generation purine nitriles exploit different binding pockets in (a) TbcatB, (b) cathepsin L, and (c) rhodesain. The phenyl substituted compound, **3a**, adopts the most common binding mode observed for each protease (see Supporting Information for additional models).

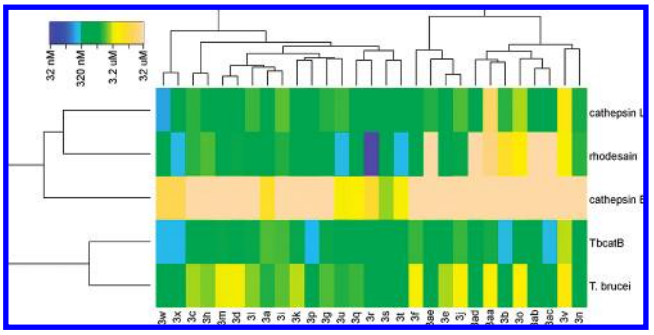


Figure 3. Heatmap of EC_{50} values for the second generation purine nitriles. Compounds are hierarchically clustered by Tanimoto chemical similarity along the *x*-axis, whereas the assays are hierarchically clustered along the *y*-axis according to activity profile. Compounds **3y** and **3z** were excluded from analysis because they incorporate an unusually large phenylsulfonyl indole substitution at the Ar position.

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 TbcatB and the S3 pocket of cathepsin L. It may also be productive to adjust the placement of the distal aryl ring by replacing the 6-amino benzyl linker.

Protease Inhibition and Trypanocidal Activity. Although the purine nitrile scaffold is thought to kill *T. brucei* by inhibition of TbcatB, the contribution of rhodesain inhibition to trypanocidal activity remains poorly understood. 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 analogues 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 TbcatB, 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* EC_{50} as a function of the following independent variables: log-normalized TbcatB and rhodesain EC_{50}, 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.[4] An iterative model building process yielded an equation with three terms (standard errors in parentheses, see Supporting Information for regression details):

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

Increasing potency (decreasing EC_{50}) 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,[13] 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 ΔRMSE is [−0.012 to −0.008] at the 95% confidence interval) to justify inclusion.

It is clear from the plot of experimental versus predicted *T. brucei* EC_{50} 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 antiparasitic 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 (Table 4).

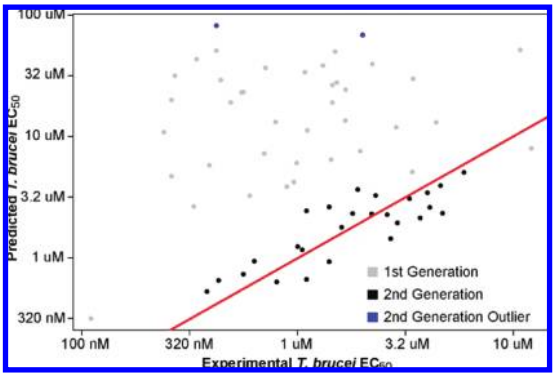


Figure 4. Experimental trypanocidal activity vs trypanocidal activity predicted from the three-term linear model. The red line has unity slope and represents perfect predictions.

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 C_{18} and YMC Pro C_4 columns (MeOH:H_2O, 1% HCOOH). Intermediate nitrile **2** was synthesized as previously described.[4]

Synthesis of 3a−3ae. Compound **2** (23 mg, 0.06 mmol), Na_2CO_3 (25 mg, 0.24 mmol, 4.0 equiv), $Pd(PPh_3)_4$ (14 mg, 0.012 mmol, 0.2 equiv), and the appropriate aryl boronic acid (0.12 mmol, 2.0 equiv) 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 min in a Biotage Initiator microwave until the reaction was complete as shown by LC/MS, a total of 20−90 min. The crude product was purified by reverse-phase HPLC (Waters XTerra preparative C_{18} column, ACN:H_2O, 1% HCOOH, linear gradient from 40 to 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.[4]

Cellular and Biochemical Assays. Inhibitors were screened for protease inhibition, trypanocidal activity, and toxicity against mammalian cell lines as previously described.[4]

Molecular Docking Studies. The previously reported TbcatB homology model,[4] 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 PDB2PQR[8] and PropKa[9] programs. All molecular dynamic simulations were performed using AMBER10 in conjunction with AMBER Tools version 1.0[10] and visualized in VMD 1.8.6 (http://www.ks.uiuc.edu/Research/vmd/).[11] Diagnostics plots and statistics were calculated using custom PERL and R scripts.[12] Default parameters and settings were used unless otherwise noted.

Acknowledgment. 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. We also thank Cindy Choy and David Smithson for their advice and feedback.

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.

References

(1) WHO. *African Trypanosomiasis* (*Sleeping Sickness*); World Health Organization: Geneva, **2006**; http://www.who.int/mediacentre/factsheets/fs259/en/.
(2) Abdulla, M. H.; O'Brien, T.; Mackey, Z. B.; Sajid, M.; Grab, D. J.; McKerrow, J. H. RNA Interference of *Trypanosoma brucei* Cathepsin B and L Affects Disease Progression in a Mouse Model. *PLoS Neglected Trop. Dis*. **2008**, *2* (9), e298.
(3) Mackey, Z. B.; O'Brien, T. C.; Greenbaum, D. C.; Blank, R. B.; McKerrow, J. H. A cathepsin B-like protease is required for host protein degradation in *Trypanosoma brucei*. *J. Biol. Chem*. **2004**, *279* (46), 48426–48433.
(4) Mallari, J. P.; Shelat, A. A.; Obrien, T.; Caffrey, C. R.; Kosinski, A.; Connelly, M.; Harbut, M.; Greenbaum, D.; McKerrow, J. H.; Guy, R. K. Development of potent purine-derived nitrile inhibitors of the trypanosomal protease TbcatB. *J. Med. Chem*. **2008**, *51* (3), 545–552.
(5) Deussing, J.; Roth, W.; Saftig, P.; Peters, C.; Ploegh, H. L.; Villadangos, J. A. 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.
(6) Felbor, U.; Kessler, B.; Mothes, W.; Goebel, H. H.; Ploegh, H. L.; Bronson, R. T.; Olsen, B. R. Neuronal loss and brain atrophy in mice lacking cathepsins B and L. *Proc. Natl. Acad. Sci. U.S.A* **2002**, *99* (12), 7883–7888.
(7) Altmann, E.; Cowan-Jacob, S. W.; Missbach, M. Novel purine nitrile derived inhibitors of the cysteine protease cathepsin K. *J. Med. Chem*. **2004**, *47* (24), 5833–5836.
(8) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson−Boltzmann electrostatics calculations. *Nucleic Acids Res*. **2004**, *32* (Web Server issue), W665–W667.
(9) Li, H.; Robertson, A. D.; Jensen, J. H. Very fast empirical prediction and rationalization of protein pKa values. *Proteins* **2005**, *61* (4), 704–721.
(10) Case, D. A. D., T.A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M; Walker,R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California: San Francisco, **2008**.
(11) Humphrey, W.; Dalke, A.; Schulten, K., VMD: visual molecular dynamics. *J. Mol. Graph*. **1996**, *14*, (1), 27−28; 33−38.
(12) R Development Core Team *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, 2005; http://www.R-project.org.
(13) Schuurmann, G.; Ebert, R. U.; Chen, J.; Wang, B.; Kuhne, 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.