

## Development of Potent Purine-Derived Nitrile Inhibitors of the Trypanosomal Protease TbcabB

Jeremy P. Mallari,<sup>†,‡</sup> Anang A. Shelat,<sup>‡</sup> Terri O'Brien,<sup>§</sup> Conor R. Caffrey,<sup>§</sup> Aaron Kosinski,<sup>‡</sup> Michele Connelly,<sup>‡</sup> Michael Harbut,<sup>||</sup> Doron Greenbaum,<sup>||</sup> James H. McKerrow,<sup>§</sup> and R. Kiplin Guy<sup>\*,†,‡</sup>

Graduate Program in Chemistry and Chemical Biology and Department of Cellular and Molecular Pharmacology, 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 Pharmacology, University of Pennsylvania, Philadelphia, Pennsylvania 19146

Received June 26, 2007

Human African trypanosomiasis (HAT), a major health concern in sub-Saharan Africa, is caused by the protozoan parasite *Trypanosoma brucei*. Recent studies have shown that a cathepsin B like protease, TbcabB, is essential to the survival of *T. brucei* in vitro (Mackey, Z. B.; O'Brien, T. C.; Greenbaum, D. C.; Blank, R. B.; McKerrow, J. H. J. Biol. Chem. 2004, 279, 48426–48433). Herein, we describe the first inhibitors of TbcabB, a series of purine nitriles. The compounds are potent trypanocides, killing the parasite with a high degree of selectivity over a panel of three human cell lines. In addition, a predictive model of trypanocidal activity was developed on the basis of potency against TbcabB and various calculated physical property descriptors.

### Introduction

It is estimated that more than 50 000 people are infected with human African trypanosomiasis (HAT<sup>a</sup>) and that 60 million more are at risk of infection.<sup>2</sup> The drugs currently in clinical use are associated with toxicity, emerging resistance, and impractical dosing regimens.<sup>3–5</sup> Effective vaccination is unlikely because of the antigenic variation of the parasite.<sup>6</sup>

One potential strategy for discovering small-molecule therapeutics is to target *T. brucei*'s cysteine proteases. Early research demonstrated that irreversible peptidyl protease inhibitors are trypanocidal,<sup>7</sup> and over the past 2 decades there has been significant progress toward validation of the parasite's cysteine protease activity as a therapeutic target.<sup>8–12</sup> Until recently, the presumed target of these inhibitors was a protease known interchangeably as brucipain, trypanopain, and rhodesain, which accounts for the majority of the parasite's total cathepsin activity.<sup>13</sup> However, recently a cathepsin B like protease known as TbcabB was shown to be essential for parasite survival in both in vitro<sup>1</sup> and in vivo models (Jim McKerrow, personal communication). Although the function of TbcabB is not well understood, RNAi studies showed that even a partial loss of TbcabB activity leads to accumulation of transferrin, enlargement of the endosome, and death of the parasite.<sup>1</sup> In contrast, RNAi knockdown of brucipain in vitro produces no phenotype, suggesting that TbcabB may be the relevant protease target.<sup>1</sup> In this report, we describe the initial testing and optimization of a purine-derived nitrile scaffold designed to target TbcabB.

Although alkyl nitriles and aryl nitriles are well established inhibitors of cysteine proteases,<sup>14–20</sup> the activity of purine-derived nitriles is relatively unexplored. *T. brucei* expresses

several transporters to take up purines and purine analogues from its host.<sup>21–24</sup> The purine nitrile scaffold was chosen for development as a trypanocide based on its established activity as a protease inhibitor<sup>25</sup> and its potential to take advantage of purine-scavenging mechanisms to improve potency and selective toxicity against the parasite. In this report, purine nitriles were synthesized and evaluated for activity against TbcabB. Additionally, specificity against human cathepsins L and B was assessed and the general toxicity of these compounds was evaluated against a panel of mammalian cell lines (HEK 293, HEP G2, and Raji). Finally, a predictive model of trypanocidal activity was developed on the basis of potency against TbcabB and various calculated physical property descriptors.

### Design

In the initial report of the scaffold, the authors described a crystal structure that indicated that the 6-amino substituent binds at the S2 site of cathepsin K.<sup>25</sup> Because of the structural similarity within the papain-like protease family, we reasoned that the scaffold might similarly orient within the active site of TbcabB. The sequence homology of the TbcabB S2 site to the S2 sites of the human cathepsins (B, C, F, H, K, L, O, S, V, W, and X) was determined. These studies indicated the S2 site of TbcabB was most similar to that of cathepsin L, largely because of the residue forming the bottom of the S2 pocket (Gly245 in TbcabB and Ala245 in cathepsin L). These studies suggested that larger hydrophobic alkyl and aromatic ligands are most likely to form productive interactions at the S2 site,<sup>26–29</sup> which is consistent with combinatorial substrate-profiling studies of TbcabB (Jim McKerrow, personal communication). The first generation of inhibitors was designed to explore this hypothesis.

### Chemistry

Purine nitriles were synthesized by the general route (Scheme 1) previously described.<sup>25</sup> Briefly, **1** was reacted with the appropriate amines (**X**) in 2-butanol to install the 6-amino substituent. The crude reaction mixture was concentrated in vacuo and resuspended in DMF with K<sub>2</sub>CO<sub>3</sub>. Alkylation at N9 was accomplished by heating this mixture with the desired alkyl bromides (**Y**) to give intermediates **2{X, Y}**. Purification was accomplished by flash chromatography, and overall yield for

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

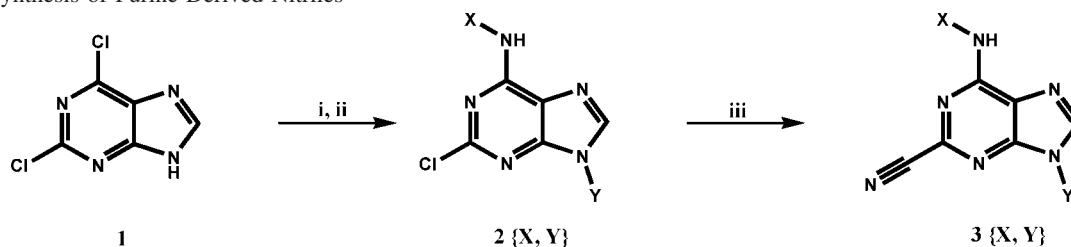
<sup>†</sup> Graduate Program in Chemistry and Chemical Biology, University of California.

<sup>‡</sup> St. Jude Children's Research Hospital.

<sup>§</sup> Department of Cellular and Molecular Pharmacology, University of California.

<sup>||</sup> University of Pennsylvania.

<sup>a</sup> Abbreviations: ANOVA, analysis of variance; DTT, dithiothreitol; FBS, fetal bovine serum; HAT, human African trypanosomiasis; HIC, hydrophobic interaction chromatography; MHC, major histocompatibility complex; TbcabB, cathepsin B like protease in *Trypanosoma brucei*.

**Scheme 1.** Synthesis of Purine-Derived Nitriles<sup>a</sup>

<sup>a</sup> Conditions: (i) X-NH<sub>2</sub>, 2-butanol, 60°C, 3–12 h; (ii) Y-Br, K<sub>2</sub>CO<sub>3</sub>, DMF, 60°C, 10–16 h; (iii) NaCN, DMSO,  $\mu$ W, 180°C, 160 W, 5–20 min.

the two reactions was 25–60%. Subsequent reaction with sodium cyanide in DMSO with microwave acceleration afforded the target nitriles **3{X, Y}**, which were purified by preparative C<sub>18</sub> chromatography with a resulting yield of 30–70%. Purity of target compounds was confirmed by LCMS on both C<sub>4</sub> and C<sub>18</sub> columns.

**Determining the Activity of Purine Inhibitors 3{X, Y}**

Chemset **3{X, Y}** was tested in vitro to determine inhibitory activity against TbcacB, cathepsin L, and cathepsin B (Table 1). All compounds were also assayed to determine antiproliferative activity against the cultured bloodstream form of *T. brucei*. Finally, cytotoxicity was evaluated in HEK 293, Raji, and HEP G2 cell lines to determine a therapeutic index. A subset of compounds was tested for time-dependent activity against TbcacB (Figure S1 in Supporting Information). These experiments indicate these inhibitors' activity is not time-dependent, suggesting that these compounds act as reversible inhibitors. This observation is further supported by activity-based probe labeling experiments (data not shown) and is consistent with previous reports indicating that nitrile-based inhibitors may display either time-dependent or -independent behavior.<sup>17,20</sup>

**Results and Discussion**

This initial series produced several compounds with moderate potency against TbcacB that inhibited parasite proliferation (Table 2). Although modestly active in the proliferation assay, compounds with a 2-methylbutyl, hexyl, or  $\alpha$ -methylbenzyl substituent at the 6-amino position were not inhibitors of TbcacB. This discrepancy may be due to off-target effects in the parasite or insolubility under the protease assay conditions. Cyclohexyl and benzyl moieties at the 6-amino position were well tolerated by the protease, and all displayed trypanocidal activity in the low micromolar range. As the most amenable to medicinal chemistry, the benzyl moiety was chosen for further optimization. A second compound series was synthesized to optimize interactions at the 6-aminobenzyl ring and to explore structural requirements at the N9 position.

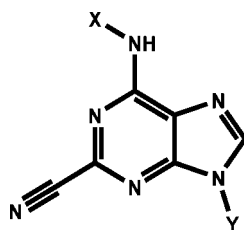
The first round of optimization produced several compounds active against TbcacB and that were trypanocidal (Table 3). Replacing the benzyl group with a phenylethyl group led to decreased potency against the protease. Various substitution patterns at the benzyl position were tolerated by the protease, with the 4-Cl and 3,4-Cl<sub>2</sub> substitutions resulting in potent compounds in both the enzyme and proliferation assays. For several compounds, the introduction of a hydroxyl group at the end of an alkyl linker at the N9 position led to significantly improved potency against TbcacB. On the basis of this observation, compound **3{12, 1}** was selected for further optimization at N9.

In this second round of optimization, various oxygen-containing moieties were introduced at the N9 position (Table 4). It was observed that a variety of functional groups (alcohols,

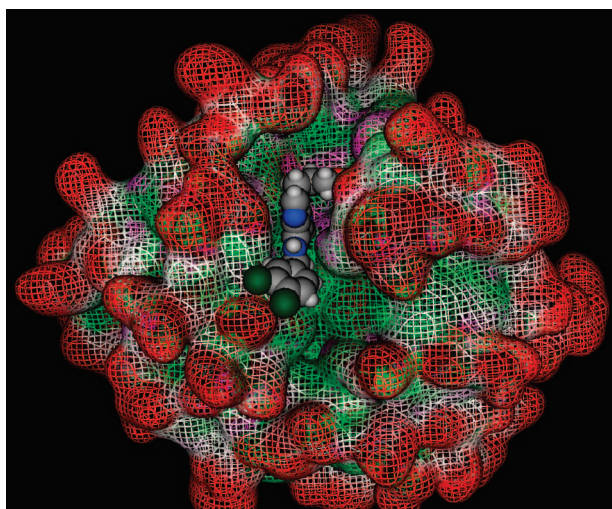
**Table 1.** Compound Numbering

Chemset #	X	Y
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		

Table 2. First-Generation Inhibitors



compd	X	Y	IC <sub>50</sub> vs TbcatB, $\mu$ M	EC <sub>50</sub> vs <i>T. brucei</i> , $\mu$ M
3{1, 13}	2-methylbutyl	cyclopentyl	14 $\pm$ 3	1.1 $\pm$ 0.2
3{1, 17}	2-methylbutyl	benzyl	5.0 $\pm$ 0.6	3.4 $\pm$ 0.2
3{2, 11}	cyclohexyl	propyl	8.3 $\pm$ 0.8	12 $\pm$ 1
3{2, 14}	cyclohexyl	cyclohexyl	8.3 $\pm$ 0.5	1.4 $\pm$ 0.2
3{2, 17}	cyclohexyl	benzyl	2.3 $\pm$ 0.2	0.33 $\pm$ 0.05
3{3, 12}	hexyl	butyl	>25	0.44 $\pm$ 0.06
3{3, 13}	hexyl	cyclopentyl	>25	1.5 $\pm$ 0.5
3{3, 15}	hexyl	cyclohexylmethyl	>25	3.6 $\pm$ 0.7
3{3, 17}	hexyl	benzyl	>25	0.7 $\pm$ 0.1
3{4, 12}	$\alpha$ -methylbenzyl	butyl	>25	1.5 $\pm$ 0.3
3{4, 13}	$\alpha$ -methylbenzyl	cyclopentyl	>25	3.4 $\pm$ 0.6
3{4, 17}	$\alpha$ -methylbenzyl	benzyl	>25	1.3 $\pm$ 0.2
3{5, 13}	4-MeO-benzyl	cyclopentyl	7 $\pm$ 1	0.8 $\pm$ 0.2
3{7, 12}	3-Me-benzyl	butyl	11 $\pm$ 2	0.27 $\pm$ 0.04
3{7, 13}	3-Me-benzyl	cyclopentyl	>25	1.1 $\pm$ 0.2
3{7, 15}	3-Me-benzyl	cyclohexylmethyl	>25	2.2 $\pm$ 0.7
3{7, 17}	3-Me-benzyl	benzyl	3.3 $\pm$ 0.7	0.48 $\pm$ 0.06
3{8, 12}	benzyl	butyl	14 $\pm$ 4	1.7 $\pm$ 0.1
3{8, 13}	benzyl	cyclopentyl	>25	1.7 $\pm$ 0.2
3{8, 17}	benzyl	benzyl	6 $\pm$ 1	2.0 $\pm$ 0.2
3{10, 12}	4-Cl-benzyl	butyl	10 $\pm$ 1	0.85 $\pm$ 0.09
3{10, 13}	4-Cl-benzyl	cyclopentyl	19 $\pm$ 4	1.2 $\pm$ 0.2
3{10, 15}	4-Cl-benzyl	cyclohexylmethyl	>25	>10
3{10, 17}	4-Cl-benzyl	benzyl	>25	1.5 $\pm$ 0.3



**Figure 1.** Lowest energy pose of 3{12, 2} (space filling representation, colored by atom type) bound to a homology model of TbcatB. Hydrophobic pockets are colored green, polar pockets are purple, and exposed surfaces are red.

ethers, diols) were tolerated by the protease and that the length of the alkyl linker could also be varied without strongly affecting activity. This strategy produced the best TbcatB inhibitors of the series, with 3{12, 2} being the most potent (270 nM, Table 4). Figure 1 shows the lowest energy pose of 3{12, 2} bound to a homology model of TbcatB (see Supporting Information for modeling details). This subset of inhibitors also yielded the three most potent trypanocides, with compound 3{12, 4} as the most active (60 nM, Table 4). These potencies are comparable with those of suramin (130–360 nM) and berenil (20–100 nM), which

are used to treat African trypanosomiasis in humans and cattle, respectively.<sup>30</sup>

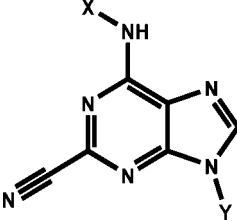
Although the most potent TbcatB inhibitors are generally the most potent trypanocides, there were discrepancies that complicated a direct correlation between IC<sub>50</sub> values in the protease and parasite proliferation assays (Figure 2). This complication was not unexpected, as antiproliferative activity is likely to be a function of multiple variables. To evaluate the relationship between inhibition of TbcatB and parasite proliferation and to determine the potential significance of other parameters, a regression model of trypanocidal activity was developed. An unbiased analysis was carried out to identify parameters displaying a statistically significant correlation with trypanocidal activity (see Supporting Information). Following an iterative model-building process, the final model contained three descriptors (standard errors in parentheses):

$$T. brucei \log_{10}(\text{IC}_{50}) = [0.158(0.030)][\text{TbcatB } \log_{10}(\text{IC}_{50})] - [0.286(0.031)][\text{SlogP\_VSA9}] - [0.202(0.040)][\text{a\_count}] - [0.159(0.028)]$$

where a\_count is the total number of atoms including hydrogens and SlogP\_VSA9 is the sum of van der Waals surface area from atoms with an SlogP parameter greater than 0.4. SlogP is an atom-based method to estimate the log(octanol/water) partition coefficient.<sup>31</sup> Atoms assigned SlogP values of >0.4 are highly hydrophobic. All calculated physical property descriptors were generated in MOE (Chemical Computing Group, version 2006.08).

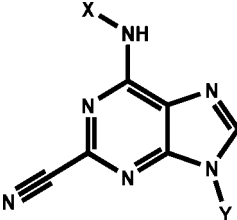
In the final model, 7 of 56 compounds were flagged as outliers (3{1, 17}, 3{2, 11}, 3{2, 17}, 3{3, 15}, 3{7, 15}, 3{8, 9}, 3{10, 15}) (Figure 3 and Supporting Information Table S1). Removal of these points followed by model optimization increased

Table 3. Second-Generation Inhibitors



compd	X	Y	IC <sub>50</sub> vs TbcatB, $\mu$ M	IC <sub>50</sub> vs <i>T. brucei</i> , $\mu$ M
3{13, 10}	phenylethyl	pyridine-3-ylmethyl	10 $\pm$ 2	4.4 $\pm$ 0.4
3{13, 12}	phenylethyl	butyl	16 $\pm$ 4	1.5 $\pm$ 0.1
3{13, 17}	phenylethyl	benzyl	>25	2.4 $\pm$ 0.2
3{5, 12}	4-MeO-benzyl	butyl	6.0 $\pm$ 0.8	0.39 $\pm$ 0.05
3{5, 17}	4-MeO-benzyl	benzyl	4.5 $\pm$ 0.6	1.0 $\pm$ 0.1
3{6, 2}	4-Me-benzyl	3-hydroxypropyl	4.2 $\pm$ 0.5	0.9 $\pm$ 0.2
3{6, 12}	4-Me-benzyl	butyl	>25	0.55 $\pm$ 0.08
3{6, 17}	4-Me-benzyl	benzyl	>25	0.56 $\pm$ 0.06
3{8, 2}	benzyl	3-hydroxypropyl	12 $\pm$ 3	2.9 $\pm$ 0.2
3{8, 9}	benzyl	tetrahydropyran-2-yl methyl	11 $\pm$ 1	0.24 $\pm$ 0.03
3{8, 16}	benzyl	4-MeO-benzyl	11 $\pm$ 2	0.40 $\pm$ 0.06
3{8, 18}	benzyl	3-CF <sub>3</sub> -benzyl	2.8 $\pm$ 0.5	0.26 $\pm$ 0.03
3{9, 3}	4-F-benzyl	4-hydroxybutyl	18 $\pm$ 2	0.8 $\pm$ 0.2
3{9, 12}	4-F-benzyl	butyl	7.1 $\pm$ 0.6	0.7 $\pm$ 0.2
3{9, 17}	4-F-benzyl	benzyl	3.3 $\pm$ 0.3	0.9 $\pm$ 0.1
3{10, 10}	4-Cl-benzyl	pyridine-3-ylmethyl	>25	0.5 $\pm$ 0.1
3{10, 2}	4-Cl-benzyl	3-hydroxypropyl	0.57 $\pm$ 0.03	0.31 $\pm$ 0.06
3{10, 3}	4-Cl-benzyl	4-hydroxybutyl	4.8 $\pm$ 0.5	0.60 $\pm$ 0.09
3{10, 16}	4-Cl-benzyl	4-MeO-benzyl	23 $\pm$ 4	0.27 $\pm$ 0.06
3{10, 18}	4-Cl-benzyl	3-CF <sub>3</sub> -benzyl	>25	0.34 $\pm$ 0.06
3{11, 12}	4-Br-benzyl	butyl	8.4 $\pm$ 0.9	0.54 $\pm$ 0.08
3{11, 17}	4-Br-benzyl	benzyl	>25	1.1 $\pm$ 0.2
3{11, 19}	4-Br-benzyl	phenylpropyl	>25	0.42 $\pm$ 0.05
3{12, 1}	3,4-Cl <sub>2</sub> -benzyl	2-hydroxyethyl	1.4 $\pm$ 0.2	0.6 $\pm$ 0.1
3{12, 12}	3,4-Cl <sub>2</sub> -benzyl	butyl	4.1 $\pm$ 0.5	0.18 $\pm$ 0.02
3{12, 17}	3,4-Cl <sub>2</sub> -benzyl	benzyl	15 $\pm$ 4	0.26 $\pm$ 0.04

Table 4. Optimization at N9



compd	X	Y	IC <sub>50</sub> vs TbcatB, $\mu$ M	IC <sub>50</sub> vs <i>T. brucei</i> , $\mu$ M
3{12, 1}	3,4-Cl <sub>2</sub> -benzyl	2-hydroxyethyl	1.4 $\pm$ 0.2	0.6 $\pm$ 0.1
3{12, 2}	3,4-Cl <sub>2</sub> -benzyl	3-hydroxypropyl	0.27 $\pm$ 0.02	0.11 $\pm$ 0.01
3{12, 5}	3,4-Cl <sub>2</sub> -benzyl	2,3-dihydroxypropyl	1.2 $\pm$ 0.2	0.45 $\pm$ 0.1
3{12, 4}	3,4-Cl <sub>2</sub> -benzyl	5-hydroxypentyl	0.93 $\pm$ 0.03	0.06 $\pm$ 0.01
3{12, 8}	3,4-Cl <sub>2</sub> -benzyl	2-(methoxymethoxy)ethyl	0.8 $\pm$ 0.1	0.32 $\pm$ 0.04
3{12, 6}	3,4-Cl <sub>2</sub> -benzyl	3-methoxypropyl	2.5 $\pm$ 0.3	0.15 $\pm$ 0.02
3{12, 7}	3,4-Cl <sub>2</sub> -benzyl	3,3-dimethoxypropyl	3.4 $\pm$ 0.3	0.22 $\pm$ 0.03

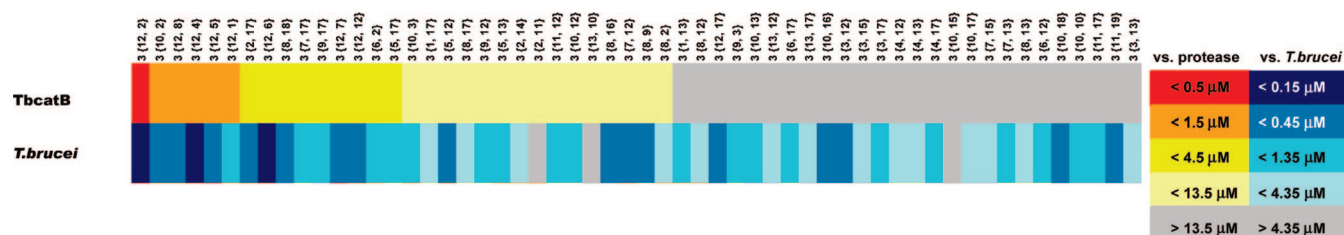
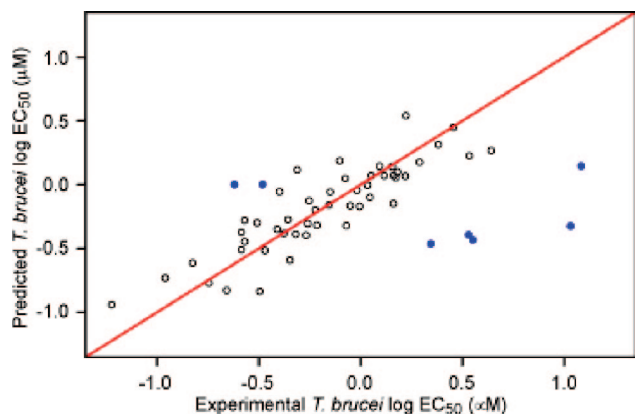


Figure 2. TbcatB inhibition and trypanocidal activity.

adjusted  $R^2$  from 0.42 to 0.75 and decreased residual standard error (rse, in log units) from 0.35 to 0.19 (df = 45). Cross-

validated root-mean-squared error (rmse, in log units) was 0.20, approximately equal to the final model rse, suggesting the





**Figure 3.** Experimental trypanocidal activity versus predicted trypanocidal activity. Outliers are plotted in blue. The red line has unity slope and represents perfect predictions.

absence of overfitting. All covariates were statistically significant by ANOVA ( $p < 10^{-5}$ ) and predictive as assessed by cross-validation.

SlogP\_VSA9 is a measure of hydrophobicity and may act as a proxy for membrane permeability. The *a\_count* descriptor is more difficult to interpret. In general, increasing the number of atoms augments molecular weight and surface area, thereby decreasing membrane permeability. On the other hand, *a\_count* includes hydrogen atoms, so the parameter would increase by substituting a heteroatom for unsaturated carbon. In light of the compounds studied in this series, it is reasonable to interpret *a\_count* as capturing another aspect of hydrophobicity. In the proposed model, then, activity against the parasite increases with both hydrophobicity and potency vs TbcAtB, suggesting that trypanocidal activity is determined by membrane permeability and activity against the protease target.

However, this does not rule out activity due to off-target effects. There are compounds in this series inactive against TbcAtB that still display trypanocidal activity. This observation is not necessarily inconsistent with the proposed model, considering that increased membrane permeability may also increase activity against other intracellular targets. We hypothesize that although the best trypanocides are generally the best TbcAtB inhibitors, the high membrane permeability of some inhibitors in this series may allow compounds with poor activity against TbcAtB to produce off-target effects in the cultured parasites. It is interesting to note that in this inhibitor series, six compounds contained a cyclohexyl moiety, and of these six, five were outliers in this model. This may suggest that the cyclohexyl ring is involved in an off-target effect or that there is something peculiar about this group in the context of the calculations used for this model. The discrepancies between the two activities may also be due to differences in intracellular inhibitor concentrations in the parasite assay. Because these inhibitors are based on a purine template, significant differences

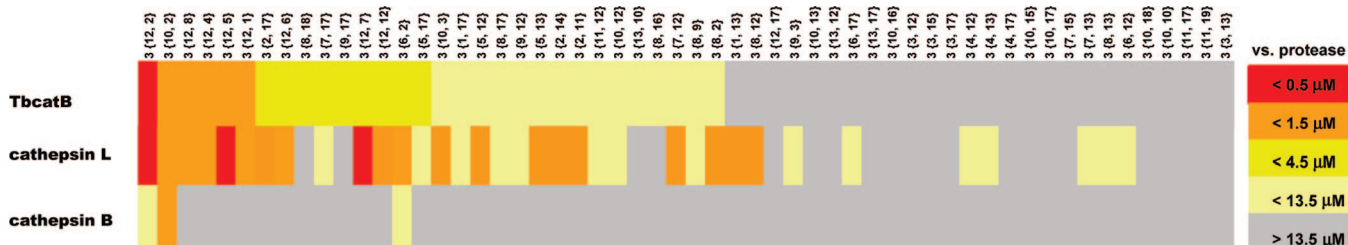
in intracellular accumulation could result from transporter mediated uptake.

To confirm that the purine nitrile scaffold is able to access and inhibit TbcAtB in an intact parasite, activity-based probe experiments were performed (data not shown). However, the reversible nature of the scaffold and the relatively fast off-rate of these compounds precluded the use of this method, which relies on competition with a nonspecific irreversible inhibitor to detect direct interaction between TbcAtB and the inhibitor *in vivo*.

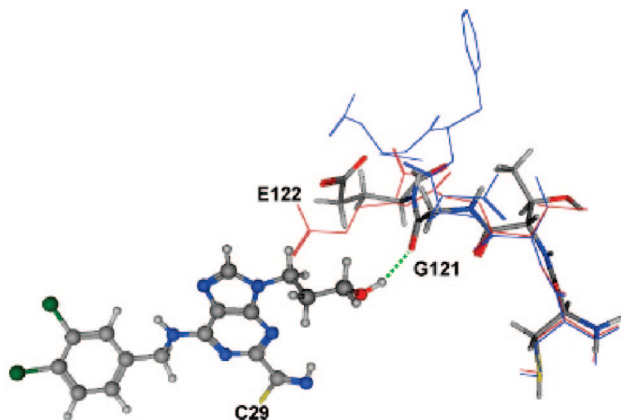
**Protease Specificity.** Cathepsins B and L are widely expressed in a variety of tissues in mammals.<sup>32</sup> Although cathepsins B and L are not vital in a mouse model, cathepsin L knockout mice display problems that include abnormal skin morphology and impaired MHC class II mediated antigen presentation.<sup>33,34</sup> The S2 site of TbcAtB is most similar to that of cathepsin L, whereas the complete sequence of TbcAtB is most homologous to that of cathepsin B. This suggests that the two proteases are likely candidates for off-target effects. To evaluate the potential utility of TbcAtB directed inhibitors, we determined the specificity of this compound series for TbcAtB relative to human cathepsins B and L.

The preference of TbcAtB's S2 binding site for large hydrophobic ligands is quite similar to those of cathepsins B and L. Therefore, it is not surprising that there was no significant selectivity between TbcAtB and cathepsin L (Figure 4). However, this compound series was generally highly selective for TbcAtB relative to human cathepsin B with only three compounds displaying IC<sub>50</sub> values below 10  $\mu$ M against the human protease (Figure 4). It is interesting to note that all three of these compounds have a 3-hydroxypropyl substituent at the N9 position. Changing the propyl linker to an ethyl or pentyl completely abolished activity against cathepsin B, as did replacement of the hydroxyl group with a methyl. This result suggests that the 3-hydroxypropyl moiety is making a specific interaction necessary for effective inhibition of cathepsin B, although TbcAtB and cathepsin L do not share this requirement.

Molecular modeling was used to further explore this hypothesis. Like cathepsin B, the prime side of TbcAtB's active site cleft is occluded by a peptide loop that is unique among members of the papain superfamily of proteases. This region adopts nearly identical conformations in the homology model of TbcAtB and two reported crystal structures of human cathepsin B except around residues 121–123 (Gly-Glu-Gly in human cathepsin B). In TbcAtB, a phenylalanine insertion and the lack of flanking glycine residues were predicted to force the main chain (Gln-Phe-Asn-Phe) away from the active site, thereby opening space around the putative location of the purine N9 position (Figure 5). The cathepsin B sequence affords multiple conformations: in 1gmy, Glu122 clashes with the N9 substituent, whereas the same residue in 1huc is directed into the solvent and the Gly121 carbonyl projects into the prime side of the active site cleft. We hypothesized that the 3-hy-



**Figure 4.** Inhibitor specificity between TbcAtB and human cathepsins B and L.



**Figure 5.** A hydrogen bond (dotted green line) between the hydroxyl moiety on the 3-hydroxypropyl substituent of **3{12, 2}** and the Gly121 carbonyl stabilizes the 1huc S' loop conformation of human cathepsin B. The oxygen–oxygen distance is 2.81 Å. The 1gmy and Tbcab conformations are depicted in brown and blue wire representations, respectively.

droxylpropyl group stabilizes the open conformation observed in 1huc by hydrogen-bonding to this carbonyl. Smaller or larger N9 substituents lacking this interaction would be unfavorable for cathepsin B binding. Indeed, a hydrogen bond between the hydroxyl moiety in **3{12, 2}** and the Gly121 carbonyl was observed in predicted structures following minimization of one conformation of the ligand within the 1huc active site, although an exhaustive enumeration of starting conformers and the role of bound waters in the active site were not explored.

Although good selectivity was observed for most inhibitors against cathepsin B, specificity against cathepsin L was not achieved in this series. However, a comparison of the S2 pockets from the homology model of Tbcab and the 1mhw structure of human cathepsin L suggests steric and electronic differences that can be exploited to achieve specificity (Figure 6). The

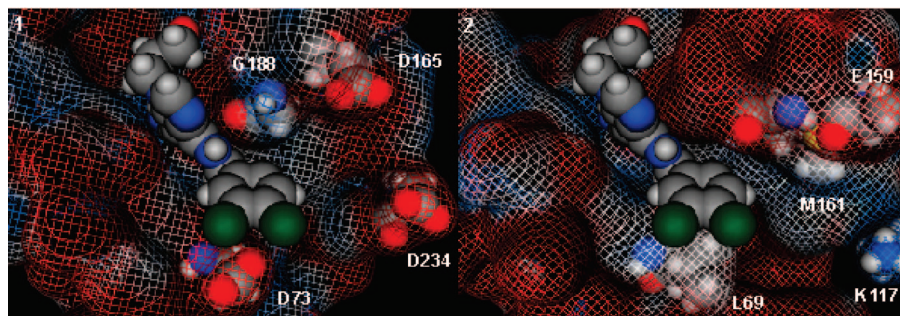
deeper S2 site of Tbcab should be able to accommodate substituents at the 6-amino position with more steric bulk and a positively ionizable functional group. This hypothesis will be explored in future studies.

It is possible that the trends in inhibition observed for Tbcab and cathepsin L may be due to differences in the inhibitors' reactivities at the nitrile warhead. However, there are several examples of compounds with a wide range of potencies in the protease assays that would be expected to be similarly electrophilic at the nitrile moiety. This suggests that binding affinity plays a primary role in determining potency, although electrophilicity cannot be ruled out as a contributing factor.

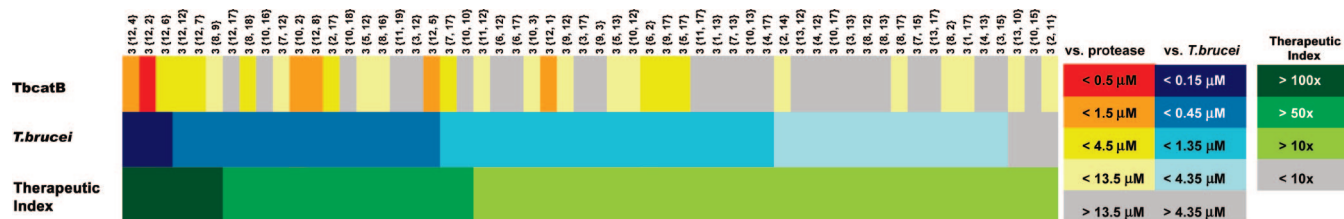
**Toxicity Studies.** General cytotoxicity of each inhibitor was evaluated by screening each at 25  $\mu\text{M}$  in cultures of Raji (a lymphoblastoid cell line derived from a Burkitt's lymphoma), HEK 293 (a human embryonic kidney cell line), and HEP G2 (a liver cell line derived from a human hepatoblastoma). Of the human cell lines, HEK 293 was by far the most sensitive. Therefore, a cellular therapeutic index for each inhibitor was determined by calculating the ratio of the  $\text{EC}_{50}$  values against HEK 293 and *T. brucei* (Figure 7). At 25  $\mu\text{M}$  the first generation of inhibitors caused some toxicity, but second-generation compounds produced almost none. No inhibitors displayed  $\text{EC}_{50}$  values below 20  $\mu\text{M}$  in HEK 293 cultures, with six compounds exhibiting greater than a 100-fold therapeutic index for *T. brucei* over any of the mammalian cell lines. These data suggest an adequate therapeutic index. It is notable that as potency against Tbcab increased, parasites were killed more selectively relative to the HEK293 line. This suggests that increasing potency against Tbcab may reduce off-target effects. Together, these data further support the proposal that targeting Tbcab is a viable strategy to develop therapies for HAT.

## Conclusions

This paper describes the development of a series of purine-derived nitriles representing the first reported inhibitors of Tbcab. Homology modeling techniques were used to guide



**Figure 6.** Electrostatic potential surfaces of the S2 pockets in Tbcab and catL reveal exploitable steric and electronic differences. (1) The lowest energy pose of **3{12, 2}** bound to the homology model of Tbcab shows the 3,4-dichlorobenzyl side chain packed against solvent exposed Asp73. The deep S2 pocket is hydrophobic in the vicinity of Gly188 and bounded by an electronegative region marked by Asp165 and Asp234. (2) The S2 pocket of catL is occluded by Met161 and tends to be more electropositive. In addition, Leu69 replaces Asp73 on the other side of the pocket. Compound **3{12, 2}** is superimposed for reference.



**Figure 7.** Activity summary sorted by potency against Tbcab. Selective toxicity is defined as  $(\text{EC}_{50} \text{ HEK 293})/(\text{EC}_{50} \text{ T. brucei})$ .



inhibitor design at the 6-amino position, and substituents at the 6-amino and N9 positions were optimized. This process yielded several inhibitors with submicromolar potency against TbcAtB and *T. brucei*. A regression model of trypanocidal activity was developed on the basis of potency against TbcAtB and various measures of hydrophobicity. Several compounds with a high degree of selectivity for TbcAtB versus human cathepsin B were discovered. Modeling offered insight into achieving specificity against cathepsin L. In addition, the general toxicity of these compounds was evaluated in three mammalian cell lines, and a high degree of selectivity against the parasite was demonstrated.

## Experimental Section

**Synthesis.** Amines and alkyl bromides were purchased from Sigma Aldrich, TCI America, and Matrix Scientific. Sodium cyanide, potassium carbonate, and 2,6-dichloropurine **1** were purchased from Sigma Aldrich. All reaction solvents were anhydrous. LCMS data were collected on a Waters system using XTerra C<sub>18</sub> columns (MeOH/H<sub>2</sub>O, 1% HCOOH).

Inhibitors were synthesized as previously described<sup>25</sup> with minor modifications. Briefly, 2,6-dichloropurine **1** (0.25 g, 1.3 mmol) was suspended in 3.9 mL of 2-butanol, followed by the dropwise addition of the appropriate amine (**X**) (3.9 mmol). The mixture was stirred for 3–12 h at 60 °C and monitored by thin layer chromatography (silica gel 60). The crude product mixture was then concentrated under vacuum and resuspended in 5 mL of anhydrous DMF. K<sub>2</sub>CO<sub>3</sub> (solid, 0.54 g, 3.9 mmol) was added, followed by dropwise addition of the appropriate alkyl bromide (**Y**) (5.2 mmol). The mixture was stirred for 12–18 h at 60 °C and monitored by LCMS. The crude mixture was concentrated under vacuum and purified by flash chromatography (30:1 DCM/MeOH v/v) using 40M columns (Biotage) on a Biotage SP1 flash chromatography instrument to give intermediates **2{X, Y}** with an overall yield of 25–60%.

An amount of 0.05 mmol of **2{X, Y}** was dissolved in 0.5 mL of DMSO in a 1 mL microwave reaction vial (Biotage). Sodium cyanide (0.35 mmol) was added. The vial was sealed and heated to 180 °C for 10 min in a Biotage Initiator microwave until the reaction was complete as shown by LCMS. The crude product was purified by reverse-phase HPLC (Waters XTerra preparative C<sub>18</sub> column, ACN/H<sub>2</sub>O, 1% HCOOH) to afford target nitrile **3{X, Y}**. The yield ranged from 30% to 70%.

**Cloning of TbcAtB.** The gene encoding the TbcAtB open reading frame was isolated from *T. brucei* cDNA and cloned into the TOPO TA vector (Invitrogen) as described.<sup>1</sup> By use of this as a PCR template, the cDNA encoding the TbcAtB zymogen was amplified, purified, and ligated into the pPICZalpha B vector (Invitrogen) for eventual methanol-induced expression in *Pichia pastoris* using methods previously described.<sup>13</sup> Cloning was unidirectional, using the following primers: forward 5'-ATACTCGAGAAAAGAG-TAAACGCCGCCCTCGTTGCT-3', containing a restriction endonuclease site for Xho I and a Kexin II cleavage site, and reverse 5'-ATAGCGGCCGCTCACGCCGTGTTGGGTGCAAGAGG-3', containing a restriction site for Not I immediately preceding a translation termination codon.

**Expression of TbcAtB in the Yeast *Pichia pastoris*.** One liter of yeast–peptone–dextrose (YPD) medium with 100 µg/mL Zeocin (Invitrogen) was inoculated with 2.0 mL of *P. pastoris* recombinant for TbcAtB. The culture was expanded at 29 °C until an A<sub>600</sub> value of 2–3 was reached. The culture was then centrifuged at 1000g for 10 min at room temperature (Eppendorf A-4-62 rotor). The resulting pellet was resuspended in 200 mL of buffered minimal methanol (BMM) induction medium (Invitrogen) containing 0.5% methanol and incubated at 29 °C for 48 h with the addition of methanol after 24 h. The medium was centrifuged at 1000g for 10 min, and the supernatant was decanted and lyophilized. The crude lyophilized solid was dissolved in HIC loading buffer (1.5 M ammonium sulfate, 50 mM K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub>, pH 7.0) and loaded onto a Tricorn 10/100 column packed with Phe15 resin

(Amersham) in an Akta Explorer Box-900 chromatography system (Amersham). HIC loading buffer was run through the column, and protein was eluted by a linear decrease in ammonium sulfate from 0.7 to 0 M over 90 min at a flow rate of 1 mL/min. Fractions were assayed for proteolytic activity as described below, and active fractions were pooled. The pooled fractions were buffer-exchanged into anion-exchange loading buffer (50 mM MES, pH 6.5) and concentrated in an Amicon centrifugal filtration device (Millipore). This solution was then loaded onto a Mono Q 5/50 GL anion exchange column (Amersham) in an Akta Explorer Box-900 chromatography system. Anion exchange loading buffer was run through the column, and protein was eluted by a linear increase in NaCl concentration from 0 to 0.5 M over 55 min at a flow rate of 1 mL/min. Fractions were assayed for proteolytic activity, and active fractions were checked for purity by SDS–PAGE and pooled. Protease concentration was determined by active site titration with CA074.

**Enzyme Inhibition Assays.** All IC<sub>50</sub> values were determined by sigmoidal curve-fitting with GraphPad Prism software.

**TbcAtB.** An amount of 100 µL of TbcAtB (final assay concentration, 35 nM) in assay buffer (0.05 M sodium acetate, pH 5.5, 4 mM DTT, 0.5 mM EDTA) was added to each well of a black flat-bottom polystyrene 96-well plate (Corning). An amount of 2 µL of inhibitor in DMSO was added, and the mixture was incubated at room temperature for 20 min before the addition of 100 µL of Z-Phe-Arg-AMC (final assay concentration, 10 µM; Bachem) substrate in assay buffer. V<sub>0</sub> was determined by monitoring fluorescence (λ<sub>ex</sub> = 350 nm, λ<sub>em</sub> = 460 nm) on an Envision plate reader (Perkin-Elmer) for 10 min. All data were collected in triplicate.

**Cathepsin L.** An amount of 12 µL of human liver cathepsin L (final assay concentration, 0.21 ng/mL; Biomol) in assay buffer (0.05 M sodium acetate, pH 5.5; 4 mM DTT; 0.01% Triton X-100; 0.5 mM EDTA) was added to each well of a black polystyrene 384-well assay plate (Corning). An amount of 240 nL of inhibitor in DMSO was added, and the mixture was incubated at room temperature for 20 min before the addition of 12 µL of Z-Phe-Arg-AMC substrate (final assay concentration, 10 µM) in assay buffer. V<sub>0</sub> was determined by monitoring fluorescence (λ<sub>ex</sub> = 350 nm, λ<sub>em</sub> = 460 nm) on an Envision plate reader (Perkin-Elmer) for 10 min. All data were collected in quadruplicate.

**Cathepsin B.** An amount of 12 µL of human liver cathepsin L (final assay concentration, 0.56 ng/mL; Biomol) in assay buffer (0.05 M sodium acetate pH 5.5; 4 mM DTT; 0.01% Triton X-100; 0.5 mM EDTA) was added to each well of a black polystyrene 384-well plate (Corning). An amount of 240 nL of inhibitor (240 nL) in DMSO was added, and the mixture was incubated at room temperature for 20 min before the addition of 12 µL of Z-Arg-Arg-AMC substrate (final assay concentration, 10 µM; Bachem) in assay buffer. V<sub>0</sub> was determined by monitoring fluorescence (λ<sub>ex</sub> = 350 nm, λ<sub>em</sub> = 460 nm) on an Envision plate reader (Perkin-Elmer) for 10 min. All data were collected in quadruplicate.

**Cell Proliferation Assays.** All IC<sub>50</sub> values were determined by curve-fitting with GraphPad Prism software.

***T. b. brucei*.** Culture-adapted *T. brucei* were grown at 37 °C in 5% CO<sub>2</sub> in HMI-9 medium (HyClone) supplemented with penicillin/streptomycin (50 units/mL), 10% heat-inactivated FBS (Omega Scientific, lot no. 3137), and 10% Serum Plus (JHR Biosciences) to a density of 1 × 10<sup>6</sup> cells/mL and then diluted to 1 × 10<sup>4</sup> cells/mL. An amount of 100 µL of the diluted culture was added to each well of a white flat-bottom sterile 96-well plate (Greiner), and 1 µL of inhibitor in DMSO was added. Plates were incubated for 48 h at 37 °C, in 5% CO<sub>2</sub>, and then equilibrated at room temperature for 1 h before the addition of 100 µL of Cell Titer Glo (Promega) to each well. Plates were then shaken on an orbital shaker for 2 min at 500 rpm. Luminescence was read after 8 min on an Envision plate reader (Perkin-Elmer). All data were collected in quadruplicate. (Note: For this inhibitor series it was noted that IC<sub>50</sub> values varied by as much as 10-fold depending on the lot of FBS used in the medium.)

**Toxicity Screens.** Raji, HEK 293, and HEP G2 cell lines were purchased from the American Type Culture Collection (ATCC, Manassas, VA) and were cultured according to recommendations. Cell culture media were purchased from ATCC. Cells were routinely tested for mycoplasma contamination using the MycoAlert mycoplasma detection kit (Cambrex). Exponentially growing cells were plated in Costar 96-well custom assay plates, which were black with clear bottoms (Corning), and incubated overnight at 37 °C in a humidified 10% CO<sub>2</sub> incubator. DMSO inhibitor stock solutions were added the following day to a final concentration of 25 μM, 0.5% DMSO. Cytotoxicity was determined following a 72 h incubation using the Alamar blue assay. Fluorescence ( $\lambda_{\text{ex}} = 510$  nm,  $\lambda_{\text{em}} = 590$  nm) was measured on an Envision plate reader (Perkin-Elmer) 4 h after the addition of Alamar blue.

**Acknowledgment.** This work was supported by the American Lebanese Syrian Associated Charities (ALSAC) and 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 TbcAtB TOPO TA vector was a kind gift from Zachary Mackey, and we thank him for all of his additional help during the *T. brucei* studies. We also thank Joey Hansell for her assistance with the TbcAtB assays.

**Supporting Information Available:** Spectroscopic data, LCMS data, enzyme kinetic data, complete protease assay data for all listed compounds, and experimental details for modeling studies. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- Mackey, Z. B.; O'Brien, T. C.; Greenbaum, D. C.; Blank, R. B.; McKerrow, J. H. *J. Biol. Chem.* **2004**, *279*, 48426–48433.
- Sleeping Sickness*; WHO: Geneva, 2007; <http://www.who.int/mediacentre/factsheets/fs259/en/>.
- Sleeping Sickness Treatment Schedule*; WHO: Geneva, 2007. [whqlibdoc.who.int/hq/2001/WHO\\_CDS\\_CSR\\_EPH\\_2001.3\\_fre.pdf](http://whqlibdoc.who.int/hq/2001/WHO_CDS_CSR_EPH_2001.3_fre.pdf).
- Brun, R.; Schumacher, R.; Schmid, C.; Kunz, C.; Burri, C. *Trop. Med. Int. Health* **2001**, *6*, 906–914.
- Legros, D.; Evans, S.; Maiso, F.; Enyaru, J. C.; Mbulamberi, D. *Trans. R. Soc. Trop. Med. Hyg.* **1999**, *93*, 439–442.
- Vickerman, K. *Nature* **1978**, *273*, 613–617.
- Ashall, F.; Angliker, H.; Shaw, E. *Biochem. Biophys. Res. Commun.* **1990**, *170*, 923–929.
- Scory, S.; Stierhof, Y. D.; Caffrey, C. R.; Steverding, D. *Kinetoplastid Biol. Dis.* **2007**, *6*, 2.
- Nkemgu, N. J.; Grande, R.; Hansell, E.; McKerrow, J. H.; Caffrey, C. R.; Steverding, D. *Int. J. Antimicrob. Agents* **2003**, *22*, 155–159.
- Scory, S.; Caffrey, C. R.; Stierhof, Y. D.; Ruppel, A.; Steverding, D. *Exp. Parasitol.* **1999**, *91*, 327–333.
- Troeberg, L.; Morty, R. E.; Pike, R. N.; Lonsdale-Eccles, J. D.; Palmer, J. T.; McKerrow, J. H.; Coetzer, T. H. *Exp. Parasitol.* **1999**, *91*, 349–355.
- Fujii, N.; Mallari, J. P.; Hansell, E. J.; Mackey, Z.; Doyle, P.; Zhou, Y. M.; Gut, J.; Rosenthal, P. J.; McKerrow, J. H.; Guy, R. K. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 121–123.
- Caffrey, C. R.; Hansell, E.; Lucas, K. D.; Brinen, L. S.; Alvarez Hernandez, A.; Cheng, J.; Gwaltney, S. L., 2nd; Roush, W. R.; Stierhof, Y. D.; Bogyo, M.; Steverding, D.; McKerrow, J. H. *Mol. Biochem. Parasitol.* **2001**, *118*, 61–73.
- Hanzlik, R. P.; Zygmunt, J.; Moon, J. B. *Biochim. Biophys. Acta* **1990**, *1035*, 62–70.
- Greenspan, P. D.; Clark, K. L.; Tommasi, R. A.; Cowen, S. D.; McQuire, L. W.; Farley, D. L.; van Duzer, J. H.; Goldberg, R. L.; Zhou, H.; Du, Z.; Fitt, J. J.; Coppa, D. E.; Fang, Z.; Macchia, W.; Zhu, L.; Capparelli, M. P.; Goldstein, R.; Wigg, A. M.; Doughty, J. R.; Bohacek, R. S.; Knap, A. K. *J. Med. Chem.* **2001**, *44*, 4524–4534.
- Greenspan, P. D.; Clark, K. L.; Cowen, S. D.; McQuire, L. W.; Tommasi, R. A.; Farley, D. L.; Quadros, E.; Coppa, D. E.; Du, Z.; Fang, Z.; Zhou, H.; Doughty, J.; Toscano, K. T.; Wigg, A. M.; Zhou, S. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 4121–4124.
- Robichaud, J.; Oballa, R.; Prasad, P.; Falgueyret, J. P.; Percival, M. D.; Wesolowski, G.; Rodan, S. B.; Kimmel, D.; Johnson, C.; Bryant, C.; Venkatraman, S.; Setti, E.; Mendonca, R.; Palmer, J. T. *J. Med. Chem.* **2003**, *46*, 3709–3727.
- Palmer, J. T.; Bryant, C.; Wang, D. X.; Davis, D. E.; Setti, E. L.; Rydzewski, R. M.; Venkatraman, S.; Tian, Z. Q.; Burrill, L. C.; Mendonca, R. V.; Springman, E.; McCarter, J.; Chung, T.; Cheung, H.; Janc, J. W.; McGrath, M.; Somoza, J. R.; Enriquez, P.; Yu, Z. W.; Strickley, R. M.; Liu, L.; Venuti, M. C.; Percival, M. D.; Falgueyret, J. P.; Prasad, P.; Oballa, R.; Riendeau, D.; Young, R. N.; Wesolowski, G.; Rodan, S. B.; Johnson, C.; Kimmel, D. B.; Rodan, G. *J. Med. Chem.* **2005**, *48*, 7520–7534.
- Altmann, E.; Aichholz, R.; Betschart, C.; Buhl, T.; Green, J.; Lattmann, R.; Missbach, M. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2549–2554.
- Loser, R.; Schilling, K.; Dimmig, E.; Gutschow, M. *J. Med. Chem.* **2005**, *48*, 7688–7707.
- Carter, N. S.; Fairlamb, A. H. *Nature* **1993**, *361*, 173–176.
- Barrett, M. P.; Zhang, Z. Q.; Denise, H.; Giroud, C.; Baltz, T. *Mol. Biochem. Parasitol.* **1995**, *73*, 223–229.
- Maser, P.; Sutterlin, C.; Kralli, A.; Kaminsky, R. *Science* **1999**, *285*, 242–244.
- De Koning, H. P. *Mol. Pharmacol.* **2001**, *59*, 586–592.
- Altmann, E.; Cowan-Jacob, S. W.; Missbach, M. *J. Med. Chem.* **2004**, *47*, 5833–5836.
- Maciewicz, R. A.; Etherington, D. J. *Biochem. J.* **1988**, *256*, 433–440.
- Choe, Y.; Leonetti, F.; Greenbaum, D. C.; Lecaille, F.; Bogyo, M.; Bromme, D.; Ellman, J. A.; Craik, C. S. *J. Biol. Chem.* **2006**, *281*, 12824–12832.
- Gosalia, D. N.; Salisbury, C. M.; Ellman, J. A.; Diamond, S. L. *Mol. Cell. Proteomics* **2005**, *4*, 626–636.
- Puzer, L.; Cotrin, S. S.; Alves, M. F.; Egborge, T.; Araujo, M. S.; Juliano, M. A.; Juliano, L.; Bromme, D.; Carmona, A. K. *Arch. Biochem. Biophys.* **2004**, *430*, 274–283.
- Merschjohann, K.; Sporer, F.; Steverding, D.; Wink, M. *Planta Med.* **2001**, *67*, 623–627.
- Wildman, S. A.; Crippen, G. M. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 868–873.
- Xing, R.; Addington, A. K.; Mason, R. W. *Biochem. J.* **1998**, *332* (Part 2), 499–505.
- Deussing, J.; Roth, W.; Saftig, P.; Peters, C.; Ploegh, H. L.; Villadangos, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4516–4521.
- Nakagawa, T.; Roth, W.; Wong, P.; Nelson, A.; Farr, A.; Deussing, J.; Villadangos, J. A.; Ploegh, H.; Peters, C.; Rudensky, A. Y. *Science* **1998**, *280*, 450–453.

JM070760L