

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Cysteine-reactive chemical probes: tools to identify and validate potential therapeutic targets

Permalink

<https://escholarship.org/uc/item/4ck666fr>

Author

Choy, Jonathan Wymond

Publication Date

2013

Peer reviewed|Thesis/dissertation

Cysteine-reactive chemical probes: tools to identify and validate
potential therapeutic targets

by

Jonathan Wymond Choy

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry and Chemical Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Copyright 2013
by
Jonathan Wymond Choy

Dedicated to my family, friends, and teachers

For their support, encouragement, and guidance.

Acknowledgements

Many people have contributed to the work described in this dissertation and to my success and development as a scientist.

First, I would like to acknowledge my advisors Jack Taunton and Adam Renslo for providing me the opportunity to work on these exciting projects. I thank them for their support, tutelage, and patience as we navigated the often frustrating, but fascinating complexities of *T. brucei* kinase biology. I would also like to thank my other thesis committee members, Kevan Shokat and Danica Fujimori, for their valuable insight and support.

Next, I would like to acknowledge all my collaborators whose contributions were essential to the success and breadth of these projects.

Nathan Gushwa, Mari Nishino, Kyriacos Koupparis, Juan Oses-Prieto,
Daniel Santi, Arvin Dar, Jiri Gut, Clifford Bryant, Punitha Vedantham,
Mindy Davis, Min Shen

In particular, I would like to thank Nathan Gushwa for his mentorship and friendship.

I would also like to thank my other colleagues in the SMDC, the Taunton Lab, the Wells Lab, and others from UCSF who provided guidance, support, and companionship.

Sumit Mahajan, Jesse McFarland, Charlie Morgan, Jason Porter, Juan
Diaz, Shyam Krishnan, Marja Tarr, Beatrice Wang, Brian Hearn, Priya
Janikar, Erica Lauterwasser, Shaun Fontaine, Sam Pfaff, Julie Zorn, James
Wells, Ingrid Chen, Mark Burlingame, Michelle Arkin, Valerie Ohman, Ville
Paavilainen, Julia Rumpf, Rand Miller, Rebecca Maglathlin, Iana
Serafimova, Amada Silber, Sumit Prakash, Jeff Henise, Robin Agiletti,
Cheryl Tajon, Christine Olson, Julia Molla, Nicole Flowers

Finally, I want to thank all my family and friends for their love, support, and motivation that made this all possible.

My parents, Wymond and Terry; my brother and sister, Philip and
Christina; my aunt and uncle, Yvonne and Javier; my cousins, Christopher
and Nicholas; my grandparents and extended family; the Lee family; and
my friends, Michael Sun, Christine Sutu, Jason Chan, and the Gilliberts

Part of this thesis is a reproduction of material published previously or in the publication process.

Chapter 1 is reproduced in part with permission from: Nishino, M., Choy, J.W., Gushwa, N.N., Koupparis, K., Oses-Prieto, J.A., Burlingame, A.L., Renslo, A.R., McKerrow, J.H., and Taunton, J. Therapeutic targets in *Trypanosoma brucei* revealed by hypothemycin-based chemoproteomics and RNAi. Manuscript submitted.

Chapter 3 is reproduced in part with permission from: Choy, J.W., Bryant, C., Calvet, C.M., Doyle, P.S., Gunatilleke, S.S., Leung, S.S.F., Ang, K.K.H., Chen, S., Gut, J., Oses-Prieto, J.A., Johnston, J.B., Arkin, M.R., Burlingame, A.L., Taunton, J., Jacobson, M.P., McKerrow, J.M., Podust, L.M., Renslo, A.R. Chemical-biological characterization of a cruzain inhibitor reveals a second target and a mammalian off-target. *Beilstein J. Org. Chem* **9**, 15–25 (2013).

Abstract

Activity-based protein profiling with chemical probes derived from cysteine-reactive small molecule inhibitors provides a mechanism for prioritizing target lists for further studies and lead optimization. Thousands of proteins remain uncharacterized for the parasitic organisms that are responsible for a variety of tropical diseases. This makes target selection for the development of new, safer therapeutic agents difficult. The current deficit in the understanding of parasitic protein function has contributed to the slow progress toward the eradication of these diseases. Chemoproteomic profiling of cysteine-reactive small molecules with anti-parasitic activity can help identify and validate potential drug targets, while also providing useful tools to study the biological roles of these proteins in cells.

Protein kinases have emerged as potential therapeutic targets for many diseases, including African trypanosomiasis caused by *Trypanosoma brucei*. However, identifying relevant kinase drug targets is difficult because most of the 182 *T. brucei* protein kinases are still uncharacterized. Hypothemycin, an electrophilic kinase inhibitor shown to react selectively with kinases bearing a “CDXG” motif, was found to potently kill *T. brucei* parasites in tissue culture and in infected mice. Using quantitative chemoproteomics with propargyl-hypothemycin, eight CDXG kinases inhibited by hypothemycin were identified; four of these exhibited greater than 50% inhibition by 200 nM hypothemycin *in vitro*. Of those four, only *TbCLK1* was both essential by RNAi and inhibited in the cells at concentrations of hypothemycin sufficient to kill the parasite in tissue culture. We

later found that expression of an allele of *TbCLK1* resistant to hypothemycin was insufficient to rescue the parasites in a cellular proliferation assay. This result suggests that the inhibition of other cellular targets may contribute to hypothemycin's trypanocidal activity, but does not diminish the value of *TbCLK1* as a potential therapeutic target.

In a continued effort to validate *TbCLK1* as a therapeutic target and to study its role in cells, we hypothesized a chemical genetics (pharmacological) approach would be most appropriate. To generate selective inhibitors of *TbCLK1*, a second non-conserved active site cysteine, found in only two *T. brucei* kinases (*TbCLK1/2*), was exploited for the development of covalent inhibitors. Our most promising series of electrophilic compounds, based on AD57, are potent ($IC_{50} < 200$ nM), allele-selective (> 50 -fold) inhibitors of wild-type *TbCLK1* *in vitro* with potent ($EC_{50} < 1$ μ M) anti-proliferative activity. These molecules, when used in combination with cells expressing compound-resistant *TbCLK1*, have the potential to be powerful tools to elucidate the cellular functions of *TbCLK1* kinase activity.

Activity-based protein profiling was also used to identify targets of the cysteine protease inhibitor K777 and related analogs. These molecules were unexpectedly found to inhibit both cruzain and *TcCYP51*, a cysteine protease and a 14- α -demethylase enzyme, respectively, in *Trypanosoma cruzi*. Specifically, we designed and synthesized a chemical probe, propargyl-K777, with the aim of characterizing the cysteine protease related effects of K777 and the putative dual-targeted analogs. Attempts to label and identify

parasite derived targets of propargyl-K777 proved challenging due to competitive reaction with an abundant host-cell protein. Chemoproteomics using propargyl-K777 identified this host cell off-target as cathepsin B. The probe was also used to show that some of our non-basic analogs do not significantly inhibit cathepsin B at anti-trypanosomal concentrations. Thus, this new probe can serve as a tool to assay future compounds for inhibition of cathepsin B in a cellular context in order to minimize this off-target activity.

Table of Contents

Introduction: Cysteine-reactive chemical probes: tools to identify and validate potential therapeutic targets	1
References	7
Chapter 1: Quantitative chemical proteomics with a “clickable” hypothemycin probe reveals new therapeutic targets in <i>Trypanosoma brucei</i>	
Abstract	9
Introduction	10
Results	13
Discussion	28
Methods	32
References	41
Chapter 2: Development of chemical tools to study <i>Tb</i>CLK1	
Abstract	49
Introduction	50
Section 1: Testing the reactivity of <i>Tb</i>CLK1 C215	54
Section 2: Identification of AD57 and development of electrophilic analogs targeting C215 in <i>Tb</i>CLK1	60
Discussion	73
Methods	78
References	127

**Chapter 3: A Chemical-Biological Characterization of a Cruzain Inhibitor
Reveals a Second Target and a Mammalian Off-Target.**

Abstract	131
Introduction	132
Results and Discussion	136
Conclusion	153
References	155
Appendix 1: Supplemental information for Chapter 1	160
Appendix 2: Supplemental information for Chapter 2	171
Appendix 3: Supplemental information for Chapter 3	179

List of Tables

Table 3.1: Activity of anilinopyrimidines against <i>Tb</i> CLK1 <i>in vitro</i>	58
Table 3.2: Activity of acrylamide analogs of AD57 against <i>Tb</i> CLK1 <i>in vitro</i>	68
Table 3.3: Activity of analogs of 3.41 against <i>Tb</i> CLK1 <i>in vitro</i>	73
Table 4.1: <i>In vitro</i> biochemical and cellular activities of test compounds and controls	137
Table 4.2: <i>In vitro</i> inhibition of important mammalian CYP enzymes	141
Table 4.3: Mechanism of inhibition studies for compounds 4.1 and 4.4	142
Appendix 1 Table 1: iTRAQ mass spectrometry results	160
Appendix 1 Table 2: Summary of <i>T. brucei</i> CDXG kinases and RNAi results	167
Appendix 1 Table 3: Primers used in this work described in Chapter 2	169
Appendix 2 Table 1: Invitrogen SelectScreen Kinase Profiling Results at 1 μ M AD57 (3.25)	177
Appendix 2 Table 2: <i>Tb</i> CLK modified cell lines	178
Appendix 3 Table 1: Rank ordering of experimentally and computationally determined binding affinities for <i>Tc</i> CYP51	179

List of Figures and Schemes

Equation 1.1: Irreversible binding equation	4
Figure 1.1: Schematic for Activity-Based Protein Profiling	5
Figure 1.2: Chemical genetics	6
Figure 2.1: Potent trypanocidal activity of hypothemycin	14
Figure 2.2: Design and validation of a hypothemycin-based affinity probe	16
Figure 2.3: Quantification of hypothemycin binding to CDXG kinases	19
Figure 2.4: Functional analysis of CDXG kinases by RNAi	21
Figure 2.5: Recombinant <i>Tb</i> CLK1 is highly phosphorylated	23
Figure 2.6: Preferential inhibition of <i>Tb</i> CLK1 by hypothemycin	24
Figure 2.7: <i>In vitro</i> kinase assay with recombinant <i>Tb</i> CLK1	26
Figure 2.8: The catalytic activity of C288A <i>Tb</i> CLK1 is unaffected in cells treated with hypothemycin	27
Figure 2.9: Expression of C288A <i>Tb</i> CLK1 is insufficient to rescue trypanosomes from hypothemycin treatment	28
Figure 3.1: Sequence alignment of <i>Tb</i> CLK1 and human CLK1	50
Figure 3.2: Features of the <i>Tb</i> CLK1 active site revealed by bioinformatics	53
Figure 3.3: PCI-32765 inhibits <i>Tb</i> CLK1 in a C215 dependent manner	55
Figure 3.4: Electrophilic analogs of PP2 inhibit <i>Tb</i> CLK1 in a C215 dependent manner	56
Figure 3.5: Chemical structures of anilinopyrimidine inhibitors	57
Figure 3.6: <i>T. brucei</i> cell proliferation assay with compounds 3.5 and 3.6	59
Figure 3.7: <i>Tb</i> CLK1 screening hits from the NCGC	62
Figure 3.8: <i>Tb</i> CLK1 screening hits from UCSF	63
Figure 3.9: Inhibition of <i>Tb</i> CLK1 by analogs of AD57	64
Figure 3.10: The trypanocidal effects of AD57 analogs mirror <i>in vitro</i> inhibition of <i>Tb</i> CLK1 after 48h of treatment	65
Figure 3.11: Chemical structures of electrophilic AD57 analogs with a variety of two-carbon linkers	67

Figure 3.12: Inhibition of <i>Tb</i> CLK1 autophosphorylation	70
Figure 3.13: <i>T. brucei</i> cell proliferation assay with 3.41 and 3.43	71
Figure 3.14: Second generation electrophilic analogs targeting C215 in <i>Tb</i> CLK1	72
Figure 4.1: Chemical structures of cruzain inhibitors	133
Figure 4.2: Computational docking model of inhibitors bound to <i>Tc</i> CYP51	139
Figure 4.3: Synthesis of the reduced control compounds for cruzain	143
Figure 4.4: GC-MS analysis of lipid extracts from <i>T. cruzi</i> parasites treated with test compounds	146
Figure 4.5: Chemical structures of compound 4.9 and ‘clickable’ activity-based probe azido-tags	147
Figure 4.6: Competitive labeling of host cell (C2C12) proteins using 4.9	149
Figure 4.7: MS/MS spectrum of the tryptic peptide S264-R281 from mouse cathepsin B	150
Appendix 2 Figure 1: NCGC HTS results for <i>Tb</i> CLK1 - 4-2011	171
Appendix 2 Figure 2: NCGC HTS results for <i>Tb</i> CLK1 - 4-2012	172
Appendix 2 Figure 3: In-house <i>Tb</i> CLK1 screen	173
Appendix 2 Figure 4: Growth curves comparing unmodified cells and <i>Tb</i> CLK2 KO cells	175
Appendix 2 Figure 5: Growth curves comparing <i>Tb</i> CLK1 modified cell lines	176
Appendix 3 Figure 1: GC/MS spectra showing sterol composition of intracellular <i>T. cruzi</i> parasites treated with test compounds	180
Appendix 3 Figure 2: Time-course of reaction between compounds 4.1 or 4.6 with 10 mM glutathione	181
Appendix 3 Scheme 1: Synthesis of compound 4.4	182
Appendix 3 Scheme 2: Synthesis of activity-based probe 4.9	182
Appendix 3 Scheme 3: Synthesis of cleavable biotin azide reagent 4.11	182
Appendix 3 Scheme 4: Synthesis of analogs 4.12 and 4.13	183

Introduction:

Cysteine-reactive chemical probes: tools to identify and validate potential therapeutic targets.

A component of Chemical Biology is the design and application of chemical probes to elucidate biological pathways. In this thesis, we focus on the use of covalent active site inhibitors as chemical tools to identify and validate targets that may aid in the development of new, safer therapies for two neglected tropical diseases.

African trypanosomiasis and Chagas disease require safer, more effective therapeutics.

African trypanosomiasis (sleeping sickness), a disease endemic to sub-Saharan Africa, is caused by two subspecies of *Trypanosoma brucei*. There are approximately 10 thousand cases of infection reported per year.¹ The disease is characterized by an initial hemolymphatic stage, which is followed by a neurological stage of infection that, if left untreated, is fatal. There are currently four drugs (pentamidine, suramine, melarsoprol, and eflornithine) approved for treating the disease, all of which are not orally bioavailable. When the disease has progressed to the neurological stage, only melarsoprol and eflornithine are effective for treating the disease. Melarsoprol, an organoarsine, results in lethal side-effects in 5-10% of patients.¹ Eflornithine, which is not lethally toxic, requires a strict and difficult dosing regimen and is ineffective against the more aggressive form of the disease caused by the *rhodesiense* sub-species.

A related parasitic disease, Chagas disease, is caused by *Trypanosoma cruzi*. Approximately 10 million people in Central and South America are living with Chagas disease, which kills about 10 thousand of those infected annually.² This trypanosome

disease is also characterized by two stages of progression: acute infection of the circulatory system and chronic infection of the heart and digestive muscles. The chronic stage can last decades and usually results in heart failure. There are currently two drugs (benznidazole and nifurtimox) approved for treating Chagas disease. Both are effective curative-agents in the acute phase of the disease, but their efficacy declines significantly with disease progression.² Because many cases go undetected in the acute phase, clinical failure rates for the drug are quite high.

Both of these diseases will benefit from safer, more effective, orally bioavailable drugs. However, most of proteins in these parasites remain uncharacterized, making the selection of targets a first barrier for drug development. Here, we focus primarily on protein kinases.

Protein kinases as drug targets.

Protein phosphorylation is critical to cellular signaling in all organisms, including trypanosomes. Of the 182 protein kinases in *Trypanosoma brucei*, most remain uncharacterized.³ Identifying those kinases that regulate essential pathways and demonstrating that inhibition with small molecules reduces parasite viability could yield novel therapeutic targets. However, the development of potent, selective inhibitors as potential drugs and chemical tools can be challenging. Protein kinases catalyze the transfer of the terminal phosphate of ATP to a serine, threonine, or tyrosine on their substrates. As a result, their active sites contain a highly conserved hydrophobic pocket

designed to accommodate ATP. This pocket is also where most kinase inhibitors bind. Unfortunately, the conserved nature of this pocket also presents two major challenges for inhibitor development: selectivity in the presence of hundreds of other ATP binding proteins and potency while competing with high intracellular concentrations of ATP. Despite these challenges, protein kinases remain a promising class of drug targets that has yielded many orally bioavailable chemotherapeutics.⁴

Covalent active site inhibitors (irreversible inhibitors) as drugs

Covalent inhibitors address some of the issues that make potent and selective inhibition of kinases challenging. These inhibitors incorporate electrophiles that react with nucleophilic residues in the active site, typically non-conserved cysteines. Covalent binding of these molecules to a kinase is defined by (Equation 1.1), which includes a reversible (thermodynamic) binding constant (K_i) as well as a kinetic reaction constant for covalent bond formation (k_{inact}). By deriving inhibitory activity from two aspects (reversible binding affinity and covalent bond formation), electrophilic inhibitors can be developed to potently and selectively modulate the activity of kinases that contain an specific active site cysteine or even specific kinases within such a class.



Equation 1.1: Irreversible binding equation.

In the context of the cellular environment, the activity of irreversibly inhibited kinases can be restored only via protein synthesis, and thus the effect of inhibition is more likely to be sustained over time.⁵ However, the potential to irreversibly inhibit off-targets

makes selectivity a crucial issue for covalent inhibitors, and is a reason why reversible inhibitors are often preferred as therapeutics.

Covalent active site inhibitors as chemical tools

Covalent inhibitors make exceptional chemical tools for target identification. activity-based protein profiling (ABPP) is a proteomic strategy that uses covalent active site inhibitors with a tag (fluorescent reporter, affinity tag, or alkyne/azide for copper catalyzed Huisgen 1, 3-dipolar cycloaddition (“click chemistry”)) to identify the targets of the probe and quantitatively characterize their activity (Fig. 1.1).^{6,7} This technique can also be used to quantify the activity of a competitor compound against targets of interest if it is applied prior to or during labeling with the probe.

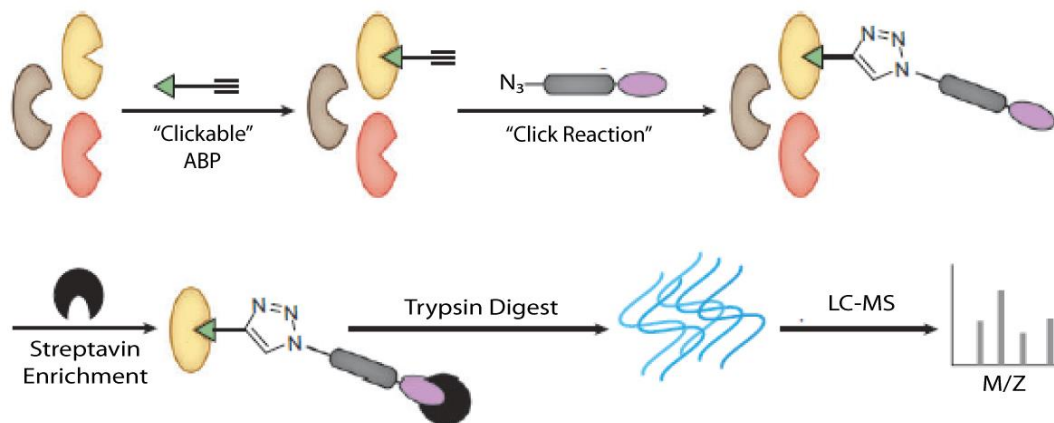


Figure 1.1: Schematic for ABPP with a “Clickable” activity-based probe (reactive group: green triangle; tag: alkyne) and biotin-azide (biotin: lavender oval) with LC-MS analysis. (Figure from Cravatt, *et. al.* (2008)).

Covalent inhibitors are also useful for target validation. Although traditional genetic techniques such as knockouts and RNAi are proven methods for generating loss-of-

function phenotypes, pharmacologic modulation is more relevant for therapeutic target validation. However, traditional pharmacologic agents lack the single-gene specificity of genetic techniques. In the case of covalent inhibitors with a known mechanism, mutation of the targeted nucleophilic residue to a non-nucleophilic residue should confer resistance to the inhibitor (Fig. 1.2). When cells express these compound-resistant alleles, these cells can reveal the effects of off-target inhibition. This strategy of combining chemical probes and genetic engineering provides a method for isolating the effects of inhibition of the protein of interest for target validation and for determining its biological role.



Figure 1.2: Chemical genetics. Using genetic engineering to confer resistance to electrophilic inhibitors.

The work described in this thesis focuses primarily around the development of selective covalent kinase inhibitors for use as chemical tools to identify and validate therapeutic targets in *Trypanosoma brucei*. In Chapter 1, we describe our work using quantitative chemoproteomics with propargyl-hypothemycin to identify *TbCLK1*, a previously uncharacterized and essential protein kinase. Chapter 2 describes the identification and development of covalent inhibitors of *TbCLK1* for use as chemical tools. Chapter 3 describes the development of non-basic analogs of K777 and the identification of their target in *T. cruzi* as well as an off-target interaction in host cells.

References

1. WHO. *African Trypanosomiasis*, <http://www.who.int/topics/trypanosomiasis_african/en/> (2013).
2. WHO. *Chagas Disease*, <http://www.who.int/topics/chagas_disease/en/> (2013).
3. Nett, I. R. *et al.* The phosphoproteome of bloodstream form *Trypanosoma brucei*, causative agent of African sleeping sickness. *Mol Cell Proteomics* **8**, 1527-1538 (2009).
4. Zhang, J., Yang, P. L. & Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer* **9**, 28-39 (2009).
5. Smith, A. J., Zhang, X., Leach, A. G. & Houk, K. N. Beyond picomolar affinities: quantitative aspects of noncovalent and covalent binding of drugs to proteins. *J Med Chem* **52**, 225-233 (2009).
6. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* **77**, 383-414 (2008).
7. Speers, A. E. & Cravatt, B. F. Profiling enzyme activities in vivo using click chemistry methods. *Chem Biol* **11**, 535-546 (2004).

CHAPTER 1

Therapeutic targets in *Trypanosoma brucei* revealed by hypothemycin-based chemoproteomics and RNAi

Chapter 1 is reproduced with permission from: Nishino, M., Choy, J.W., Gushwa, N.N., Koupparis, K., Osés-Prieto, J.A., Burlingame, A.L., Renslo, A.R., McKerrow, J.H., and Taunton, J. Therapeutic targets in *Trypanosoma brucei* revealed by hypothemycin-based chemoproteomics and RNAi. Manuscript submitted.

ABSTRACT

Protein kinases are attractive therapeutic targets for neglected parasitic diseases, including African trypanosomiasis caused by the protozoan, *Trypanosoma brucei*. How to prioritize *T. brucei* kinases and quantify their intracellular engagement by small-molecule inhibitors remain unsolved problems. Here, we combine chemoproteomics and RNA interference to interrogate trypanosome kinases bearing a Cys-Asp-Xaa-Gly motif (CDXG kinases). We discovered that hypothemycin, a fungal polyketide previously shown to covalently inactivate a subset of human CDXG kinases, kills *T. brucei* in culture and in infected mice. Quantitative chemoproteomic analysis with a hypothemycin-based probe revealed the relative sensitivity of endogenous CDXG kinases, including *TbGSK3short* and a previously uncharacterized kinase, *TbCLK1*. RNAi-mediated knockdown demonstrated that both kinases are essential, but only *TbCLK1* was fully occupied in intact cells at cytotoxic concentrations of hypothemycin. Our study identifies *TbCLK1* as a therapeutic target for African trypanosomiasis and establishes a new chemoproteomic tool for interrogating CDXG kinases in their native context.

INTRODUCTION

Human African trypanosomiasis, or African sleeping sickness, is a debilitating and fatal parasitic disease endemic to sub-Saharan Africa (Fevre et al., 2008; Simarro et al., 2008). Caused by two subspecies of *Trypanosoma brucei*, the disease begins in the hemolymphatic system and later crosses the blood-brain barrier resulting in sleep disturbances that deteriorate to coma and death. During the course of infection, the parasites evade the host immune system through periodic switching of the variable surface glycoprotein coat, making vaccination-based approaches unlikely to succeed (Horn and McCulloch, 2010; Taylor and Rudenko, 2006). Currently, only four chemotherapeutics are approved for sleeping sickness. These therapies suffer from poor oral bioavailability, severe toxicity, and/or emerging resistance. Eflornithine, the only drug with a known mechanism of action, is also the only therapeutic for sleeping sickness approved in the last 50 years (Fairlamb, 2003). Although a combination of eflornithine and nifurtimox, a drug for American trypanosomiasis (Chagas disease, caused by *T. cruzi*), was introduced in 2009 for *T. brucei* infection, no new drugs are near the clinic (Priotto et al., 2009).

Protein kinases have been intensely pursued as therapeutic targets for cancer and autoimmune disease, and more than 15 kinase inhibitors have received FDA approval during the past decade. As in humans, protein phosphorylation is an essential regulatory mechanism in *T. brucei*, and has critical but poorly understood roles in controlling its unique life cycle (Nett et al., 2009; Parsons et al., 2005). High-throughput screening and medicinal chemistry recently led to the discovery of SCYX-5070, the first reported protein kinase inhibitor with efficacy in a murine model of *T. brucei* infection (Mercer et

al., 2011). Affinity chromatography with an immobilized derivative suggested that SCYX-5070 binds to at least six trypanosome kinases related to mammalian mitogen activated protein kinases (MAPK) and cyclin-dependent kinases, although the extent to which SCYX-5070 engages these kinases *in vivo* is not known. Genetic manipulation in *T. brucei*, largely by RNA interference (RNAi), has defined essential roles for orthologs of several human kinases, including cyclin-dependent kinases (Gourguechon and Wang, 2009; Tu and Wang, 2004), Aurora kinase (Tu et al., 2006), polo-like kinase (Li et al., 2010), glycogen synthase kinase-3 (Ojo et al., 2008), casein kinase-1 (Urbaniak, 2009), and two nuclear DBF-2-related kinases (Ma et al., 2010). In addition, high-throughput phenotyping studies based on parallel RNAi suggest that additional protein kinases are required for optimal *T. brucei* fitness (Alsford et al., 2011; Mackey et al., 2011). However, the vast majority of the 182 protein kinases (Nett et al., 2009) in *T. brucei* remain poorly characterized, and few have been interrogated in their native cellular context with small-molecule inhibitors.

Hypothemycin is a fungal polyketide that bears an electrophilic *cis*-enone (**1**, Fig. 2.1a) and covalently inhibits a diverse subset of human kinases, including MEK, ERK, PDGFR, VEGFR2, and FLT3 (Schirmer et al., 2006). Hypothemycin (Tanaka et al., 1999) and related compounds (Barluenga et al., 2010) have antitumor activity in mouse xenograft models, and one variant has entered clinical trials (Kumar et al., 2011). All hypothemycin-sensitive kinases have a shared cysteine immediately preceding the catalytic DXG motif (where X is usually Phe or Leu) (Schirmer et al., 2006). A homologous CDXG motif is present in 48 of 518 human protein kinases (Leproult et al., 2011). These CDXG kinases are functionally diverse, encompassing Tyr, Ser/Thr, and dual specificity kinases distributed throughout the kinome phylogenetic tree. The

nucleophilic thiol of hypothemycin-sensitive CDXG kinases undergoes conjugate addition to the *cis*-enone, as seen by a crystal structure of the ERK2/hypothemycin complex (Rastelli et al., 2008). While the CDXG motif is necessary for potent inhibition by hypothemycin, it does not appear to be sufficient. One-third of human CDXG kinases were unaffected by hypothemycin variants screened at a concentration of 1 μ M (Barluenga et al., 2010). Unbiased affinity-based identification of hypothemycin-binding proteins has not been reported, and the selectivity of hypothemycin toward endogenous kinases in living cells remains unknown.

In this study, we exploited a semi-synthetic derivative of hypothemycin to identify targets of therapeutic interest in *T. brucei*. Hypothemycin is potently trypanocidal, both in cell culture and in mice. To identify its molecular targets, we synthesized the propargyl-hypothemycin derivative **2** (Fig. 2.2b), an equipotent cell-permeable variant amenable to click chemistry conjugation. This affinity probe provided a means to purify and identify covalent targets of hypothemycin. More importantly, probe **2** enabled us to quantify the extent of target engagement upon exposure to a defined concentration of hypothemycin. We identified 11 CDXG kinases that bound probe **2**, three of which were competed by nanomolar concentrations of hypothemycin. RNAi-mediated knockdown of all 21 CDXG kinases demonstrated that two of the kinases most sensitive to hypothemycin, *TbGSK3short* and *TbCLK1*, are also the only essential CDXG kinases. Thus, a combined genetic and quantitative chemoproteomic approach has revealed two hypothemycin-sensitive kinases, including one (*TbCLK1*) with little sequence similarity to human orthologs, as potential therapeutic targets for African trypanosomiasis.

RESULTS

Hypothemycin has potent trypanocidal activity

Similar to humans, approximately ten percent of *T. brucei* kinases (21 of 182) have a CDXG motif. Hypothemycin's selectivity toward CDXG kinases prompted us to test its effects on *T. brucei* grown in culture. We reasoned that hypothemycin would enable pharmacological interrogation of a focused subset of kinases that are distributed throughout the kinome. Treatment of bloodstream form (BSF) parasites with increasing concentrations of hypothemycin for 24 h resulted in a sharp reduction in cell viability ($EC_{50} \sim 170$ nM, Fig. 2.1b). Cells treated for 5 h with 500 nM hypothemycin were round and swollen, with the flagellum detached from the cell body. They died within 2 h and did not recover even when transferred to media lacking hypothemycin (data not shown).

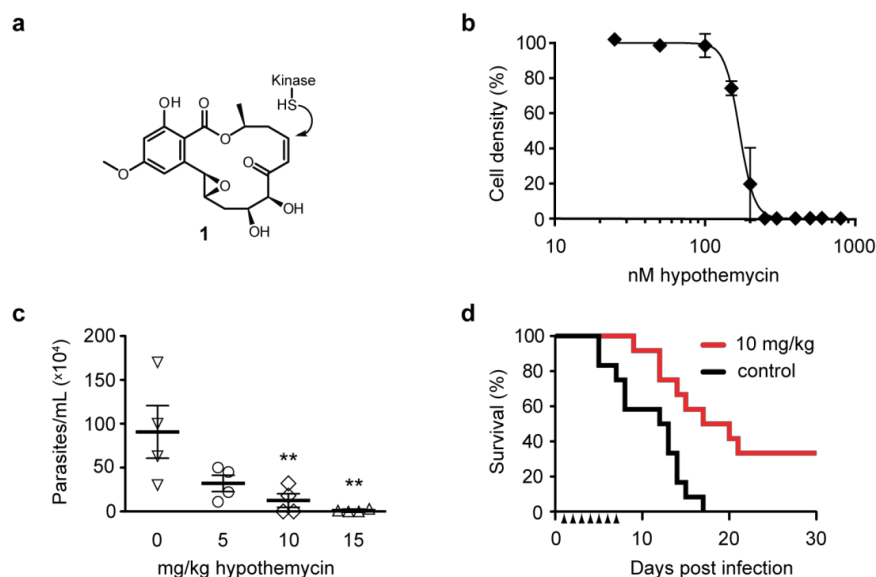


Figure 2.1: Potent trypanocidal activity of hypothemycin.

(a) Hypothemycin (**1**) covalently inhibits CDXG kinases via conjugate addition to the *cis*-enone. (b) Bloodstream form *T. brucei* were treated with hypothemycin and cell density was measured after 24 h (mean \pm s.d., $n=3$). (c) Parasitemia in *T. brucei* infected mice. Mice received once daily intraperitoneal injections of hypothemycin and parasitemia was measured 5 d post infection (mean \pm s.d., $n=4$, ** $p < 0.05$). (d) Kaplan-Meier analysis of *T. brucei* infected mice. Hypothemycin or vehicle was administered by intraperitoneal injection once daily for 7 d post infection (arrowheads, $n=12$). Data were accumulated from three studies, $p < 0.01$.

Previous studies with hypothemycin demonstrated activity in murine tumor xenograft models (Tanaka et al., 1999). We therefore tested hypothemycin in mice infected with *T. brucei*. Infected mice showed a dose-dependent reduction in parasitemia (Fig. 2.1c), and a 10 mg/kg dose administered once daily for seven days prolonged survival of infected mice over 30 days, with a cure rate of 33% (Fig. 2.1d). At doses greater than 10 mg/kg, signs of toxicity were evident (weight loss, lethargy), precluding more aggressive dosing regimens. Although hypothemycin's therapeutic window is narrow, likely due to inhibition of essential mammalian CDXG kinases (e.g., VEGFR2, MEK1/2), its potent trypanocidal activity *in vitro* and *in vivo* motivated us to identify the targets responsible.

Design and validation of a hypothemycin-based affinity probe

We hypothesized that hypothemycin's trypanocidal effects were mediated by covalent inhibition of one or more *T. brucei* CDXG kinases and sought to identify its targets in an unbiased manner. In the crystal structure of hypothemycin bound to ERK2, the C4 methyl ether is solvent exposed (Rastelli et al., 2008), suggesting that this and other kinases would accommodate a larger group at this position (Fig. 2.2a). We therefore replaced the methyl ether with a propargyl ether to enable copper-promoted ("click") conjugation to a biotin- or rhodamine-linked azide. Alkylation of 4-*O*-desmethyl hypothemycin (Wee et al., 2006) with propargyl bromide was promoted by cesium carbonate to afford the C4 propargyl ether **2**, (Fig. 2.2b). Consistent with our hypothesis, probe **2** was equipotent to hypothemycin against cultured BSF *T. brucei* (Fig. 2.2c).

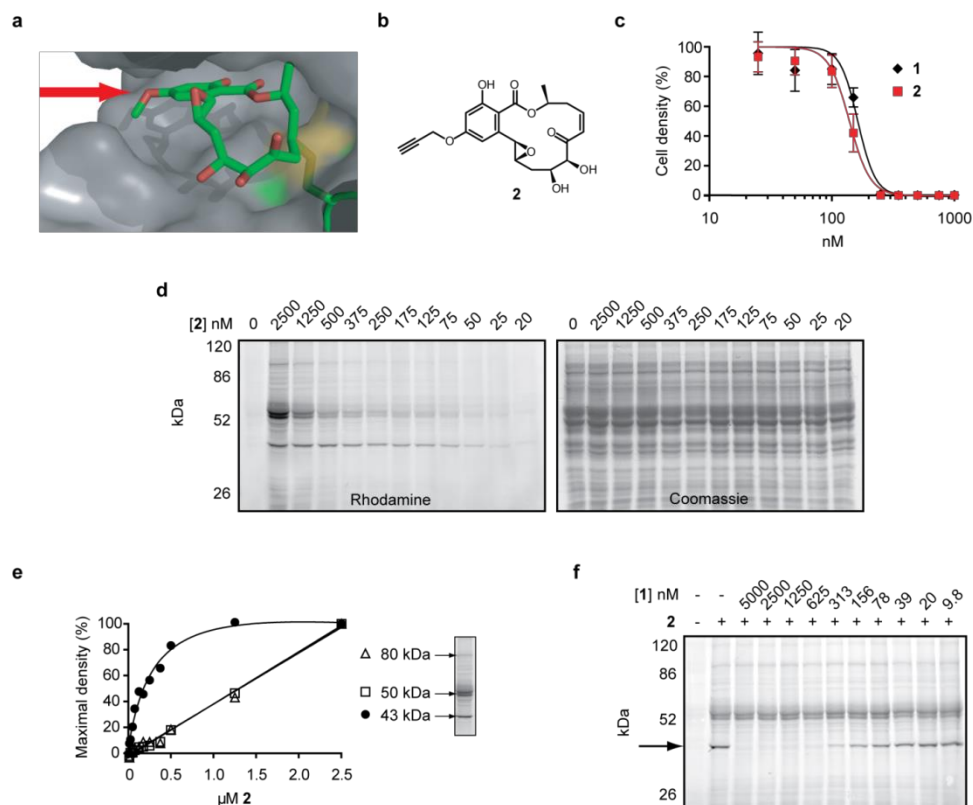


Figure 2.2: Design and validation of a hypothemycin-based affinity probe.

(a) Crystal structure of hypothemycin (**1**) bound to ERK2 (PDB: 3C9W), indicating the solvent-exposed C4 methyl ether. (b) Structure of probe **2**. (c) Effect of **2** on proliferation of cultured BSF *T. brucei* (mean \pm s.d., $n=3$). (d) *T. brucei* whole-cell lysates were treated with **2** for 30 min. Labeled proteins were visualized after click conjugation to rhodamine-azide and separation by SDS-PAGE. (e) Fluorescence quantification of bands at 90, 50, and 43 kDa demonstrating saturation of the 43 kDa band, but not the 90 or 50 kDa bands. (f) Lysates were treated with the indicated concentrations of **1** for 30 min, followed by 500 nM **2** for 30 min. Labeled proteins were visualized as above. The saturable 43 kDa protein is indicated by the arrow.

We next tested the ability of **2** to covalently modify *T. brucei* proteins. Treatment of whole-cell lysates with **2**, followed by click conjugation to rhodamine-azide, revealed multiple fluorescent bands by SDS-PAGE. Labeling of a 43 kDa protein was detected with 25 nM **2**, reaching maximum intensity at 500 nM (Figs. 2.2d, e). Labeled bands at 50 and 90 kDa were apparent at higher concentrations, but they failed to saturate (Fig.

2.2e) and are likely nonspecific adducts. Pretreating lysates with increasing concentrations of hypothemycin abolished subsequent labeling of the 43 kDa band, demonstrating that it is a specific and saturable hypothemycin target (Fig. 2.2f).

To identify the 43 kDa protein, we used a biotin-azide in the click reaction and affinity purified covalently modified proteins. After SDS-PAGE, analysis of the 43 kDa band by mass spectrometry identified the dominant protein (based on total peptide counts) as *TbGSK3short* (Tb927.10.13780), a CDXG kinase previously found to be essential in BSF *T. brucei* (Ojo et al., 2008). Peptides from three other CDXG kinases were also detected in the same gel band (Appendix 1, Table 1). Despite revealing four distinct CDXG kinases, this experiment provided no information about their relative sensitivity to hypothemycin. Moreover, we suspected that abundant and/or hyperreactive proteins may have bound **2** nonspecifically, obscuring specific, low-abundance targets in the gel-based analysis (Fig. 2.2d).

Hypothemycin sensitivity of CDXG kinases revealed by quantitative mass spectrometry

To identify hypothemycin-binding proteins in a more comprehensive and quantitative manner, we employed a gel-free method with hypothemycin competition and isobaric mass tags, similar to previously reported methods for identifying small-molecule targets (Bantscheff et al., 2007; Huang et al., 2009; Ong et al., 2009). Four lysate samples were treated in parallel with increasing concentrations of hypothemycin for 30 min (0, 20, 200, or 1000 nM), followed by probe **2** (500 nM, 30 min). As before, covalently modified proteins were conjugated to biotin-azide and affinity purified. After trypsinization of the

eluted proteins, each sample was derivatized with a unique iTRAQ reagent (Isobaric Tag for Relative and Absolute Quantitation) (Ross et al., 2004) and then pooled for fractionation and analysis by mass spectrometry. Using this method, peptides corresponding to ten protein kinases were identified, including three of the four kinases identified by in-gel digest. Four peptides correspond to two nearly identical *Tb*CLK (*C*dc2-*L*ike *K*inase) genes, *Tb*CLK1 (Tb927.11.12410) and *Tb*CLK2 (Tb927.11.12420), which could not be differentiated on the basis of the identified peptides. In total, we identified 11 kinase targets of **2** (Appendix 1, Table 1). Of note, every protein kinase identified by probe **2** has a CDXG motif, accounting for over half of the 21 CDXG kinases encoded in the *T. brucei* genome.

Quantification of iTRAQ reporter ions revealed the extent of CDXG kinase labeling by **2** after pretreatment with increasing concentrations of hypothemycin (Fig. 2.3). This experiment thus provided an estimate of each kinase's sensitivity to hypothemycin. Fragmentation spectra of peptides derived from eight kinases provided iTRAQ reporter ions of sufficient intensity to quantify kinase recovery (i.e., labeling by **2**) as a function of hypothemycin pretreatment (Fig. 2.3b, c). In addition to CDXG kinases, several other proteins (e.g., tubulin, heat shock proteins, ribosomal proteins) were identified. However, hypothemycin pretreatment did not affect their recovery (Fig. 2.3b; Appendix 1, Table 1), which likely resulted from nonspecific, low-level modification by **2** and/or nonspecific adsorption to the avidin-agarose beads.

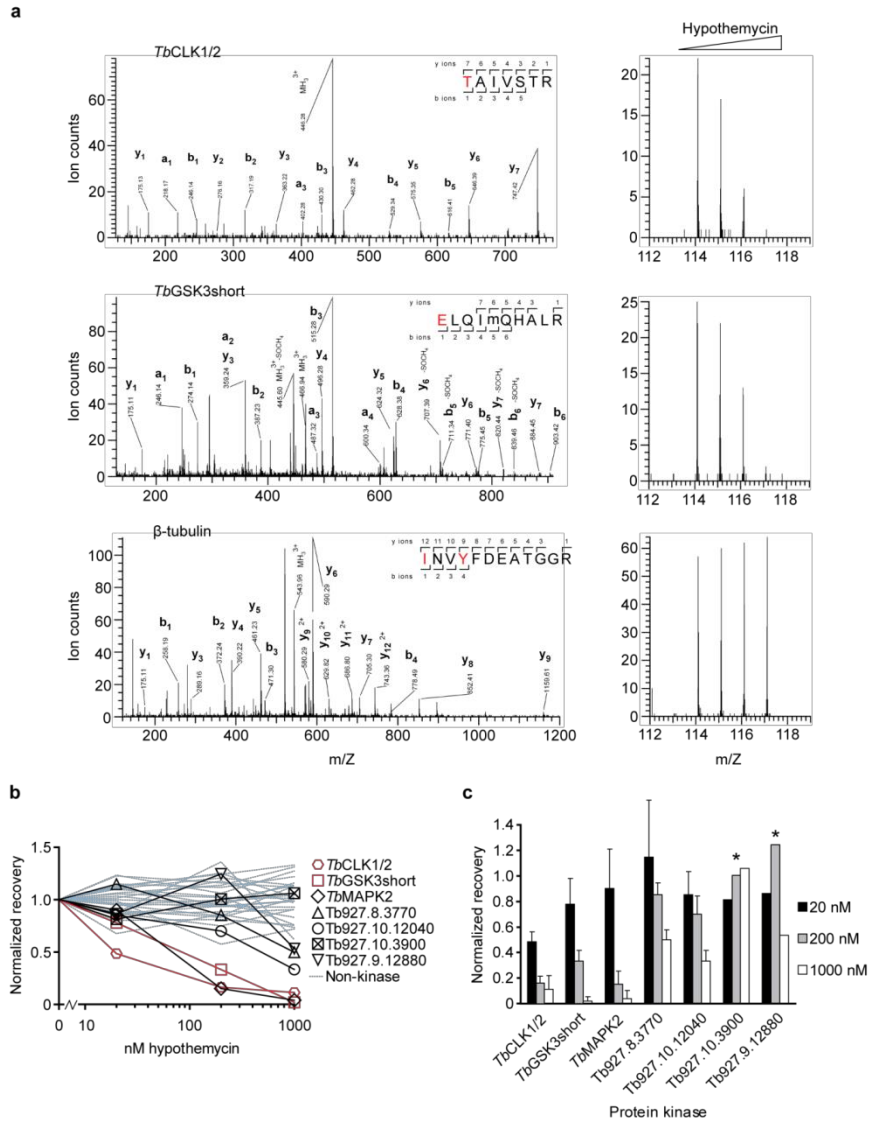


Figure 2.3: Quantification of hypothemycin binding to CDXG kinases.

(a) MS/MS fragmentation spectra of peptides from *TbCLK1/2*, *TbGSK3short*, and β -tubulin after labeling with **2**, affinity purification, and trypsinization. Red text: residues bearing an iTRAQ tag; “m”: oxidized Met. iTRAQ reporter ions for each peptide are shown to the right (m/Z 114, 115, 116, and 117 correspond to lysate samples pretreated with 0, 20, 200, or 1000 nM hypothemycin, respectively). (b) Normalized recovery values, based on iTRAQ quantification, for all identified protein kinases and 25 non-kinases (blue lines) with ≥ 3 unique peptides. (c) Recovery of protein kinases as a function of hypothemycin pretreatment. Values for each hypothemycin pretreatment condition are derived from mean iTRAQ reporter ion counts normalized to vehicle across all peptides identified (\pm s.d.). Asterisks indicate values derived from a single spectrum.

Among the CDXG kinases quantified, recovery of *TbCLK1/2* was uniquely affected by pretreatment with 20 nM hypothemycin (Fig. 2.3). Recovery of two additional kinases, *TbGSK3short* and *TbMAPK2* (Tb927.10.16030), was reduced more than 50% by 200 nM hypothemycin, and recovery of all but one CDXG kinase was reduced by 1 μ M. These data indicate a wide range of hypothemycin sensitivity, with *TbCLK1/2* exhibiting greater sensitivity than *TbGSK3short* and *TbMAPK2*. *TbMAPK2* is not essential in BSF *T. brucei*, but is required for proliferation of the procyclic (insect host) form (Muller et al., 2002). Because *TbCLK1* and *TbCLK2* have identical kinase domains and only diverge in their N-terminal regions, they are presumed to be equally sensitive to hypothemycin.

RNAi analysis of CDXG kinases in bloodstream form *T. brucei*

We assessed the requirement for each of the 21 CDXG kinases for cell viability using RNA interference. This required the creation of 21 cell lines, each containing a stably integrated cassette under tetracycline control (Wang et al., 2000), designed to silence a unique CDXG kinase. After induction of RNAi, cell proliferation was followed for six days. Consistent with previous results (Ojo et al., 2008), we observed reduced viability after knockdown of *TbGSK3short* (Fig. 2.4a). Also consistent with previous studies, knockdown of MAP kinases *TbMAPK2* (Muller et al., 2002), Tb927.6.1780 (Guttinger et al., 2007), and Tb927.6.4220 (Domenicali Pfister et al., 2006) had no effect on cell viability (Fig. 2.4b; Appendix 1, Table 2).

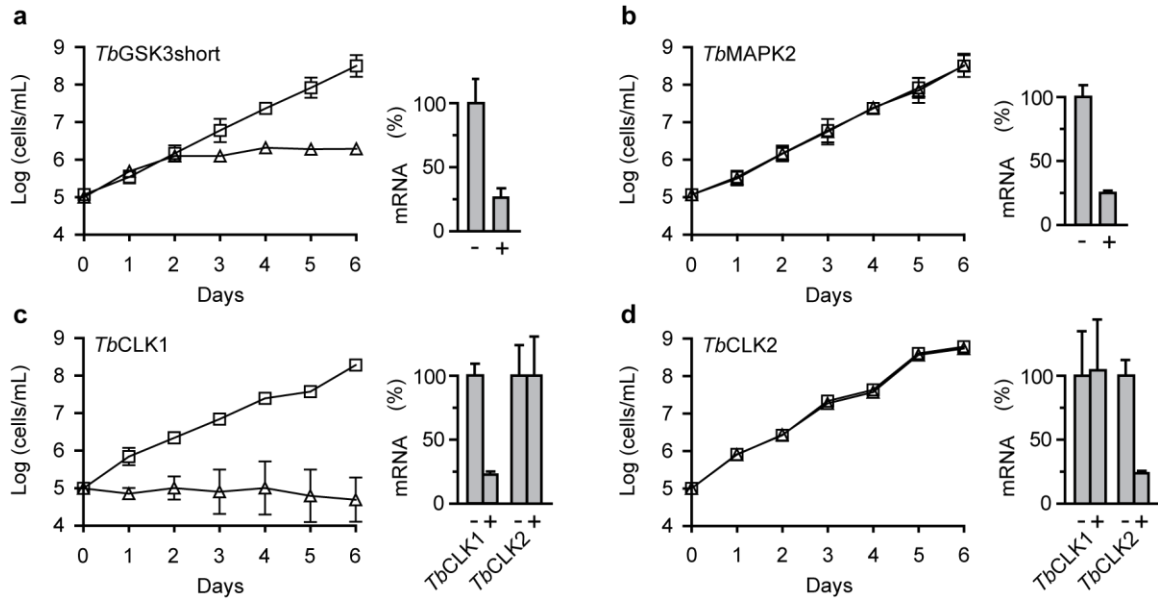


Figure 2.4: RNAi analysis of CDXG kinases.

(a–d) BSF *T. brucei* were stably transfected with the indicated RNAi constructs and induced with tetracycline (triangles) or left uninduced (squares) starting on day 0. Cell density was measured every 24 h and cumulative cell growth was plotted on a log scale (mean \pm s.d., n=3). mRNA levels were measured 48 h after induction by quantitative RT-PCR (bar graphs) in the absence (–) and presence (+) of tetracycline.

Our initial RNAi construct did not distinguish between *TbCLK1* and *TbCLK2*. Induction with tetracycline resulted in reduced proliferation and increased cell death (data not shown). To test for nonredundant functions of *TbCLK1* and *TbCLK2*, we designed a set of RNAi constructs targeting the unique 5'-UTR of each gene. A lethal phenotype was only observed after knockdown of *TbCLK1*; knockdown of *TbCLK2* had no obvious effect (Figs. 2.4c, d). Individual knockdown of the remaining CDXG kinases similarly resulted in no effect on cell proliferation. In each case, knockdown of the corresponding mRNA was confirmed by quantitative PCR (55-80% reduction across 21 mRNAs; Appendix 1, Table 2). With the caveat that insufficient knockdown may underlie the lack of a proliferative defect in certain cases, these results suggest that *TbGSK3short* and *TbCLK1* are the

only essential CDXG kinases in BSF *T. brucei*. Remarkably, these kinases also appeared to be among the most sensitive to hypothemycin (Fig. 2.3).

Hypothemycin preferentially targets *TbCLK1* in intact parasites

Because both *TbGSK3short* and *TbCLK1* reacted with nanomolar concentrations of hypothemycin in cell lysates and exhibited strong RNAi phenotypes, we focused on these CDXG kinases and assessed their sensitivity to hypothemycin in enzymatic assays using recombinant full-length *TbCLK1* and *TbGSK3short* expressed in *E. coli*, and in living parasites. Initially, a gel-based autophosphorylation assay was developed for *TbCLK1* by Mari Nishino. However, low signal limited utility of this assay (data not shown). We later determined using mass spectrometry that this recombinantly expressed *TbCLK1* was highly phosphorylated (Fig. 2.5), thus reducing the available sites for subsequent phosphorylation and detection. To amplify the signal from *TbCLK1* kinase activity, several common kinase substrate proteins (i.e., myelin basic protein, dephosphorylated α -casein, dephosphorylated β -casein) were tested in a radiometric ($\gamma^{32}\text{P}$ -ATP) dot blot assay, and myelin basic protein was found to be the most effective. A similar assay was developed for *TbGSK3short* using GSM, a substrate peptide for GSK3 homologs.

MDSISSGLTHRAGLQQYPSGLVNVHVNNPKINNGACSNMIPRESTNSARRSGSKRDRETADTA
DDGRNKNNQSITGTTATTTSTAKHSEMQPPPKKKKVTYALPNQSREEGHFYVVLGEDIDASTGR
FKILSLLGEGTFGKVVVEAWDRKRKEYCAVKIVRNVPKYTRDAKIEIQFMERVRLSDVEDRFPLM
KIQRYFQNETGHMCIVMPKYGPCLLDWIMKHGPFNHRHLAQIIFQVGAALDYFHTELHLMHTDL
KPENILMESGDTSVDPMTHRALPPEPCRVRICDLGGCCDERHSRTAIVSTRHYRSPVVLSLGW
MYSTDLWSMGCIIYELYTGKLLYDTHDNLEHLHLMKTLGRLPADWSVRCGTQEARDLFTAAGT
LQPCKDPKHIARIARARPVREVITEPLLCDLILNLLHYDRQRRLNARQMMSHAYVHKYFPECRO
HPNHVDNRKLPPTPVM

Figure 2.5: Recombinant *Tb*CLK1 is highly phosphorylated.

Phosphorylated residues on *Tb*CLK1 expressed in *E. coli* were identified by mass spectrometry. Phosphoserines and threonines are highlighted in red. Phosphotyrosines are highlighted in yellow.

Using these more robust assays for kinase activity, we tested hypothemycin's effects on recombinant *Tb*CLK1 and *Tb*GSK3short *in vitro*. Hypothemycin inhibited *Tb*CLK1 with an IC₅₀ of 150 nM when pre-incubated for 30 min in the presence of 100 μM ATP, whereas comparable inhibition of *Tb*GSK3short required 30-fold higher concentrations (Fig. 2.6a). This is consistent with our chemoproteomic experiment in which 20 nM hypothemycin reduced subsequent labeling of *Tb*CLK1 in lysates, while having little effect on *Tb*GSK3short (Fig. 2.3). The greater difference in IC₅₀ values in the enzymatic assays is likely due to the presence of competing ATP, which had been depleted from cell lysates prior to the labeling experiments.

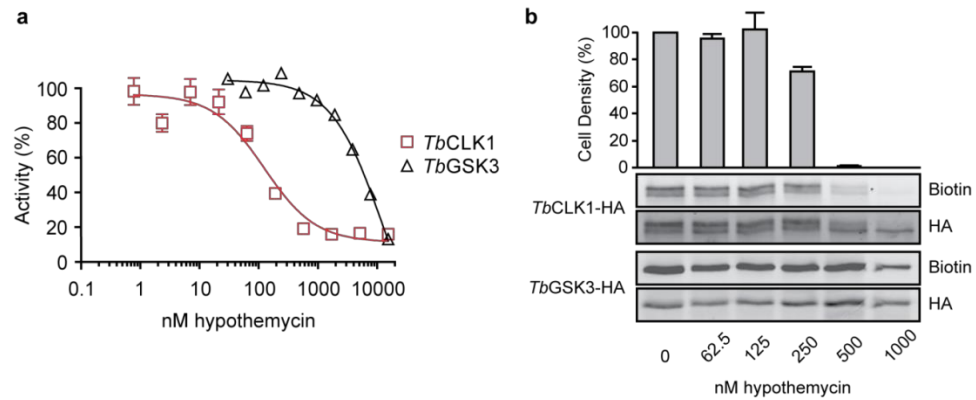


Figure 2.6: Preferential inhibition of *TbCLK1* by hypothemycin.

(a) *In vitro* assays with recombinant *TbGSK3short* and *TbCLK1*. Kinases were incubated with hypothemycin and 100 μ M ATP for 30 min before initiating reactions with substrate and γ^{32} P-ATP. Substrate phosphorylation was quantified in triplicate and normalized to DMSO control (mean \pm s.d.). (b) *T. brucei* expressing HA-tagged *TbGSK3short* and *TbCLK1* from their endogenous loci were incubated with hypothemycin for 5 h, and then were either harvested or diluted 1:30 into drug-free media and counted after 24 h (bar graph, mean \pm s.d., n=3). Harvested cells were lysed, labeled with **2**, and submitted to click conjugation with biotin-azide. HA-tagged proteins were immunoprecipitated, resolved by SDS-PAGE, and analyzed by Western blotting for biotin and HA.

The significant difference in hypothemycin sensitivity of recombinant *TbGSK3short* and *TbCLK1*, along with their differential sensitivity in lysates, prompted us to ask whether this difference is also observed in intact cells. To quantify hypothemycin binding without mass spectrometry, we generated a strain of BSF *T. brucei* that contains a C-terminal hemagglutinin (HA) tag inserted into the endogenous loci of both *TbGSK3short* and *TbCLK1* (hemizygous for each HA-tagged variant). HA-tagged *TbGSK3short* and *TbCLK1* are easily distinguished on Western blots by their molecular weight. Cells in log-phase growth were treated with increasing concentrations of hypothemycin. After 5 h, an aliquot of cells from each treatment condition was diluted 30-fold into fresh media, and cell density was measured after 24 h. The remaining cells were harvested after hypothemycin treatment, and lysates from these cells were treated with 500 nM **2**,

followed by click conjugation with biotin-azide. After immunoprecipitation, the HA-tagged kinases were analyzed for covalently attached **2** by Western blot with streptavidin detection.

In control cells, *TbCLK1* migrated as two major species (likely corresponding to differentially phosphorylated forms, discussed later), which collapsed to a single band upon treatment with hypothemycin (Fig. 2.6b). Probe **2** labeled both *TbCLK1* bands in the DMSO-treated cells, but labeling was sharply reduced in cells pretreated with hypothemycin. Moreover, the concentration dependence of this effect (reduction of labeling) correlated with the loss of cell viability. In contrast to its potent effect on *TbCLK1*, hypothemycin had a negligible effect on HA-tagged *TbGSK3short*, even at the highest concentration tested. Taken together, our results suggest that inhibition of *TbCLK1*, rather than *TbGSK3short*, plays a more dominant role in mediating the cytotoxic effects of hypothemycin.

Mutation of the CDXG cysteine in *TbCLK1* confers resistance to hypothemycin

Data obtained using probe **2** along with previous studies (Schirmer et al., 2006) suggested that the covalent reaction between the CDXG cysteine and the cis-enone in hypothemycin is necessary for efficient *TbCLK1* inhibition. We therefore predicted that mutation of this critical residue (C288) to non-nucleophilic alanine would confer resistance to hypothemycin. C288A *TbCLK1* was expressed in *E. coli* as previously described and possessed similar activity as the wild-type kinase. As expected, C288A was unaffected by ten micromolar hypothemycin, more than 100-fold higher than the IC_{50} of the wild-type kinase (Fig. 2.7).

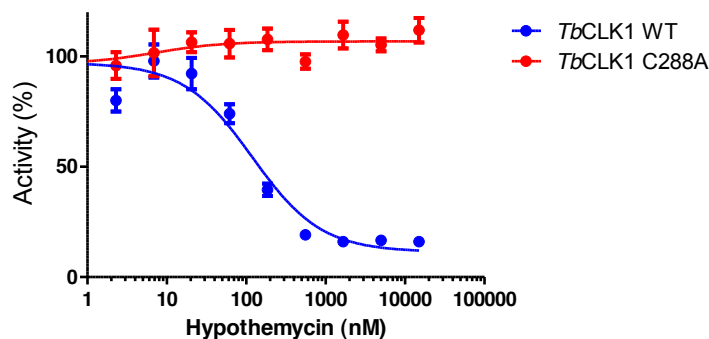


Figure 2.7: *In vitro* kinase assay with recombinant *TbCLK1*.

Kinases were incubated with hypothemycin and 100 μ M ATP for 30 min before initiating reactions with substrate and γ ³²P-ATP. Phosphotransfer was quantified in triplicate and normalized to DMSO control (mean \pm s.d.). (wild-type kinase: blue, “ADXG” mutant: red)

To test for *TbCLK1* inhibition in living cells we used *in vivo* autophosphorylation to report *TbCLK1* activity. As described previously, HA-*TbCLK1* expressed in BSF *T. brucei* ran as a doublet when resolved on a denaturing gel. When cells were incubated with hypothemycin prior to lysis, the doublet collapsed to a single band (Fig. 2.8, wild-type). The same effect was observed when the samples were treated with lambda phosphatase. There was also no further decrease in apparent molecular weight when both treatments were applied implying that hypothemycin inhibits phosphorylation of *TbCLK1*. In cells expressing C288A HA-*TbCLK1*, a doublet was also observed, but it was unaffected when the cells were treated with hypothemycin (Fig. 2.8, C288A). This result suggested that C288A HA-*TbCLK1* autophosphorylates in cells, activity that likely extends to all active forms of *TbCLK1*, and that the catalytic activity of the C288A *TbCLK1* was not significantly affected in cells treated with hypothemycin as also observed *in vitro*.

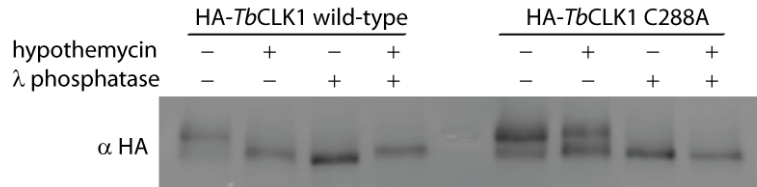


Figure 2.8: The catalytic activity of C288A *TbCLK1* is unaffected in cells treated with hypothemycin.

Cells expressing mutant or wild-type HA-tagged *TbCLK1* were treated $\pm 2 \mu\text{M}$ hypothemycin 2 h then lysed. Lysates were split into two portions: one of which was treated with lambda phosphatase. Samples were resolved by SDS-PAGE, transferred to nitrocellulose, and probed for HA.

Hypothemycin-resistant *TbCLK1* does not rescue BSF *T. brucei* from hypothemycin treatment

We next tested whether the C288A mutation was sufficient to confer hypothemycin resistance to cells. Cell lines expressing WT or C288A HA-*TbCLK1* were generated as previously described but the remaining copy of the gene was knocked out. In contrast to *in vitro* and cellular assays of kinase activity, C288A *TbCLK1* expressing cells were not resistance to hypothemycin (Fig. 2.9). These results suggested hypothemycin does not kill trypanosomes solely through inhibition of *TbCLK1*. One potential explanation is that *TbCLK2*, predicted to be equally sensitive to hypothemycin, serves a compensatory role in the hemizygous cell lines and thus essential in this context. Alternatively, inhibition of *TbCLK2* may have a dominant negative effect on C288A *TbCLK1* activity in a heterodimer. Preliminary results in pull-down assays suggest there is an intermolecular interaction between *TbCLK1* and *TbCLK2*. A third possibility is that hypothemycin toxicity may be mediated through the cumulative inhibition of multiple CDXG kinases through an as-yet unidentified target.

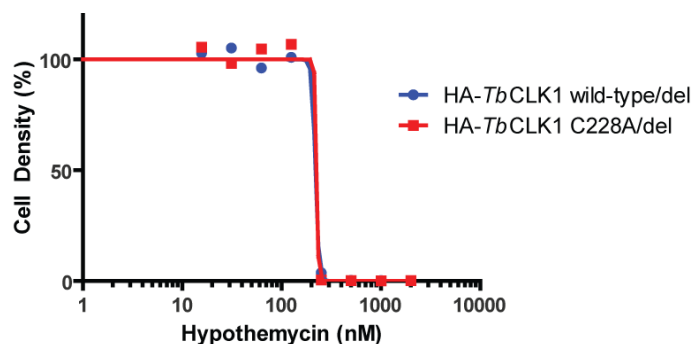


Figure 2.9: Expression of C288A *TbCLK1* is insufficient to rescue trypanosomes from hypothemycin treatment.

Cells expressing just a single copy HA-tagged *TbCLK1* (HA-*TbCLK1*/deletion) as wild-type or C288A mutant were incubated with hypothemycin. After 24 h, cell density was measured and normalized to untreated cells. (n=1).

DISCUSSION

New drugs are urgently needed to treat human African trypanosomiasis. While protein kinases are attractive targets, elucidating which of the 182 *T. brucei* kinases to prioritize for drug discovery efforts is nontrivial. Many essential *T. brucei* kinases show high sequence identity with human orthologs, frustrating attempts to obtain selective inhibitors. Moreover, once a kinase target has been validated genetically and subjected to high-throughput biochemical screens, demonstrating that a given inhibitor engages the kinase in intact cells is difficult or impossible. This has not been accomplished for any *T. brucei* kinase inhibitor to our knowledge. To begin to address these challenges, we have applied a combination of pharmacology, quantitative chemoproteomics, and reverse genetics to the CDXG kinases, a focused yet phylogenetically diverse subset of all eukaryotic kinomes.

We initially discovered that hypothemycin and the propargyl ether derivative **2** are potently trypanocidal. Although hypothemycin reduced parasitemia and cured one-third

of infected mice, it was toxic at higher doses, and we do not consider it to be a drug lead. Hypothemycin-mediated toxicity may be caused by inactivation of MEK and VEGFR2, as previous studies have found toxic effects associated with selective inhibitors of these kinases (Adjei et al., 2008; Rugo et al., 2005). Trypanosomes lack orthologs of most hypothemycin-sensitive mammalian kinases, including MEK and VEGFR2. Conversely, the majority of CDXG kinases in *T. brucei* lack clear human orthologs. It was therefore imperative to develop an unbiased and quantitative method for identifying direct targets of hypothemycin in *T. brucei*.

Of 21 predicted CDXG kinases, 11 were identified by probe **2** after affinity purification from cell lysates. It is likely that the remaining CDXG kinases are not efficiently labeled by **2** or are not expressed in BSF *T. brucei*. Mass spectrometry-based quantification revealed the extent of CDXG kinase occupancy as a function of hypothemycin concentration. We observed a wide spectrum of sensitivity, with most CDXG kinases requiring micromolar concentrations of hypothemycin to block labeling by probe **2** in ATP depleted lysates. Two of the kinases showing the greatest reduction in probe labeling after exposure to hypothemycin, *TbGSK3short* and *TbCLK1*, are also the only CDXG kinases whose knockdown by RNAi significantly reduced cell growth.

TbGSK3short was previously shown to be essential for proliferation and survival of BSF *T. brucei* (Ojo et al., 2008). High-throughput screening with recombinant *TbGSK3short* has provided inhibitors with high biochemical potency, yet modest selectivity when tested against human GSK3 β (Oduor et al., 2011; Urbaniak et al., 2012), consistent with their high sequence similarity. Although many of these compounds block *T. brucei*

proliferation, it is unclear whether they actually inhibit *TbGSK3short* in cells. Indeed, hypothemycin was able to saturate *TbGSK3short* at nanomolar concentrations in cell lysates, yet it had little effect on this kinase in intact cells, even after prolonged treatment at cytotoxic concentrations. Thus, despite *TbGSK3short* being the most prominent protein labeled by **2** in lysates, it is unlikely that hypothemycin kills *T. brucei* by inhibiting *TbGSK3short*. Nevertheless, we cannot exclude the possibility that partial inhibition of *TbGSK3short* or other CDXG kinases contributes to hypothemycin's cytotoxic effects.

Treatment of intact trypanosomes with cytotoxic concentrations of hypothemycin resulted in full occupancy of *TbCLK1*. Together with the demonstration that *TbCLK1* silencing by RNAi is sufficient to kill *T. brucei*, these results validate *TbCLK1* as a druggable target. *T. brucei* and other kinetoplastids, including the human pathogens *T. cruzi* and *Leishmania* species, are the only organisms in which a CLK ortholog bears a CDXG motif. The kinase domains of *TbCLK1* and human CLK1-4 share only 30% sequence identity. By contrast, *T. brucei* and human GSK3 share 52% sequence identity, including the CDXG motif. Obtaining inhibitors that discriminate between *T. brucei* and human CLK may therefore be more straightforward as compared to GSK3. In addition, *TbCLK1* autophosphorylation and competitive labeling with probe **2** can be used to evaluate future *TbCLK1* inhibitors in cells as potential therapeutic leads and chemical tools for studying the kinase. CLKs have been shown to regulate pre-mRNA splicing in fission yeast (Tang et al., 2011), flies (Du et al., 1998), and human cells (Fedorov et al., 2011; Karlas et al., 2010; Muraki et al., 2004), and recent studies have revealed important splicing-independent roles (Lee et al., 2012; Rodgers et al., 2010; Rodgers et al., 2011). Further work will be required to elucidate the functions and substrates of *TbCLK1* in BSF *T. brucei*.

By employing **2** as an affinity probe, it should now be possible to quantify engagement of both *TbGSK3short* and *TbCLK1* by compounds derived from high-throughput screening campaigns, even without knowledge of their downstream substrates or signaling pathways. We note that *TbCLK1* was not among the 57 kinases identified in *T. brucei* cell lysates using broad-spectrum kinase inhibitors immobilized on “Kinobeads” (Urbaniak et al., 2012). Although Kinobeads and other powerful chemoproteomic platforms (Patricelli et al., 2011) have the advantage of profiling a larger swath of the kinome, a potential disadvantage is that the probes only work in cell lysates and cannot be used to quantify kinase occupancy in intact cells. More broadly, probe **2** and the chemoproteomic methodology described herein can be used to explore the CDXG kinomes of related kinetoplastids (i.e., *T. cruzi* and *Leishmania*) as well as other disease-causing eukaryotes. Finally, many human CDXG kinases are validated or potential drug targets (e.g., MEK, ERK, VEGFR2, c-KIT, FLT3, PDGFR, TAK1, MNK), and probe **2** may prove valuable in quantifying their intracellular engagement by candidate inhibitors.

METHODS

Cell culture. BSF *T. brucei* (strain 221) was cultured at 37 °C and 5% CO₂ in HMI-9 medium with 10% Fetal Bovine Serum, 10% Serum Plus (Sigma-Aldrich, St. Louis, MO), 100 units/mL penicillin, and 100 µg/mL streptomycin. The 90-13 cell line, expressing T7 RNA polymerase and tetracycline repressor, was similarly cultured with addition of 2.5 µg/mL G418, 5.0 µg/mL hygromycin. Transgenic cell lines were maintained in medium supplemented with 5.0 µg/mL hygromycin, 2.5 µg/mL phleomycin, and/or 0.1 µg/mL puromycin.

Cell density assay. The endpoint luciferase-based assay for quantifying cell density was performed as previously described (Mackey et al., 2011) with the following modifications. BSF *T. brucei* were treated with hypothemycin in media lacking antibiotics at a density of 5×10^5 cells/mL. Cell density was measured after 24 h using CellTiter-Glo Luminescent Cell Viability Assay (Promega, Madison, WI). Luminescence was measured using a SpectraFluor Plus luminometer (Tecan, San Jose, CA). Results were plotted using GraphPad Prism (GraphPad Software, La Jolla, CA).

Mouse infection model. Adult female Balb/c mice (Charles River Laboratories, Wilmington, VA) weighing 18–22 g were infected via intraperitoneal injection with 10^3 BSF parasites in 100 µL of PBS containing 1% glucose. One day post-infection, hypothemycin (ChemieTek, Indianapolis, IN) in 60% DMSO/water was administered once daily via intraperitoneal injection for 7 d. Mice were monitored every 48 h for parasitemia in tail blood and visually inspected for general health. Surviving aparasitemic mice at day 30 were considered cured. Data from tail blood collected on day five post-

injection was analyzed using the Student's t-test for significant differences between treatments and vehicle control. Survival data was processed using a Kaplan-Meier analysis and log-rank test. Experiments were carried out in accordance with protocols approved by the Institutional Animal Care and Use Committee at the University of California, San Francisco.

Synthesis of probe 2. 4-*O*-desmethylhypothemycin (Wee et al., 2006) (10 mg, 0.0274 mmol) was added to a dried glass reaction vessel equipped with a magnetic stir bar and dissolved in dry acetonitrile (6 mL). Propargyl bromide (69 μ L 80% in toluene, 0.55 mmol) and Cs₂CO₃ (10.7 mg, 0.033 mmol) were added. After 6 h at rt, the mixture was concentrated, dissolved in minimal methylene chloride, and purified by preparative silica TLC using 2% then 3.5% methanol in methylene chloride. Pure **2** (4.5 mg, 41% yield) was eluted from the silica with 10% methanol in methylene chloride. NMR (¹H, DMSO, 400MHz): 0.98 (dd, *J*=14.9 Hz, *J*=9.4 Hz, 1H), 1.36 (d, *J*=6.2 Hz, 3H), 1.86 (dd, *J*=14.0, 9.8 Hz, 1H), 2.55 (m, 1H), 2.76 (m, 1H), 2.94 (dt, *J*=17.3, 10.9 Hz, 1H), 3.61 (t, *J*=2.4 Hz, 1H), 3.87 (m, 1H), 4.32 (d, *J*=1.7 Hz, 1H), 4.45 (dd, *J*=5.1, 1.5 Hz, 1H), 4.83 (d, *J*=2.4 Hz, 2H), 4.93 (d, *J*=5.1 Hz, 1H), 5.15 (d, *J*=6.6 Hz, 1H), 5.39 (m, 1H), 6.11 (dt, *J*=11.2, 2.6 Hz, 1H), 6.32 (d, *J*=2.8 Hz, 1H), 6.42 (dd, *J*=11.7, 2.7 Hz, 1H), 6.51 (d, *J*=2.8 Hz, 1H), 11.90 (s, 1H). NMR (¹³C, DMSO, 400MHz): 21.02, 34.24, 36.77, 56.39, 57.29, 63.72, 69.54, 74.83, 79.12, 79.51, 81.84, 102.37, 104.00, 105.66, 128.59, 143.10, 143.21, 162.82, 165.09, 171.00, 201.85. HRMS: predicted [M+H⁺] m/Z 403.1387; measured 403.1383.

Preparation of lysates. BSF *T. brucei* at a density of $10^6 - 5 \times 10^6$ cells/mL were collected and washed twice with PBS. The pellet was suspended in PBS containing Complete Protease Inhibitor Cocktail (Roche, Basel, Switzerland) and PhosStop Phosphatase Inhibitor Cocktail (Roche) and sonicated on ice. Cellular debris was pelleted, and the supernatant was passed through a NAP-5 column (GE Healthcare, Buckinghamshire, United Kingdom) equilibrated with lysis buffer.

Labeling with probe 2 and click chemistry. Whole-cell lysates (18.75 μ L, 1-3 mg/mL protein) were treated with increasing concentrations of hypothemycin for 30 min, followed by the indicated concentration of **2** for 30 min. Samples were denatured with SDS (1.25 μ L, 20%), followed by addition of TAMRA-azide (0.5 μ L, 5 mM), TCEP (0.5 μ L, 50 mM, pH ~7.0), TBTA ligand in 1:4 DMSO:*tert*-butyl alcohol (1.5 μ L, 1.67 mM), and CuSO₄ (0.5 μ L, 50 mM). Reactions were incubated at rt for 1 h, resolved by SDS-PAGE, scanned for fluorescence (Typhoon Imaging System, Molecular Dynamics, Sunnyvale, CA), and stained with Coomassie.

Preparation of iTRAQ mass spectrometry samples. *T. brucei* whole-cell lysates (14 mg, 3 mg/mL) were treated with 0 (DMSO control), 20, 200, or 1000 nM hypothemycin for 30 min, followed by **2** (500 nM, 30 min). Samples were then denatured by adding 20% SDS to a final concentration of 1%. Base-cleavable biotin-azide (Choy et al., 2013) was added (100 μ M), followed by the remaining click reagents at the concentrations described above. After 1 h at rt, proteins were precipitated with cold acetone (80% v/v final). Protein pellets were washed three times with acetone and dried. Pellets were resuspended in a minimal volume of 1% SDS in 50 mM Tris pH 8.0. The solution was

then diluted with 9 volumes of 1% NP-40 in PBS and passed through a PD-10 column (GE Healthcare), eluting with 1% NP-40 and 0.1% SDS in PBS. Avidin-agarose (30 μ L, Sigma-Aldrich) was added and the samples were rotated overnight at 4 °C. The beads were washed twice for 1 h at rt with 1% NP-40, 0.1% SDS in PBS; twice for 1 h at 4 °C with 6 M urea in PBS; 1 h with PBS at 4 °C; and 1 h with PBS at rt. To facilitate the quantitative release of captured proteins, the ester linkage (Choy et al., 2013) was hydrolyzed with NaOH (0.4 N, 20 μ L). After 20 min at room temperature, the solution was neutralized with HCl (0.8 N, 10 μ L), brought to 1% SDS, and heated to 90 °C for 3 min before collecting the supernatant.

Eluted proteins were acetone-precipitated and pellets were washed twice with additional cold acetone. Dried pellets were resuspended in 8 M guanidinium HCl, reduced (5 mM TCEP, 50 mM ammonium bicarbonate, 6 M guanidinium HCl), and alkylated with 10 mM iodoacetamide. The solution was adjusted to 1M guanidinium HCl, 100 mM ammonium bicarbonate, 10% acetonitrile and trypsinized overnight. Volatiles were removed, and the samples were resuspended in 0.1% formic acid, extracted with C18 OMIXtips (Varian, Palo Alto, CA), and eluted with 50% acetonitrile, 0.1% formic acid. Volatiles were removed and peptides were resuspended in 0.5 M triethylammonium bicarbonate pH 8.5 and labeled with iTRAQ Reagents (Applied Biosystems, Foster City, CA). Samples were mixed, dried, and resuspended in 30% acetonitrile, 5 mM potassium phosphate (pH 2.7). The samples were then fractionated on a 50.0 \times 1.0 mm 5 μ M 200Å Polysulfoethyl A column (PolyLC, Columbia, MD) with a 1-40% gradient of 350 mM potassium chloride in 30% acetonitrile, 5 mM potassium phosphate (pH 2.7). Approximately thirty peptide-containing fractions were dried, resuspended in 0.1% formic acid, extracted using C18

ZipTips (Millipore, Billerica, MA), eluted with 50% acetonitrile, dried, resuspended in 0.1% formic acid, and analyzed by LC/MS/MS.

Mass spectrometry. Fractionated tryptic peptides were separated by nano-flow liquid chromatography using a 75 μm \times 150 mm reverse phase C18 PepMap column (Dionex-LC-Packings, Sunnyvale, CA) at a flow rate of 350 nL/min in a NanoLC-1D Proteomics high-performance liquid chromatography system (Eksigent Technologies, Dublin, CA) equipped with a FAMOS autosampler (Dionex-LC-Packings). Peptides were eluted using a 2-30% gradient of acetonitrile/0.1% formic acid over 40 min, followed by 50% acetonitrile for 3 min. The eluate was coupled to a microionspray source attached to a QSTAR Elite mass spectrometer (Applied Biosystems/MDS Sciex, Framingham, MA). Peptides were analyzed in positive ion mode. MS spectra were acquired for 1 sec in the m/Z range between 350 and 1500. MS acquisitions were followed by 4 \times 2.5 sec collision-induced dissociation (CID) experiments in information-dependent acquisition mode. For each MS spectrum, the two most intense multiple charged peaks over a threshold of 25 counts were selected for generation of CID mass spectra. Two MS/MS spectra of each were acquired, first on the m/Z range 119-1500, with Q1 resolution set at "low", and then on the m/Z range 112-119, with resolution set at "unit". The CID collision energy was automatically set according to mass to charge (m/Z) ratio and charge state of the precursor ion. A dynamic exclusion window was applied which prevented the same m/Z from being selected for 1 min after its acquisition.

Mass spectrometry data analysis. A peak list was generated using the mascot.dll script. The list was searched against the Trypanosoma subset of the NCBI nr database

using ProteinProspector. A minimal ProteinProspector protein score of 15, a peptide score of 15, a maximum expectation value of 0.1, and a minimal discriminant score threshold of 0.0 were used for initial identification criteria. iTRAQ modification of the amino terminus or the epsilon-amino group of lysines, carbamidomethylation of cysteine, acetylation of the N-terminus of the protein and oxidation of methionine were allowed as variable modifications. Peptide tolerance in searches was 100 ppm for precursor and 0.2 Da for product ions, respectively. Peptides containing two miscleavages were allowed. The number of modifications was limited to two per peptide.

Quantification was based on the relative areas of the reporter ions ($m/z = 114, 115, 116$ and 117) generated by the isobaric iTRAQ reagents during CID experiments. Abundance ratios of individual peptides between the different samples were calculated by dividing the areas of their respective iTRAQ reporter ions using the DMSO-treated sample as a reference. Peptides with peak areas lower than 30 for the most intense reporter ion were discarded. For changes in relative abundance at the protein level, all MS/MS spectra for the different peptides belonging to a particular protein were used to calculate the average and SD of the abundance ratios.

Generation of transgenic *T. brucei* cell lines. Plasmid DNA (10 μg) was linearized (MfeI, *TbGSK3short-HA*; Bsu36I, *TbCLK1-HA*; NotI, pZJM constructs) and resuspended to 1 $\mu\text{g}/\mu\text{L}$ in water. Cells (2×10^7) were electroporated using the Human T Cell Nucleofector kit and the Amaxa program X-001 (Lonza, Basel, Switzerland), then diluted in media and allowed to recover for 24 h before selection for 5-7 d with the appropriate antibiotic.

RNA interference. A 250-500 bp fragment of each CDXG kinase gene was amplified from genomic DNA by PCR (primers listed in Appendix 1 (Table 3)) and cloned into the XhoI/HindIII site of pZJM (Wang et al., 2000). Plasmid DNA was stably transfected into the 90-13 strain of BSF *T. brucei* as described above. RNAi was induced with 1 µg/mL tetracycline. Parasites were counted every 24 h using a Coulter Counter (Beckman Coulter, Brea, CA), and cultures were diluted to maintain cell density between 10^5 and 2×10^6 cells/mL. RNA was extracted 48 h after tetracycline induction using TRIzol (Invitrogen) and RNeasy kit (Qiagen, Hilden, Germany). For qRT-PCR quantification (relative to *TbGAPDH*, Tb927.6.4300), cDNA was generated with the High Capacity RNA-to-cDNA Kit (AppliedBiosystems) and quantified using the PowerSYBR kit (AppliedBiosystems).

In vitro kinase assays. Full-length *TbGSK3short* and *TbCLK1* were amplified from genomic DNA using primers listed in Appendix 1 (Table 3) and cloned into the pET100 expression vector (Invitrogen). Proteins were expressed in *E. coli* ArcticExpress(DE3) (Stratagene, La Jolla, CA) in ZY5052 medium at 20 °C for 60 h with agitation. Cells were lysed using a microfluidizer into lysis buffer (50 mM Tris pH 8, 300 mM NaCl, 10 mM imidazole, 5% glycerol, 1 mM CaCl₂, 1 mM MgCl₂, 500 nM PMSF, 1× Protease inhibitor cocktail (Roche), DNase (Sigma-Aldrich), lysozyme (Sigma-Aldrich)). The soluble fraction was isolated by centrifugation and incubated overnight at 4 °C with Ni-NTA beads. The recombinant protein was eluted with lysis buffer containing 250 mM imidazole and dialyzed into storage buffer (30 mM Tris pH 7.5, 50 mM NaCl, 50%

glycerol, 5 mM DTT, 1 mM EDTA, 0.03% Brij35). Aliquots of the proteins were flash frozen with liquid nitrogen and stored at -80°C .

Kinases (5 nM *TbGSK3short*, 10 nM WT *TbCLK1*, 2.5 nM C288A *TbCLK1*) were incubated with hypothemycin in reaction buffer (50 mM Tris pH7.4, 10 mM MgCl_2 , 0.2 mM EGTA, 0.2 mg/mL BSA, 1 mM DTT) and 100 μM ATP for 30 min at rt. A solution of reaction buffer with $\gamma^{32}\text{P}$ -ATP (70-150 $\mu\text{Ci}/\text{mL}$, Perkin Elmer, Waltham, MA) and GSM peptide substrate (*TbGSK3short*, 0.05 mg/mL, Millipore) or myelin basic protein (*TbCLK1*, 0.5 mg/mL, Sigma-Aldrich) was added to initiate the kinase reaction. After 15 (*TbGSK3short*) or 30 min (*TbCLK1*), reactions were spotted onto phosphocellulose paper, washed once with 10% acetic acid, twice with 1% phosphoric acid, and once with methanol, then dried. Kinase activity was quantified using a Typhoon Imaging System (Molecular Dynamics) and ImageQuant 5.2 (Molecular Dynamics). IC_{50} values and dosing curves were generated using GraphPad Prism 5 (GraphPad Software).

Cellular labeling assays. *TbGSK3short* or *TbCLK1* lacking start and stop codons were amplified from genomic DNA using primers listed in Appendix 1 (Table 3) and inserted between the KpnI and XhoI sites of pC-PTP-PURO, which was modified to contain a C-terminal HA-tag (Gourguechon and Wang, 2009; Schimanski et al., 2005). For the *TbCLK1*-HA construct, the puromycin resistance cassette was replaced with a hygromycin resistance gene. *T.brucei* 221 cells were first transfected with the *TbGSK3short*-HA construct to establish a puromycin resistant cell line in which the vector stably integrated into the endogenous *TbGSK3short* locus via homologous recombination. This cell line was then transfected with the *TbCLK1*-HA construct and

selected for hygromycin and puromycin resistance. Cells stably expressing HA-tagged *TbGSK3short* and *TbCLK1* (5×10^6 cells/mL) were treated with the indicated concentrations of hypothemycin. After 5 h, aliquots of the cultures were diluted 30-fold with fresh media and cell density was quantified after an additional 24 h. The remaining cells were pelleted, washed with trypanosome dilution buffer (5 mM KCl, 80 mM NaCl, 1 mM MgSO₄, 20 mM glucose, 22 mM sodium phosphate pH 7.7) and lysed by sonication in 50 μ L PBS containing 0.25% NP-40, 1 \times Complete EDTA-free protease inhibitor cocktail, and 1 \times PhosStop phosphatase inhibitors (Roche). Lysates were clarified by centrifugation, normalized for protein content, and treated with 500 nM **2** for 30 min. Samples were subjected to click conjugation with biotin-azide as described above and diluted 10-fold with 1.1% NP-40 in PBS. HA-tagged proteins were immunoprecipitated using 12CA5 anti-HA antibody (Roche) and Protein A Dynabeads (Invitrogen), eluted with sample buffer, resolved by SDS-PAGE, and transferred to nitrocellulose membranes. HA and biotin were detected with anti-HA (1:1000, Sigma Aldrich, H6908), IRDye(680)-conjugated anti-rabbit (1:10,000), IRDye(800)-conjugated streptavidin (1:20,000) using the Odyssey Imaging System (Li-Cor Biosciences, Lincoln, NE).

References

Adjei, A.A., Cohen, R.B., Franklin, W., Morris, C., Wilson, D., Molina, J.R., Hanson, L.J., Gore, L., Chow, L., Leong, S., *et al.* (2008). Phase I pharmacokinetic and pharmacodynamic study of the oral, small-molecule mitogen-activated protein kinase kinase 1/2 inhibitor AZD6244 (ARRY-142886) in patients with advanced cancers. *J Clin Oncol* 26, 2139-2146.

Alsford, S., Turner, D.J., Obado, S.O., Sanchez-Flores, A., Glover, L., Berriman, M., Hertz-Fowler, C., and Horn, D. (2011). High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res* 21, 915-924.

Bantscheff, M., Eberhard, D., Abraham, Y., Bastuck, S., Boesche, M., Hobson, S., Mathieson, T., Perrin, J., Raida, M., Rau, C., *et al.* (2007). Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat Biotechnol* 25, 1035-1044.

Barluenga, S., Jogireddy, R., Koripelly, G.K., and Winssinger, N. (2010). In vivo efficacy of natural product-inspired irreversible kinase inhibitors. *Chembiochem* 11, 1692-1699.

Choy, J.W., Bryant, C., Calvet, C.M., Doyle, P.S., Gunatilleke, S.S., Leung, S.S.F., Ang, K.K.H., Chen, S., Gut, J., and Oses-Prieto, J.A. (2013). Chemical-biological characterization of a cruzain inhibitor reveals a second target and a mammalian off-target. *Beilstein Journal of Organic Chemistry* 9, 15-25.

Domenicali Pfister, D., Burkard, G., Morand, S., Renggli, C.K., Roditi, I., and Vassella, E. (2006). A Mitogen-activated protein kinase controls differentiation of bloodstream forms of *Trypanosoma brucei*. *Eukaryot Cell* 5, 1126-1135.

Du, C., McGuffin, M.E., Dauwalder, B., Rabinow, L., and Mattox, W. (1998). Protein phosphorylation plays an essential role in the regulation of alternative splicing and sex determination in *Drosophila*. *Mol Cell* **2**, 741-750.

Fairlamb, A.H. (2003). Chemotherapy of human African trypanosomiasis: current and future prospects. *Trends Parasitol* **19**, 488-494.

Fedorov, O., Huber, K., Eisenreich, A., Filippakopoulos, P., King, O., Bullock, A.N., Szklarczyk, D., Jensen, L.J., Fabbro, D., Trappe, J., *et al.* (2011). Specific CLK inhibitors from a novel chemotype for regulation of alternative splicing. *Chem Biol* **18**, 67-76.

Fevre, E.M., Wissmann, B.V., Welburn, S.C., and Lutumba, P. (2008). The burden of human African trypanosomiasis. *PLoS Negl Trop Dis* **2**, e333.

Gourguechon, S., and Wang, C.C. (2009). CRK9 contributes to regulation of mitosis and cytokinesis in the procyclic form of *Trypanosoma brucei*. *BMC Cell Biol* **10**, 68.

Guttinger, A., Schwab, C., Morand, S., Roditi, I., and Vassella, E. (2007). A mitogen-activated protein kinase of *Trypanosoma brucei* confers resistance to temperature stress. *Mol Biochem Parasitol* **153**, 203-206.

Horn, D., and McCulloch, R. (2010). Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. *Curr Opin Microbiol* **13**, 700-705.

Huang, S.M., Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., *et al.* (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* **461**, 614-620.

Karlas, A., Machuy, N., Shin, Y., Pleissner, K.P., Artarini, A., Heuer, D., Becker, D., Khalil, H., Ogilvie, L.A., Hess, S., *et al.* (2010). Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* *463*, 818-822.

Kumar, V., Schuck, E.L., Pelletier, R.D., Farah, N., Condon, K.B., Ye, M., Rowbottom, C., King, B.M., Zhang, Z.Y., Saxton, P.L., *et al.* (2011). Pharmacokinetic characterization of a natural product-inspired novel MEK1 inhibitor E6201 in preclinical species. *Cancer Chemother Pharmacol* *69*, 229-237.

Lee, J., Moir, R.D., McIntosh, K.B., and Willis, I.M. (2012). TOR signaling regulates ribosome and tRNA synthesis via LAMMER/Clk and GSK-3 family kinases. *Mol Cell* *45*, 836-843.

Leproult, E., Barluenga, S., Moras, D., Wurtz, J.M., and Winssinger, N. (2011). Cysteine mapping in conformationally distinct kinase nucleotide binding sites: application to the design of selective covalent inhibitors. *J Med Chem* *54*, 1347-1355.

Li, Z., Umeyama, T., and Wang, C.C. (2010). Polo-like kinase guides cytokinesis in *Trypanosoma brucei* through an indirect means. *Eukaryot Cell* *9*, 705-716.

Ma, J., Benz, C., Grimaldi, R., Stockdale, C., Wyatt, P., Frearson, J., and Hammarton, T.C. (2010). Nuclear DBF-2-related kinases are essential regulators of cytokinesis in bloodstream stage *Trypanosoma brucei*. *J Biol Chem* *285*, 15356-15368.

Mackey, Z.B., Koupparis, K., Nishino, M., and McKerrow, J.H. (2011). High-throughput analysis of an RNAi library identifies novel kinase targets in *Trypanosoma brucei*. *Chem Biol Drug Des* *78*, 454-463.

Mercer, L., Bowling, T., Perales, J., Freeman, J., Nguyen, T., Bacchi, C., Yarlett, N., Don, R., Jacobs, R., and Nare, B. (2011). 2,4-Diaminopyrimidines as potent inhibitors of

Trypanosoma brucei and identification of molecular targets by a chemical proteomics approach. *PLoS Negl Trop Dis* 5, e956.

Muller, I.B., Domenicali-Pfister, D., Roditi, I., and Vassella, E. (2002). Stage-specific requirement of a mitogen-activated protein kinase by Trypanosoma brucei. *Mol Biol Cell* 13, 3787-3799.

Muraki, M., Ohkawara, B., Hosoya, T., Onogi, H., Koizumi, J., Koizumi, T., Sumi, K., Yomoda, J., Murray, M.V., Kimura, H., *et al.* (2004). Manipulation of alternative splicing by a newly developed inhibitor of Clks. *J Biol Chem* 279, 24246-24254.

Nett, I.R., Martin, D.M., Miranda-Saavedra, D., Lamont, D., Barber, J.D., Mehlert, A., and Ferguson, M.A. (2009). The phosphoproteome of bloodstream form Trypanosoma brucei, causative agent of African sleeping sickness. *Mol Cell Proteomics* 8, 1527-1538.

Oduor, R.O., Ojo, K.K., Williams, G.P., Bertelli, F., Mills, J., Maes, L., Pryde, D.C., Parkinson, T., Van Voorhis, W.C., and Holler, T.P. (2011). Trypanosoma brucei glycogen synthase kinase-3, a target for anti-trypanosomal drug development: a public-private partnership to identify novel leads. *PLoS Negl Trop Dis* 5, e1017.

Ojo, K.K., Gillespie, J.R., Riechers, A.J., Napuli, A.J., Verlinde, C.L., Buckner, F.S., Gelb, M.H., Domostoj, M.M., Wells, S.J., Scheer, A., *et al.* (2008). Glycogen synthase kinase 3 is a potential drug target for African trypanosomiasis therapy. *Antimicrob Agents Chemother* 52, 3710-3717.

Ong, S.E., Schenone, M., Margolin, A.A., Li, X., Do, K., Doud, M.K., Mani, D.R., Kuai, L., Wang, X., Wood, J.L., *et al.* (2009). Identifying the proteins to which small-molecule probes and drugs bind in cells. *Proc Natl Acad Sci U S A* 106, 4617-4622.

Parsons, M., Worthey, E.A., Ward, P.N., and Mottram, J.C. (2005). Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*. *BMC Genomics* 6, 127.

Patricelli, M.P., Nomanbhoy, T.K., Wu, J., Brown, H., Zhou, D., Zhang, J., Jagannathan, S., Aban, A., Okerberg, E., Herring, C., *et al.* (2011). In situ kinase profiling reveals functionally relevant properties of native kinases. *Chem Biol* 18, 699-710.

Priotto, G., Kasparian, S., Mutombo, W., Ngouama, D., Ghorashian, S., Arnold, U., Ghabri, S., Baudin, E., Buard, V., Kazadi-Kyanza, S., *et al.* (2009). Nifurtimox-eflornithine combination therapy for second-stage African *Trypanosoma brucei gambiense* trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. *Lancet* 374, 56-64.

Rastelli, G., Rosenfeld, R., Reid, R., and Santi, D.V. (2008). Molecular modeling and crystal structure of ERK2-hypothemycin complexes. *J Struct Biol* 164, 18-23.

Rodgers, J.T., Haas, W., Gygi, S.P., and Puigserver, P. (2010). Cdc2-like kinase 2 is an insulin-regulated suppressor of hepatic gluconeogenesis. *Cell Metab* 11, 23-34.

Rodgers, J.T., Vogel, R.O., and Puigserver, P. (2011). Clk2 and B56beta mediate insulin-regulated assembly of the PP2A phosphatase holoenzyme complex on Akt. *Mol Cell* 41, 471-479.

Ross, P.L., Huang, Y.N., Marchese, J.N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., *et al.* (2004). Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 3, 1154-1169.

Rugo, H.S., Herbst, R.S., Liu, G., Park, J.W., Kies, M.S., Steinfeldt, H.M., Pithavala, Y.K., Reich, S.D., Freddo, J.L., and Wilding, G. (2005). Phase I trial of the oral antiangiogenesis agent AG-013736 in patients with advanced solid tumors: pharmacokinetic and clinical results. *J Clin Oncol* 23, 5474-5483.

Schimanski, B., Nguyen, T.N., and Gunzl, A. (2005). Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. *Eukaryot Cell* 4, 1942-1950.

Schirmer, A., Kennedy, J., Murli, S., Reid, R., and Santi, D.V. (2006). Targeted covalent inactivation of protein kinases by resorcylic acid lactone polyketides. *Proc Natl Acad Sci U S A* 103, 4234-4239.

Simarro, P.P., Jannin, J., and Cattand, P. (2008). Eliminating human African trypanosomiasis: where do we stand and what comes next? *PLoS Med* 5, e55.

Tanaka, H., Nishida, K., Sugita, K., and Yoshioka, T. (1999). Antitumor efficacy of hypothemycin, a new Ras-signaling inhibitor. *Jpn J Cancer Res* 90, 1139-1145.

Tang, Z., Luca, M., Portillio, J., Ngo, B., Chang, C., Wen, T., Murray, J., and Carr, A. (2011). LAMMER kinase Kic1 is involved in pre-mRNA processing. *Exp Cell Res* 317, 2308-2320.

Taylor, J.E., and Rudenko, G. (2006). Switching trypanosome coats: what's in the wardrobe? *Trends Genet* 22, 614-620.

Tu, X., Kumar, P., Li, Z., and Wang, C.C. (2006). An aurora kinase homologue is involved in regulating both mitosis and cytokinesis in *Trypanosoma brucei*. *J Biol Chem* 281, 9677-9687.

Tu, X., and Wang, C.C. (2004). The involvement of two cdc2-related kinases (CRKs) in *Trypanosoma brucei* cell cycle regulation and the distinctive stage-specific phenotypes caused by CRK3 depletion. *J Biol Chem* 279, 20519-20528.

Urbaniak, M.D. (2009). Casein kinase 1 isoform 2 is essential for bloodstream form *Trypanosoma brucei*. *Mol Biochem Parasitol* 166, 183-185.

Urbaniak, M.D., Mathieson, T., Bantscheff, M., Eberhard, D., Grimaldi, R., Miranda-Saavedra, D., Wyatt, P., Ferguson, M.A., Frearson, J., and Drewes, G. (2012). Chemical Proteomic Analysis Reveals the Drugability of the Kinome of *Trypanosoma brucei*. *ACS Chem Biol* 7, 1858-1865.

Wang, Z., Morris, J.C., Drew, M.E., and Englund, P.T. (2000). Inhibition of *Trypanosoma brucei* gene expression by RNA interference using an integratable vector with opposing T7 promoters. *J Biol Chem* 275, 40174-40179.

Wee, J.L., Sundermann, K., Licari, P., and Galazzo, J. (2006). Cytotoxic hypothemycin analogues from *Hypomyces subiculosus*. *J Nat Prod* 69, 1456-1459.

CHAPTER 2

Development of chemical tools to study *TbCLK1*

Abstract

TbCLK1 is an essential protein kinase with unknown function in *Trypanosoma brucei*, the parasite that causes Human African Trypanosomiasis (HAT). Within the active site of *TbCLK1* are two non-conserved cysteines that can be exploited for the development of covalent inhibitors. The first, targeted by the natural product hypothemycin, is found in approximately 10% of both trypanosome and human kinases. The second cysteine, C215, is unique in the trypanosome kinome and is analogous to the EGFR cysteine exploited by a number of clinical kinase inhibitors. We sought to develop C215-targeted electrophilic inhibitors of *TbCLK1* for use as chemical tools to study the roles of *TbCLK1* kinase activity. We first demonstrated that C215 is accessible to and reactive with electrophilic compounds that covalently inhibit human EGFR and BTK covalently through the EGFR cysteine. We next screened over 10,000 compounds from kinase focused libraries against *TbCLK1* and identified AD57 (**3.25**), a potent type II inhibitor of human c-Src, as a potent *TbCLK1* inhibitor and trypanocidal agent. By testing close structural variants of AD57, we established structure activity relationships correlating the compounds' trypanocidal effects with *TbCLK1* inhibition. To generate selective chemical genetic tools of *TbCLK1*, we synthesized electrophilic analogs of AD57 targeting C215. One of these, **3.41**, was a potent covalent inhibitor of *TbCLK1* showing more than a 70-fold preference for wild-type *TbCLK1* over the C215S mutant. In combination with cells expressing C215S *TbCLK1* and C234S *TbCLK2* (in development), **3.41** and several second-generation electrophilic analogs will provide potent chemical tools to study the roles of these kinases in cells.

Introduction

Our previous analysis of *Tb*CLK1 and *Tb*CLK2 showed that these kinases are most closely related to the LAMMER family of protein kinases. Members of this family include the human CLKs¹, *Drosophila* DOA², and *Arabidopsis* AFC³. All members of this family contain the signature sequence motif “EHLLAMMERILG” in the C-terminal lobe of their catalytic domain. A similar sequence, “EHLHLMEKTG,” is found near the C-terminus of *Tb*CLK1/2 (Fig. 3.1 boxed in red). Additionally, the amino acid sequences of the kinase domains of *Tb*CLK1 and human CLKs and share ~34% identity, higher than any other group of human kinases. Serine-arginine (SR) protein kinases (SRPKs), relatives of the LAMMER family, are the next most closely related in sequence (~30% identity). The remainder of the human kinome shares ~25% identity or less with *Tb*CLK1.

```

Identities = 119/351 (34%), Positives = 190/351 (55%), Gaps = 39/351 (11%)
TbCLK1 110 EEGHFYVVLGEDIIDASTGRFKILSLLGEGTFGKVVVEAWDRKRK-EYCAVKIVRNPKYTR 168
hCLK1 187 EEGH G D + R++I+ LGEF FGKVV E D K + AVKIV+NV +Y
EEGH LICQSG---DVL SARYEIVDTL GEGAFGKVV E C IDHKAGGRHVAVKIVKNVDRYCE 243
TbCLK1 169 DAKIEIQFMERVRLSDVEDRFPLMKIQRYFQNETGHMCIVMPKYGPCLLDWIMKHG--PF 226
hCLK1 244 AARSEIQVLEHLNTDPNSTFRVQMLEWFEHH-GHICIVFELLGLSTYDFIKENGFLPF 302
TbCLK1 227 NHRHLAQIIFQVGAALDYFHTELHMTDLKPENILMESGD--TSVDPMTHR----ALPP 280
hCLK1 303 RLDHIRKMAYQICKSVNFLHSN-KLTHDTLKPENILFVQSDYTEAYNPKIKRDERTLINP 361
TbCLK1 281 EPCRVRICDLGGCCDERHSRTAIVSTRHYRSPEVVL SLGWMYSTDLWSMGCIIYELYTGK 340
hCLK1 362 D---IKVDFG SATYDDEHHSTLVSTRHYRAPEVILALGWSQPCDVWSIGILIEYYLGF 418
TbCLK1 341 LLYDTHDNL EHLHLM EKTG RLPADWSVRCGTQEA-----RDLFTAAGTLQPCKDKPH 393
hCLK1 419 TVFPTHDSK EHLAMMER ILG LPKHMIQKTRKRKYFHHDRLDWEHSSAG-----RY 470
TbCLK1 394 IARIARARPVREV-----ITEPLLCDLILNLLHYDRQRRLNARQMMSHAYV 439
hCLK1 471 VSR--RCKPLKEFMLSQDVEHERLFDLIQKMLEYDPAKRITLREALKHFFF 519

```

Figure 3.1: Sequence alignment of *Tb*CLK1 and human CLK1. The “LAMMER” motifs in both *Tb*CLK1 and human CLK1 are boxed in red. *Tb*CLK1 was aligned to human CLK1 using NCBI-BLAST (Altschul, et. al. 1990).

The LAMMER kinases and SRPKs are important regulators of alternative splicing in yeast and higher eukaryotes, acting through phosphorylation of serine-arginine rich (SR)

proteins.^{4,5} In trypanosomes, only two genes (poly(A) polymerase⁶ and a RNA helicase⁷) are known to be *cis*-spliced; however, the related process of *trans*-splicing ligates a 22 base splice leader (SL) RNA, which carries a 7-methylguanosine cap, to pre-mRNAs from polycistronic transcripts, a process essential for generating stable, translatable mRNA.⁸ TSR1, a *T. brucei* SR protein, has been shown to specifically recognize SL-RNA in yeast three-hybrid screens.⁹ Members of the LAMMER family also exhibit dual-specificity kinase activity, phosphorylating both serine/threonine and tyrosine residues *in vitro*.¹⁰ Trypanosomes do not encode any canonical tyrosine kinases, yet contain tyrosine phosphatases and phosphotyrosine residues, including at least one found in *TbCLK1* (pY102).¹¹ We previously witnessed *TbCLK1*'s ability to autophosphorylate on tyrosine residues *in vitro* including Y102 (Chapter 1, Fig. 2.5) suggesting that trypanosome CLKs may act as a dual-specificity kinase in the organism.

Given its potentially critical roles in trypanosome biology and potential as a drug target, we sought to develop selective inhibitors and apply chemical genetic methodologies to study *TbCLK1*. Although traditional genetic techniques (gene deletion and RNA interference) are well-established methods for generating loss-of-function phenotypes, pharmacological modulation is more relevant for the validation of therapeutic targets. Small molecules also act rapidly on their cellular targets, which reduces the likelihood of cellular compensation, an important consideration when elucidating a kinase's biological function.¹² Selective kinase inhibitors can, however, be difficult to obtain due

to the highly conserved nature of the kinase active site. *TbCLK1* contains two non-conserved active site cysteines (C288 and C215, Fig. 3.2), homologs of which have been successfully targeted in human kinases with electrophilic inhibitors. C288, the CDXG cysteine, is found in approximately 10% of both the human and *T. brucei* kinomes, and is targeted by the natural product hypothemycin.¹³ C215, a cysteine found in C-terminal portion of the “hinge” region connecting the N- and C-lobes of the kinase domain, is targeted by electrophilic inhibitors of BTK and EGFR (EGFR cysteine).¹⁴ In contrast to the CDXG cysteine, the EGFR cysteine is found in just 11 human kinases and only *TbCLK1/2* in *T. brucei*.¹⁴ The rare nature of this residue is especially attractive for the development of selective covalent inhibitors as both therapeutic agents and chemical tools to study *TbCLK1/2*. Electrophilic inhibitors targeting C215 can be used in conjunction with cell lines expressing resistant alleles to elucidate the effects of *TbCLK1* inhibition and results of off-target interactions of the inhibitors. Through this combination of genetic engineering and pharmacology, we hope to generate a method to study *TbCLK1* and the results of its inhibition.

Identities = 119/351 (34%), Positives = 190/351 (55%), Gaps = 39/351 (11%)

```

TbCLK1 110 EEGHFYVVLGEDIIDASTGRFKILSLLGEGTFGKVVVEAWDRKRK-EYCAVKIVRNVPKYTR 168
EEGH G D + R++I+ LGEF FGKVVVE D K + AVKIV+NV +Y
hCLK1 187 EEGHLICQSG---DVL SARYEIVDTL GEGAFGKVVVECIDHKAGGRHVAVKIVKNVDRYCE 243

TbCLK1 169 DAKIEIQFMERVRLSDVEDRFPLMKIQRYFQNETGHMCIVMPKYGFCLLDWIMKHG--PF 226
|A+ EIQ +E + +D F +++ +F++ GH+CIV G D+I ++G PF
hCLK1 244 AARSEIQVLEHLNNTDPNSTFRVQMLEWFEHH-GHICIVFELLGLSTYDFIKENGFLPF 302

TbCLK1 227 NHRHLAQIIFQVGAALDYFHTELHMHDTLKPENILMESGD--TSVDPMTHR----ALPP 280
H+ ++ +Q+ ++++ H+ L HTDLKPENIL D + +P R + P
hCLK1 303 RLDHIRK MAYQICKSVNFLHSN-KLHTDLKPENILFVQSDYTEAYNPKIKRDERTLINP 361

TbCLK1 281 EPCRVRICDLGGCCDERHSRTAIVSTRHYRSPVVLSLGMVSTDLWMSGCIYELYTGK 340
+ +++ D G + + +VSTRHYR+PEV+L+LGW D+WS+GCI+ E Y G
hCLK1 362 D---IKVVDGFSATYDDEHHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYYLGF 418

TbCLK1 341 LLYDTHDNLEHLHMEKTLGRLPADWSVRCGTQEA-----RDLFTAAGTLQPCKDPKH 393
++ THD+ EHL +ME+ LG LP + ++ D ++AG ++
hCLK1 419 TVFPTHDSKEHLAMMERILGPLPKHMIQKTRKRYFHHDRLDWDEHSSAG-----RY 470

TbCLK1 394 IARIARARPVREV-----ITEPLLCDLILNLLHYDRQRRLNARQMMSHAYV 439
++R R +P++E + L DLI +L YD +R+ R+ + H +
hCLK1 471 VSR--RCKPLKEFMLSQDVEHERLFDLIQKMLEYDPAKRITLREALKHFFF 519

```

Figure 3.2: Features of the *TbCLK1* active site revealed by bioinformatics. The “EGFR cysteine” in *TbCLK1* (C215) is boxed in orange. The “CDXG cysteine” in *TbCLK1* (C288) is boxed in yellow. The “gatekeeper” residue in *TbCLK1* (M209) is boxed in green. *TbCLK1* was aligned to human CLK1 using NCBI-BLAST (Altschul, et. al. 1990).

This chapter is divided into two sections. As detailed in the first section, “Testing the reactivity of *TbCLK1* C215,” several inhibitors known to target the EGFR cysteine were used to verify the positioning and reactivity of C215. In the second section, “Investigation of AD57 and development of electrophilic analogs targeting C215 *TbCLK1*,” we present the results of screening kinase directed libraries against *TbCLK1* and subsequent optimization and of our best hit AD57. Based on our results in the first section, electrophilic analogs of AD57 were developed. These analogs exhibited potent, allele-selective *TbCLK1* inhibition and potent trypanocidal activity. These new covalent inhibitors will serve as useful tools to study the precise function of *TbCLK1* in cells and to validate it as a potential therapeutic target.

Section 1: Testing the reactivity of *Tb*CLK1 C215

Before undertaking development of targeted covalent inhibitors of *Tb*CLK1, we sought to verify the reactivity of C215 toward acrylamide electrophiles and determine its relative position in the active site using molecules developed to target the analogous cysteine in human kinases. These inhibitors exhibited weak yet C215-dependent inhibition of *Tb*CLK1. This demonstrated the feasibility of targeting C215 with covalent inhibitors; however, the compounds' poor activity in cellular assays led us to pursue alternative scaffolds.

Pyrazolopyrimidines bearing a two-carbon linker to an acrylamide on the N1 position inhibit *Tb*CLK1 in a C215 dependent manner.

PCI-32765 (**3.1**) is an irreversible inhibitor of Bruton's tyrosine kinase (BTK) that reacts covalently with the EGFR cysteine (C481 in BTK).¹⁵ When **3.1** was tested against *Tb*CLK1 *in vitro* (Fig. 3.3b), it was weakly active against the wild-type ($IC_{50} = 75 \mu\text{M}$) and C288A kinase (CDXG cysteine, $IC_{50} = 34 \mu\text{M}$), but inactive ($IC_{50} > 100 \mu\text{M}$) against the C215A mutant. The weak activity against the wild-type kinase was not surprising based on sequence alignments showing a large methionine residue at the gatekeeper position in *Tb*CLK1 (M209; Fig. 3.2, green box). Previous studies have shown that large gatekeeper residues sterically hinder pyrazolopyrimidines bearing bulky aryl substituents like the phenoxyphenyl in **3.1**. Despite the extremely high IC_{50} of **3.1**, a clear C215-dependent

effect was observed prompting further investigation into similar aminopyrazolopyrimidines bearing N1 acrylamides.

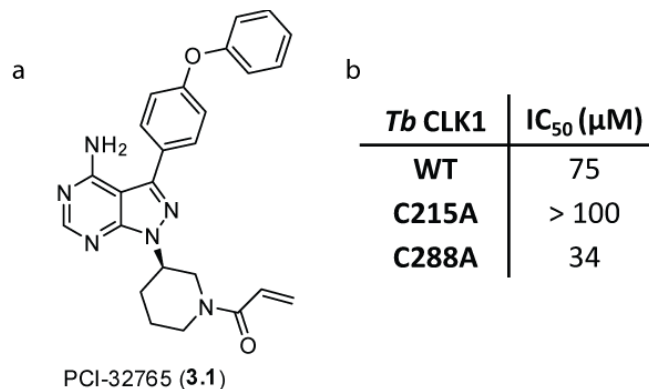


Figure 3.3: PCI-32765 inhibits *Tb*CLK1 in a C215 dependent manner. **(a)** Chemical structure of PCI-32765 (**3.1**). **(b)** Table of IC₅₀ values for **3.1** tested *in vitro* against recombinant *Tb*CLK1. Kinases were incubated with **3.1** and 100 μM ATP for 30 min before initiating reactions with substrate and γ³²P-ATP. Phosphotransfer was quantified and normalized to DMSO control (n=2).

PP2 (**3.2**) is a reversible aminopyrazolopyrimidine inhibitor of the Src-family of tyrosine kinases¹⁶ that bears a *para*-chlorophenyl substituent at the C3 position. Compared to **3.1**, the aryl substitution in PP2 is significantly smaller than the *para*-phenoxyphenyl. We synthesized two electrophilic PP2 analogs: one with a two-carbon ethyl linker (**3.3**) and a second with a more rigid 3-piperidinyl linker (**3.4**, racemic mixture) connecting the pyrazolopyrimidine core to the acrylamide. As expected, these inhibitors were significantly more active against wild-type *Tb*CLK1 (**3.4**: IC₅₀ = 1.4 μM; **3.3**: IC₅₀ = 20 μM; Fig. 3.4b) than **3.1** (IC₅₀ = 75 μM). The C215A mutation reduced the sensitivity of *Tb*CLK1 to these inhibitors, in the case of **3.4** by more than 14-fold, consistent with a covalent mode of action. The significantly higher potency of **3.4** also suggested that the more

constrained piperidinyll linker positioned the acrylamide more favorably for reaction with the C215.

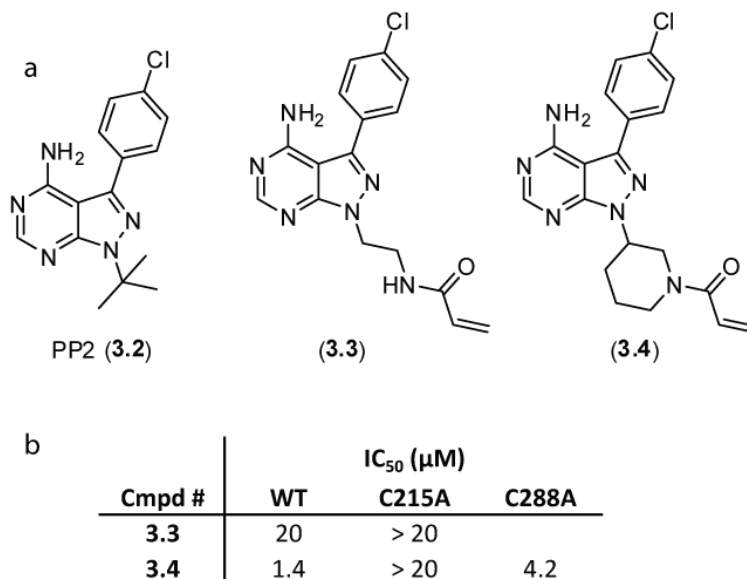


Figure 3.4: Electrophilic analogs of PP2 inhibit *Tb*CLK1 in a C215 dependent manner. **(a)** Chemical structure of PP2 (**3.2**), JC063-006 (**3.3**), and JC063-017 (**3.4**) **(b)** Table of *In vitro* IC₅₀ values of **3.3** and **3.4** for recombinant *Tb*CLK1. Kinases were incubated with compound and 100 μM ATP for 30 min before initiating reactions with substrate and γ³²P-ATP. Phosphotransfer was quantified and normalized to DMSO control (n=2).

The chloro-anilinopyrimidine WZ3146 inhibits *Tb*CLK1 in a C215 dependent manner.

WZ3146 (**3.5**) is a selective, covalent inhibitor of the EGFR T790M “gatekeeper” mutant that reacts with the analog of C215 in EGFR (C797).¹⁷ Our sequence alignments and the higher potency of the *para*-chlorophenyl versus *para*-phenoxyphenyl pyrazolopyrimidines suggested binding of these compounds to *Tb*CLK1 was heavily influenced by its methionine gatekeeper. We therefore synthesized WZ3146 which preferentially targets the EGFR T790M over the wild-type kinase along with the propionamide **3.6** (Fig. 3.5). Additionally we synthesized a structurally related

electrophilic inhibitor designed to inhibit BTK, I-182 (**3.7**)¹⁸, along with the non-electrophilic propionamide **3.8** and two structural hybrids of WZ1346 and I-182 (**3.9** and **3.10**).

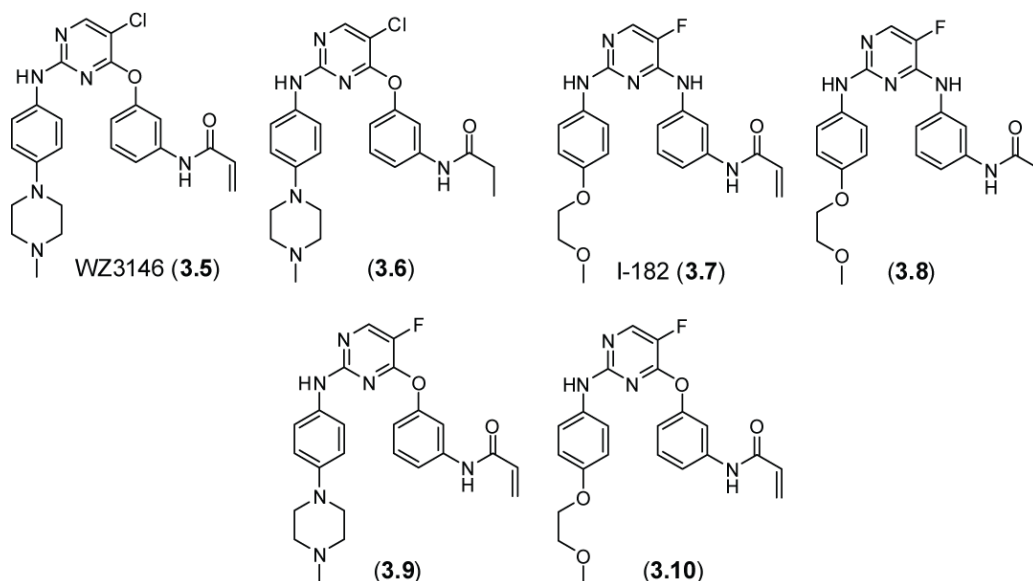


Figure 3.5: Chemical structures of anilinopyrimidine inhibitors.

In our initial *in vitro* tests of the anilinopyrimidine series of inhibitors, only **3.5** exhibited activity below 20 μM ($\text{IC}_{50} = 1.8 \mu\text{M}$, Table 3.1) against the wild-type kinase. As predicted, the C288A mutation did not affect the activity of **3.5** ($\text{IC}_{50} = 2.1 \mu\text{M}$, Table 3.1). When **3.5** was tested against the C215A mutant of the kinase, there was no measurable inhibition even at the highest concentration tested (20 μM , Table 3.1). This dramatic drop in potency combined with the inactivity of propionamide **3.6** suggested that C215 is required for *Tb*CLK1 inhibition by **3.5**, but the anilinopyrimidine scaffold has very low reversible affinity for the kinase. Somewhat surprisingly, neither **3.7** nor hybrids (**3.9** and **3.10**) exhibited activity against *Tb*CLK1 *in vitro* below 20 μM . This implied that the seemingly small change from chlorine in **3.5** to fluorine in **3.10** resulted

in the loss of activity. A crystal structure of a related chloro-anilinopyrimidine bound to EGFR showed the chlorine making a favorable hydrophobic contact with the methionine gatekeeper.¹⁷ Our results seemed to indicate a similar interaction is made in the active site of *Tb*CLK1, and disrupting this interaction has a strong effect on binding and potency of the anilinopyrimidine scaffold.

Table 3.1: Activity of anilinopyrimidines against *Tb*CLK1 *in vitro*

Cmpd #	IC ₅₀ (μM)		
	WT	C215A	C288A
3.5	1.8	> 20	2.1
3.6	> 20	> 20	> 20
3.7	> 20	> 20	
3.8	> 20		
3.9	> 20		
3.10	> 20	> 20	
3.11	> 20	> 20	

Kinases were incubated with compound and 100 μM ATP for 30 min before initiating reactions with substrate and $\gamma^{32}\text{P}$ -ATP. Phosphotransfer was quantified and normalized to DMSO control (n=2).

Despite its relatively low *in vitro* potency, we tested **3.5** and its non-electrophilic analog **3.6** in culture with BSF *T. brucei*. In a cellular proliferation assay, **3.5** and **3.6** demonstrated nearly equivalent activity (EC₅₀ ~6.5 μM, Fig. 3.6). The absence of a difference suggested that these molecules likely confer their trypanocidal effect through targets other than *Tb*CLK1 in an electrophile-independent mechanism. However, the C215-dependent *in vitro* activity of both the pyrazolopyrimidine and anilinopyrimidine series of compounds validated targeting C215 in the next generation of *Tb*CLK1 inhibitors.

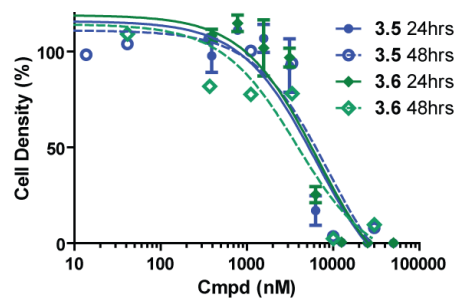


Figure 3.6: *T. brucei* cell proliferation assay with compounds **3.5** and **3.6**. Cells were incubated with compound at 37 °C. After 24 h (solid shapes, solid lines) or 48 h (hollow shapes, dashed lines), cell density was measure and normalized to untreated cells. (24 h: mean \pm s.d., n=3; 48 h: mean, n=2).

Using electrophilic inhibitors known to target the EGFR cysteine, we were quickly able to verify that reaction with C215 could be exploited in future inhibitors. In particular, it appeared that an acrylamide attached to a 4-aminopyrazolopyrimidine via a two-carbon linker could achieve the proper orientation in the active site to react with C215.

Section 2: Identification of AD57 and development of electrophilic analogs targeting C215 in *Tb*CLK1

To identify new potential scaffolds for developing *Tb*CLK1 C215-targeted electrophilic inhibitors, we screened more than 10,000 molecules from kinase-focused libraries against recombinant *Tb*CLK1. The most potent hit, AD57 (**3.25**), exhibited both potent *in vitro* inhibition of *Tb*CLK1 and potent cellular activity. Based on the structure of AD57, we developed a series of potent, C215 targeting covalent *Tb*CLK1 inhibitors.

Screening of kinase-focused libraries revealed AD57, a potent inhibitor of *Tb*CLK1.

The pursuit of kinase-directed therapeutics has resulted in thousands of potential inhibitors with a wide variety of chemical scaffolds for screening. By testing a subset of these inhibitors against *Tb*CLK1, we hoped to identify leads or general scaffolds for further development as *Tb*CLK1 inhibitors. These hits could then be further refined through the addition of a C215-directed electrophile into potent and selective inhibitors of trypanosome CLK kinases.

In collaboration with the NIH Chemical Genomics Center (NCGC), we conducted a high-throughput screen against *Tb*CLK1 using a kinase-focused library comprised of over 10,000 compounds. The NCGC adopted the luminescence-based ADP-Glo Kinase Assay (Promega) to measure kinase activity in the presence of the compounds. Surprisingly,

relatively few hits were identified for *Tb*CLK1 compared to previous studies with other kinases using the same library. Of the molecules that exhibited activity in the assay, it was not surprising that several of the most potent were electrophilic inhibitors developed to target the EGFR cysteine (**3.22**, **3.23**, and others; Appendix, 2 Figs. 1 and 2). We selected five compounds (NVEP-BEZ-235 (**3.20**), AZD-732 (**3.21**), neratinib (**3.22**), afatinib (**3.23**), and ponatinib (**3.24**); Fig. 3.7a) using AC_{50} (EC_{50} , NCGC) value, structural class, and commercial availability as filters for retesting in-house. In our more stringent assay, which included 100 μ M ATP as compared to only 5 μ M in the HTS format, these molecules exhibited between two- and 300- fold reductions in inhibitory activity (Fig. 3.7b). These molecules offered useful starting points for further development; however, a smaller in-house screen of select kinase inhibitors identified a more potent scaffold that we elected to pursue in developing electrophilic *Tb*CLK1 inhibitors.

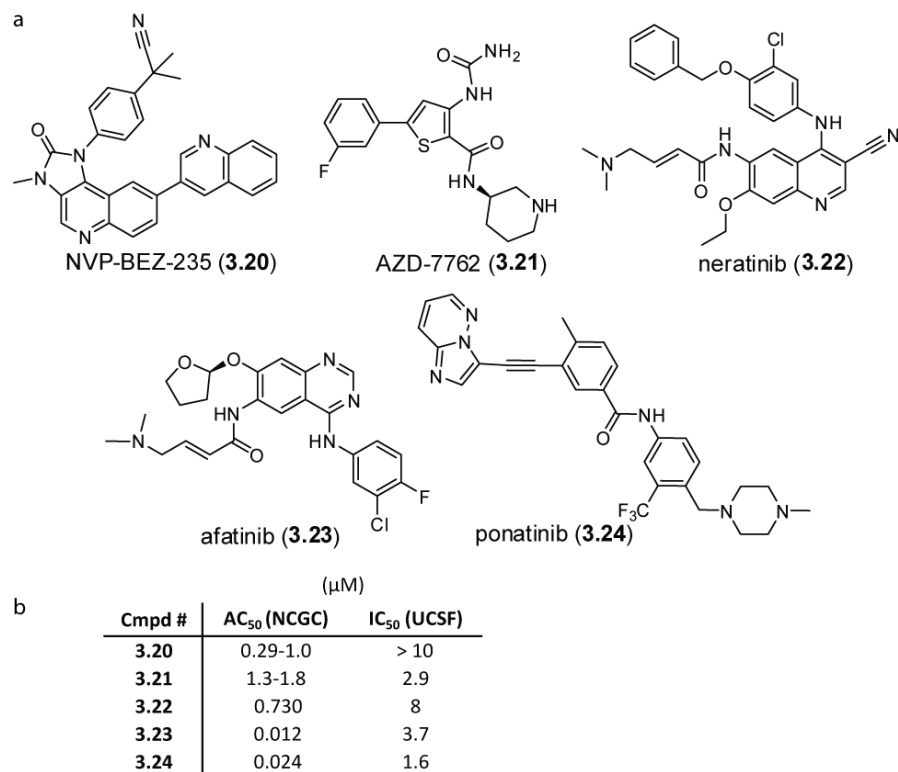


Figure 3.7: *Tb*CLK1 screening hits from the NCGC. **(a)** Chemical structures of kinase inhibitors selected for retesting in-house. **(b)** Table of AC₅₀ (EC₅₀, NCGC) and IC₅₀ (UCSF) values for the NCGC hits against recombinant *Tb*CLK1. [ATP] = 5 μM (NCGC), 100 μM (UCSF).

In parallel with the NCGC high-throughput assay, we screened a small library comprised of 23 kinase inhibitors synthesized in laboratories at UCSF¹ and six clinical-stage kinase inhibitors (Appendix 2, Fig. 3). This library surveyed a variety of ATP-competitive scaffolds and contained a number of known promiscuous inhibitors. At a concentration of 1 μM, only three molecules (AD57 (**3.25**), DJM1177 (**3.26**), and DJM2004 (**3.27**); Fig. 3.8a) exhibited greater than 50% inhibition in this single-point assay (Appendix 2, Fig. 3). This activity was confirmed when tested across a range of concentrations (Fig. 3.8b). AD57 was comparable in activity to hypothemycin (**1**, IC₅₀ = 150 nM), a covalent

¹ UCSF sourced molecules were synthesized by Dustin Maly (Shokat Laboratory), Arvin Dar (Shokat Laboratory), and Jeff Henise (Taunton Laboratory).

inhibitor, when tested under the same assay conditions. This lead compound was especially interesting in light of work with electrophilic analogs of PP2 presented in Section 1 of this chapter, suggesting that C215 could be accessed by electrophiles projected from the N1 position of the 4-aminopyrazolopyrimidine core of AD57.

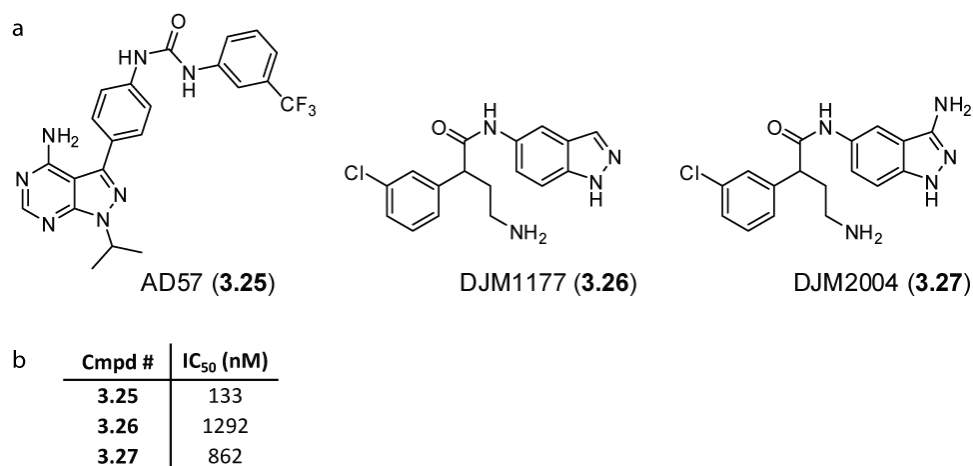


Figure 3.8: *Tb*CLK1 screening hits from UCSF. **(a)** Chemical structures of kinase inhibitors exhibiting greater than 50% inhibition of *Tb*CLK1 at 1 μ M (single point, n=1). **(b)** Table of IC₅₀ values for these hits against recombinant *Tb*CLK1 in dose response (n=2). Recombinant *Tb*CLK1 was incubated with compound and 100 μ M ATP for 30 min before initiating reactions with substrate and γ ³²P-ATP. Phosphotransfer was quantified and normalized to DMSO control.

To further explore 4-aminopyrazolopyrimidine scaffolds similar to AD57, we tested a series of molecules containing a variety of diaryl substitutions at the C3 position of the heterocycle (Fig. 3.9a, b) against *Tb*CLK1 *in vitro*. These assays revealed that the *para*-substitution of the *meta*-trifluoromethyl phenyl urea is essential for the activity exhibited by **3.25**. Deletion of the trifluoromethyl group (AD58, **3.28**; IC₅₀ = 4.1 μ M) reduced activity by more than 10-fold. Even more striking was that substitution of an amide for the urea (AD56, **3.29**) or *meta*-substitution of the aryl urea (AD53, **3.30**) resulted in molecules with negligible activity at the highest concentration tested (10

μM). This tight SAR was suggestive of a specific interaction between the diaryl urea in **3.25** and *Tb*CLK1. A similar trend is observed with these compounds and Src kinase and results from interaction with an allosteric pocket formed when the DXG motif flips “out” to accommodate the aryl urea. Though we cannot be certain without a crystal structure, the activity of **3.25** and SAR are consistent with this type II binding mode.

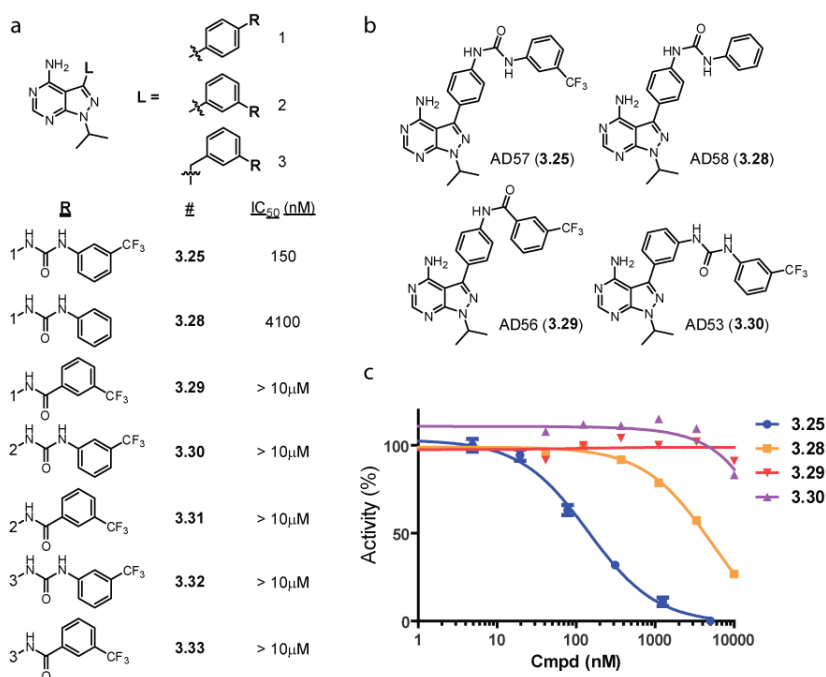


Figure 3.9: Inhibition of *Tb*CLK1 by analogs of AD57. **(a)** Structures of compounds and associated IC_{50} values for *in vitro* inhibition of *Tb*CLK1 **(b)** Complete structures of compounds studied in detail in this section, with **(c)** representative dose-response curves. Recombinant wild-type *Tb*CLK1 was incubated with compound and 100 μM ATP for 30 min before initiating reactions with substrate and $\gamma^{32}\text{P}$ -ATP. Phosphotransfer was quantified and normalized to DMSO control (mean, $n = 2$).

Structure-activity relationships suggest that AD57 confers its trypanocidal activity through *Tb*CLK1 inhibition.

Based on our results from RNAi knockdown of *Tb*CLK1 (Chapter 1, Fig. 2.4), we expected that AD57 would have strong trypanocidal activity in culture. We assayed the

compound's effect on BSF *T. brucei* using a Cell-Titer Glo based proliferation assay after 24h of treatment. In contrast to its potent *in vitro* activity, AD57 showed a relatively weak effect on cell proliferation (Fig. 3.10a, magenta circles; EC₅₀ ~4 μM). However, upon visual inspection, there was a clear cytokinetic defect in cells at concentrations lower than the 24 h EC₅₀ for proliferation/survival, with multiple cells still attached to together though a central point like a pinwheel (data not shown). The number of these cellular clusters was similar to the initial seeding density for the assay. The minimal difference between treated and untreated cells using Cell-Titer Glo assay, which measures ATP content to report cell number, suggested that the number of cell bodies was equivalent to what would be expected from normal growth and replication. We therefore extended the incubation time with **3.25** to 48 h. The additional 24 h of treatment was sufficient to cause cell death at much lower concentrations of **3.25** (Fig. 3.11a, blue squares; EC₅₀ ~200 nM).

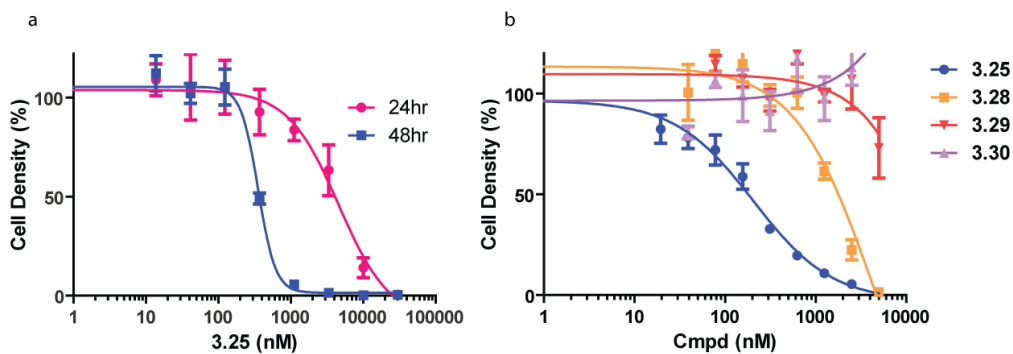


Figure 3.10: The trypanocidal effects of AD57 analogs mirror *in vitro* inhibition of *TbCLK1* after 48h of treatment. BSF *T. brucei* cells were incubated with compound. After 24 or 48 h, cell density was measured and normalized to untreated control (mean ± s.d., n=3). **(a)** Untransfected cells) were incubated with AD57 (**3.25**) for 24 or 48 h. **(b)** Cells haploid for *TbCLK1* (HA-*TbCLK1*/del) were incubated with compound for 48 h.

When AD58 (**3.28**), AD56 (**3.29**), and AD53 (**3.30**) were tested in the 48 h assay, the compounds' effect on cell proliferation/survival mirrored that seen previously in *in vitro* kinase assays with *Tb*CLK1 (Fig 3.9c, 3.10b). The analog lacking the *m*-trifluoromethyl group (**3.28**) exhibited moderate activity relative to AD57 (EC₅₀ ~2 μM), while the arylamide **3.29** and AD57 regioisomer **3.30** had negligible effects at the highest tested concentration of 10 μM (Fig. 3.10b). The correlation between *in vitro* and cellular data suggested that AD57 confers its potent trypanocidal effects through inhibition of cellular *Tb*CLK1.

Electrophilic analogs of AD57 potently inhibit *Tb*CLK1 in a C215 dependent manner.

Our earlier experiments described in Section 1 with electrophilic pyrazolopyrimidines **3.1**, **3.3**, and **3.4** suggested that an acrylamide attached to the pyrazolopyrimidine at N1 by a two carbon linker can react with C215 in the active site of *Tb*CLK1. To determine the optimal linker for an AD57-based scaffold, we synthesized a series of acrylamides linked through two carbons to the N1 of the pyrazolopyrimidine (Fig. 3.11).

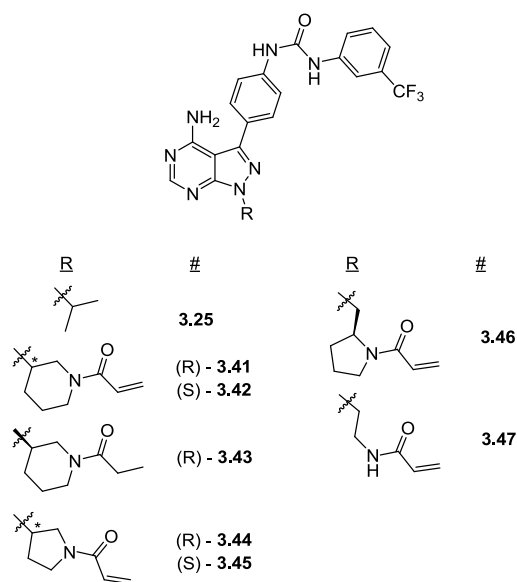


Figure 3.11: Chemical structures of electrophilic AD57 analogs with a variety of two-carbon linkers

We evaluated this new series of electrophilic compounds on the basis of two qualities important for their future utility in studying *TbCLK1*. The first was *in vitro* potency. Highly potent compounds can be used at low concentrations thereby reducing (non-covalent) inhibition of off-target kinases and suppressing non-specific alkylation of other cellular nucleophiles. The second criteria we used in evaluating compound efficacy was differential inhibition of wild-type and C215S *TbCLK1*. We hypothesized that the C215S mutation would confer cellular resistance to electrophilic C215-targeting *TbCLK1* inhibitors, thus providing a powerful chemical genetic tools to evaluate *TbCLK* specific functions in living cells

Table 3.2: Activity of acrylamide analogs of AD57 against *Tb*CLK1 *in vitro*

Cmpd #	IC ₅₀ (nM)		Selectivity (WT/C215S)	Chemical Structures			
	WT	C215S		B	#	B	#
3.25 (AD57)	143	68	< 1				
3.41	20	1486	74.3				
3.42	871	4391	5				
3.43	850	1150	1.4				
3.44	208	170	< 1				
3.45	196	68	< 1				
3.46	396	481	1.2				
3.47	108	506	4.7				

Kinases were incubated with compound and 100 μ M ATP for 30 min before initiating reactions with substrate and γ -³²P-ATP. Phosphotransfer was quantified and normalized to DMSO control (n=2).

Based on *in vitro* kinase assays, electrophilic AD57 derivative **3.41** met both of our criteria for further development and testing. **3.41** exhibited the greatest *in vitro* potency against wild-type *Tb*CLK1 (Table 3.2; IC₅₀ = 20 nM) and the best C215-specificity of all the electrophilic analogs in the series. The geometrically less constrained ethyl-linked **3.47** was also active against wild-type kinase (IC₅₀ = 108 nM), but was far less selective for the wild-type kinase than **3.41** (approximately five- vs. 75-fold more selective, respectively). The enantiomer of **3.41** (**3.42**) was also less potent (IC₅₀ = 396 nM) and less selective for the wild-type kinase than **3.41**. Interestingly, the pyrrolidine-linked compounds **3.44** and **3.45** did not exhibit cysteine dependent inhibition despite having the same linker length as other compounds. This emphasized the importance of the electrophile's spatial positioning within the active site for covalent bond formation. From this series, the 3(*R*)-piperidinyl linker was the most suitable for positioning the acrylamide to react with

C215. The non-electrophilic **3.43** was an important control and further demonstrated the importance of the covalent interaction between C215 and the acrylamide in **3.41**. There was no significant difference in activity of **3.43** against the wild-type and C215S mutant kinase (IC_{50} = 850 nM and 1.2 μ M, respectively).

We next tested compound **3.41** for its effects on *Tb*CLK1 expressed in cells. We found that incubation with 2 μ M **3.41** for 4 h was sufficient to suppress autophosphorylation of HA-tagged wild-type *Tb*CLK1 expressed from its endogenous locus in BSF *T. brucei* (discussed in Chapter 1) such that HA-*Tb*CLK1 appeared as a single band on a denaturing gel (Fig. 3.12). When the treated cells were washed and allowed to recover in drug-free media, the level of autophosphorylation returned to control levels within 1.5 h, presumably through resynthesis of *Tb*CLK1. In cells expressing HA-tagged C215S *Tb*CLK1, 2 μ M **3.41** was unable to completely suppress auto-phosphorylation as observed with the wild-type kinase. After washout, the C215S kinase returned to normal levels of phosphorylation within approximately 30 min, at least three-times faster than wild-type kinase. The reduced sensitivity of the C215S mutant to **3.41** is consistent with the C215-dependent covalent mechanism observed *in vitro*.

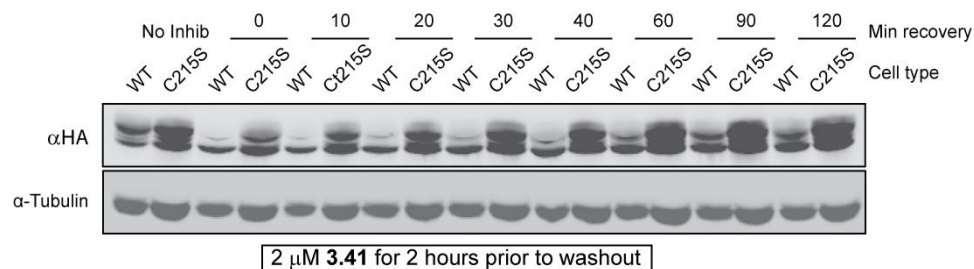


Figure 3.12: Inhibition of *Tb*CLK1 autophosphorylation. Cells expressing wild-type or C215S HA-tagged *Tb*CLK1 were treated with 2 μ M **3.41** for 4 h and then allowed to recover in fresh media without inhibitor for up to 2 h. Lysates were generated, resolved on a denaturing gel, blotted, and probed for HA and α -tubulin. (Nathan Gushwa)

To determine if *Tb*CLK1 C215S was sufficient to rescue cells treated with **3.41**, cells expressing a single copy of HA-tagged wild-type or C215S *Tb*CLK1 (HA-*Tb*CLK1/deletion) were treated with **3.41** or its non-electrophilic analog **3.43**. Acrylamide **3.41** was potently cytotoxic to cells expressing wild-type *Tb*CLK1 (EC_{50} = 123 nM; Fig. 3.13a, blue circles), while **3.43** was significantly less active (EC_{50} = 4.3 μ M; Fig. 3.13b, blue circles), indicating an electrophile-dependent mechanism consistent with *in vitro* assays (Table 3.2). However, the C215S mutation did not confer measurable resistance to **3.41** (EC_{50} = 105 nM; Fig. 3.13a, red squares). This suggested the involvement of other targets and/or a more complex interaction between *Tb*CLK1 and *Tb*CLK2. It is possible that inhibition of *Tb*CLK2 may have a dominant negative effect on C215S *Tb*CLK1 activity in a heterodimer. Preliminary results in pull-down assays suggest there is an intermolecular interaction between *Tb*CLK1 and *Tb*CLK2. To test for *Tb*CLK2-mediated toxicity of **3.41**, cell lines simultaneously expressing *Tb*CLK1 C215S and *Tb*CLK2 C234S are currently being generated.

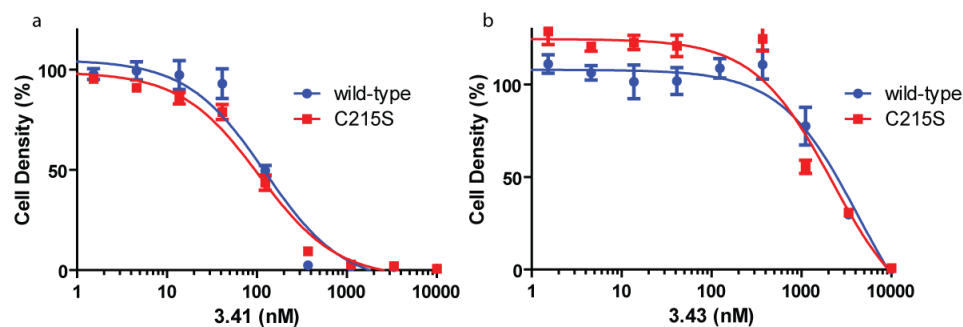


Figure 3.13: *T. brucei* cell proliferation assay with **3.41** and **3.43**. Cells expressing wild-type (blue circles) or C215S (red squares) HA-tagged *TbCLK1* (HA-*TbCLK1*/del) were incubated with **3.41** (a) or **3.43** (b). After 48 h, cell density was measured and normalized to untreated cells. (mean \pm s.d., n=3).

With the goal of generating more selective covalent inhibitors of wild-type *TbCLK1*, we synthesized analogs of **3.41** containing alternatives to the *meta*-trifluoromethyl phenyl urea. This series of compounds (Fig. 3.14a) tested whether substitutions around the phenyl urea or if urea itself were necessary for *TbCLK1* inhibition. To test whether **3.41** acts as a type II inhibitor, two substituted isoxazoles were also included. Similar 3-isoxazolyl ureas are found in clinically relevant type II inhibitors quizartinib¹⁹ and CEP-32496²⁰ (Fig. 3.14b).

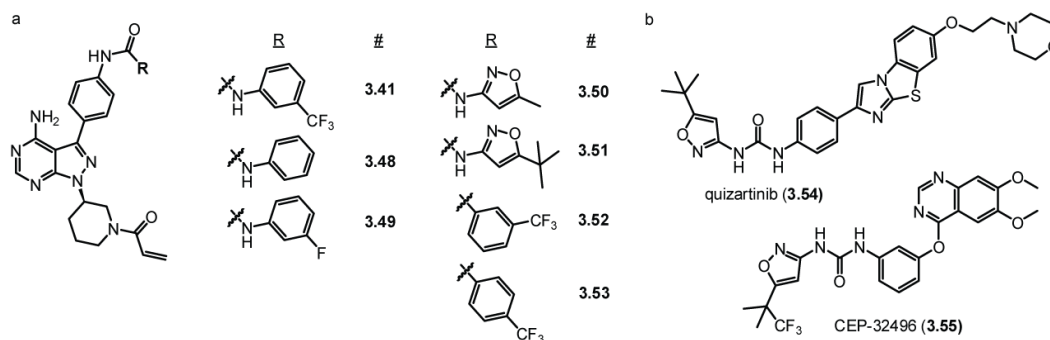
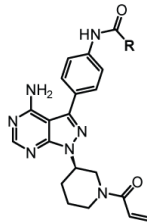
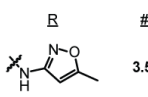
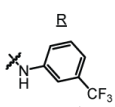
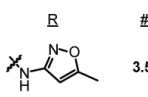
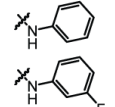
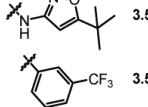

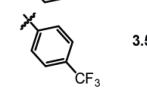
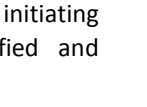


Figure 3.14: Second generation electrophilic analogs targeting C215 in *TbCLK1*. **(a)** Chemical structures of analogs of **3.41**. **(b)** Chemical structures of 3-isoxazolyl urea containing type II kinase inhibitors, quizartinib (**3.54**) and CEP-32496 (**3.55**).

In this series, only compounds with diaryl ureas displayed any significant activity ($IC_{50} < 10 \mu M$) against *TbCLK1*. In the crystal structure of AD57 (**3.25**) bound to c-Src, the urea nitrogen distal to the pyrazolopyrimidine makes a critical hydrogen bond essential for type II binding.^{21,22} This further supports a type II binding mode for AD57 and these aryl ureas in the *TbCLK1* active site. Of the ureas, all but **3.50** exhibited sub-micromolar activity against the wild-type kinase *in vitro* (Table 3.3). Though slightly less potent than **3.41**, **3.48**, and **3.49** were more selective for the wild-type than the C215S mutant. Therefore, based on our criteria for biochemical utility (potency and C215 selectivity), these new molecules may serve as better tools to study *TbCLK1* as compared to **3.41**.

Table 3.3: Activity of analogs of **3.41** against *Tb*CLK1 *in vitro*

Cmpd #	IC ₅₀ (nM)		Selectivity (WT/C215S)		#		#
	WT	C215S					
3.41	20	1486	74.3		3.41		3.50
3.48	162	> 10 μM	> 60		3.48		3.51
3.49	46	> 10 μM	> 215		3.49		3.52
3.50	5193	> 10 μM	> 2				3.53
3.51	67	3467	52				
3.52	> 5 μM	> 5 μM	-				
3.53	> 5 μM	> 5 μM	-				

Kinases were incubated with compound and 100 μM ATP for 30 min before initiating reactions with substrate and γ ³²P-ATP. Phosphotransfer was quantified and normalized to DMSO control (n=2).

Discussion

The goal of this project was to develop potent and selective inhibitors of *Tb*CLK1 to study the function of this enzyme in living cells. In Chapter 1 we used RNAi mediated knockdown of *Tb*CLK1 expression to show that the protein is essential. A potent, selective, ATP-competitive inhibitor would allow us to investigate the kinase function of endogenously expressed *Tb*CLK1 on the minute time scale rather than hours or days required for RNAi and other gene-mediated techniques. By targeting *Tb*CLK1 C215 and *Tb*CLK2 C234, the only homologs of the EGFR cysteine in the trypanosome kinome, we felt we could quickly obtain inhibitors with exquisite selectivity for this fascinating kinase. In addition to the potential boost in selectivity, we gain a powerful chemical genetic tool for biological studies. By mutating the cysteine to a non-nucleophilic residue, we anticipate that we can pharmacologically isolate and define roles for both *Tb*CLK1 and *Tb*CLK2.

In our initial experiments, we evaluated electrophilic inhibitors developed specifically to target the EGFR cysteine in human EGFR and BTK kinases. This allowed us to verify reactivity of *TbCLK1* C215 with biologically useful electrophiles and, in the absence of a crystal structure, to determine its predicted position in the active site. Two compounds (**3.4** and **3.5**) exhibited modest, but C215-dependent inhibition of *TbCLK1* *in vitro*. The modest potency was likely due to the low reversible binding affinity of these scaffolds. To address this problem, we screened a large library of compounds to identify new leads for developing new electrophilic inhibitors.

Our compound screen led us to AD57 (**3.25**), which exhibited nanomolar activity in *TbCLK1* *in vitro* kinase assays. When incubated with trypanosomes, AD57 resulted in a cytokinetic defect at 24 h and cytotoxicity at 48 h. We utilized a number of structural analogs of AD57 in the proliferation assay and *in vitro* assay to establish structure activity relationships that suggested the trypanocidal activity of AD57 is mediated by *TbCLK1*. Previous studies have shown that AD57 is a reversible type II, or “DFG-out,” inhibitor of c-Src tyrosine kinase.²¹ Our data in conjunction with unpublished SAR on AD57 analogs led us to believe that, like c-Src, *TbCLK1* is inhibited by AD57 while in a DFG-out conformation.

Despite the potent inhibition of *TbCLK1* in *T. brucei* cells, we still lacked indisputable evidence that cytotoxicity resulted from specific inhibition of *TbCLK1*. AD57 is a promiscuous inhibitor of many human kinases (Appendix 2, Table 1) at least one of which has homology to essential trypanosome kinases. Aurora kinase is highly conserved through evolution with regard to function, and inhibitors developed for the human Aurora B have shown activity against *T. brucei* Aurora kinase 1 (*TbAUK1*).^{28,29} Given the conservation in function and pharmacologic sensitivity, AD57 may also be acting on trypanosome aurora kinases in addition to *TbCLK1*. RNAi mediated knockdown of a trypanosome aurora kinase, *TbAUK1*, and inhibition by small molecules in BSF *T. brucei* resulted in defective nuclear division and cytokinesis.²⁸ In these experiments, Jetton and coworkers observed cells with multiple nuclei, flagella, and kinetoplasts, a phenotype inconsistent with the phenotype that we observed after treatment with AD57 for 24 h. Despite this difference, it is nearly impossible to rule out trypanosome Aurora kinases and other non-CLK targets of AD57, all of which may contribute to the observed phenotype. For this reason, we continued to pursue selective inhibitors that could be combined with genetically engineered cells to unequivocally prove that pharmacological inhibition of *TbCLK1* is sufficient to cause a growth defect and to study the roles of *TbCLK1* in cells.

To develop more potent and selective inhibitors, we sought to exploit the non-conserved EGFR cysteine (C215) in the *TbCLK1* active site by synthesizing electrophilic

analogs of AD57. We found that these compounds, exemplified by acrylamide **3.41**, substantially increased potency over previously tested electrophilic inhibitors and displayed excellent selectivity (>50-fold), for wild-type over C215S *TbCLK1*. Unfortunately virtually no difference was observed when cells expressing a single copy of wild-type or C215S *TbCLK1* were treated with these AD57-derived electrophiles. While the C215-targeted approach to inhibitor development resulted in potent, allele-selective inhibitors, additional work needs to be performed to optimize the compounds and biological tools for a detailed chemical genetic study of trypanosome Clk kinases.

Some of the problems associated with the *in vitro* to *in vivo* translation of our electrophilic compounds may be the result of our narrow focus on *TbCLK1* rather than on both *TbCLK1* and *TbCLK2*. Despite no evident phenotype from isoform specific RNAi mediated knockdown (Chapter 1, Fig. 2.4) or a knockout of *TbCLK2* (Appendix 2, Fig. 4), compensatory mechanisms and overlapping functions of the two Clk kinases in our stable cell lines may have been overlooked. The cellular data presented in figures 3.10b and 3.14 were derived from cell lines with modifications in at least one *TbCLK1* locus, and thus may no longer express two fully functional copies of *TbCLK1*. The kinase domains of *TbCLK1* and *TbCLK2* are completely identical raising the possibility of haplo-insufficiency and potential compensation by *TbCLK2*. In this context, *TbCLK2* may be an essential enzyme. Additionally, due to the complete sequence conservation in the catalytic domains of the trypanosome Clks, it is highly unlikely that our inhibitors

differentiate between the two kinases. To isolate their activities, we are currently developing cell lines expressing only wild-type or C215S *TbCLK1* in the context of mutant or deleted *TbCLK2*. Another possibility is that cell proliferation is not dependent on *TbCLK1* catalytic activity which is not differentiated from potential protein functions using RNAi. There are many examples where chemical inhibition does not result in the same phenotypes as genetic knockdown.¹² We will continue our efforts to validate *TbCLK1* as a therapeutic target, and believe this is possible using the allele selective inhibitors we have presented here in conjunction with new cell lines under development.

This project brought us to the cusp of a detailed exploration of *TbCLK1* biology using small molecule inhibitors and chemical genetics. We have described the development cell-active inhibitors that rely on the presence of the unique, non-catalytic cysteine C215 in *TbCLK1*. These compounds potentially represent powerful chemical tools to study the essential roles of *TbCLK1* and *TbCLK2* kinases play in this human pathogen. With one piece of the puzzle done, we await the creation of the new *TbCLK1/TbCLK2* mutant cells lines to begin this investigation. These cells will not only provide a system to isolate the activities of each of the *TbCLKs*, but may also help address some lingering questions about compensatory mechanisms in our current cell lines.

Methods

Culturing blood stream form (221) *Trypanosoma brucei*

Blood stream form parasites (strain 221) were cultured at 37 °C and 5% CO₂. Standard media consisted of HMI-9 medium (Cell Culture Facility, UCSF, CA, USA) containing 10% Fetal Bovine Serum, 10% Serum Plus™ medium supplement (SAFC, Lenexa, KS, USA), 100 units/ml penicillin, and 100 µg/mL streptomycin or HMI-9 medium (Cell Culture Facility, UCSF, CA, USA) containing 10% Fetal Bovine Serum, 100 units/ml penicillin, and 100 µg/ml streptomycin. There was no observable difference between the two culture mediums. Transgenic cell lines were maintained in the medium supplemented with appropriate antibiotics (2.5 µg/mL G418, 5.0 µg/mL hygromycin, 5.0 µg/mL blasticidin, 2.5 µg/mL phleomycin, 0.1 µg/mL puromycin; Sigma-Aldrich, St. Louis, MO, USA). Cultures were initiated with 1×10⁵ cells/mL and subcultured when cell density reached 1-2×10⁶ cell/mL.

24 h *T. brucei* cell proliferation assay

The end point luciferase-based assay for *T. brucei* cell density was done as previously described (Mackey ZB *et al.*, *Chem Biol Drug Des.* (2011)) with the following modifications. BSF parasites were introduced at 5×10⁵ cells/mL in media absent of selective antibiotics and treated with compound under standard culture conditions. Final DMSO concentration was 1% (v/v). End point cell density was measured after 24 h

using CellTiter-Glo[®] Luminescent Cell Viability Assay (Promega, Madison, WI, USA). Luminescence was measured using a Spectramax M5 plate reader (Molecular Devices, Sunnyvale CA). Results were averaged and fit to a (normalized response curve) using GraphPad Prism (GraphPad Software, Inc., La Jolla, CA, USA).

48 h *T. brucei* cell proliferation assay

Compound (400x) in DMSO was diluted with water to a 100X solution (25% DMSO). This solution was mixed with cells at 5×10^4 cells/mL (0.25% DMSO). End point cell density was measured after 48 h using CellTiter-Glo[®] Luminescent Cell Viability Assay (Promega, Madison, WI, USA). Luminescence was measured using a Spectramax M5 plate reader (Molecular Devices, Sunnyvale CA). Results were averaged and fit to a (normalized response curve) using GraphPad Prism (GraphPad Software, Inc., La Jolla, CA, USA).

Bloodstream form transfection

Transfection was done as previously described (Tu, X. & Wang, C. C. *J Biol Chem* (2004)) with some modifications. 20×10^6 cells/mL (instead of 10×10^6) were used for each transfection. 10 μ g of DNA (linearized plasmid or linear PCR product) was used instead of 100 μ g.

Primers used (listed 5'→3')

*Tb*CLK1 nucleotide sequence

ATGGATAGCATAAGCTCTGGTCTAACCACAGAGCTGGTCTCCAGCAATACCCCTCAGGACTAGTAAATGTCCACGTAAACAACAATCCCAAATCA
ACAACGGCGCATGTTCCAACATGACACCGAGGGAATCCACAAACAGCGCCCGTCTCGGGAAGCAAACGTGACAGAGAAACAGCCGACACTGCT
GACGACGCGCCGCAACAAAACAATCAATCGATCACAGGAACAACCGCCACCACGACGAGTACTGCAAAACACAGCGAAATGCAACCACCACAAA
GAAAAAGAAGGTGACGTATGCCCTGCCAAATCAGAGTCGTGAGGAGGGTCACTTCTACGTTGTGTTGGGAGAAGATATCGATGCATCAACGGGTC
GCTTTAAAATACTATCGCTGCTCGGTGAGGGTACATTCGGGAAAGTTGTGGAGGCATGGGACCGGAAGCGGAAAGAGTATTGTGCGGTAAAAAT
AGTACGTAATGTGCCAAAATCACTCGCGATGCGAAAATTGAAATTCAGTTCATGGAACGTGTTGACTCTCCGATGTGGAGGACCGTTCCCTTTA
ATGAAAATCCAGCGTTACTTCCAGAATGAAACGGGTACATATGTGTATTGTGATGCCAAGTACGGACCATTTATTAGATTGGATAATGAAACAT
GGTCCTTTCAATCATCGCCATTTGGCACAGATTATCTTCAAGTTGGTGGCGGACTGGATTACTCCACACGGAACCTCATCTCATGCATACGGACCT
GAAGCCGGAGAATATTCTAATGGAATCTGGTGATACTTCGGTAGACCCAATGACACACAGAGCGTTACCCCTGAACCATGCCGTGCGCATTG
TGATCTCGGCGGTTGTTGTGACGAGCGCCATAGCCGCACAGCAATTGTGCAACACGTCCTACTCCGAGTCCCTGAAGTTGTGTTGAGCCTTGGATG
GATGTATTCAACAGATCTGTGGTTCGATGGGATGTATCATTATGAACTTTATACTGGGAAACTGTTGTATGATACTCATGACAATTTGGAGCACCTG
CACTTAATGGAGAAGACGCTTGGACGCTTCTGCCGACTGGTCTGTGCGGTGTGGCACGAAGAGGCTCGTGACCTCTTACTGCTGCAGGTACG
TTACAACCATGCAAAGATCCCAAACATATCGCTCGCATTGCACGTGCCCGCCCGTGGCGGAAAGTTATTACTGAACCACTTTTGTGTGACCTCATTCT
AAACTGTGCTGATTATGACCGTCAGAGGCGTCTGAATGCTCGACAAATGATGAGTACGCGTATGTGCATAAATACTTTCCCGAGTGGCCGACGA
TCCCAACCACGTGGACAACCGCTCTAAACTCCCCCAACACCTGTATGTAA

M209: blue

C215: magenta

C288: yellow

C288A: GAACCATGCCGTGTGCGCATTGCTGATCTCGGCGGTTGTTGTGACGAGC

C215A: GTGATGCCCAAGTACGGACCAGCCTTATTAGATTGGATAATG

C215S: GTGATGCCCAAGTACGGACCAGCCTTATTAGATTGGATAATG

M209A: GGTCATATGTGTATTGTGGCCCAAGTACGGACCATGC

M209G: GTCATATGTGTATTGTGGCCCAAGTACGGACCATGC

M209T: GGTCATATGTGTATTGTGACCCCAAGTACGGACCATGC

M209C 1: GGTCATATGTGTATTGTGTGCCCAAGTACGGACCATGC

2: GGTCATATGTGTATTGTGTGCCCAAGTACGGACCATGC

C206T, M209A: GAAACGGGTCATATGACTATTGTGGCGCCCAAG

C206V, M209A: GAAACGGGTCATATGGTTATTGTGGCGCCCAAG

Primers for pC vector creation:

CLK1 5' KpnI with no start: CTGGTACC GATAGCATAAGCTCTGGTCTAAC

CLK2 5' primer (Kpn I with no ATG):

CTGGTACCGCTTCATACATAGCGTCAAACGTTAACTCATCC

CLK1/2 3' XhoI with no stop: GACTCGAGCATGACAGGTGTTGGGGGGAG

Primers for CLK1 knockout construction:

5' primer: CCACGGCGATTTACCCGAA

5' G418 linker: TACTACAAAACAACCACAATGATTGAACAAGATGGATTGCAC

3' G418 linker: CTTCTTGACGAGTTCTTCTGATCCGTTGTTAACTAATCTGTGAG

3' Puro-linker: GACCCGCAAGCCCGGTGCCTGATCCGTTGTTAACTAATCTGTGAG

5' purolinker: TACTACAAAACAACCACAATGACCGAGTACAAGCCCAC

5' BLA linker: TACTACAAAACAACCACAATGCCTTTGTCTCAAGAAG

3' BLA LINKER: GGTTATGTGTGGGAGGGCTAATCCGTTGTTAACTAATCTGTGAG

3' primer: CCTCCACCGTAACGGAAAGTC

Primers for CLK2 knockout construction:

5' intergenic sequence: CGGTGCTGGGTAAACTGTTATC

5' Blast linker: GGAACATTTATAGAACCACATGCCTTTGTCTCAAGAAG

3' Blast linker: GTTATGTGTGGGAGGGCTAATCCGTTGTTAACTAATC

5' Puro linker: GGAACATTTATAGAACCACATGACCGAGTACAAGCCCAC

3' Puro-linker: GTTATGTGTGGGAGGGCTGATCCGTTGTTAACTAATCTGTGAG

3' intergenic sequence: GAACAACCTGGCAGGGGAATG

***Tb*CLK1 autophosphorylation assay**

Cells (10×10^6) expressing HA (hemagglutinin)-tagged *Tb*CLK1 were treated with compound (0.1% DMSO) for 1-4 h at 37 °C, unless otherwise noted. The treated cells were then pelleted with centrifugation, washed with trypanosome diluting buffer (5 mM KCl, 80 mM NaCl, 1 mM MgSO₄, 20 mM Na₂HPO₄, 2 mM NaH₂PO₄, 20 mM glucose, pH 7.4), pelleted with centrifugation, and lysed with 2X Laemmli sample buffer (1X: 60 mM Tris-Cl pH 6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.01% bromophenol blue).

The resulting lysate was sonicated to reduce viscosity, and then was resolved on an SDS-PAGE gel and transferred to nitrocellulose. *TbCLK1* was detected using anti-HA primary antibody (1:1000 - 12CA5, Roche) or rabbit polyclonal (1:1000 - H6908, Sigma Aldrich); IRDye conjugated anti-mouse (1:10,000, Li-Cor Biosciences) or anti-rabbit (1:10,000, Li-Cor Biosciences); and the Odyssey Imaging System (Li-Cor Biosciences, Lincoln, Nebraska).

For washout experiments, the treated cells were pelleted and resuspended in fresh media not containing compound. The cells were allowed to recover at 37 °C. Collection, lysis, and detection were performed as above.

***In vitro* kinase assay (*TbCLK1*)**

TbCLK1 (WT: 0.8 nM; C288A: 2.5 nM; C215S: 0.5 nM; C215A: 0.5 nM) was incubated with compound in reaction buffer (50 mM Tris pH7.4, 10 mM MgCl₂, 0.2 mM EGTA, 0.2 mg/mL BSA, 1 mM DTT) and 100 μM ATP for 30 min at rt. A solution of kinase reaction buffer with γ ³²P-ATP (70-150 μCi/mL, Perkin Elmer) and myelin basic protein (0.5 mg/mL, Sigma-Aldrich) was added to initiate the kinase reaction. After 30 min, the reactions were spotted onto phosphocellulose paper, washed once with 10% acetic acid, twice with 1% phosphoric acid, and once with methanol, and dried. Kinase activity was quantified using a Typhoon Imaging System (Molecular Dynamics) and ImageQuant 5.2 (Molecular Dynamics). IC₅₀ values and dosing curves were generated using GraphPad Prism 5 (GraphPad Software).

Recombinant *Tb*CLK1 expression and purification

The open reading frame for *Tb*CLK1 was amplified from genomic DNA using primers listed below and cloned into the pET100 expression vector (Invitrogen). Quickchange PCR was used to insert point mutations. The primers listed above used were also used for mutating the homologous recombination vector.

*Tb*CLK1 was expressed in *E. coli* (BL21 DE3) transformed with pET100 vector in LB containing ampicillin, chloramphenicol, and isopropyl β -D-1-thiogalactopyranoside (IPTG, 0.4 μ M) at 20 °C overnight with agitation. Cells were concentrated in lysis buffer (50 mM tris pH 8.0, 300 mM NaCl, 10mM imidazole, 10% glycerol, 1 mM CaCl₂, and 1 mM MgCl₂) and lysed with mechanical shearing (microfluizder). The lysates were clarified at 10,000g. The clarified lysate was incubated at 4 °C for 3 h with rotation with Ni-NTA agarose (Qiagen). The immobilized proteins were eluted with elution buffer (50 mM tris pH 8.5, 300 mM NaCl, 350mM imidazole, 10% glycerol, 1 mM CaCl₂, and 1 mM MgCl₂). The eluent was further purified by gel filtration (Superdex 75 10/300 GL, GE Healthcare) in to a minimal storage buffer (20 mM Tris pH 7.5, 50 mM NaCl, and 5% glycerol). The protein was concentrated to by centrifugation (Amicon 10kDa MWCO centrifugal filter unit, Millipore). The concentrated solutions of protein were aliquoted, frozen with liquid nitrogen, stored at -80 °C, and thawed immediately before use.

Primers for extraction of *Tb*CLK1 open reading frame, listed 5' \rightarrow 3'.

N-terminal primer: CACCATGGATAGCATAAGCTCTGGTCTAAC

C-terminal primer: TTACATGACAGGTGTTGGGGGAGTTTAGA

Lambda phosphatase treatment

Samples were dephosphorylated with lambda protein phosphatase (New England Biosciences) according to their protocol (1x NEBuffer for PMP, 1 mM MnCl_2) for 30 min at 30 °C. When dephosphorylating recombinant kinase, bovine serum albumen was included at 0.8 mg/mL. When the dephosphorylated *TbCLK1* was used in kinase assays, phosphatase inhibitors (PhosStop Phosphatase Inhibitor Cocktail, Roche) were included in all buffers following dephosphorylation.

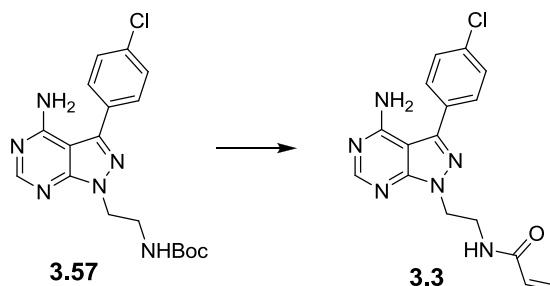
Generation to *T. brucei* whole-cell lysate

T. brucei cells were pelleted, washed with trypanosome diluting buffer (5 mM KCl, 80 mM NaCl, 1 mM MgSO_4 , 20 mM Na_2HPO_4 , 2 mM NaH_2PO_4 , 20 mM glucose, pH 7.4), pelleted with centrifugation, and lysed with lysis buffer (10 mM MgCl_2 , 20 mM Tris 7.5, 0.5 mM DTT, 0.25% NP-40, 100 mM NaCl, 1X Complete Protease Inhibitor Cocktail-EDTA (Roche)). The insoluble material was then pelleted at 25,000 g. The clarified lysate (supernatant) was removed, frozen with liquid nitrogen, and stored at -80 °C.

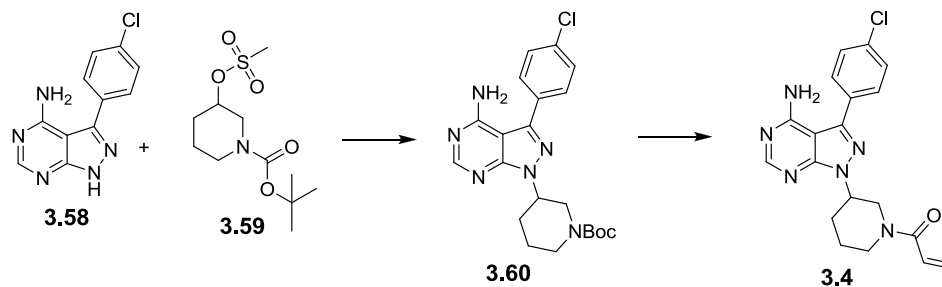
Synthetic Methods

^1H NMR spectra were recorded on a Varian INOVA-400 400 MHz spectrometer. Chemical shifts are reported in δ units (ppm) relative to residual solvent peaks. Coupling constants (J) are reported in hertz (Hz). All reagents and solvents were purchased from Aldrich Chemical, Acros Organics, Frontier Scientific, or Cambridge Isotopes Laboratory and used as received. Column chromatography was carried out using a Biotage SP1 flash chromatography system and silica gel cartridges from Silicycle, unless otherwise noted. Analytical TLC plates (Silica Gel 60 F₂₅₄) from EMD Chemicals (Gibbstown, NJ) were employed for TLC analyses and preparative TLC. Preparative HPLC purifications were performed using a Biotage Parallel Flex equipped with Waters Xbridge 19 x 50 mm C₁₈ 5 μM OBD columns and methanol/water with 0.01% trifluoroacetic acid mobile phase. Analytical HPLC separations were performed using a Waters 2795 Separations Module equipped with Waters Xbridge 4.6 x 50 mm C₁₈ 5 μM columns and continuously analyzed with a Waters 2424 Evaporative Light Scattering Detector, a Waters 2996 Photodiode Array Detector, and a Waters Micromass ZQ Mass Spectrometer. For analytical samples (LCMS), a ramping (5-95% over 3.5 min) mobile phase of methanol in water with 0.2% formic acid was used. For all compounds tested in biological assays, purity by light scattering and absorption at 254 nm was greater than 95%, unless otherwise noted.

PP2 electrophiles



3-(*p*-chlorophenyl)-4-aminopyrazolopyrimidine ethyl-linked acrylamide 3.3: The boc-protected amine **3.57** (JM4-52, 36 mg, 0.09 mmol) was dissolved in 4 M hydrochloric acid in dioxanes (1 mL). After 3 h, the reaction mixture was concentrated to a white solid. This material was then mixed with acrylic acid (6.6 mg, 6.33 μ L, 0.09 mmol), EDC-HCl (21.2 mg, 0.11 mmol), HOBT (15.0 mg, 0.11 mmol), and triethylamine (28 mg, 38.5 μ L, 0.28 mmol) in methylene chloride (1 mL). After 4 h, the reaction was diluted with additional methylene chloride, washed with water and brine, dried over sodium sulfate, and concentrated. The crude was then purified by silica flash chromatography (4 g SiliCycle silica flash column) and a gradient of methanol in ethyl acetate (0-5%). The pure fractions by LCMS containing the product were collected and dried affording 9 mg (31% yield) of a white solid. ^1H NMR (400 MHz, DMSO- d_6) δ ppm 3.61 (q, $J=6.17$ Hz, 2 H) 4.43 (t, $J=6.23$ Hz, 2 H) 5.55 (dd, $J=9.71, 2.93$ Hz, 1 H) 6.04 (dd, $J=17.03, 2.56$ Hz, 2 H) 6.12 (dd, $J=17.21, 9.71$ Hz, 2 H) 7.59 (dt, $J=8.79, 2.20$ Hz, 2 H) 7.67 (dt, $J=8.61, 2.20$ Hz, 2 H) 8.22 (t, $J=5.86$ Hz, 1 H) 8.24 (s, 1 H). LCMS: 342.99 (M+H); MW: 342.10.



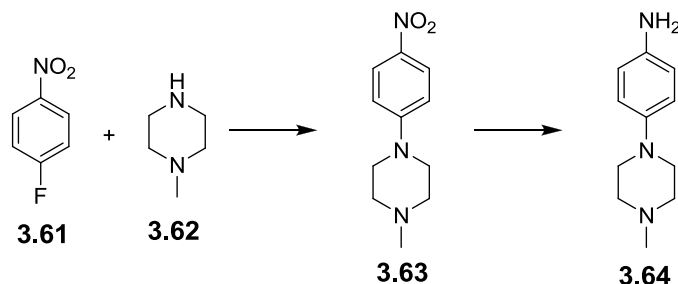
Boc-protected piperidine 3.60: 3-(*p*-chlorophenyl)-4-aminopyrazolopyrimidine 3.58

(JM2-45, 500 mg, 3.04 mmol), cesium carbonate (995 mg, 3.05 mmol), and mesylate **3.59** (853 mg, 3.05 mmol) were dissolved in DMF (5 mL). The mixture was heated to 80 °C and stirred overnight. The mixture was then diluted into ethyl acetate (80 mL), washed with water and brine. The organic layer was dried over sodium sulfate and concentrated. The crude was dry loaded on silica using a mixture of THF and methanol and purified by silica flash chromatography (25 g Biotage silica flash column) using a gradient of ethyl acetate in hexanes (0-100%). 303 mg (35% yield) of white solid was obtained. ¹H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.44 (br. s., 9 H) 1.63 - 1.78 (m, 1 H) 1.86 - 2.06 (m, 3 H) 2.25 (s, 2 H) 2.87 (td, *J*=12.64, 1.83 Hz, 2 H) 3.42 (br. s., 1 H) 4.85 (ddd, *J*=15.34, 10.76, 4.30 Hz, 1 H) 5.68 (br. s., 1 H) 7.52 (dt, *J*=8.79, 2.20 Hz, 2 H) 7.64 (dt, *J*=8.61, 2.20 Hz, 2 H) 8.37 (s, 1 H). LCMS: 429.12 (M+H); MW: 428.17.

3-(*p*-chlorophenyl)-4-aminopyrazolopyrimidine piperidinyllinked acrylamide 3.4: The boc-protected amine **3.60** (75 mg, 0.18 mmol) was dissolved in 4 M hydrochloric acid in dioxanes (2 mL). After 2 h, the reaction mixture was concentrated to a white solid. This

material was then mixed with acrylic acid (14.8 mg, 14.1 μ L, 0.21 mmol), EDC-HCl (47.2 mg, 0.25 mmol), HOBT (33.3 mg, 0.25 mmol), and triethylamine (62.3 mg, 85.9 μ L, 0.62 mmol) in methylene chloride (1 mL). After 6 h, the reaction was diluted with additional methylene chloride, washed with water and brine, dried over sodium sulfate, and concentrated. The crude was then purified by silica flash chromatography (4 g SiliCycle silica flash column) and a gradient of methanol in ethyl acetate (0-5%). Then resulting material was further purified by preparatory TLC using a 5% methanol in ethyl acetate mobile phase. 4 mg (6% yield) of a white solid was obtained. ^1H NMR (400 MHz, DMSO- d_6) δ ppm 1.49 - 2.35 (m, 2 H) 2.96 - 3.27 (m, 1 H) 3.52 - 3.77 (m, 0 H) 4.19 (d, $J=13.19$ Hz, 0 H) 4.49 - 4.78 (m, 0 H) 5.71 (d, $J=10.26$ Hz, 0 H) 6.09 (dd, $J=27.29, 17.40$ Hz, 1 H) 6.86 (dd, $J=16.48, 10.62$ Hz, 1 H) 7.59 (dt, $J=8.61, 2.01$ Hz, 2 H) 7.68 (d, $J=8.24$ Hz, 2 H) 8.26 (s, 1 H). LCMS: 382.13 (M+H); MW: 382.13.

Anilinopyrimidines



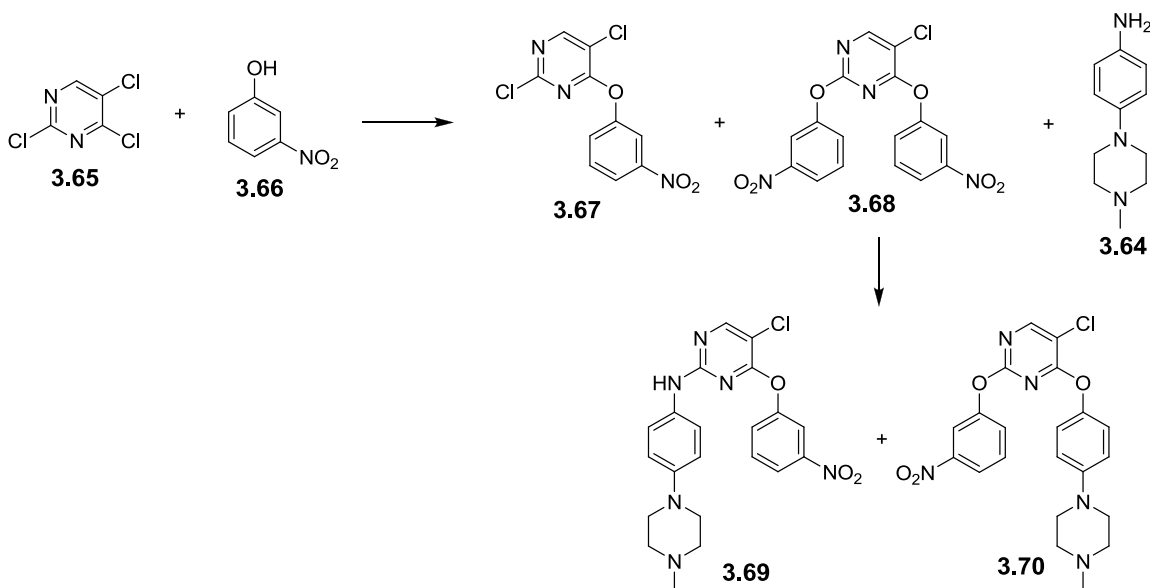
1-methyl-4-(4-nitrophenyl)piperazine 3.63: *para*-fluoro-nitrobenzene **3.61** (500 mg, 376 μL , 3.54 mmol) was dissolved in DMF (5 mL). Then N-methyl piperazine **3.62** (532 mg, 588 μL , 5.32 mmol) and diisopropylethylamine (916 mg, 1.23 mL, 7.09 mmol) were added. The reaction was heated to 55 $^{\circ}\text{C}$ and stirred for 3 h. The reaction mixture was then diluted into ethyl acetate (120 mL), washed extensively with water (10 x 5 mL) and brine, dried over sodium sulfate, and concentrated to afford 750 mg (96% yield) of orange-yellow solid. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 2.56 (s, 3 H) 3.44 (t, $J=5.13$ Hz, 2 H) 6.83 (ddd, $J=9.52, 3.30, 2.20$ Hz, 2 H) 8.12 (ddd, $J=9.52, 3.30, 2.38$ Hz, 2 H). LCMS: 221.98 (M+H); MW: 221.12.

1-methyl-4-(4-aniline)piperazine 3.64: nitrobenzene **3.63** (750 mg, 3.39 mmol) was dissolved in methanol (80 mL) and cycled through an H-Cube Continuous Flow Hydrogenation Reactor (ThalesNano, Budapest, Hungary) equipped with a 10% Pd/C H-Cube cartridge (ThalesNano) using 1 ATM of H_2 at 30 $^{\circ}\text{C}$ with a flow rate of 1 mL/min. The resulting solution was purple in color. After concentration, 640 mg of a light purple solid was obtained. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 2.39 (s, 3 H) 2.60 (t,

$J=4.94$ Hz, 2 H) 3.47 (t, $J=5.13$ Hz, 2 H) 6.84 (ddd, $J=9.52, 3.30, 2.20$ Hz, 2 H) 8.13 (ddd, $J=9.52, 3.48, 2.20$ Hz, 2 H). LCMS: 192.5 (M+H); MW: 191.14.

WZ3146

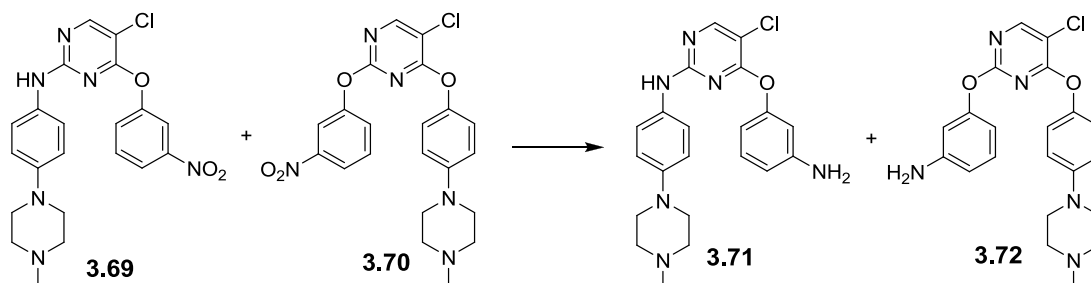
Synthesized as in [Zhou, W. *et al.* Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature* **462**, 1070-1074, (2009).] with the following modifications.



Substituted chloropyrimidines 3.67 and 3.68: 2, 4, 5-trichloropyrimidine **3.65** (1.6 g, 1 mL, 8.72 mmol) was dissolved in DMF (15 mL) and cooled on ice. Then potassium carbonate (2.41 g, 17.4 mmol) was added in one portion. Then a solution of 3-nitrophenol **3.66** (1.53 g, 10.9 mmol) in DMF (15 mL) was added slowly monitoring reaction progress. When the bis-addition product **3.68** was observed by LCMS, the reaction was diluted into ethyl acetate (400 mL), washed with water, dried over sodium

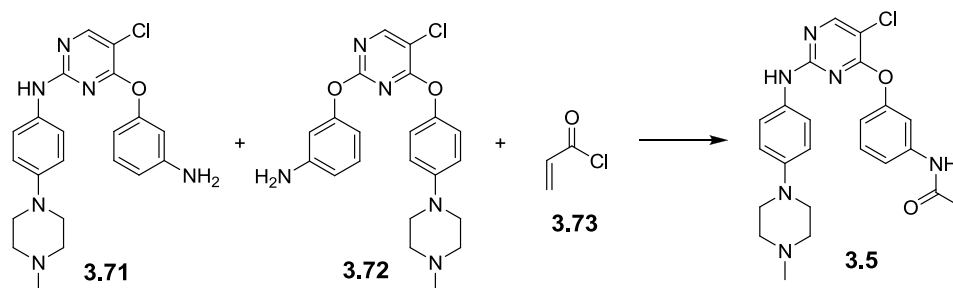
sulfate, and concentrated to afford 2.7 g of a mixture that appeared to be approximately 80% desired mono-addition product **3.67**. The two species were found to be inseparable on scale. LCMS: 285.88 and 388.93 (M+H); MW: 284.97 (**3.67**) and 388.02 (**3.68**).

Anilinopyrimidines 3.69 and 3.70: chloropyrimidines **3.67** and **3.68** (1g, 3.50 mmol) and aniline **3.64** (669 mg, 3.5 mmol) were dissolved in *sec*-butanol (8 mL). Then trifluoroacetic acid (399 mg, 268 μ L, 3.5 mmol) was added. The reaction was heated to 100 °C and stirred for 4 h. The reaction was quenched with a saturated solution of sodium bicarbonate (gas formation, consider using a different base) and extracted with ethyl acetate. The organic layers were combined, dried over sodium sulfate, and concentrated. The crude was purified by silica flash chromatography (50 g Biotage SNAP silica flash column) and a gradient of methanol in methylene chloride (0-10%) affording 700 mg of solid, a mixture of the regioisomers **3.69** and **3.70**. LCMS: 441.15 (M+H); MW: 440.14 (**3.69** and **3.70**).



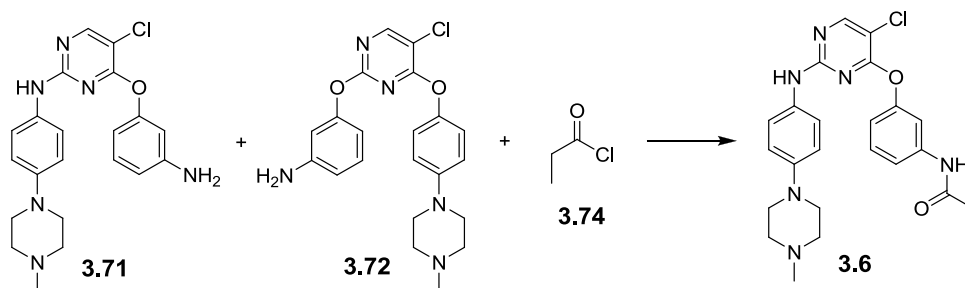
Anilinopyrimidines 3.71 and 3.72: nitrobenzenes **3.69** and **3.70** (700 mg, 1.6 mmol), ammonium chloride (606 mg, 11.3 mmol), and iron (633 mg, 11.3 mmol) were mixed in

THF (25 mL) and water (25 mL). The reaction mixture was heated to 65 °C with stirring. After the starting materials were completely consumed (12-24 h), the reaction was filtered and concentrated. The crude was then purified by silica flash chromatography (50 g Biotage SNAP silica flash column) and a gradient of methanol in methylene chloride (0-10%) affording 342 mg of solid, a mixture of the regioisomers **3.71** and **3.72**. LCMS: 411.08 (M+H); MW: 410.16 (**3.71** and **3.72**).



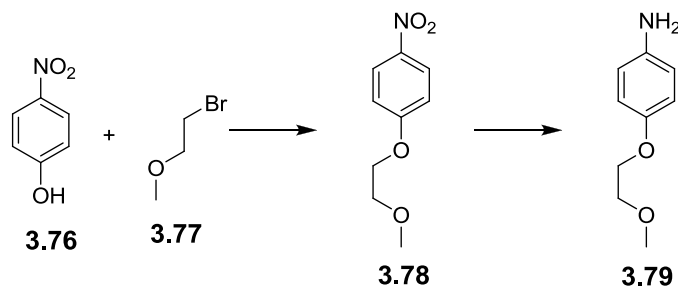
WZ3146 (3.5): anilines **3.71** and **3.72** (40 mg, 0.1 mmol) was dissolved in methylene chloride (1 mL). Then triethylamine (19.2 mg, 27 μ L, 0.2 mmol) was added, and the mixture was cooled in an ice-water bath. Then a solution of acryloyl chloride (17.6 mg, 15.8 μ L, 0.2 mmol) in methylene chloride (1 mL) was added slowly. The reaction was then allowed to slowly come to rt and stirred for an additional 24 h. The reaction mixture was then loaded directly on to silica flash column (4 g Silicycle) and purified using a gradient of methanol in methylene chloride (0-10%) affording 12 mg of a mixture of the two regioisomers. The mixture was separated by reverse-phase HPLC (methanol in water, 0.05% TFA) to afford the TFA salt of WZ3146 (**3.5**). ^1H NMR (400 MHz, DMSO- d_6) δ ppm 2.78 - 2.89 (m, 4 H) 3.13 (dt, $J=11.00, 9.30$ Hz, 2 H) 3.49 (d, $J=11.72$ Hz, 2 H)

3.63 (d, $J=13.37$ Hz, 2 H) 5.78 (dd, $J=10.07$, 1.83 Hz, 1 H) 6.27 (dd, $J=16.85$, 2.20 Hz, 1 H) 6.44 (dd, $J=16.85$, 10.26 Hz, 1 H) 6.72 (br. d, $J=7.30$ Hz, 1 H) 7.00 (dd, $J=8.06$, 2.38 Hz, 1 H) 7.29 (br. d, $J=7.20$ Hz, 1 H) 7.45 (t, $J=8.06$ Hz, 1 H) 7.60 (t, $J=2.01$ Hz, 1 H) 7.62 (t, $J=8.40$ Hz, 1 H) 8.42 (s, 1 H) 9.57 (br. s., 1 H) 9.68 (br. s., 1 H) 10.36 (s, 1 H). LCMS: 465.19 (M+H); MW: 464.17.



Propionamides 3.6: anilines **3.71** and **3.72** (50 mg, 0.12 mmol) was dissolved in methylene chloride (2 mL). Then triethylamine (24.6 mg, 32 μ L, 0.24 mmol) was added, and the mixture was cooled in an ice-water bath. Then a solution of propionyl chloride (22.5 mg, 21.3 μ L, 0.24 mmol) in methylene chloride (1 mL) was added slowly. The reaction was then allowed to slowly come to rt and stirred for an additional 3 h. The reaction mixture was washed with a solution of saturated sodium bicarbonate and brine, dried over sodium sulfate, and concentrated. The crude was then purified by reverse phase HPLC (methanol in water, 0.05% TFA) to afford the 11.5 mg of propionamide **3.6** as a TFA salt. ^1H NMR (400 MHz, DMSO- d_6) δ ppm 1.07 (t, $J=7.60$ Hz, 3 H) 2.33 (q, $J=7.63$ Hz, 3 H) 2.78 - 2.92 (m, 5 H) 3.06 - 3.23 (m, 0 H) 3.50 (br. d, $J=12.10$ Hz, 0 H) 3.65 (br. d, $J=12.60$ Hz, 0 H) 6.73 (br. d, $J=7.10$ Hz, 2 H) 6.94 (t, $J=1.83$ Hz, 0 H) 6.96

(dd, $J=2.20, 1.00$ Hz, 1 H) 7.30 (d, $J=6.40$ Hz, 2 H) 7.41 (t, $J=8.15$ Hz, 2 H) 7.49 - 7.58 (m, 2 H) 8.42 (s, 1 H) 9.55 (br. s., 1 H) 9.70 (br. s., 1 H) 10.06 (s, 1 H). LCMS: 467.08 (M+H); MW: 466.19.



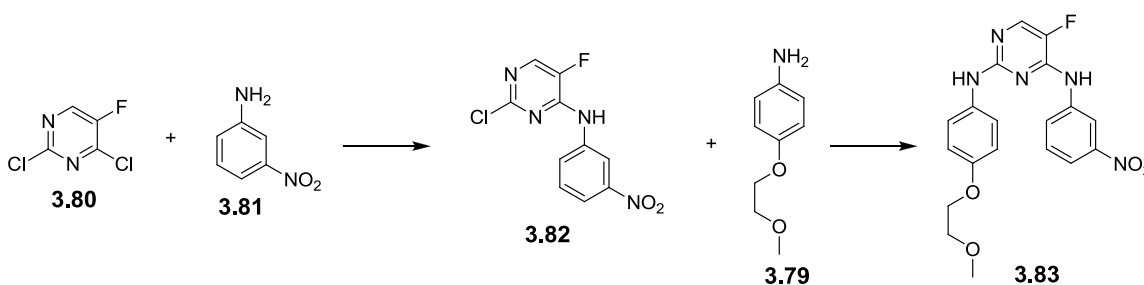
1-(2-methoxyethoxy)-4-nitrobenzene 3.78: *para*-nitrophenol **3.76** (4.17g, 30 mmol) was dissolved in DMF (15 mL). Then potassium carbonate (28.8 g, 208.7 mmol) was added, followed by methoxy bromoethane **3.77** (5g, 3.4 mL, 36 mmol). The reaction was heated to 50 °C and stirred overnight. The reaction mixture was then diluted into ethyl acetate (150 mL), washed extensively with water (10 x 10 mL), twice with 3 N NaOH (may not be necessary), and brine. The organic layer was dried over sodium sulfate and concentrated to afford 5.22 g (88% yield) of white solid. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 3.46 (s, 3 H) 3.79 (tt, $J=4.76, 1.65$ Hz, 2 H) 4.21 (tt, $J=4.58, 1.65$ Hz, 2 H) 6.99 (ddd, $J=9.34, 3.48, 2.38$ Hz, 2 H) 8.20 (ddd, $J=9.16, 3.30, 2.38$ Hz, 2 H). LCMS: 197.92 (M+H); MW: 197.07.

1-(2-methoxyethoxy)-4-aniline 3.79: nitrobenzene **3.78** (2 g, 10 mmol) was dissolved in methanol (200 mL) and cycled through an H-Cube Continuous Flow Hydrogenation Reactor (ThalesNano) equipped with a 10% Pd/C H-Cube cartridge (ThalesNano) using 1 ATM of H_2 at 30 °C with a flow rate of 1 mL/min. After concentration, 1.7 g of a brown

liquid was obtained. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 3.43 - 3.45 (m, 3 H) 3.71 (tt, $J=4.76$, 1.60 Hz, 2 H) 4.05 (tt, $J=4.76$, 1.28 Hz, 2 H) 6.64 (ddd, $J=8.42$, 3.66, 2.38 Hz, 2 H) 6.78 (ddd, $J=8.61$, 3.66, 2.20 Hz, 2 H). LCMS: 167.96 (M+H); MW: 167.09.

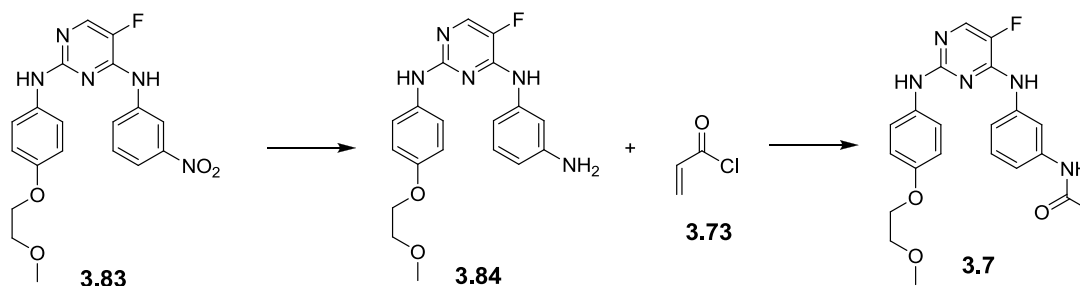
I-182

Synthesized as in [Kluge, A. *et al.* Heteroaryl compounds and uses thereof. U.S. patent WO/2009/158571 (12/30/2009).] with the following modifications.



Substituted fluoropyrimidine 3.82: 2, 4-dichloro-5-fluoropyrimidine **3.80** (1.5 g, 9.0 mmol), diisopropylethylamine (871 mg, 1.2 mL, 6.7 mmol), and 3-nitroaniline **3.81** (620 mg, 4.5 mmol) were dissolved in n-butanol (8 mL). The reaction was heated to 120 °C. After 4 h, the reaction mixture was cooled and filtered. The solid was washed with cold hexanes and dried affording 1.07 g of white solid, an 88% yield. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 7.16 (br. s., 1 H) 7.60 (t, $J=8.24$ Hz, 1 H) 8.04 (ddd, $J=8.00$, 2.20, 0.92 Hz, 1 H) 8.10 (ddd, $J=8.20$, 2.20, 0.92 Hz, 1 H) 8.18 (d, $J=2.40$ Hz, 1 H) 8.54 (t, $J=2.20$ Hz, 1 H). LCMS: 268.90 (M+H); MW: 268.02.

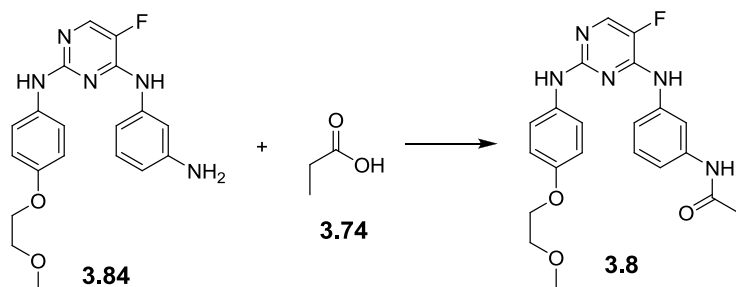
Anilinopyrimidine 3.83: fluoropyrimidine **3.82** (300 mg, 1.1 mmol) and aniline **3.79** (224 mg, 1.3 mmol) were dissolved in *sec*-butanol (10 mL). Then acetic acid (5 drops) was added. The reaction was heated to 110 °C and stirred for 24 h. The reaction was then concentrated, taken up in ethyl acetate, and washed with a saturated solution of sodium bicarbonate and brine. The organic layer was then dried over sodium sulfate and concentrated. The crude was brought up in methylene chloride, and the desired product was filtered. The solid was further washed with methylene chloride affording 310 mg (70% yield) of a yellow-brown solid. ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 3.31 (dd, *J*=2.66, 0.46 Hz, 3 H) 3.64 (ddd, *J*=5.49, 3.66, 1.65 Hz, 2 H) 4.02 (ddd, *J*=5.49, 3.85, 1.28 Hz, 2 H) 6.80 (ddd, *J*=9.16, 4.21, 2.01 Hz, 1 H) 7.48 (ddd, *J*=8.97, 4.03, 2.01 Hz, 1 H) 7.60 (t, *J*=8.24 Hz, 1 H) 7.89 (dd, *J*=7.87, 2.01 Hz, 1 H) 8.16 (d, *J*=3.30 Hz, 1 H) 8.35 (d, *J*=8.79 Hz, 1 H) 8.54 (t, *J*=2.29 Hz, 1 H) 9.11 (s, 1 H) 9.77 (s, 1 H). LCMS: 400.09 (M+H); MW: 399.13.



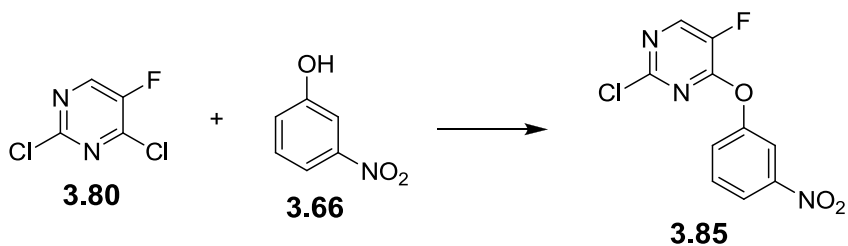
Anilinopyrimidine 3.84: nitrobenzene **3.83** (300 mg, 0.75 mmol), ammonium chloride (402 mg, 7.5 mmol), and iron (420 mg, 7.5 mmol) were mixed in THF (16 mL) and water (16 mL). The reaction mixture was heated to 65 °C with stirring. After 3 h, the reaction was filtered and concentrated. The crude was brought up in methylene chloride, and the

desired product was filtered. The solid was further washed with methylene chloride affording 200 mg (72% yield) of **3.82**, a yellow solid. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 3.46 (s, 3 H) 3.65 (br. s., 2 H) 3.75 (tt, $J=4.58$, 1.28 Hz, 2 H) 4.11 (tt, $J=4.94$, 1.65 Hz, 1 H) 6.42 (ddd, $J=7.87$, 2.20, 0.73 Hz, 1 H) 6.66 (br. d, $J=2.70$ Hz, 1 H) 6.73 (d, $J=2.00$ Hz, 1 H) 6.74 - 6.77 (m, 1 H) 6.93 (ddd, $J=8.97$, 3.48, 2.20 Hz, 1 H) 7.09 (t, $J=7.97$ Hz, 1 H) 7.24 - 7.28 (m, 1 H) 7.43 (ddd, $J=9.16$, 3.66, 2.20 Hz, 1 H) 7.91 (d, $J=3.30$ Hz, 1 H). LCMS: 370.06 (M+H); MW: 369.16.

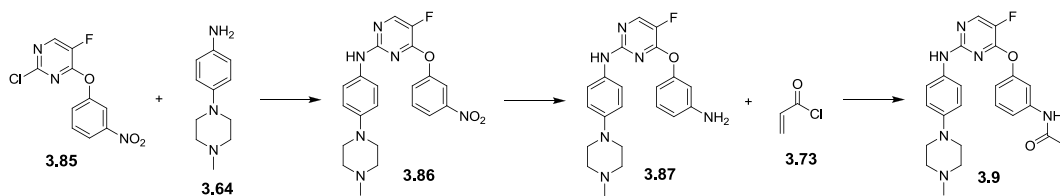
I-182 (3.7): aniline **3.84** (20 mg, 0.05 mmol) was dissolved in methylene chloride (2 mL). Then triethylamine (11 mg, 15 μL , 0.1 mmol) was added, and the mixture was cooled in an ice-water bath. Then a solution of acryloyl chloride (10 mg, 8.8 μL , 0.1 mmol) in methylene chloride (1 mL) was added slowly. The reaction was then allowed to slowly come to rt and stirred for an additional 2 h. The reaction mixture was washed with water and brine, dried over sodium sulfate, and concentrated. The crude was first purified by silica chromatography using a gradient of ethyl acetate in hexanes (40-70%). Then reverse phase HPLC was used to separate the desired product, 4.5 mg of a white solid, from poly-acylated species. ^1H NMR (400 MHz, DMSO-*d*₆) δ ppm 3.30 (d, $J=0.55$ Hz, 3 H) 3.63 (tt, $J=4.58$, 1.47 Hz, 2 H) 4.00 (tt, $J=5.30$, 1.28 Hz, 2 H) 5.76 (dd, $J=10.07$, 2.01 Hz, 1 H) 6.26 (dd, $J=17.03$, 1.83 Hz, 1 H) 6.46 (dd, $J=16.85$, 10.07 Hz, 1 H) 6.78 (ddd, $J=9.71$, 3.85, 1.65 Hz, 2 H) 7.29 (dd, $J=8.61$, 7.87 Hz, 1 H) 7.41 (d, $J=8.42$ Hz, 1 H) 7.46 (ddd, $J=9.50$, 2.70, 1.50 Hz, 2 H) 7.95 (t, $J=2.01$ Hz, 1 H) 8.12 (d, $J=4.21$ Hz, 1 H) 9.31 (br. s., 1 H) 9.76 (br. s., 1 H) 10.16 (s, 1 H). LCMS: 424.10 (M+H); MW: 423.17.



Propionamide 3.8: aniline **3.84** (20 mg, 0.05 mmol) was dissolved in methylene chloride (2 mL). Then triethylamine (11 mg, 15 μ L, 0.1 mmol) was added, and the mixture was cooled in an ice-water bath. Then a solution of propionyl chloride (10 mg, 9.5 μ L, 0.1 mmol) in methylene chloride (1 mL) was added slowly. The reaction was then allowed to slowly come to rt and stirred for an additional 2 h. The reaction mixture was washed with water and brine, dried over sodium sulfate, and concentrated. The crude was first purified by silica chromatography using a gradient of ethyl acetate in hexanes (40-70%). Then reverse phase HPLC was used to separate the desired product, 11 mg of a white solid, from poly-acylated species. ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ ppm 1.07 (t, $J=7.60$ Hz, 3 H) 2.32 (q, $J=7.57$ Hz, 2 H) 3.30 (s, 3 H) 3.63 (tt, $J=4.94, 1.28$ Hz, 2 H) 4.02 (tt, $J=4.94, 1.46$ Hz, 2 H) 6.79 (ddd, $J=9.34, 3.30, 2.20$ Hz, 2 H) 7.24 (t, $J=8.06$ Hz, 1 H) 7.32 (ddd, $J=8.42, 1.65, 0.92$ Hz, 1 H) 7.42 (d, $J=8.20$ Hz, 1 H) 7.48 (ddd, $J=9.16, 3.30, 2.38$ Hz, 2 H) 7.87 (t, $J=1.92$ Hz, 1 H) 8.09 (d, $J=4.21$ Hz, 0 H) 9.21 (br. s., 1 H) 9.61 (br. s., 1 H) 9.84 (s, 1 H). LCMS: 426.08 (M+H); MW: 425.19.



Substituted fluoropyrimidine 3.85: 2, 4-dichloro-5-fluoropyrimidine **3.80** (1.5g, 9.0 mmol) was dissolved in DMF (15 mL) and cooled on ice. Then potassium carbonate (2.5 g, 18.0 mmol) was added in one portion. Then a solution of 3-nitrophenol **3.66** (1.57 g, 11.2 mmol) in DMF (15 mL) was added in three portions. The reaction was allowed to come to rt and was stirred for an additional 3 h. The reaction was then diluted into ethyl acetate (400 mL), washed with water, dried over sodium sulfate, and concentrated to afford 2.65 g of white solid that >90% **3.85**. This material was carried forward without further purification. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 7.59 (ddd, $J=8.24, 2.38, 1.10$ Hz, 1 H) 7.67 (td, $J=8.20, 0.37$ Hz, 1 H) 8.13 (td, $J=2.38, 0.37$ Hz, 1 H) 8.22 (ddd, $J=8.15, 2.11, 1.10$ Hz, 1 H) 8.44 (d, $J=1.83$ Hz, 1 H). LCMS: 469.9 (M+H); MW: 469.0.



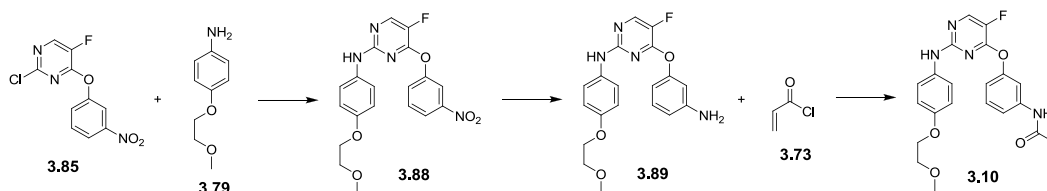
Anilinopyrimidine 3.86: fluoropyrimidine **3.85** (1.0 g, 3.7 mmol) and aniline **3.64** (709 mg, 3.7 mmol) were dissolved in *sec*-butanol (8 mL). Then trifluoroacetic acid (423 mg, 284 μL , 3.7 mmol) was added. The reaction was heated to 100 $^\circ\text{C}$ and stirred for 24 h.

The reaction was then quenched with a saturated solution of sodium bicarbonate (gas formation, consider using a different base) and extracted with ethyl acetate. The organic layers were combined, dried over sodium sulfate, and concentrated. The crude was purified by silica chromatography (50 g Biotage SNAP silica flash column) and a gradient of methanol in methylene chloride (0-10%) affording 300 mg of solid with some impurities. This material was carried forward without further purification. LCMS: 425.15 (M+H); MW: 424.17.

Anilinopyrimidine 3.87: nitrobenzene **3.86** (300 mg, 0.7 mmol) was dissolved in methanol (175 mL) and cycled through an H-Cube Continuous Flow Hydrogenation Reactor (ThalesNano) equipped with a 10% Pd/C H-Cube cartridge (ThalesNano) using 1 ATM of H₂ at 30 °C with a flow rate of 1 mL/min. After concentration, the crude was purified by silica chromatography (25 g Silicycle silica flash column) and a gradient of methanol in methylene chloride (0-10%). 25 mg of impure material was obtained. LCMS: 395.05 (M+H); MW: 394.19.

Acrylamide 3.9: aniline **3.87** (25 mg, 0.063 mmol) was dissolved in methylene chloride (2 mL). Then triethylamine (12.8 mg, 17.7 µL, 0.126 mmol) was added, and the mixture was cooled in an ice-water bath. Then a solution of acryloyl chloride (11.5 mg, 10.3 µL, 0.126 mmol) in methylene chloride (1 mL) was added slowly. The reaction was then allowed to slowly come to rt and stirred for an additional 2 h. The crude reaction mixture was loaded directly onto a silica flash column (4 g Silicycle). Prior to loading on silica, consider using an aqueous work-up to remove triethylammonium chloride (TEA-

HCl). The product was eluted with a gradient of methanol in methylene chloride (0-10%) affording 14.8 mg of solid contaminated with TEA-HCl. This material was brought up in ethyl acetate and washed with a saturated sodium bicarbonate. The organic layer was then concentrated affording 4 mg of a clear gum (~90% pure). ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 2.21 (dd, *J*=3.85, 2.38 Hz, 3 H) 2.43 (s, 3 H) 2.93 - 3.12 (m, 3 H) 5.77 (dd, *J*=10.26, 2.01 Hz, 1 H) 6.25 (dd, *J*=17.40, 2.20 Hz, 1 H) 6.43 (dd, *J*=16.66, 10.26 Hz, 1 H) 6.74 (dt, *J*=9.16, 2.56 Hz, 1 H) 6.86 (ddd, *J*=7.69, 2.38, 0.92 Hz, 1 H) 7.36 (t, *J*=8.42 Hz, 1 H) 7.39 (ddd, *J*=8.42, 2.01, 0.92 Hz, 1 H) 7.52 (s, 1 H) 7.55 - 7.58 (m, 1 H) 8.11 (d, *J*=3.48 Hz, 1 H) 9.51 (s, 1 H) 10.27 (s, 1 H). LCMS: 449.2 (M+H); MW: 448.2.



Anilinopyrimidine 3.88: fluoropyrimidine **3.85** (750 mg, 2.8 mmol) and aniline **3.79** (465 mg, 2.8 mmol) were dissolved in *sec*-butanol (6 mL). Then trifluoroacetic acid (317 mg, 213 μL, 2.8 mmol) was added. The reaction was heated to 100 °C and stirred for 24 h. The reaction was then quenched with a saturated solution of sodium bicarbonate (gas formation, consider using a different base) and extracted with ethyl acetate. The organic layers were combined, dried over sodium sulfate, and concentrated. The crude was purified by silica chromatography (50 g Biotage SNAP silica flash column) and a gradient of ethyl acetate in hexanes (0-100%) affording 50 mg of pure product. ¹H NMR (400

MHz, CHLOROFORM-*d*) δ ppm 3.46 (s, 3 H) 3.75 (tt, $J=4.58$, 1.65 Hz, 1 H) 4.09 (tt, $J=4.94$, 1.65 Hz, 1 H) 6.82 (ddd, $J=9.16$, 3.66, 2.20 Hz, 3 H) 7.32 (ddd, $J=9.16$, 3.30, 2.20 Hz, 2 H) 7.52 (dtd, $J=8.50$, 2.10, 2.10, 0.70 Hz, 1 H) 7.56 (t, $J=7.87$ Hz, 1 H) 7.98 (d, $J=2.75$ Hz, 1 H) 8.07 - 8.14 (m, 2 H). LCMS: 401.07 (M+H); MW: 400.12.

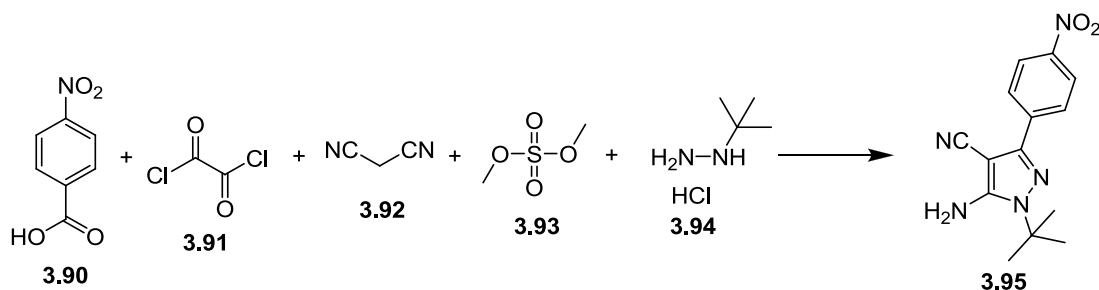
Anilinopyrimidine 3.89: nitrobenzene **3.88** (50 mg, 0.12 mmol) was dissolved in methanol (15 mL) and cycled through an H-Cube Continuous Flow Hydrogenation Reactor (ThalesNano) equipped with a 10% Pd/C H-Cube cartridge (ThalesNano) using 1 ATM of H₂ at 30 °C with a flow rate of 1 mL/min. After concentration, 38 mg of aniline was obtained. ¹H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 3.39 - 3.51 (m, 4 H) 3.74 (tt, $J=4.58$, 1.10 Hz, 2 H) 4.09 (tt, $J=4.94$, 1.28 Hz, 2 H) 6.52 (t, $J=2.20$ Hz, 1 H) 6.58 (dt, $J=7.97$, 2.61 Hz, 2 H) 6.76 (br. s., 1 H) 6.84 (ddd, $J=9.16$, 3.48, 2.20 Hz, 2 H) 7.17 (t, $J=8.06$ Hz, 1 H) 7.41 (dt, $J=8.61$, 2.20 Hz, 2 H) 7.97 (d, $J=2.93$ Hz, 1 H). LCMS: 371.96 (M+H); MW: 370.14.

Acrylamide 3.10: aniline **3.89** (25 mg, 0.063 mmol) was dissolved in methylene chloride (1 mL). Then triethylamine (10.9 mg, 15 μ L, 0.108 mmol) was added, and the mixture was cooled in an ice-water bath. Then a solution of acryloyl chloride (9.8 mg, 8.8 μ L, 0.108 mmol) in methylene chloride (1 mL) was added slowly. The reaction was then allowed to slowly come to rt and stirred for an additional 2 h. The crude reaction mixture was loaded directly onto a silica flash column (4 g Silicycle). The product was eluted with a gradient of methanol in methylene chloride (0-5%) affording 13.5 mg of solid. ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 3.30 (s, 3 H) 3.63 (ddd, $J=5.31$, 3.85, 1.65 Hz,

2 H) 4.01 (ddd, $J=5.40, 3.75, 1.28$ Hz, 2 H) 5.76 (dd, $J=9.89, 2.01$ Hz, 1 H) 6.25 (dd, $J=17.40, 1.83$ Hz, 1 H) 6.43 (dd, $J=16.85, 10.26$ Hz, 1 H) 6.74 (ddd, $J=9.71, 3.30, 2.01$ Hz, 1 H) 6.87 (dd, $J=7.87, 2.38$ Hz, 1 H) 7.37 (t, $J=8.06$ Hz, 1 H) 7.41 - 7.52 (m, 2 H) 7.59 (t, $J=2.01$ Hz, 1 H) 8.14 (d, $J=3.30$ Hz, 1 H) 9.58 (s, 1 H) 10.27 (s, 1 H). LCMS: 425.2 (M+H); MW: 424.15.

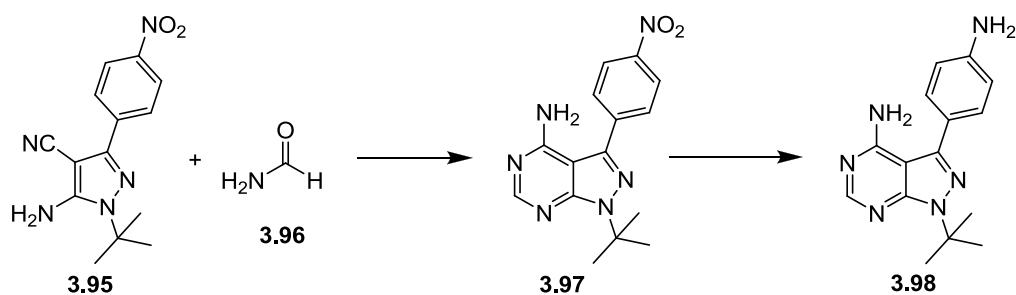
Electrophilic analogs of AD57 (3.25)

N1-linker series



Pyrazole 3.95: 4-nitrobenzoic acid (15g, 89.8 mmol, 1.0 equiv.) was dissolved in THF (141 mL) with DMF (200 μ L) and cooled in an ice-water bath. Then the oxalyl chloride (13.7 g, 9.1 mL, 107.7 mmol, 1.2 equiv.) was added slowly over 10 min via addition funnel. The reaction was allowed to warm to rt and was then heated to 40 °C. After 3 h, the reaction was concentrated under reduced pressure. This mixture was dissolved in THF (100 mL). Malonitrile (6.53 g, 98.7 mmol, 1.1 mmol) was added in 1 portion, and the resulting mixture was cooled in an ice-water bath. Triethylamine (22.72 g, 31.29 mL, 224.5 mmol, 2.5 equiv.) was added over 1 h via addition funnel. The reaction mixture

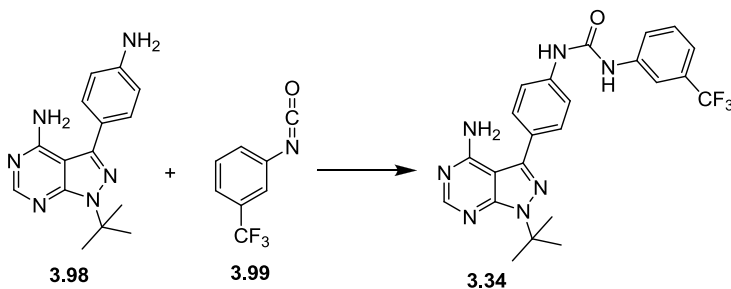
was then allowed to come to rt. After 2 h, dimethyl sulfate (33.98 g, 25.5 mL, 269.4 mmol, 3 equiv.) was added. The reaction was heated to 70 °C for 2 h and then cooled to rt. After stirring overnight, the reaction went from a yellow solution to a red/brown solution. Then ethanol (100 mL) was added. The mixture was cooled in an ice-water bath, and then *tert*-butyl hydrazine – HCl (33.54 g, 269.4 mmol, 3.0 equiv.) and triethylamine (29.96 g, 41.3 mL, 296 mmol, 3.3 equiv.) were added. The reaction was heated to reflux. After 4 h, the reaction was cooled, and a precipitate (triethylammonium chloride) formed. Ethyl acetate was added, and the salt was removed by filtration. The solution was concentrated and then redissolved in ethyl acetate. The solution was washed with water and brine, dried over sodium sulfate, and concentrated onto silica with THF. The crude was purified using a gradient of ethyl acetate in hexanes (10-40%) to afford 9.2 g (35.9% yield) of **3.95**, a yellow solid. ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 1.42 - 1.75 (m, 9 H) 6.55 (s, 2 H) 7.89 - 8.17 (m, 2 H) 8.23 - 8.40 (m, 2 H). LCMS: 283.65 (M-H); MW: 285.12.



Pyrazolopyrimidine 3.97: pyrazole **3.95** (9.3 g, 32.6 mmol, 1.0 equiv.) was dissolved in formamide (70 mL) with N-methyl pyrrolidone (NMP, 5 mL). The mixture was heated to 80 °C for 6 h, until complete by LCMS. The reaction was then cooled and mixed with

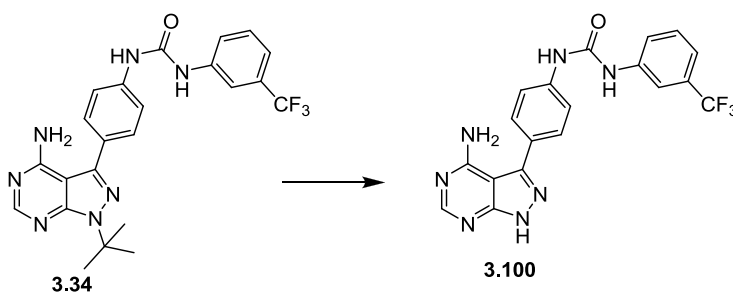
water (200 mL). The resulting precipitate was filtered and washed with water. The brown solid was dried under reduced pressure to afford 8.4 g (82.5% yield) of **3.97**. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm 1.77 (s, 9 H) 7.92 (ddd, $J=8.79, 3.11, 2.00$ Hz, 2 H) 8.27 (s, 1 H) 8.37 (ddd, $J=8.42, 3.11, 1.80$ Hz, 2 H). LCMS: 313.03 (M+H); MW: 312.13.

Aniline 3.98: pyrazolopyrimidine **3.97** (2 g, 6.4 mmol, 1.0 equiv.) was dissolved in THF (25 mL) and mixed with granular zinc (2.03 g, 32.0 mmol, 5 equiv.). Then acetic acid (15 mL) was added. The mixture was heated with stirring to 55 °C. After 24 h, the reaction was cooled, diluted with ethyl acetate, and filtered to remove the solids. The aqueous layer was extracted with two portions of ethyl acetate. The organic layers were combined, washed with brine, dried over sodium sulfate, and concentrated to afford 1.4 g (78% yield) of **3.98**, an orange solid. ^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ ppm 1.73 (s, 3 H) 5.39 (s, 1 H) 6.69 (ddd, $J=8.24, 3.11, 2.40$ Hz, 1 H) 7.29 (ddd, $J=8.24, 2.93, 1.80$ Hz, 1 H) 8.19 (s, 1 H). LCMS: 283.06 (M+H); MW: 282.16.



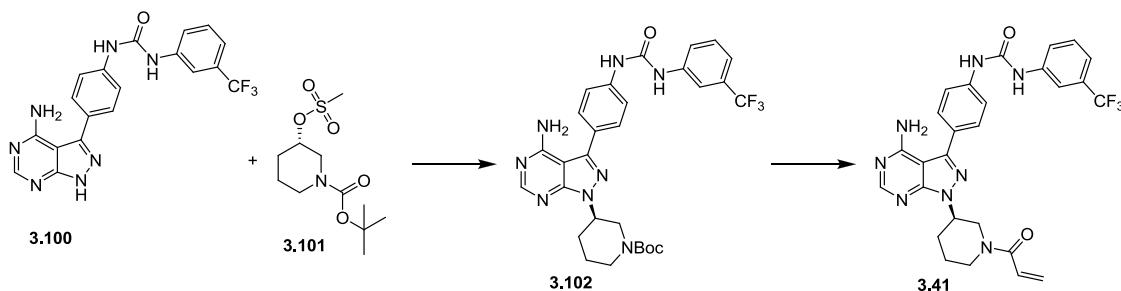
Pyrazolopyrimidine 3.34: aniline **3.98** (100 mg, 0.354 mmol, 1.0 equiv.) was dissolved in methylene chloride (5 mL) and cooled in an acetonitrile-dry ice bath. Then a solution of isocyanate **3.99** (59.6 mg, 43.8 μL , 0.319 mmol, 0.9 equiv.) in methylene chloride (1 mL)

was added in small portion over 1 h. The reaction was allowed to proceed for 3 h and was then concentrated. The crude was dissolved in ethyl acetate and concentrated onto silica. The impurities were eluted with 30% ethyl acetate in hexanes. Then the desired product was eluted with ethyl acetate. After concentration, 93 mg (62% yield) of **3.34**, a yellow-brown solid, was obtained. $^1\text{H NMR}$ (400 MHz, $\text{DMSO-}d_6$) δ ppm 1.75 (s, 9 H) 7.33 (d, $J=7.69$ Hz, 1 H) 7.49 - 7.69 (m, 6 H) 8.04 (t, $J=2.20$ Hz, 1 H) 8.23 (s, 1 H) 9.02 (s, 1 H) 9.11 (s, 1 H). LCMS: 470.13 (M+H); MW: 469.18.



Pyrazolopyrimidine 3.100: pyrazolopyrimidine **3.34** (30 mg, 0.064 mmol, 1.0 equiv.) was dissolved in formic acid (2 mL). Concentrated hydrochloric acid (200 μL) was added. The reaction was heated to 110 $^\circ\text{C}$. After 3 h, the reaction was concentrated to dryness and carried forward without further purification. $^1\text{H NMR}$ (400 MHz, $\text{DMSO-}d_6$) δ ppm 7.00 (br. d, $J=7.90$ Hz, 1 H) 7.04 (br. s., 1 H) 7.12 (br. s., 1 H) 7.33 (t, $J=7.69$ Hz, 1 H) 7.54 (dd, $J=7.69, 3.66$ Hz, 2 H) 7.61 (dt, $J=8.79, 2.01$ Hz, 1 H) 7.70 (ddd, $J=8.79, 2.93, 2.20$ Hz, 1 H) 8.02 (t, $J=1.83$ Hz, 1 H) 8.52 (d, $J=2.75$ Hz, 1 H) 9.76 (br. s., 1 H) 9.86 (br. s., 1 H). LCMS: 414.00 (M+H); MW: 413.12.

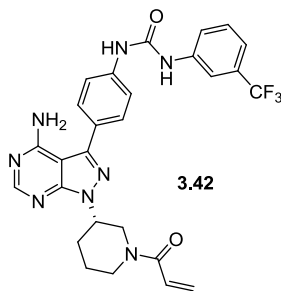
Synthesis of JC080-009R (3.41) - General procedure for N1 modification



Boc-protected N1-substituted pyrazolopyrimidine 3.102: pyrazolopyrimidine **3.100** (148 mg, 0.36 mmol, 1.0 equiv.), mesylate **3.101** (250 mg, 0.9 mmol, 2.5 equiv.), and cesium carbonate (292 mg, 0.9 mmol, 2.5 equiv.) were mixed in DMF (4 mL). The reaction was heated to 60 °C and stirred overnight. The reaction was monitored for progress by LCMS, and the reaction does not proceed to completion even with additional mesylate and cesium carbonate. Using a higher temperature can drive the reaction to completion, see below. The reaction was then diluted into ethyl acetate and washed with water and brine. The organic layer was dried over sodium sulfate and concentrated. The crude was purified by silica flash chromatography (12 g SiliCycle silica flash column) and a gradient of methanol in ethyl acetate (0-5%). 72 mg (34% yield) of **3.102** was obtained. LCMS: 597.2 (M+H); MW: 596.25.

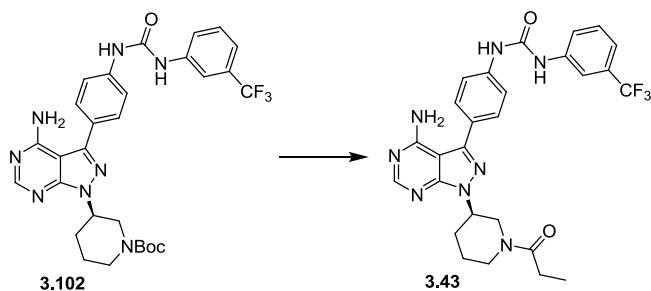
Acrylamide 3.41: boc-protected N1-substituted pyrazolopyrimidine **3.102** (72 mg, 0.12 mmol, 1.0 equiv) was deprotected in 4 N HCl in 1, 4-dioxanes (4 mL) at rt for 1 h. After concentration, the deprotected material was combined with acrylic acid (9.3 mg, 8.9 μ L, 0.13 mmol, 1.1 equiv.), EDC-HCl (27.8 mg, 0.14 mmol, 1.2 equiv.), hydroxybenzotriazole (19.6 mg, 0.14 mmol, 1.2 equiv.), and triethylamine (36.6 mg, 50.5 μ L 0.14 mmol, 1.2

equiv.) in DMF (2 mL). After 2 h, the reaction mixture was diluted into ethyl acetate and washed with water and brine. The organic layer was then dried over sodium sulfate and concentrated. The crude was then purified by silica flash chromatography (12 g SiliCycle silica flash column) using a gradient of methanol in ethyl acetate (0-5%) affording 7 mg (12.6% yield) of **3.41**. ^1H NMR (400 MHz, DMSO- d_6) δ ppm 1.52 - 1.66 (m, 2 H) 1.93 (dt, $J=14.00$, 3.66 Hz, 1 H) 2.11 (d, $J=11.54$ Hz, 1 H) 2.28 (qd, $J=11.50$, 3.30 Hz, 1 H) 3.00 (t, $J=11.54$ Hz, 1 H) 3.22 (t, $J=10.60$ Hz, 1 H) 3.64 - 3.75 (m, 1 H) 4.08 (d, $J=12.64$ Hz, 1 H) 4.22 (t, $J=13.90$ Hz, 1 H) 4.56 (d, $J=11.54$ Hz, 1 H) 4.70 (s, 1 H) 5.60 (d, $J=10.44$ Hz, 1 H) 5.71 (d, $J=10.62$ Hz, 1 H) 6.10 (dd, $J=24.91$, 17.21 Hz, 2 H) 6.73 (dd, $J=17.95$, 11.17 Hz, 1 H) 6.87 (dd, $J=16.21$, 10.16 Hz, 1 H) 7.33 (d, $J=7.69$ Hz, 1 H) 7.53 (t, $J=8.20$ Hz, 1 H) 7.60 (br. d, $J=6.80$ Hz, 3 H) 7.66 (dt, $J=8.79$, 2.01 Hz, 2 H) 8.04 (s, 1 H) 8.25 (s, 1 H) 9.09 (s, 1 H) 9.18 (s, 1 H). LCMS: 551.2 (M+H); MW: 550.21.



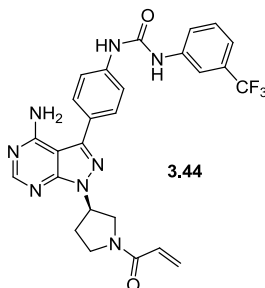
Acrylamide 3.42: synthesized as described for **3.41**. 2% three-step yield from **3.100**. ^1H NMR (400 MHz, DMSO- d_6) δ ppm 1.59 (td, $J=13.28$, 6.59 Hz, 1 H) 1.93 (dt, $J=13.73$, 4.00 Hz, 1 H) 2.12 (d, $J=5.49$ Hz, 1 H) 2.27 (qd, $J=12.10$, 3.48 Hz, 1 H) 2.99 (t, $J=12.50$ Hz, 1 H) 3.14 (dd, $J=6.50$, 5.40 Hz, 1 H) 3.21 (t, $J=11.20$ Hz, 1 H) 3.70 (t, $J=11.72$ Hz, 1 H) 4.08 (d, $J=12.45$ Hz, 1 H) 4.16 - 4.26 (m, 1 H) 4.55 (d, $J=11.72$ Hz, 1 H) 4.69 (d, $J=11.54$ Hz, 1 H)

5.59 (d, $J=10.26$ Hz, 1 H) 5.70 (d, $J=9.89$ Hz, 1 H) 6.10 (dd, $J=24.17, 16.66$ Hz, 1 H) 6.72 (dd, $J=16.66, 12.09$ Hz, 1 H) 6.86 (dd, $J=17.03, 10.07$ Hz, 1 H) 7.31 (d, $J=7.69$ Hz, 1 H) 7.52 (t, $J=7.97$ Hz, 1 H) 7.56 - 7.63 (m, 3 H) 7.66 (dt, $J=8.61, 2.01$ Hz, 2 H) 8.04 (t, $J=1.65$ Hz, 1 H) 8.23 - 8.27 (m, 1 H) 9.26 (br. s., 1 H) 9.34 (br. s., 1 H). LCMS: 551.2 (M+H); MW: 550.21.



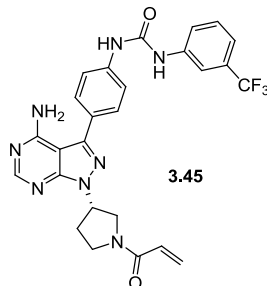
Propionamide 3.43: boc-protected N1-substituted pyrazolopyrimidine **3.102** (50 mg, 0.084 mmol, 1.0 equiv) was deprotected in 4 N HCl in 1, 4-dioxanes (2 mL) at rt for 2 h. After concentration, the deprotected material was dissolved in DMF (2 mL). Then sodium propionate (12.1 mg, 0.13 mmol, 1.5 equiv.) and DIPEA (54.2 mg, 73 μ L, 0.42 mmol, 5.0 equiv.) were added, and the mixture was cooled in an ice/water bath. Then the PyBOP (65.4 mg, 0.13 mmol, 1.5 equiv.) was added. The reaction was allowed to slowly warm to rt and stirred for an additional 2 h. Brine (9 mL) was added, and the organics were extracted with ethyl acetate (4 x 5 mL). The combined organic layers were washed with brine, dried over sodium sulfate, and concentrated. The crude was purified first by silica flash chromatography (12 g SiliCycle silica flash column) and a gradient of methanol in ethyl acetate (0-10%) to afford 47 mg of impure material. This was further purified by preparative TLC using 10% methanol in ethyl acetate as the mobile phase to

afford 30 mg of 85-90% pure material contaminated with Tris(pyrrolidinophosphine) oxide. A second preparative TLC purification using 5:1 toluene/ isopropanol as the mobile phase afforded 20 mg (43% yield) of **3.43**, a white solid. ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ ppm 0.99 (dt, $J=16.48, 9.00$ Hz, 3 H) 1.58 (d, $J=44.32$ Hz, 1 H) 1.89 (t, $J=12.10$ Hz, 1 H) 2.11 (br. s., 1 H) 2.16 - 2.32 (m, 1 H) 2.38 (qd, $J=6.80, 3.66$ Hz, 1 H) 2.84 (t, $J=12.50$ Hz, 1 H) 3.11 (q, $J=12.30$ Hz, 1 H) 3.57 (t, $J=10.30$ Hz, 1 H) 3.88 (d, $J=13.55$ Hz, 1 H) 4.02 (d, $J=12.82$ Hz, 1 H) 4.11 - 4.15 (m, 1 H) 4.24 (d, $J=14.10$ Hz, 1 H) 4.52 (d, $J=13.92$ Hz, 1 H) 4.58 - 4.68 (m, 1 H) 4.69 - 4.79 (m, 1 H) 7.33 (d, $J=7.33$ Hz, 2 H) 7.53 (t, $J=7.87$ Hz, 1 H) 7.60 (d, $J=8.06$ Hz, 3 H) 7.63 - 7.68 (m, 2 H) 8.04 (s, 1 H) 8.26 (d, $J=9.71$ Hz, 1 H) 9.03 (s, 1 H) 9.07 - 9.15 (m, 1 H). LCMS: 553.2 (M+H); MW: 552.22.

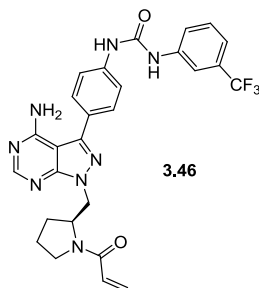


Acrylamide 3.44: synthesized as described for **3.41**. 7% three-step yield from **3.100**. ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ ppm 2.39 (ddd, $J=14.65, 7.87, 2.01$ Hz, 1 H) 2.70 (dt, $J=19.96, 6.96$ Hz, 1 H) 3.44 - 4.38 (m, 3 H) 5.51 (ddd, $J=24.17, 12.27, 6.04$ Hz, 0 H) 5.68 (ddd, $J=19.73, 10.30, 2.38$ Hz, 1 H) 5.89 (td, $J=10.80, 1.50$ Hz, 1 H) 6.16 (ddd, $J=18.68, 6.23, 2.56$ Hz, 1 H) 6.29 (ddd, $J=17.31, 11.08, 1.65$ Hz, 1 H) 6.58 (dd, $J=16.76, 10.35$ Hz, 1 H) 6.66 (dd, $J=16.85, 10.07$ Hz, 1 H) 7.33 (d, $J=7.69$ Hz, 1 H) 7.53 (t, $J=7.69$ Hz, 1 H) 7.56 -

7.62 (m, 3 H) 7.65 (dq, $J=8.79$, 1.83 Hz, 2 H) 8.04 (s, 1 H) 8.25 - 8.28 (m, 1 H) 9.07 (d, $J=2.20$ Hz, 1 H) 9.15 (s, 1 H). LCMS: 537.1 (M+H); MW: 536.19.

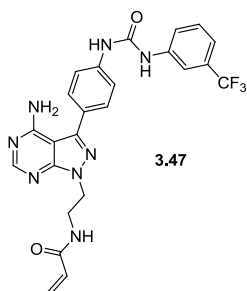


Acrylamide 3.45: synthesized as described for **3.41**. 4% three-step yield from **3.100**. ^1H NMR (400 MHz, DMSO- d_6) δ ppm 2.39 (qd, $J=7.14$, 2.20 Hz, 1 H) 2.70 (dt, $J=19.96$, 6.50 Hz, 1 H) 3.52 - 4.37 (m, 3 H) 5.51 (ddq, $J=30.40$, 7.14, 5.49, 5.49, 5.49 Hz, 0 H) 5.68 (ddd, $J=19.73$, 10.30, 2.38 Hz, 1 H) 5.89 (td, $J=10.30$, 2.56 Hz, 1 H) 6.16 (ddd, $J=18.68$, 5.86, 2.75 Hz, 1 H) 6.29 (ddd, $J=17.21$, 10.99, 1.65 Hz, 1 H) 6.59 (dd, $J=16.48$, 10.62 Hz, 1 H) 6.66 (dd, $J=16.85$, 10.62 Hz, 1 H) 7.33 (d, $J=6.96$ Hz, 1 H) 7.53 (t, $J=8.06$ Hz, 1 H) 7.56 - 7.62 (m, 3 H) 7.65 (dq, $J=9.34$, 1.65 Hz, 2 H) 8.04 (s, 1 H) 8.24 - 8.28 (m, 1 H) 9.07 (d, $J=2.20$ Hz, 1 H) 9.15 (s, 1 H). LCMS: 537.1 (M+H); MW: 536.19.



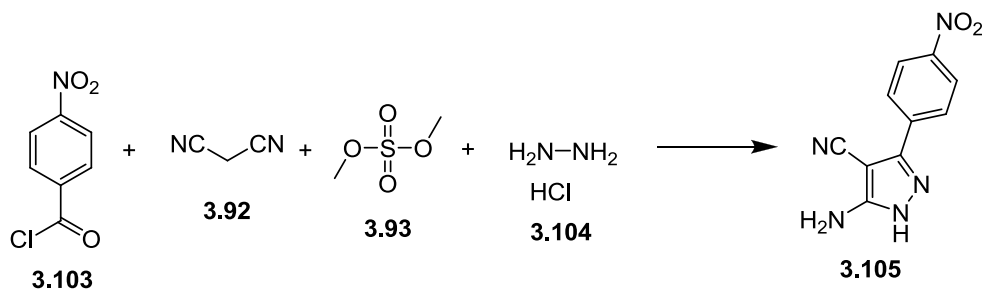
Acrylamide 3.46: synthesized as described for **3.41**. 3% three-step yield from **3.100**. ^1H NMR (400 MHz, DMSO- d_6) δ ppm 1.54 - 1.97 (m, 2 H) 3.38 - 3.54 (m, 1 H) 4.33 - 4.59 (m,

2 H) 5.46 (dd, $J=9.89, 2.56$ Hz, 1 H) 5.69 (dd, $J=10.07, 2.93$ Hz, 1 H) 5.98 (dd, $J=16.85, 2.56$ Hz, 1 H) 6.17 (dd, $J=16.48, 2.56$ Hz, 1 H) 6.44 (dd, $J=16.57, 10.16$ Hz, 1 H) 6.59 (dd, $J=16.76, 10.35$ Hz, 1 H) 7.33 (d, $J=7.69$ Hz, 1 H) 7.53 (t, $J=8.24$ Hz, 1 H) 7.57 - 7.63 (m, 3 H) 7.66 (dd, $J=9.34, 2.75$ Hz, 2 H) 8.04 (t, $J=1.83$ Hz, 1 H) 8.25 (d, $J=7.51$ Hz, 1 H) 9.06 (d, $J=3.48$ Hz, 1 H) 9.14 (s, 1 H). LCMS: 551.3 (M+H); MW: 550.21.

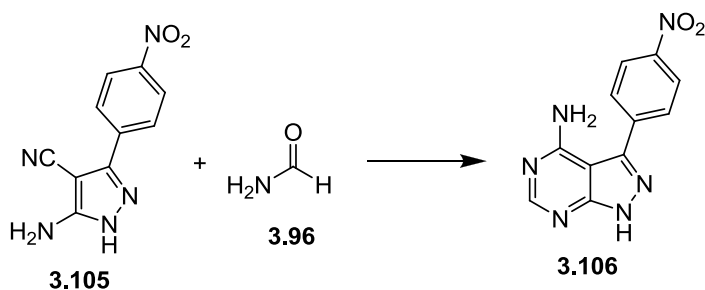


Acrylamide 3.47: synthesized as described for **3.41**. 10% three-step yield in ~70-75% purity from **3.100**. ^1H NMR (400 MHz, DMSO- d_6) δ ppm 2.39 (t, $J=6.41$ Hz, 1 H) 3.54 (q, $J=5.86$ Hz, 1 H) 3.61 (q, $J=6.23$ Hz, 1 H) 4.25 (t, $J=6.41$ Hz, 1 H) 4.40 (dt, $J=15.93, 6.60$ Hz, 2 H) 5.56 (dd, $J=9.89, 2.56$ Hz, 1 H) 5.89 (dd, $J=10.26, 1.65$ Hz, 1 H) 6.01 - 6.18 (m, 2 H) 6.28 (dd, $J=17.21, 1.65$ Hz, 1 H) 7.33 (d, $J=7.69$ Hz, 1 H) 7.53 (t, $J=8.06$ Hz, 1 H) 7.60 (dd, $J=9.34, 1.65$ Hz, 3 H) 7.66 (dt, $J=8.79, 1.47$ Hz, 2 H) 8.04 (t, $J=1.83$ Hz, 1 H) 8.10 (t, $J=5.86$ Hz, 1 H) 8.23 (td, $J=5.86, 1.83$ Hz, 1 H) 9.03 (s, 1 H) 9.12 (s, 1 H). LCMS: 511.2 (M+H); MW: 510.17.

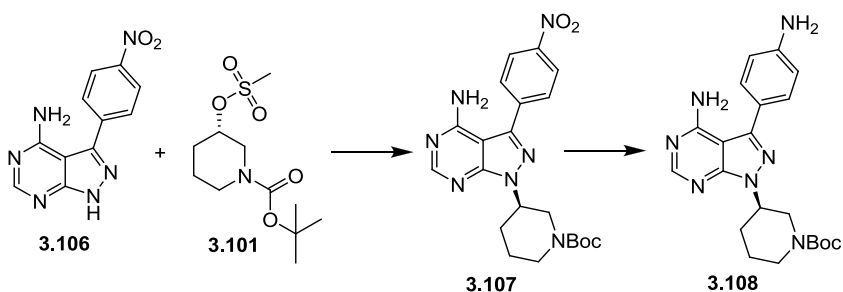
3.41 *meta*-trifluoromethyl phenyl urea replacement series



Pyrazole 3.105: 4-nitrobenzoyl chloride (100 g, 538 mmol, 1.0 equiv.) and malononitrile (39.1 g, 586 mmol, 1.11 equiv.) were dissolved in tetrahydrofuran (500 mL) and cooled in an ice-water bath. Then triethylamine (187.8 mL, 1.35 mol, 2.5 equiv.) was added to the solution dropwise over 1 h. The reaction was then warmed to rt and stirred from an additional 2 h. The dimethyl sulfate (153 mL, 1.62 mol, 3 equiv.) was added in one portion. The reaction was stirred for 3 h at 70 °C and then 15 h at rt. Ethanol (300 mL) was added, and the mixture was cooled in an ice-water bath. Hydrazine monohydrate (80.93 g, 1.62 mol, 3 equiv.) was added, and the reaction was allowed to warm to rt and stirred for 2 h. Then additional hydrazine monohydrate (26.98 g, 530 mmol, 1 equiv.) was added. After 2 h, 1.5 L of water was added dropwise and stirred overnight. An additional 250 mL of water was then added, and the precipitate was filtered. The filtered solid was washed with 500 mL of 1:1 mixture of methanol and water followed by 500 mL of water to obtain 52 g of pyrazole **3.105**, a yellow solid, with some minor impurities. This material was carried forward without further purification. LCMS: 229.8 (M+H); MW: 229.06.



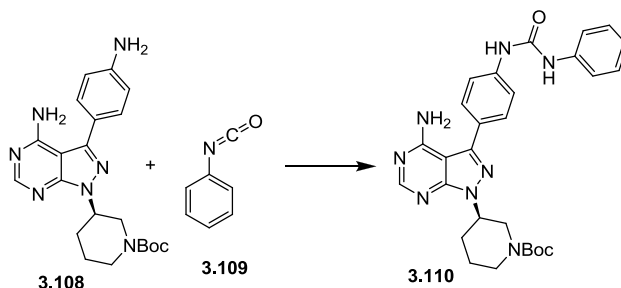
Pyrazolopyrimidine 3.106: pyrazole **3.105** (52 g, 226 mmol, 1.0 equiv.) was dissolved in formamide (300 mL). The resulting solution was heated to 180 °C for 4 h and then cooled to 80 °C. Then water (150 mL) was added, and after 2 h, the mixture was cooled to rt. The solid precipitate was filtered out and washed with water (300 mL) and ethyl acetate (300 mL) resulting in 48 g (34.7% yield over two steps) of pyrazolopyrimidine **3.105**, a yellow solid. ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm 7.94 (ddd, *J*=8.79, 3.11, 2.01 Hz, 2 H) 8.26 (s, 1 H) 8.38 (ddd, *J*=8.97, 2.56, 1.65 Hz, 2 H). LCMS: 256.8 (M+H); MW: 256.07.



Pyrazolopyrimidine 3.107: pyrazolopyrimidine **3.106** (2 g, 7.8 mmol, 1.0 equiv), mesylate **3.101** (5.9 g, 15.6 mmol, 2 equiv.), and cesium carbonate (10.2 g, 31.2 mmol, 4 equiv.) were mixed in DMF (40 mL). The reaction was heated to 105 °C and stirred for

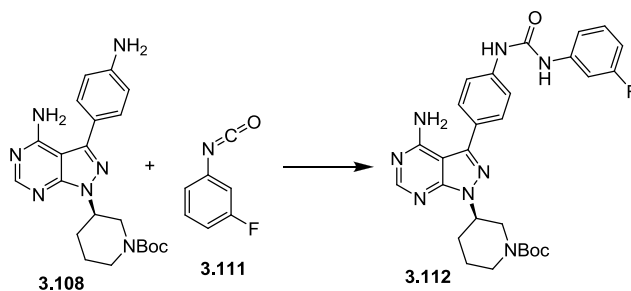
6h. The reaction appeared complete by LCMS. The reaction diluted into ethyl acetate (250 mL) and washed extensively with water and then with brine. The organic layer was dried over magnesium sulfate and concentrated. This material was carried forward without further purification. LCMS: 440.2 (M+H); MW: 439.20.

Pyrazolopyrimidine 3.108: pyrazolopyrimidine **3.107** (3.4 g, 7.8 mmol, 1.0 equiv) was dissolved in THF (40 mL) and acetic acid (20 mL). Then granular zinc (4.94 g, 78 mmol, 10 equiv.) was added. The reaction was heated to 65 °C and stirred vigorously overnight. The reaction was then diluted into ethyl acetate (150 mL); washed with water (2x), 10% ammonium chloride, and brine; dried over magnesium sulfate, and concentrated. The crude was purified by silica flash chromatography (40 g SiliCycle silica flash column) with a gradient of ethyl acetate in hexanes (60-95%). 1.2 g of **3.108**, a yellow solid, was obtained in 37.5% yield over two steps. ¹H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.45 (br. s., 9 H) 1.66 - 1.77 (m, 1 H) 1.83 - 1.96 (m, 1 H) 2.09 - 2.34 (m, 2 H) 2.84 (td, *J*=13.20, 3.11 Hz, 1 H) 3.88 (s, 2 H) 4.77 - 4.87 (m, 2 H) 5.47 (br. s., 2 H) 6.82 (dt, *J*=8.61, 2.01 Hz, 2 H) 7.48 (ddd, *J*=8.42, 2.56, 1.83 Hz, 2 H) 8.36 (s, 1 H). LCMS: 410.2 (M+H); MW: 409.22.



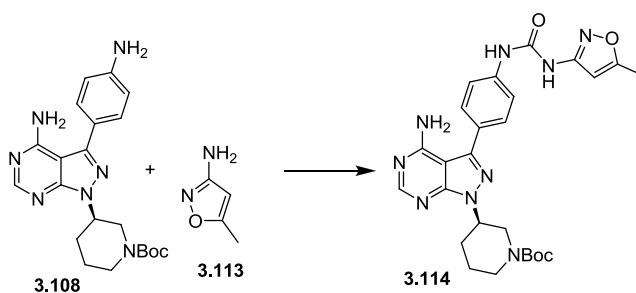
Pyrazolopyrimidine 3.110: pyrazolopyrimidine **3.108** (200 mg, 0.49 mmol, 1.0 equiv.) was dissolved in methylene chloride (4 mL). This solution was chilled in a dry ice-

isopropanol bath. Then isocyanate **3.109** (58.5 μ L, 0.54 mmol, 1.1 equiv.) was added to the solution dropwise. The reaction was allowed to warm to rt over 3 h. The reaction had not proceeded to completion, so the reaction was re-cooled and an additional portion of isocyanate **3.109** (10.6 μ L, 0.10 mmol, 0.2 equiv.) was added. After slow warming, the reaction appeared mostly complete and was concentrated to dryness. The crude was purified by silica flash chromatography (12 g SiliCycle silica flash column) using a mixture of ethyl acetate in hexanes (80-100%, 100% isocratic) to obtain 165 mg (64% yield) of diaryl urea **3.110**, an off-white solid. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.46 (br. s., 9 H) 1.69 - 1.77 (m, 2 H) 1.93 (br. s., 1 H) 2.19 (br. s., 1 H) 2.28 (qd, $J=11.90, 3.66$ Hz, 1 H) 2.94 (td, $J=12.50, 3.10$ Hz, 1 H) 4.10 (d, $J=13.92$ Hz, 1 H) 4.79 - 4.90 (m, 1 H) 5.56 (br. s., 2 H) 7.10 (t, $J=7.23$ Hz, 1 H) 7.32 (t, $J=8.20$ Hz, 3 H) 7.38 (dd, $J=8.42, 1.28$ Hz, 2 H) 7.51 (br. s., 4 H) 8.33 (br. s., 1 H). LCMS: 529.2 (M+H); MW: 528.61.

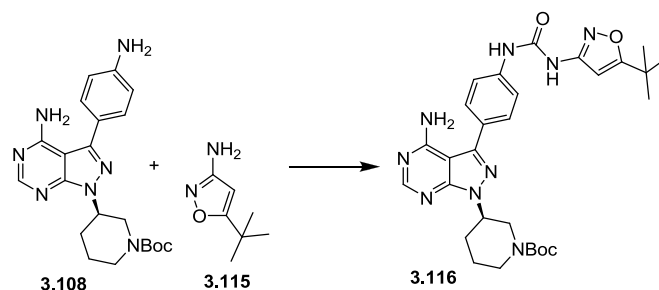


Pyrazolopyrimidine 3.112: pyrazolopyrimidine **3.108** (200 mg, 0.49 mmol, 1.0 equiv.) was dissolved in methylene chloride (4 mL). The resulting solution was chilled in a dry ice-isopropanol bath. Then isocyanate **3.111** (61.4 μ L, 0.54 mmol, 1.1 equiv.) was added to the solution dropwise. The reaction was allowed to warm to rt over 3 h. The reaction appeared mostly complete and was concentrated to dryness. The crude was purified by

silica flash chromatography (12 g SiliCycle silica flash column) using a mixture of ethyl acetate in hexanes (70-100%, 100% isocratic) to obtain 170 mg (64% yield) of diaryl urea **3.112**, an off-white solid. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.48 (br. s., 9 H) 1.72 (br. s., 3 H) 1.96 (br. s., 1 H) 2.20 (br. s., 1 H) 2.29 (qd, $J=11.50, 3.30$ Hz, 1 H) 3.00 (t, $J=12.50$ Hz, 1 H) 4.07 (d, $J=13.19$ Hz, 1 H) 4.86 (quin, $J=5.30$ Hz, 1 H) 5.57 (br. s., 2 H) 6.76 (td, $J=7.69, 2.01$ Hz, 1 H) 7.06 (d, $J=8.24$ Hz, 1 H) 7.23 (t, $J=6.80$ Hz, 1 H) 7.32 (dt, $J=10.62, 2.56$ Hz, 1 H) 7.52 (d, $J=11.17$ Hz, 4 H) 8.33 (br. s., 1 H). LCMS: 547.2 (M+H); MW: 546.25.

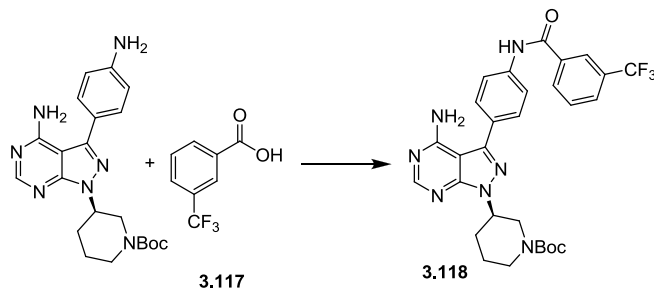


Pyrazolopyrimidine 3.114: 1, 1'-carbonyldiimidazole (CDI, 99 mg, 0.61 mmol, 2.5 equiv.) was dissolved in methylene chloride (1.5 mL) and cooled in an ice-water bath. Isoxazole **3.113** (50 mg, 0.51 mmol, 2.1 equiv.) was dissolved in methylene chloride (0.5 mL) and added dropwise to the solution of CDI. The reaction was stirred for 1 h at rt. This solution was then added in portions to pyrazolopyrimidine **3.108** (100 mg, 0.24 mmol, 1.0 equiv.) in methylene chloride (1 mL) until the reaction was >95% complete. 32 mg (25% yield) of the diaryl urea **3.114**, a white solid, precipitated from solution and was filtered. LCMS: 534.2 (M+H); MW: 533.25.



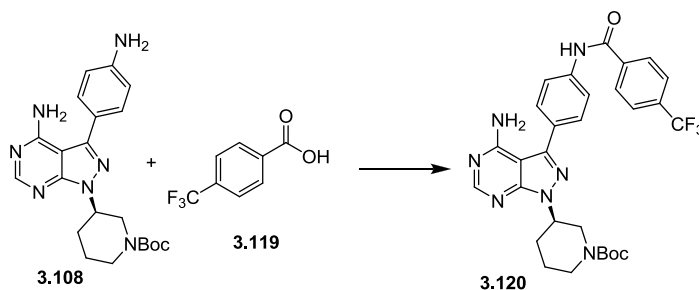
Pyrazolopyrimidine 3.116: triphosgene (68 mg, 0.23 mmol, 0.47 equiv.) was dissolved in methylene chloride (1 mL) and cooled in an ice-water bath. Isoxazole **3.115** (89 mg, 0.63 mmol, 1.3 equiv.) and triethyl amine (204 μ L, 1.47 mmol, 3.0 equiv.) were dissolved in methylene chloride (1 mL) and added dropwise to the triphosgene solution. The reaction was stirred for 1 h, and then added dropwise to a solution of pyrazolopyrimidine **3.108** (200 mg, 0.49 mmol, 1.0 equiv.) in methylene chloride (2 mL) chilled in a dry ice-isopropanol bath. The reaction was allowed to warm to rt over 2 h and was then checked for completion. Two more rounds of isocyanate formation and addition starting with 2.3 equivalents of isoxazole **3.115** were used to further reaction progress. The reaction solution was then washed with 2M NaOH (2 x 4 mL), 1 M HCl (2 x 4 mL), solution of saturated sodium bicarbonate (4 mL), and solution of saturated sodium chloride (4 mL). The organic layer containing some precipitate was dried over magnesium sulfated and concentrated. The crude was purified by silica flash chromatography (12 g SiliCycle silica flash column) using a mixture of ethyl acetate in hexanes (60-90%, 90% isocratic) to obtain 65 mg (23% yield) of diaryl urea **3.116**, a light yellow solid. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.38 (s, 9 H) 1.63 - 1.79 (m, 2 H) 1.92 (br. d, $J=11.70$ Hz, 1 H) 2.15 - 2.34 (m, 2 H) 2.87 (t, $J=11.81$ Hz, 1 H) 4.85 (s, 1 H)

5.64 (br. s., 2 H) 5.86 (br. s., 1 H) 7.67 (dt, $J=8.61$, 1.65 Hz, 2 H) 7.74 (dt, $J=8.79$, 1.83 Hz, 2 H) 8.39 (s, 1 H) 9.61 (br. s., 2 H). LCMS: 576.3 (M+H); MW: 575.30.



Pyrazolopyrimidine 3.118: benzoic acid **3.117** (136 mg, 0.71 mmol, 1.3 equiv.) was dissolved in methylene chloride (2 mL) and reacted with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (148 mg, 0.77 mmol, 1.4 equiv.) and hydroxybenzotriazole hydrate (118 mg, 0.77 mmol, 1.4 equiv.) for 5 min. Then pyrazolopyrimidine **3.108** (225 mg, 0.55 mmol, 1.0 equiv.) dissolved in methylene chloride (2 mL) was added. After 1 h of stirring at rt, the reaction showed minimal progress, and was heated to 40 °C for 1 h. The reaction mixture was then washed with saturated solution of sodium bicarbonate (2 x 4 mL), 0.1M HCl (2 x 4 mL, results in slight precipitation), saturated solution of sodium bicarbonate (1 x 4 mL, resolubilizes the precipitate), and saturated solution of sodium chloride (1 x 4 mL). The organic layer was dried over magnesium sulfate and concentrated. The crude was purified by silica flash chromatography (12 g SiliCycle silica flash column) using a mixture of ethyl acetate in hexanes (60-90%, 90% isocratic) to obtain 40 mg (13% yield) of aryl amide **3.118**, an orange-yellow solid. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.45 (br. s., 9 H) 1.73 (t, $J=12.60$ Hz, 2 H) 1.93 (d, $J=12.09$ Hz, 1 H) 2.16 - 2.34 (m, 2 H) 2.88 (td, $J=13.00$, 2.75 Hz, 0

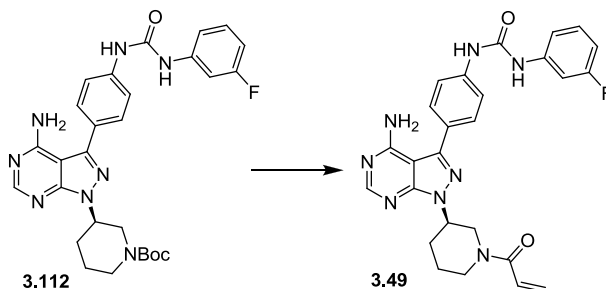
H) 4.81 - 4.91 (m, 1 H) 5.59 (br. s., 2 H) 7.69 (t, $J=7.87$ Hz, 1 H) 7.75 (dt, $J=8.42, 1.83$ Hz, 2 H) 7.86 (dt, $J=8.42, 1.83$ Hz, 2 H) 8.01 (s, 1 H) 8.12 (d, $J=7.87$ Hz, 1 H) 8.18 (s, 1 H) 8.40 (s, 1 H). LCMS: 582.1 (M+H); MW: 581.24.



Pyrazolopyrimidine 3.120: benzoic acid **3.119** (136 mg, 0.71 mmol, 1.3 equiv.) was dissolved in methylene chloride (2 mL) and reacted with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (148 mg, 0.77 mmol, 1.4 equiv.) and hydroxybenzotriazole hydrate (118 mg, 0.77 mmol, 1.4 equiv.) for 5 min. Then pyrazolopyrimidine **3.108** (225 mg, 0.55 mmol, 1.0 equiv.) dissolved in methylene chloride (2 mL) was added. After 1 h of stirring at rt, the reaction showed minimal progress, and was heated to 40 °C for 1 h. The reaction mixture was then washed with saturated solution of sodium bicarbonate (2 x 4 mL), 0.1M HCl (2 x 4 mL, results in slight precipitation), saturated solution of sodium bicarbonate (1 x 4 mL, resolubilizes the precipitate), and saturated solution of sodium chloride (1 x 4 mL). The organic layer was dried over magnesium sulfate and concentrated. The crude was purified by silica flash chromatography (12 g SiliCycle silica flash column) using a mixture of ethyl acetate in hexanes (60-90%, 90% isocratic) to obtain 50 mg (16% yield) of aryl amide **3.120**, an orange-yellow solid. ^1H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.45 (br. s., 9 H) 1.65 -

1.77 (m, 3 H) 1.93 (d, $J=11.35$ Hz, 1 H) 2.15 - 2.34 (m, 2 H) 2.88 (td, $J=12.82, 2.75$ Hz, 1 H) 4.81 - 4.92 (m, 1 H) 5.54 (br. s., 2 H) 7.74 (dt, $J=8.42, 1.83$ Hz, 0 H) 7.81 (d, $J=8.24$ Hz, 1 H) 7.85 (ddd, $J=8.61, 2.38, 1.65$ Hz, 1 H) 8.02 (d, $J=6.23$ Hz, 1 H) 8.05 (s, 1 H) 8.40 (s, 1 H).
LCMS: 582.1 (M+H); MW: 581.24.

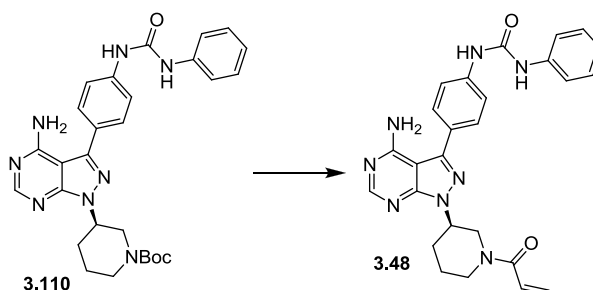
Synthesis of JC080-096 (3.49) - General procedure for acrylamide formation of *meta*-trifluoromethyl urea replacement series



Acrylamide 3.49: pyrazolopyrimidine **3.112** (72 mg, 0.12 mmol, 1.0 equiv) was deprotected in 4 N HCl in 1, 4-dioxanes (4 mL) at rt for 30 min. After concentration, the deprotected material was combined with diisopropylethylamine (24 mg, 32 μ L, 0.18 mmol, 5 equiv.), acrylic acid (3.95 mg, 3.8 μ L, 0.055 mmol, 1.5 equiv.), and PyBOP (28.6 mg, 0.055 mmol, 1.5 equiv.) in anhydrous DMF (1.5 mL). After 30 min – 1 h, brine (9 mL) was added. The mixture was extracted with ethyl acetate (3 x 4-5 mL). The organic layers were combined, dried over sodium sulfate, and concentrated. The crude was purified by preparatory TLC (10% methanol in ethyl acetate). 12.5 mg (68% yield) of **3.49**, a white solid, was obtained. ^1H NMR (400 MHz, DMSO- d_6) δ ppm 1.60 (d, $J=12.64$ Hz, 1 H) 1.93 (dt, $J=13.55, 2.70$ Hz, 1 H) 2.12 (d, $J=12.27$ Hz, 1 H) 2.28 (qd, $J=12.30, 3.85$

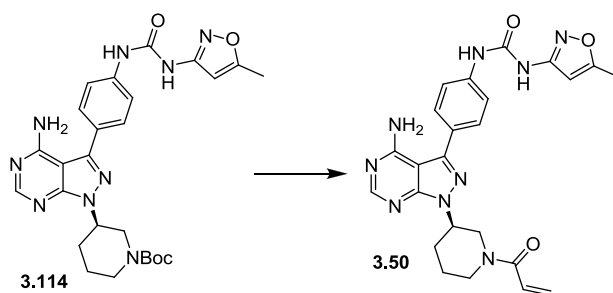
Hz, 1 H) 2.92 - 3.06 (m, 1 H) 3.14 - 3.28 (m, 1 H) 3.70 (t, $J=12.27$ Hz, 1 H) 4.08 (d, $J=13.19$ Hz, 1 H) 4.22 (t, $J=12.82$ Hz, 1 H) 4.55 (d, $J=9.16$ Hz, 1 H) 4.69 (br. s., 1 H) 5.60 (d, $J=10.62$ Hz, 1 H) 5.71 (d, $J=9.89$ Hz, 1 H) 6.10 (dd, $J=23.99, 16.48$ Hz, 1 H) 6.74 (dd, $J=19.96, 12.09$ Hz, 1 H) 6.77 (td, $J=8.47, 2.47$ Hz, 1 H) 6.87 (dd, $J=17.76, 11.90$ Hz, 1 H) 7.10 - 7.10 (m, 0 H) 7.19 (d, $J=8.61$ Hz, 1 H) 7.30 (q, $J=7.30$ Hz, 1 H) 7.53 (dt, $J=12.04, 2.31$ Hz, 1 H) 7.57 (d, $J=8.42$ Hz, 2 H) 7.68 (d, $J=8.61$ Hz, 2 H) 8.25 (s, 1 H) 9.66 (br. s., 1 H). LCMS: 501.2 (M+H); MW: 500.21.

In some cases, the product can be contaminated with tris(pyrrolidinophosphine) oxide after purification with the methanol/ethyl acetate solvent system. This can be separated from the desired products with a second purification step by preparatory TLC using mobile phase comprised of a mixture of isopropanol (1:5 – 1:3) in toluene.



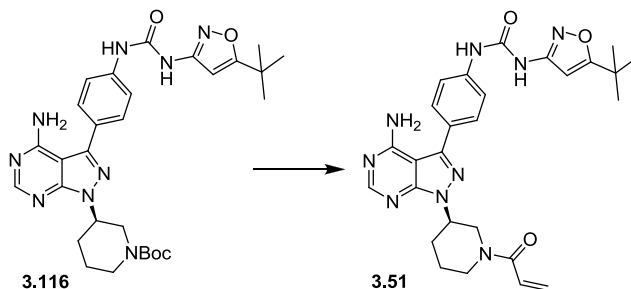
Acrylamide 3.48: synthesized as described for **3.49** from boc-protected amine **3.110** except purification. The crude was first purified by preparative TLC using 10% methanol in ethyl acetate as the mobile phase affording product contaminated with tris(pyrrolidinophosphine) oxide. A second preparative TLC step using 3:1 toluene/isopropanol as the mobile phase afforded the desired acrylamide, a white solid, in 28% yield. ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ ppm 1.60 (d, $J=13.92$ Hz, 1 H) 1.94 (dt,

$J=13.55$, 4.03 Hz, 1 H) 2.14 (d, $J=14.83$ Hz, 1 H) 2.28 (qd, $J=11.50$, 3.80 Hz, 1 H) 3.01 (t, $J=12.60$ Hz, 1 H) 3.22 (t, $J=11.50$ Hz, 1 H) 3.71 (t, $J=10.99$ Hz, 1 H) 4.09 (d, $J=11.35$ Hz, 1 H) 4.22 (t, $J=13.90$ Hz, 1 H) 4.56 (d, $J=11.54$ Hz, 1 H) 4.70 (d, $J=6.23$ Hz, 1 H) 5.60 (d, $J=10.99$ Hz, 1 H) 5.72 (d, $J=9.34$ Hz, 1 H) 6.11 (dd, $J=25.27$, 15.93 Hz, 1 H) 6.73 (dd, $J=16.39$, 11.08 Hz, 1 H) 6.87 (dd, $J=16.66$, 10.07 Hz, 1 H) 6.99 (tt, $J=7.33$, 1.28 Hz, 1 H) 7.30 (tt, $J=7.78$, 1.83 Hz, 2 H) 7.49 (dt, $J=7.51$, 1.83 Hz, 2 H) 7.58 (d, $J=8.42$ Hz, 2 H) 7.65 (dt, $J=8.42$, 2.20 Hz, 2 H) 8.26 (s, 1 H) 8.84 (s, 1 H) 9.01 (s, 1 H). LCMS: 483.1 (M+H); MW: 482.2.

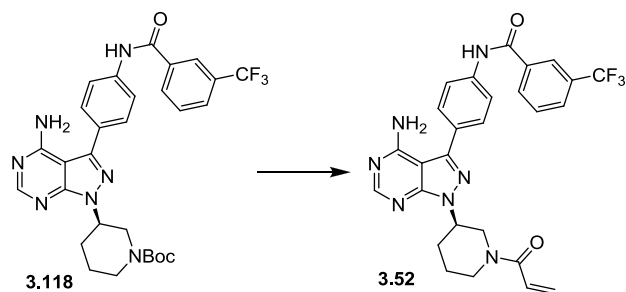


Acrylamide 3.50: synthesized as described for **3.49** from boc-protected amine **3.114** except purification. The crude was first purified by preparative TLC using 10% methanol in ethyl acetate as the mobile phase affording product contaminated with tris(pyrrolidinophosphine) oxide. A second preparative TLC step using 5:1 toluene/isopropanol as the mobile phase afforded the desired acrylamide, a white solid, in 38% yield. ^1H NMR (400 MHz, DMSO- d_6) δ ppm 1.60 (d, $J=11.17$ Hz, 1 H) 1.93 (d, $J=13.92$ Hz, 1 H) 2.12 (d, $J=10.99$ Hz, 1 H) 2.27 (qd, $J=11.70$, 3.30 Hz, 1 H) 2.37 (s, 3 H) 3.00 (t, $J=12.10$ Hz, 1 H) 3.23 (d, $J=10.62$ Hz, 1 H) 3.68 (s, 1 H) 4.02 - 4.26 (m, 1 H) 4.55 (d, $J=12.45$ Hz, 1 H) 4.70 (br. s., 1 H) 5.60 (d, $J=9.71$ Hz, 1 H) 5.71 (d, $J=11.72$ Hz, 1 H) 6.10

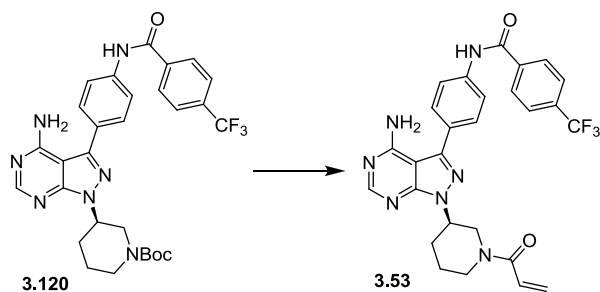
(dd, $J=24.36, 17.95$ Hz, 1 H) 6.56 (s, 1 H) 6.72 (dd, $J=16.48, 10.80$ Hz, 1 H) 6.87 (dd, $J=16.66, 10.80$ Hz, 1 H) 7.62 (q, $J=8.61$ Hz, 4 H) 8.25 (s, 1 H) 9.18 (br. s., 1 H) 9.64 (br. s., 1 H). LCMS: 487.21 (M+H); MW: 488.1.



Acrylamide 3.51: synthesized as described for **3.49** from boc-protected amine **3.116** except purification. The crude was first purified by preparative TLC using 10% methanol in ethyl acetate as the mobile phase affording product contaminated with tris(pyrrolidinophosphine) oxide. A second preparative TLC step using 5:1 toluene/isopropanol as the mobile phase afforded the desired acrylamide, a white solid, in 24% yield. ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ ppm 1.30 (s, 9 H) 1.60 (d, $J=9.89$ Hz, 1 H) 1.93 (d, $J=14.28$ Hz, 1 H) 2.13 (d, $J=12.64$ Hz, 1 H) 2.27 (qd, $J=12.09, 5.13$ Hz, 1 H) 3.01 (s, 1 H) 3.21 (dd, $J=13.19, 11.35$ Hz, 1 H) 3.70 (dd, $J=12.64, 11.17$ Hz, 1 H) 4.10 (dd, $J=14.47, 12.09$ Hz, 1 H) 4.22 (dd, $J=15.57, 12.64$ Hz, 1 H) 4.55 (d, $J=11.54$ Hz, 1 H) 4.70 (br. s., 1 H) 5.60 (d, $J=10.62$ Hz, 1 H) 5.71 (d, $J=9.34$ Hz, 1 H) 6.02 - 6.17 (m, 1 H) 6.53 (s, 1 H) 6.73 (dd, $J=15.90, 9.50$ Hz, 1 H) 6.87 (dd, $J=17.95, 11.35$ Hz, 1 H) 7.62 (q, $J=8.55$ Hz, 4 H) 8.25 (s, 1 H) 9.14 (br. s., 1 H) 9.66 (br. s., 1 H). LCMS: 530.2 (M+H); MW: 529.25.



Acrylamide 3.52: synthesized and purified by preparative TLC (10% methanol in ethyl acetate mobile phase) as described for **3.49** from boc-protected amine **3.118** affording the desired acrylamide, a white solid, in 34% yield. ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ ppm 1.60 (br. s., 1 H) 1.94 (d, $J=15.20$ Hz, 1 H) 2.13 (d, $J=10.80$ Hz, 1 H) 2.29 (q, $J=11.00$ Hz, 1 H) 3.02 (q, $J=7.00$ Hz, 1 H) 3.22 (q, $J=10.40$ Hz, 1 H) 3.34 (t, $J=7.30$ Hz, 1 H) 3.53 - 3.76 (m, 1 H) 4.11 (t, $J=14.65$ Hz, 1 H) 4.23 (br. s., 1 H) 4.57 (d, $J=13.92$ Hz, 1 H) 4.71 (t, $J=7.30$ Hz, 1 H) 5.53 - 5.78 (m, 2 H) 6.02 - 6.19 (m, 2 H) 6.57 (ddd, $J=16.85, 10.44, 1.83$ Hz, 1 H) 6.74 (dd, $J=18.13, 12.27$ Hz, 1 H) 6.88 (dd, $J=15.93, 9.52$ Hz, 1 H) 7.69 (d, $J=8.97$ Hz, 2 H) 7.82 (t, $J=8.06$ Hz, 1 H) 7.95 - 8.04 (m, 2 H) 8.23 - 8.39 (m, 2 H) 10.67 (s, 1 H). LCMS: 536.2 (M+H); MW: 535.19.



Acrylamide 3.53: synthesized and purified by preparative TLC (10% methanol in ethyl acetate mobile phase) as described for **3.49** from boc-protected amine **3.120** affording the desired acrylamide, a white solid, in 45% yield. ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ ppm 1.60 (br. s., $J=11.40, 11.40$ Hz, 1 H) 1.94 (m, $J=15.20$ Hz, 1 H) 2.13 (d, $J=12.82$ Hz, 1 H) 2.29 (qd, $J=13.00, 3.11$ Hz, 1 H) 3.22 (q, $J=12.50$ Hz, 2 H) 3.62 - 3.76 (m, 1 H) 4.09 (d, $J=12.27$ Hz, 1 H) 4.23 (t, $J=10.60$ Hz, 1 H) 4.57 (d, $J=11.72$ Hz, 1 H) 4.71 (br. s., 2 H) 5.52 - 5.77 (m, 2 H) 6.00 - 6.20 (m, 2 H) 6.57 (dd, $J=16.85, 10.26$ Hz, 1 H) 6.74 (dd, $J=15.93, 10.62$ Hz, 1 H) 6.88 (dd, $J=16.66, 11.17$ Hz, 1 H) 7.69 (d, $J=8.06$ Hz, 1 H) 7.97 (dd, $J=16.48, 8.61$ Hz, 4 H) 8.19 (d, $J=8.61$ Hz, 2 H) 8.27 (s, 1 H) 10.67 (s, 1 H).. LCMS: 536.1 (M+H); MW: 535.19.

References

1. Hanes, J., von der Kammer, H., Kludiny, J. & Scheit, K. H. Characterization by cDNA cloning of two new human protein kinases. Evidence by sequence comparison of a new family of mammalian protein kinases. *J Mol Biol* **244**, 665-672 (1994).
2. Yun, B., Farkas, R., Lee, K. & Rabinow, L. The Doa locus encodes a member of a new protein kinase family and is essential for eye and embryonic development in *Drosophila melanogaster*. *Genes Dev* **8**, 1160-1173 (1994).
3. Bender, J. & Fink, G. R. AFC1, a LAMMER kinase from *Arabidopsis thaliana*, activates STE12-dependent processes in yeast. *Proc Natl Acad Sci U S A* **91**, 12105-12109 (1994).
4. Tang, Z. *et al.* LAMMER kinase Kic1 is involved in pre-mRNA processing. *Exp Cell Res* **317**, 2308-2320 (2011).
5. Giannakouros, T., Nikolakaki, E., Mylonis, I. & Georgatsou, E. Serine-arginine protein kinases: a small protein kinase family with a large cellular presence. *FEBS J* **278**, 570-586 (2011).
6. Mair, G. *et al.* A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. *RNA* **6**, 163-169 (2000).
7. Jae, N. *et al.* Essential role of a trypanosome U4-specific Sm core protein in small nuclear ribonucleoprotein assembly and splicing. *Eukaryot Cell* **9**, 379-386 (2010).
8. Gunzl, A. The pre-mRNA splicing machinery of trypanosomes: complex or simplified? *Eukaryot Cell* **9**, 1159-1170 (2010).
9. Ismaili, N. *et al.* Characterization of a SR protein from *Trypanosoma brucei* with homology to RNA-binding cis-splicing proteins. *Mol Biochem Parasitol* **102**, 103-115 (1999).
10. Lee, K., Du, C., Horn, M. & Rabinow, L. Activity and autophosphorylation of LAMMER protein kinases. *J Biol Chem* **271**, 27299-27303 (1996).
11. Nett, I. R., Davidson, L., Lamont, D. & Ferguson, M. A. Identification and specific localization of tyrosine-phosphorylated proteins in *Trypanosoma brucei*. *Eukaryot Cell* **8**, 617-626 (2009).

12. Knight, Z. A. & Shokat, K. M. Features of selective kinase inhibitors. *Chem Biol* **12**, 621-637 (2005).
13. Schirmer, A., Kennedy, J., Murli, S., Reid, R. & Santi, D. V. Targeted covalent inactivation of protein kinases by resorcylic acid lactone polyketides. *Proc Natl Acad Sci U S A* **103**, 4234-4239 (2006).
14. Singh, J., Petter, R. C. & Kluge, A. F. Targeted covalent drugs of the kinase family. *Curr Opin Chem Biol* **14**, 475-480 (2010).
15. Pan, Z. *et al.* Discovery of selective irreversible inhibitors for Bruton's tyrosine kinase. *ChemMedChem* **2**, 58-61 (2007).
16. Hanke, J. H. *et al.* Discovery of a novel, potent, and Src family-selective tyrosine kinase inhibitor. Study of Lck- and FynT-dependent T cell activation. *J Biol Chem* **271**, 695-701 (1996).
17. Zhou, W. *et al.* Novel mutant-selective EGFR kinase inhibitors against EGFR T790M. *Nature* **462**, 1070-1074 (2009).
18. Kluge, A. *et al.* Heteroaryl compounds and uses thereof. U.S. patent WO/2009/158571 (12/30/2009).
19. Zarrinkar, P. P. *et al.* AC220 is a uniquely potent and selective inhibitor of FLT3 for the treatment of acute myeloid leukemia (AML). *Blood* **114**, 2984-2992 (2009).
20. James, J. *et al.* CEP-32496: a novel orally active BRAF(V600E) inhibitor with selective cellular and in vivo antitumor activity. *Mol Cancer Ther* **11**, 930-941 (2012).
21. Dar, A. C., Lopez, M. S. & Shokat, K. M. Small molecule recognition of c-Src via the Imatinib-binding conformation. *Chem Biol* **15**, 1015-1022 (2008).
22. Liu, Y. & Gray, N. S. Rational design of inhibitors that bind to inactive kinase conformations. *Nat Chem Biol* **2**, 358-364 (2006).
23. Gourguechon, S., Savich, J. M. & Wang, C. C. The multiple roles of cyclin E1 in controlling cell cycle progression and cellular morphology of *Trypanosoma brucei*. *J Mol Biol* **368**, 939-950 (2007).
24. Tu, X. & Wang, C. C. The involvement of two cdc2-related kinases (CRKs) in *Trypanosoma brucei* cell cycle regulation and the distinctive stage-specific phenotypes caused by CRK3 depletion. *J Biol Chem* **279**, 20519-20528 (2004).

25. Li, Z., Umeyama, T. & Wang, C. C. Polo-like kinase guides cytokinesis in *Trypanosoma brucei* through an indirect means. *Eukaryot Cell* **9**, 705-716 (2010).
26. Tu, X., Kumar, P., Li, Z. & Wang, C. C. An aurora kinase homologue is involved in regulating both mitosis and cytokinesis in *Trypanosoma brucei*. *J Biol Chem* **281**, 9677-9687 (2006).
27. Ojo, K. K. *et al.* Glycogen synthase kinase 3 is a potential drug target for African trypanosomiasis therapy. *Antimicrob Agents Chemother* **52**, 3710-3717 (2008).
28. Jetton, N. *et al.* The cell cycle as a therapeutic target against *Trypanosoma brucei*: Hesperadin inhibits Aurora kinase-1 and blocks mitotic progression in bloodstream forms. *Mol Microbiol* **72**, 442-458 (2009).
29. Li, Z., Umeyama, T. & Wang, C. C. The Aurora Kinase in *Trypanosoma brucei* plays distinctive roles in metaphase-anaphase transition and cytokinetic initiation. *PLoS Pathog* **5**, e1000575 (2009).

CHAPTER 3

Chemical-Biological Characterization of a Cruzain Inhibitor Reveals a Second Target and a Mammalian Off-Target

Chapter 3 is reproduced in part with permission from: Choy, J.W., Bryant, C., Calvet, C.M., Doyle, P.S., Gunatilleke, S.S., Leung, S.S.F., Ang, K.K.H., Chen, S., Gut, J., Osés-Príteo, J.A., Johnston, J.B., Arkin, M.R., Burlingame, A.L., Taunton, J., Jacobson, M.P., McKerrow, J.M., Podust, L.M., Renslo, A.R. Chemical–biological characterization of a cruzain inhibitor reveals a second target and a mammalian off-target. *Beilstein J. Org. Chem* **9**, 15–25 (2013).

Abstract

Inhibition of the *Trypanosoma cruzi* cysteine protease cruzain has been proposed as a therapeutic approach for the treatment of Chagas' disease. Among the best-studied cruzain inhibitors to date is the vinylsulfone K777 (**4.1**), which has proven effective in animal models of Chagas' disease. Recent structure-activity studies aimed at addressing potential liabilities of **4.1** have now produced analogs such as N-[(2S)-1-[[[E,3S)-1-(benzenesulfonyl)-5-phenylpent-1-en-3-yl]amino]-3-(4-methylphenyl)-1-oxopropan-2-yl]pyridine-4-carboxamide (**4.4**), which is trypanocidal at ten-fold lower concentrations than for **4.1**. We now find that the trypanocidal activity of **4.4** derives primarily from inhibition of *T. cruzi* 14 α -demethylase (*TcCYP51*), a cytochrome P450 enzyme involved in the biosynthesis of ergosterol in the parasite. Compound **4.4** also inhibits mammalian CYP isoforms but is trypanocidal at concentrations below those required to significantly inhibit mammalian CYPs *in vitro*. A chemical proteomics approach employing an activity-based probe derived from **4.1** was used to identify mammalian cathepsin B as a potentially important off-target of **4.1** and **4.4**. Computational docking studies and the evaluation of truncated analogs of **4.4** reveal structural determinants for *TcCYP51* binding, information that will be useful in further optimizing this new class of inhibitors.

Introduction

The kinetoplastid protozoan *Trypanosoma cruzi* is the causative agent of Chagas' disease, a leading cause of heart failure in endemic regions of Latin America [1]. The parasite is transmitted by the reduviid bug and disease is manifest in an initial acute phase, followed by a chronic phase that can last decades and typically culminates in heart failure. Existing therapy for Chagas' disease involves extended therapy with nifurtimox or benznidazole, both of which are associated with undesirable side-effects and have limited efficacy against the chronic stage of disease [2, 3]. This situation has spurred the search for more effective and better tolerated therapeutics [4-6]. Among a number of drug targets being investigated are cruzain [7-10], the major cysteine protease activity in the parasite, and *T. cruzi* CYP51 (*TcCYP51*), a 14- α -demethylase enzyme of the cytochrome P450 family required for ergosterol biosynthesis [11-14]. *TcCYP51* is analogous to the fungal enzyme targeted by the azole class of antifungals, and the observation that some of these drugs (e.g., posaconazole) also inhibit *TcCYP51* has led to their preclinical and clinical investigation as potential new treatments for Chagas' disease [2, 15, 16].

Cruzain is a cathepsin L-like protease of the papain family thought to be important for intracellular replication and differentiation of the *T. cruzi* parasite [17]. A variety of small molecule cruzain inhibitors have been described, the majority of which act irreversibly via reaction with the catalytic cysteine in the enzyme active site [18-27].

One of the earliest cruzain inhibitors identified and perhaps the best studied to date is the vinylsulfone K777 (**4.1**, Figure 4.1). This irreversible inhibitor has demonstrated efficacy in animal models of Chagas' disease [28, 29] and continues to undergo preclinical evaluation leading towards a possible human clinical trial.

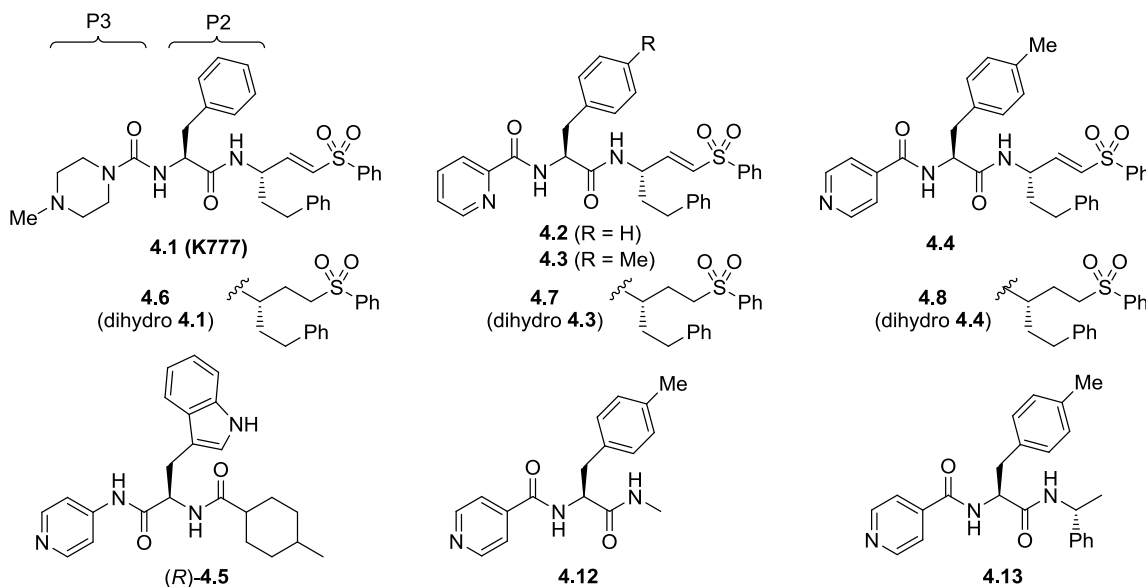


Figure 4.1. Chemical structures of vinylsulfone based cruzain inhibitors **4.1-4**, known TcCYP51 inhibitor **4.5**, dihydro controls **4.6-8**, and truncated analogs **4.12** and **4.13**. The “P2” and “P3” side chains of **4.1** are labeled and bind, respectively, in the S2 and S3 sub-sites of the cruzain active site.

Despite many favorable properties, some aspects of **4.1** are sub-optimal from a drug development perspective. For example, compound **4.1** is known to be a mechanism-based (irreversible) inhibitor of CYP3A4, an enzyme responsible for the metabolism of many drugs, including **4.1** itself [30]. In pharmacokinetic studies, compound **4.1** exhibits non-linear exposure with escalating dose and is known to be a substrate of the drug transporter P-glycoprotein (P-gp). Finally, as a basic (protonatable) drug species, **4.1** could potentially accumulate in acidic lysosomes, where mammalian cathepsins

(potential off-targets of **4.1**) are located. The issue of lysosomotropism figured prominently in the discovery and clinical development of cathepsin K inhibitors for osteoporosis. The first such inhibitor to successfully navigate human clinical trials is odanacatib, which was intentionally designed as a non-basic drug species to minimize the potential for lysosomotropic behavior [31, 32].

We sought to address the question of lysosomotropism by preparing analogs of **4.1** in which the basic piperazine substituent at “P3” (which binds the S3 subsite of cruzain) was replaced with non-basic or weakly basic heterocycles. In our initial structure-activity study [21], we found that analog **4.2** (Figure 4.1), bearing a 2-pyridyl amide at the P3 position, possessed trypanocidal activity that was on par with **4.1** (Table 4.1). However, none of the non-basic analogs examined proved superior to **4.1** and only 2-pyridyl analogs such as **4.2** and **4.3** appeared even comparable. We therefore turned to more dramatic structural alteration and successfully identified and structurally characterized a new non-peptidic cruzain inhibitor chemotype [24]. Most recently, we returned to reinvestigate non-basic analogs of **4.1** and now report that 4-pyridyl analogs such as **4.4** (Figure 4.1) are significantly more trypanocidal than **4.1** or **4.2**, and unexpectedly exert their trypanocidal effects primarily via inhibition of *TcCYP51* rather than cruzain.

Inhibition of *TcCYP51* by the 4-pyridyl analogs is predicted to occur through a reversible mechanism, but **4.4** still contains the reactive vinylsulfone “warhead” and is expected to

irreversibly inhibit cysteine proteases. Analog **4.4** was conceived as a non-basic cruzain inhibitor, but its surprising potency against parasites suggested that other targets might also be inhibited by the compound. As described below, **4.4** was indeed found to target both cruzain and the 14- α -demethylase enzyme TcCYP51. To investigate the importance of cysteine protease inhibition by **4.4** and related putatively multi-targeted analogs, we developed a propargylated analog of K777 (**4.9**) as an activity-based probe. Independently, another group also reported the synthesis of **4.9** and its use to study cysteine protease targets in *T. brucei* [41]. Unlike *T. brucei* however, culturing intracellular *T. cruzi* amastigotes requires a mammalian host cell. In treating intracellular *T. cruzi* amastigotes with **4.9**, we found that a host protein was strongly labeled. This protein, later determined to be cathepsin B, is also inhibited by **4.1** and **4.4** but not by their reduced forms **4.6** and **4.8**. With any drug candidate, off-targets in the host are of concern, especially those of irreversible mechanism-based inhibitors. In this way, this probe has proved to be and will continue to be valuable for testing the activity of current and future analogs against this target.

Results and Discussion

Structure-Activity Studies

Our exploration of the P3 position in **4.1** included the evaluation of regioisomeric 2-, 3-, and 4-pyridyl congeners in the context of various P2 side chains. In multiple such analog series, we found that regioisomeric analogs possessed similar cruzain activities *in vitro*, while the 4-pyridyl examples consistently demonstrated superior trypanocidal activity against cultured *T. cruzi* parasites. For example, 4-pyridyl analogs (e.g. **4.4**) exhibited sub-micromolar minimal trypanocidal concentration values (MTC = 0.6 μ M) while the MTC values for 2-pyridyl (e.g. **4.3**) and 3-pyridyl analogs were typically \sim 10 μ M, which was similar to the MTC of **4.1** (Table 4.1). The MTC represents the minimum concentration of test compound required to completely clear *T. cruzi* parasites from J774 macrophage host cells over a 40-day experiment, with test compound administered during the initial 28 days.

The enhanced potency of 4-pyridyl analogs as compared to **4.1** or their regioisomeric analogs was not predictable on the basis of *in vitro* cruzain activity (Table 4.1). Nor could the trends be explained as an effect of lysosomotropism, since enhanced potency was observed only for the 4-pyridyl analogs and not for 2- or 3-pyridyl analogs, which have similar pKa values. Instead, we considered that additional target(s) might be responsible for the surprising potency of the 4-pyridyl analogs. Specifically, we were aware that a 4-pyridyl ring comprises the putative heme-binding moiety in a new class

of *TcCYP51* inhibitors represented by compound **4.5** (Figure 4.1). Other structural similarities of **4.4** and **4.5** suggested that compound **4.4** could conceivably bind *TcCYP51*.

Table 4.1. *In vitro* biochemical and cellular activities of test compounds and controls.

Cmpd	Cruzain activity	<i>TcCYP51</i> activity		<i>T. cruzi</i> Growth Inhibition	
		k_{inact}/K_i ($s^{-1} M^{-1}$)	<i>In vitro</i> K_D (nM)	Cellular activity (Y/N, conc.) ^a	MTC ^b (μM)
4.1	118,000	>2,000	N (1.6 μM)	8	0.10
4.2	120,000	---	---	10	---
4.3	16,000	>2,000	N (2.0 μM)	8	1.85
4.4	67,300	≤ 5	Y (0.2 μM)	0.6	0.10
4.5	---	≤ 5	Y (5.0 μM)	$\leq 10^f$	---
4.6	N.A.	>2,000	N (2.0 μM)	>10	>10
4.7	N.A.	>2,000	N (0.1 μM)	>10	>10
4.8	N.A.	≤ 5	Y (0.1 μM)	0.25	0.11
4.9	81,500	---	---	5 ^d	0.017
4.12	N.A.	620 \pm 260	---	>10	>10
4.13	N.A.	75 \pm 26	---	1 ^e	3.9
BNZ	---	---	---	10	7.2
POSA	---	≤ 5	Y (0.1 μM)	0.003	2.7

N.A. = not active (cruzain IC₅₀ > 50 μM); BNZ = benznidazole; POSA = posaconazole.

^aCompound affects ergosterol biosynthesis at indicated concentration as determined by GC/MS analysis

^bMinimum Effective Concentration that clears J774 host cells of parasites at day 40 of the experiment, following 28 days of treatment.

^cConcentration that reduces parasite load in C2C12 cells by 90% relative to untreated controls.

^dExperiment performed in BESM host cell rather than J774 cells

^eRead at day 12 following 7 days treatment.

^fConcentrations lower than 10 μM were not examined

To test the hypothesis that **4.4** might also target *TcCYP51*, we examined the binding of this compound to *TcCYP51* using a UV-vis spectroscopic binding assay described previously [33]. Indeed, compound **4.4** bound *TcCYP51* with an estimated $K_D \leq 5$ nM, a

value comparable to the binding affinity of the known *Tc*CYP51 inhibitor **4.5** [16]. 2-Pyridyl analog **4.3** did not measurably bind *Tc*CYP51 ($K_D > 2,000$ nM, Table 4.1), whereas the corresponding 3-pyridyl congener (not shown) binds ~100-fold more weakly ($K_D \sim 500$ nM) than **4.4**. These findings were thus consistent with our hypothesis that the 4-pyridyl ring in **4.4** is involved in binding *Tc*CYP51. The 2-pyridyl ring system in **4.3** is presumably unable to chelate heme in *Tc*CYP51 due to steric hindrance from the immediately adjacent amide linkage.

Computational Docking Studies

We next employed computational docking and a model derived from the crystal structure of *Tc*CYP51 to compare predicted binding modes of **4.4** and (*R*)-**4.5**. The two ligands were docked using the induced-fit docking protocol with Glide XP [34], and the models further refined by minimizing the energies of the ligand and surrounding residues (within 5 Å of ligand) using PRIME [35]. Finally, binding scores were computed using both Glide XP and the MM/GMSA method. Compound **4.4** was predicted to bind in a similar fashion as (*R*)-**4.5**, with the 4-pyridyl ring chelating the heme-iron atom and the tolyl ring at P2 contacting many of the same residues (e.g., Try103, Phe110) predicted to interact with the tryptophan ring of (*R*)-**4.5** (Figure 4.2A). This same hydrophobic site in *Tc*CYP51 binds the fluoroaryl rings of fluconazole and posaconazole in co-crystal structures [14]. The predicted binding mode of the enantiomer (*S*)-**4.5** was described previously [16] and is distinct from that proposed for **4** and (*R*)-**4.5**.

Thus, computational docking provides a conceptual picture of how compound **4.4** – notionally a cruzain inhibitor – might also bind *TcCYP51*. Interestingly, this is not the first time that potent *TcCYP51* binding has been discovered in a molecule originally intended for a different target. Buckner and Gelb unexpectedly found that the human protein farnesyltransferase (PFT) inhibitor tipifarnib exerts its anti-trypanosomal effects via inhibition of *TcCYP51* [36]. Subsequently, these researchers succeeded in divorcing PFT activity from *TcCYP51* inhibition in the tipifarnib scaffold, producing new lead compounds with compelling properties [37-39].

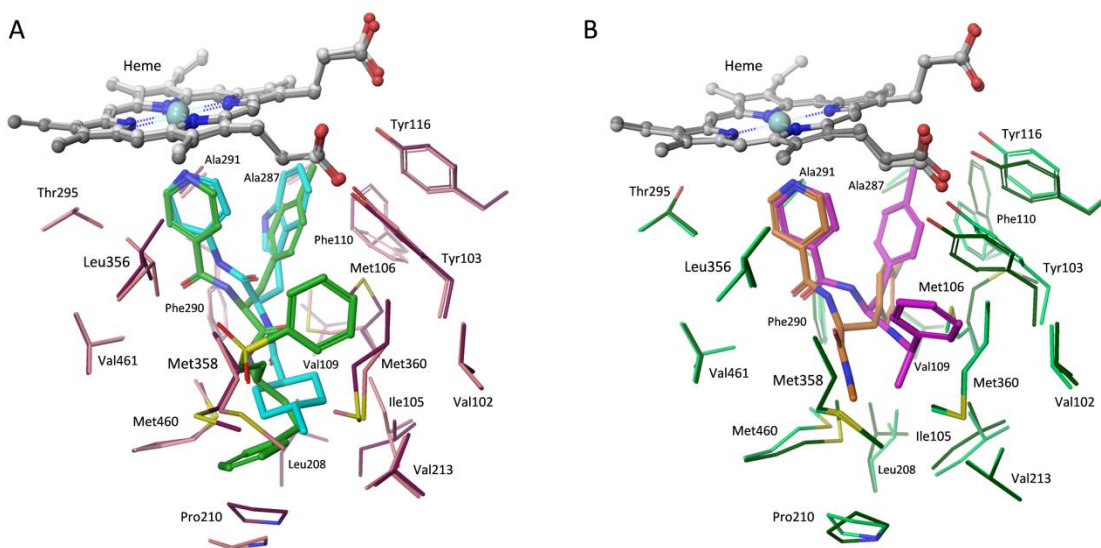


Figure 4.2. Computational docking models. (A) Predicted binding modes of **4.4** and *(R)*-**4.5** bound to *TcCYP51*. For **4.4**, the ligand, the protein, and the heme group are shown in green, pink, and grey, respectively. For *(R)*-**4.5**, the ligand, the protein, and the heme group are shown in cyan, purple, and white, respectively. Heme-iron chelation and hydrophobic binding interactions dominate in the models. (B) Predicted binding models of truncated analogs **4.12** and **4.13** to *TcCYP51*. For **4.12**, the ligand, the protein, and the heme group are shown in orange, light green, and light grey, respectively. For **4.13**, the ligand, the protein, and the heme group are shown in magenta, dark green, and dark grey, respectively.

Inhibition of Mammalian CYPs

A concern with any inhibitor of *Tc*CYP51 is the potential for cross reactivity with mammalian cytochrome P450 (CYP) enzymes, especially those CYPs involved in drug metabolism, like CYP3A4. To assess this risk, we evaluated the inhibitory activities of **4.4** and **4.1** across a panel of relevant mammalian CYP enzymes (Table 4.2). Both **4.4** and **4.1** inhibited all CYPs in the panel, with IC_{50} values generally in the low micromolar range. Although compound **4.4** did inhibit CYP3A4, the potency of inhibition ($IC_{50} = 0.8 \mu\text{M}$) was less than that exhibited by the antifungal drug ketoconazole ($IC_{50} = 0.086 \mu\text{M}$). It should be noted that the substrate-derived IC_{50} values from the CYP panel are not directly comparable to the K_D values for binding to *Tc*CYP51. What can be said is that the anti-trypanosomal effects of **4.4** are realized at concentrations ($EC_{90} = 0.1 \mu\text{M}$, MTC = $0.5 \mu\text{M}$) well below the compound's *in vitro* potency across the CYP panel (average $IC_{50} \sim 7 \mu\text{M}$). Compound **4.4** thus possesses reasonable selectivity with regard to off-target CYP inhibition, and represents a reasonable starting point from which further improvements in selectivity might be undertaken.

Table 4.2. *In vitro* inhibition of important mammalian CYP enzymes.

Compound	IC ₅₀ (μM)				
	1A2	2C9	2C19	2D6	3A4
4.1	24	32	7.6	26	1.7
4.4	22	5.5	3.4	2.7	0.8
ketoconazole	---	---	---	---	0.086

Given the similar IC₅₀ values for **4.1** and **4.4** against CYP3A4, we were curious to determine whether **4.4** is an irreversible inhibitor of this enzyme, as is the case for **4.1** [30]. Irreversible inhibition is typically assessed by measuring the activity of microsomal CYPs following pre-incubation with or without NADPH. Consistent with earlier studies [30], compound **1** exhibited irreversible inhibition of CYP3A4 as reflected in a significantly lower IC₅₀ value with NADPH pre-incubation (Table 4.3). In contrast, compound **4.4** showed behavior typical of reversible inhibition, with no NADPH-dependent shift in the IC₅₀ value. In the case of CYP2C19, both compounds were found to be reversible inhibitors. These results suggest that CYP inhibition by **4.4** involves reversible binding of the parent molecule, while the inhibition of CYP3A4 conferred by **4.1** is dependent on initial conversion to a reactive metabolite. Whatever the explanation, reversible inhibition of CYP enzymes (as with **4.4**) is clearly preferable to irreversible inhibition from a drug safety perspective.

Table 4.3. Mechanism of inhibition studies for compounds **4.1** and **4.4**. These data suggest that inhibition of CYP3A4 by **4.1** is irreversible in nature.

Compound	2C19 IC ₅₀ (μM)		3A4 IC ₅₀ (μM)	
	+ NADPH	- NADPH	+ NADPH	- NADPH
4.1	5.54	8.22	0.0059	1.08
4.4	0.300	0.170	0.117	0.046

Inhibition of *Tc*CYP51 in Live Parasites

We next sought to better define the relative importance of *Tc*CYP51 and cruzain inhibition in the anti-trypanosomal effects of compound **4.4**. Since the 2-pyridyl analog **4.3** was found to not bind *Tc*CYP51, this compound could serve as a control for the cruzain-derived (and/or other cysteine protease-derived) effects of **4.4**. To provide controls lacking activity against cysteine proteases, we reduced the vinylsulfone function in analogs **4.1**, **4.3**, and **4.4** to afford the dihydro analogs **4.6-8** (Figure 4.3). As expected, these analogs were devoid of any detectable cruzain inhibitory activity (IC₅₀ > 50 μM, Table 4.1). Compounds **4.3**, **4.4**, **4.7** and **4.8** thus comprised a set of analogs with complementary activity profiles against the two putative targets: **4.4** (cruzain and *Tc*CYP51 inhibition), **4.8** (*Tc*CYP51 inhibition only), **4.3** (cruzain inhibition only), and **4.7** (neither activity).

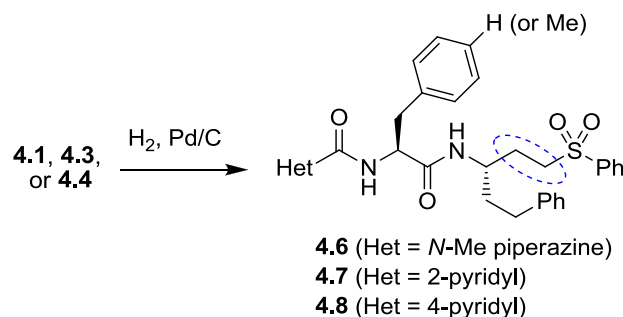


Figure 4.3. Synthesis of the additional control compound **4.6-8**, the reduced forms of analogs **4.1**, **4.3**, and **4.4** respectively.

Compounds **4.3**, **4.4**, **4.7**, and **4.8** were evaluated for potency against intracellular *T. cruzi* parasites using two different assays. The reported EC₉₀ values (Table 4.1) represent compound concentrations required to reduce parasite numbers in C2C12 host cells by 90% as compared to untreated controls, as determined using a high-content imaging-based screening (HCS) approach [33, 40]. This high-throughput assay provides a rapid measure of the initial acute effects of test compound on parasite viability. The more laborious MTC assay identifies compound concentrations that clear parasites from host cell as determined ~two weeks after the conclusion of a four-week course of treatment. This MTC assay therefore provides a measure of trypanocidal action that cannot be drawn from the more rapid HCS assay. We judge that MTC values are more representative of the therapeutic drug levels that would likely be required to produce efficacy in an animal model of Chagas' disease.

The anti-trypanosomal effects of compounds **4.3**, **4.4**, **4.7**, and **4.8** were in general agreement with their *in vitro* activities against the two putative targets (Table 4.1).

Analog **4.7**, devoid of either activity *in vitro*, showed no effects on *T. cruzi* parasites in either the HCS or MTC assay. Analog **4.3**, possessing primarily cysteine protease-derived effects, was effective in both assays and equipotent to **4.1** in the MTC assay. Putatively dual-targeted analog **4.4** was ~10-fold more potent than **4.1** in the MTC assay and equipotent by HCS. Most unexpectedly, we found that compound **4.8**, which lacks any cruzain-derived effects of **4.4**, was equipotent to **4.4** by HCS and 2-4 fold more potent than **4.4** in the MTC assay.

The *in vitro* and cell-based activities of **4.4** and **4.8** suggest *TcCYP51* as a relevant target of these compounds. To assess inhibition of *TcCYP51* in live parasites we analyzed the sterol composition of intracellular *T. cruzi* parasites treated with test compounds **4.3-8**, **4.1**, or posaconazole as a positive control. The analysis was performed employing GC/MS as reported previously for compound **4.5** [33]. The GC/MS trace for uninfected host cells establishes that the additional peaks observed in infected cells are of *T. cruzi* origin (peaks labeled **4.a-i**, Figure 4.4). Treatment with the known *TcCYP51* inhibitor posaconazole produces an increase in the relative abundance of *TcCYP51* substrates lanosterol (**f**) and eburicol (**h**) and accordingly, a reduction in the abundance of downstream sterols such as fecosterol (**e**) and cholesta-7,24-dien-3 β -ol (**a**), among others. Treatment with **4.1** had little effect on sterol composition as expected, whereas treatment with compound **4.4** or **4.8** produced effects very similar to those observed in posaconazole treated parasites (Figure 4.4; Appendix 3, Fig. 1). The other test compounds evaluated (**4.3**, **4.6**, **4.7**) produced no significant change in lipid composition

(Appendix 6, Fig. 1), as expected since these compounds do not inhibit *TcCYP51 in vitro*. Test compounds were necessarily studied at concentrations below their MTC, so as to retain a population of viable parasites for analysis.

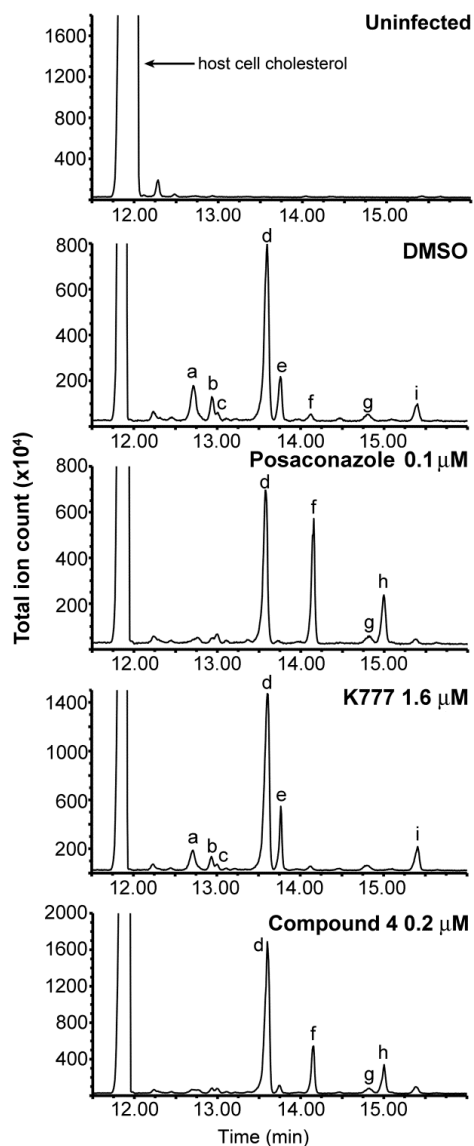


Figure 4.4. GC-MS analysis of lipid extracts from *T. cruzi* parasites treated with test compounds. DMSO and K777 (**4.1**) were used as negative controls; posaconazole served as a positive control. The analysis of **4.4** was performed concurrently with other CYP51 inhibitors described recently [33] and so the spectra for the controls shown above are reproduced from the earlier report. Spectra of lipid extracts from parasites treated with **4.3**, **4.6**, **4.7**, and **4.8** are provided in Appendix 6. Uninfected host cell panel (top) demonstrates that chromatographic peaks labeled **a** to **i** in subsequent panels are of *T. cruzi* origin. These peaks are assigned as **a** - cholesta-7,24-dien-3 β -ol, $[M]^{++} = m/z$ 454; **b** - cholesta-8,24-dien-3 β -ol (zymosterol), $[M]^{++} = m/z$ 470; **c** - 24-methyl-7-en-cholesta-en-3 β -ol, $[M]^{++} = m/z$ 472; **d** - ergosta-7,24-diene-3 β -ol (episterol), $[M]^{++} = m/z$ 470; **e** - ergosta-8,24-diene-3 β -ol (fecosterol), $[M]^{++} = m/z$ 470; **f** - lanosterol, $[M]^{++} = m/z$ 498; **g** - 4-methylepisterol, $[M]^{++} = m/z$ 484; **h** - eburicol, $[M]^{++} = m/z$ 512; **i** - 24-ethyl-7,24(24')-en-cholesta-dien-3 β -ol, $[M]^{++} = m/z$ 484.

An Activity-Based Probe Reveals an Off-Target of 1 and 4

We next sought to evaluate the cysteine protease related effects of the various test compounds in *T. cruzi* parasites. To do this, we designed and synthesized the ‘clickable’ activity-based probe **4.9** in which a propargyl group (replacing methyl in **4.1**) serves as a chemical handle for conjugation to TAMRA or biotin containing reagents (**4.10** and **4.11**, respectively, Figure 4.5). Probe **4.9** was found to be equipotent to **4.1** against cruzain *in vitro* and retained similar effects against *T. cruzi* parasites in both the HCS and MTC assays. Thus, the cysteine protease target(s) of **4.9** in parasite and host cell can reasonably be assumed to be the same as for **4.1** and close analogs like **4.4**.

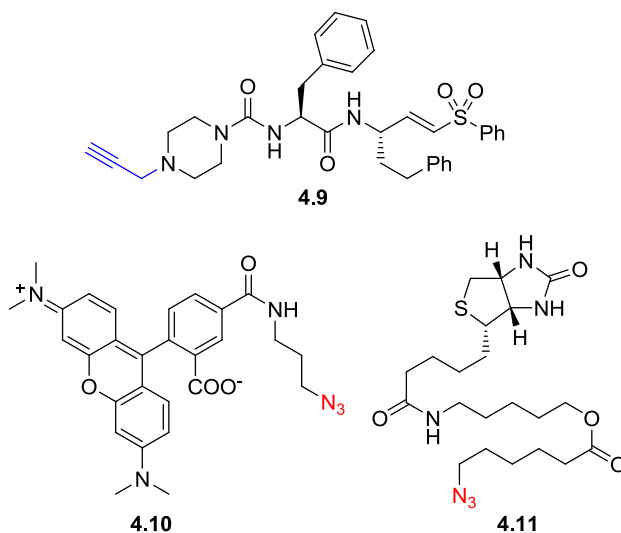


Figure 4.5. Chemical structures of compound **4.9**, a ‘clickable’ activity-based probe based on **4.1** and complementary azide-containing reagents **4.10** and **4.11**.

Independently, another group recently reported the synthesis of **4.9** and its use to identify putative targets of **4.1** in the related parasite *Trypanosoma brucei* [41]. Our efforts to similarly identify targets of **4.1** in *T. cruzi* were complicated by the presence of a host-cell protein that was apparently a major target of **4.9**. In a typical experiment,

intracellular *T. cruzi* amastigotes were treated with **4.9** for 1 hour, followed by cell lysis, 'click' reaction with TAMRA azide **4.10**, and separation/visualization by SDS-PAGE. Regardless of the host cell employed (J774 macrophage or C2C12), only one prominently labeled band at ~35 kDa was observed in these experiments. This band was attributed to a host-cell protein as it appeared also in analogous experiments employing uninfected cells. In fact, we could not conclusively identify any unique bands of parasite origin in our experiments, although such bands might well have escaped detection due to lower abundance and labeling below the limit of fluorescence detection.

The discovery of a potential mammalian off-target of probe **4.9** (and presumably also of **4.1**) was of considerable interest, so we explored this finding further. To determine if this protein was also a target of **4.1** and **4.4**, we conducted competition experiments in C2C12 cells. Hence, pre-incubation of cells with competitor compound at either 1 μ M or 10 μ M for one hour was followed by treatment for one hour with **4.9**, followed by cell lysis, conjugation to **10**, separation (SDS-PAGE), and detection by rhodamine fluorescence as before. In these experiments, pretreatment with 10 μ M of compound **4.1**, **4.3**, or **4.4** successfully blocked labeling of the ~35 kDa band by probe **4.9**, thus indicating that these compounds also react with this target (Figure 4.6). As expected, the non-electrophilic dihydro forms of **4.1** and **4.4** (i.e., compounds **4.6** and **4.8**) did not compete for labeling by **4.9**. Taken together, these results strongly suggest that

compounds **4.1**, **4.3**, and **4.4** react irreversibly with the ~35 kDa protein in a process involving the electrophilic vinylsulfone moiety.

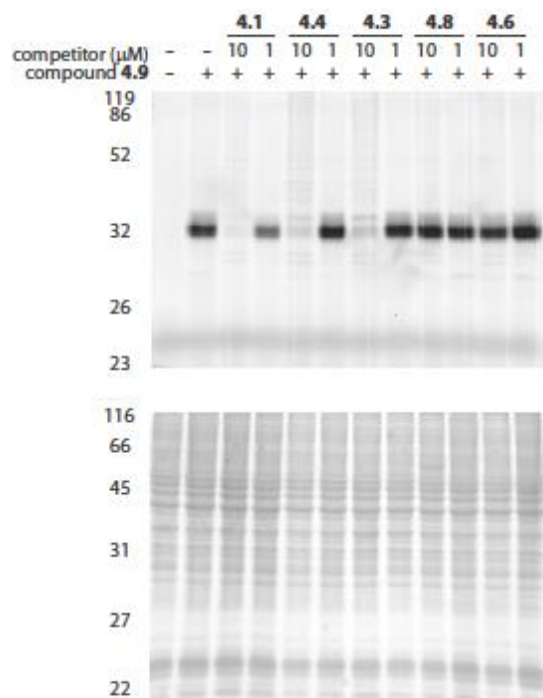


Figure 4.6. Competitive labeling of host cell (C2C12) proteins. Intact cells were labeled with probe **9** following a competitive pre-incubation step with compounds **4.1**, **4.4**, **4.3**, **4.8**, or **4.6**. After cell lysis, protein adducts of **4.9** were conjugated to rhodamide-based dye **4.10**. The gel image at top shows rhodamine fluorescence. The gel image at bottom is of the coomassie stained gel. A successfully competed band is observed at ~35 kDa and this was subsequently identified as cathepsin B.

Chemical Proteomics

We next applied mass spectrometric analysis to identify the ~35 kDa band that was an apparent target of the electrophilic inhibitors described above. To enrich for this protein, C2C12 cells were labeled with **4.9** as before and then reacted with the biotin azide reagent **4.11**, followed by biotin capture onto streptavidin beads. A base-cleavable ester function was introduced in the linker of **4.11**, and this allowed enriched

proteins to be released from beads by treatment with sodium hydroxide. The liberated proteins were separated by SDS-PAGE, and the relevant band at ~35 kDa extracted from the gel. An in-gel trypsin digest [42] was followed by UPLC separation of the tryptic peptides and MS/MS analysis using a hybrid linear ion-trap-Orbitrap mass spectrometer. Tandem mass spectra acquired were searched against the UniProtKb database employing ProteinProspector; four MS/MS spectra corresponding to the same peptide sequence were identified (Figure 4.7). This sequence was found to correspond to the tryptic peptide spanning residues S264-R281 from mouse cathepsin B (uniprot P10605). Significantly, this peptide was not found in analogous experiments where pre-incubation with **4.1** or **4.4** (at 10 μ M) preceded labeling with **4.9**, nor in an experiment in which **4.9** was not added. Thus, cathepsin B is very likely the host-cell protein target of compounds **4.1** and **4.4** identified in the competition experiments with compound **4.9**.

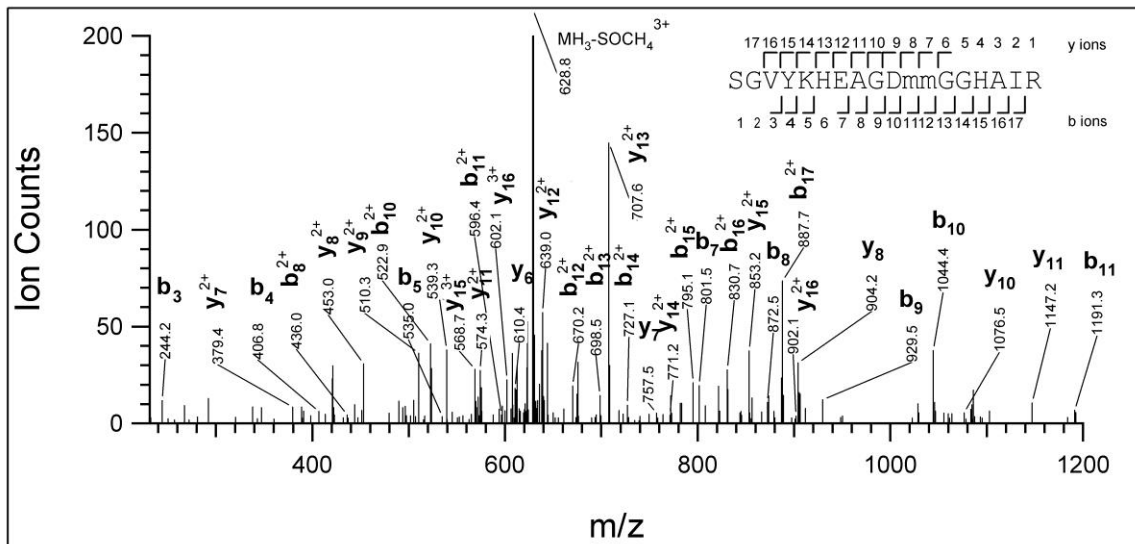


Figure 4.7. MS/MS spectrum of the tryptic peptide S264-R281 from mouse cathepsin B, identified in pull-down experiments employing compound **4.9** in C2C12 cells. Observed sequence ions are labeled (m=oxidized methionine).

The identification of cathepsin B as a relevant cellular off-target of **4.1** and **4.4** is potentially significant. On the one hand, the HCS EC₉₀ values of **4.1** and **4.4** are at least 10-fold lower than the concentrations of these compounds used in the competition experiments. Thus, one might expect to achieve effects on parasite viability before significant inhibition of cathepsin B is conferred. On the other hand, the MTC for **4.1** (8 μM) lies squarely in the range at which the compound effectively competes for cathepsin B labeling by **4.9**. Thus, if micromolar concentrations of **4.1** are indeed required to achieve therapeutic effect in animals, one might well be concerned about effects on host cathepsin B. Thus, the experiments with **4.9** identified a potential off-target while also providing an experimental means for testing the effects of new analogs on this off-target in a relevant, cellular context.

Defining a New Lead Scaffold for *Tc*CYP51 Inhibition

The similar cellular potencies of **4.4** and its reduced form **4.8** suggest that cruzain inhibition plays a relatively minor role in the trypanocidal action of **4.4**. To a first approximation, the cruzain- and *Tc*CYP51-derived effects of **4.4** should be similar to those of its close analogs **4.3** (MTC ~10 μM) and **4.8** (MTC ~0.25 μM), respectively. Unless the effects of inhibiting both targets are synergistic, which is not supported by the data, there would appear to be little benefit gained by combining a relatively weak cruzain-derived effect with a much more potent insult conferred by *Tc*CYP51 inhibition. Moreover, it now seems likely that electrophilic compounds like **4.1** and **4.4** may be

partially consumed in non-productive reactions with host-cell proteases (e.g., cathepsin B) and/or other cytosolic nucleophiles (e.g., glutathione). This possibility is supported by our competitive labeling experiments (Figure 4.6) and by *in vitro* studies employing physiological concentrations of glutathione (Appendix 3, Fig. 2). With regard to the inhibitor chemotypes covered here, there appears to be little rationale for targeting both cruzain and TcCYP51. On the other hand, the surprising potency of analog **4.8** does suggest this as a new lead scaffold for the development of novel TcCYP51 inhibitors.

We next sought to define the minimal pharmacophore within **4.8** required for inhibition of TcCYP51 *in vitro* and anti-trypanosomal effects in whole cells. We therefore synthesized truncated analogs of **4.8**, such as **4.12** and **4.13** (Figure 4.1). These compounds retain the 4-pyridyl ring and neighboring tolyl side chain of **8** while dispensing with those substituents further removed from the putative heme-binding moiety. Interestingly, the truncated analogs **4.12** and **4.13** bound TcCYP51 significantly more weakly than **4.4** or **4.8** (Table 4.1), suggesting that side chains relatively far removed from the 4-pyridyl ring nonetheless play an important role in binding.

Computational docking of **4.12** and **4.13** provided some insight into the observed binding trends. Analog **4.13** adopts a docking pose very similar to **4.4** with respect to the 4-pyridyl and tolyl ring systems. When compared to the poses for **4.4** or **4.13**, the tolyl ring in **4.12** projects much less deeply into the aromatic pocket formed by Phe110

and Tyr103 (Figure 4.2). Neither **4.12** nor **4.13** form interactions with more distal residues (e.g., Leu208, Pro210) that are predicted to form productive contacts with **4.4**. Thus, a larger number of hydrophobic contacts and better orientation of some side chains may explain the binding trends for **4.4**, **4.12**, and **4.13**. Interestingly, the rank-order binding affinities of **4.4**, **4.12**, and **4.13** were correctly predicted by the MM-GBSA method applied to the binding models of these compounds (Appendix 6, Table 1). This suggests that such models could serve to aid in the design of new *TcCYP51* inhibitors derived from this scaffold.

The anti-trypanosomal activities of analogs **4.12** and **4.13** could be correlated with their *in vitro* binding affinities for *TcCYP51* (Table 4.1). Hence, analog **4.13** ($K_D = 75$ nM) shows reduced anti-trypanosomal activity when compared to **4.8** ($K_D \sim 5$ nM). Still weaker-binding analog **4.12** ($K_D = 620$ nM) exhibited no anti-trypanosomal effect at the highest concentration examined (10 μ M). Thus, compound **4.13** can be considered to represent a ‘minimal pharmacophore’ that retains reasonable affinity for *TcCYP51* *in vitro* while also conferring an effect on *T. cruzi* parasites in culture. Future work will focus on further refining the *in vitro* and cellular potency of this scaffold, with compound **4.13** serving as a chemical departure point.

Conclusion

Structure-activity studies are often conducted with the underlying assumption that molecular mechanisms are the same within congeneric analog series. This assumption

is reinforced when activity in biochemical assays can be correlated with cell-based activity. Of course perfect correlation is rarely observed, even when a series is in fact “on-target”. Especially perilous is the construction of mechanistic hypotheses based solely on the correlation of *in vitro* biochemical assay data with gross phenotypic endpoints such as parasite growth inhibition or cell death. As demonstrated here, even seemingly trivial structural changes within a congeneric SAR series can produce analogs with disparate molecular mechanisms of action. Advisable approaches to deal with these uncertainties include the use of cell-based counter assays that can detect action at specific targets or signaling pathways of interest. Activity-based probes can serve as useful tools to verify on-target action during the course of chemical optimization campaigns.

Methods and Supporting Information (Appendix 3)

References

1. de Souza, W. *Microbes. Infect.* **2007**, *9*, 544-545.
2. Urbina, J. A. *Mem. I. Oswaldo Cruz* **2009**, *104*, 311-318.
3. Castro, J. A.; de Mecca, M. M.; Bartel, L. C. *Hum. Exp. Toxicol.* **2006**, *25*, 471-479.
4. Renslo, A. R.; McKerrow, J. H. *Nat. Chem. Biol.* **2006**, *2*, 701-710.
5. McKerrow, J. H.; Doyle, P. S.; Engel, J. C.; Podust, L. M.; Robertson, S. A.; Ferreira, R.; Saxton, T.; Arkin, M.; Kerr, I. D.; Brinen, L. S.; Craik, C. S. *Mem. I. Oswaldo Cruz* **2009**, *104 Suppl 1*, 263-269.
6. Urbina, J. A. *J. Mol. Med.* **1999**, *77*, 332-338.
7. McKerrow, J. H.; Engel, J. C.; Caffrey, C. R. *Bioorg. Med. Chem.* **1999**, *7*, 639-644.
8. Eakin, A. E.; McGrath, M. E.; McKerrow, J. H.; Fletterick, R. J.; Craik, C. S. *J. Biol. Chem.* **1993**, *268*, 6115-6118.
9. McKerrow, J. H.; McGrath, M. E.; Engel, J. C. *Parasit. Today* **1995**, *11*, 279-282.
10. McGrath, M. E.; Eakin, A. E.; Engel, J. C.; McKerrow, J. H.; Craik, C. S.; Fletterick, R. *J. J. Mol. Biol.* **1995**, *247*, 251-259.
11. Buckner, F. S.; Joubert, B. M.; Boyle, S. M.; Eastman, R. T.; Verlinde, C. L.; Matsuda, S. P. *Mol. Biochem. Parasitol.* **2003**, *132*, 75-81.
12. Hankins, E. G.; Gillespie, J. R.; Aikenhead, K.; Buckner, F. S. *Mol. Biochem. Parasitol.* **2005**, *144*, 68-75.

13. Podust, L. M.; von Kries, J. P.; Eddine, A. N.; Kim, Y.; Yermalitskaya, L. V.; Kuehne, R.; Ouellet, H.; Warriar, T.; Altekoster, M.; Lee, J. S.; Rademann, J.; Oschkinat, H.; Kaufmann, S. H.; Waterman, M. R. *Antimicrob. Agents. Chemother.* **2007**, *51*, 3915-3923.
14. Chen, C. K.; Leung, S. S.; Guilbert, C.; Jacobson, M. P.; McKerrow, J. H.; Podust, L. M. *PLoS Negl. Trop. Dis.* **2010**, *4*, e651
15. Urbina, J. A. *Acta. Trop.* **2010**, *115*, 55-68.
16. Doyle, P. S.; Chen, C. K.; Johnston, J. B.; Hopkins, S. D.; Leung, S. S.; Jacobson, M. P.; Engel, J. C.; McKerrow, J. H.; Podust, L. M. *Antimicrob. Agents. Chemother.* **2010**, *54*, 2480-2488
17. Harth, G.; Andrews, N.; Mills, A. A.; Engel, J. C.; Smith, R.; McKerrow, J. H. *Mol. Biochem. Parasit.* **1993**, *58*, 17-24.
18. Roush, W. R.; Cheng, J.; Knapp-Reed, B.; Alvarez-Hernandez, A.; McKerrow, J.H.; Hansell, E.; Engel, J. C. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2759-2762.
19. Huang, L.; Lee, A.; Ellman, J. A. *J. Med. Chem.* **2002**, *45*, 676-684.
20. Greenbaum, D. C.; Mackey, Z.; Hansell, E.; Doyle, P.; Gut, J.; Caffrey, C. R.; Lehrman, J.; Rosenthal, P. J.; McKerrow, J. H.; Chibale, K. *J. Med. Chem.* **2004**, *47*, 3212-3219.
21. Jaishankar, P.; Hansell, E.; Zhao, D. M.; Doyle, P. S.; McKerrow, J. H.; Renslo, A. R. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 624-628.
22. Brak, K.; Doyle, P. S.; McKerrow, J. H.; Ellman, J. A. *J. Am. Chem. Soc.* **2008**, *130*, 6404-6410.

23. Chen, Y. T.; Lira, R.; Hansell, E.; McKerrow, J. H.; Roush, W. R. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5860-5863.
24. Bryant, C.; Kerr, I. D.; Debnath, M.; Ang, K. K.; Ratnam, J.; Ferreira, R. S.; Jaishankar, P.; Zhao, D.; Arkin, M. R.; McKerrow, J. H.; Brinen, L. S.; Renslo, A. R. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6218-6221.
25. Brak, K.; Kerr, I. D.; Barrett, K. T.; Fuchi, N.; Debnath, M.; Ang, K.; Engel, J. C.; McKerrow, J. H.; Doyle, P. S.; Brinen, L. S.; Ellman, J. A. *J. Med. Chem.* **2010**, *53*, 1763-1773.
26. Mott, B. T.; Ferreira, R. S.; Simeonov, A.; Jadhav, A.; Ang, K. K.; Leister, W.; Shen, M.; Silveira, J. T.; Doyle, P. S.; Arkin, M. R.; McKerrow, J. H.; Inglese, J.; Austin, C. P.; Thomas, C. J.; Shoichet, B. K.; Maloney, D. J. *J. Med. Chem.* **2010**, *53*, 52-60.
27. Beaulieu, C.; Isabel, E.; Fortier, A.; Masse, F.; Mellon, C.; Methot, N.; Ndao, M.; Nicoll-Griffith, D.; Lee, D.; Park, H.; Black, W. C. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 7444-7449.
28. Engel, J. C.; Doyle, P. S.; Hsieh, I.; McKerrow, J. H. *J. Exp. Med.* **1998**, *188*, 725-734.
29. Barr, S. C.; Warner, K. L.; Kornreic, B. G.; Piscitelli, J.; Wolfe, A.; Benet, L.; McKerrow, J. H. *Antimicrob. Agents Chemother.* **2005**, *49*, 5160-5161.
30. Jacobsen, W.; Christians, U.; Benet, L. Z. *Drug Metab. Dispos.* **2000**, *28*, 1343-1351.

31. Falgueyret, J. P.; Desmarais, S.; Oballa, R.; Black, W. C.; Cromlish, W.; Khougaz, K.; Lamontagne, S.; Masse, F.; Riendeau, D.; Toulmond, S.; Percival, M. D. *J. Med. Chem.* **2005**, *48*, 7535-7543.
32. Robichaud, J.; Black, W. C.; Therien, M.; Paquet, J.; Oballa, R. M.; Bayly, C. I.; McKay, D. J.; Wang, Q.; Isabel, E.; Leger, S.; Mellon, C.; Kimmel, D. B.; Wesolowski, G.; Percival, M. D.; Masse, F.; Desmarais, S.; Falgueyret, J. P.; Crane, S. N. *J. Med. Chem.* **2008**, *51*, 6410-6420.
33. Gunatilleke, S. S.; Calvet, C. M.; Johnston, J. B.; Chen, C. K.; Erenburg, G.; Gut, J.; Engel, J. C.; Ang, K. K.; Mulvaney, J.; Chen, S.; Arkin, M. R.; McKerrow, J. H.; Podust, L. M. *PLoS Negl. Trop. Dis.* **2012**, *6*, e1736.
34. Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. *J. Med. Chem.* **2006**, *49*, 534-553.
35. Zhu, K.; Shirts, M. R.; Friesner, R. A.; Jacobson, M. P. *J. Chem. Theory. Comput.* **2007**, *3*, 640-648.
36. Hucke, O.; Gelb, M. H.; Verlinde, C. L. M. J.; Buckner, F. S. *J. Med. Chem.* **2005**, *48*, 5415-5418.
37. Kraus, J. M.; Verlinde, C. L.; Karimi, M.; Lepesheva, G. I.; Gelb, M. H.; Buckner, F. S. *J. Med. Chem.* **2009**, *52*, 1639-1647.
38. Chennamaneni, N. K.; Arif, J.; Buckner, F. S.; Gelb, M. H. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6582-6584.
39. Kraus, J. M.; Tatipaka, H. B.; McGuffin, S. A.; Chennamaneni, N. K.; Karimi, M.; Arif, J.; Verlinde, C. L.; Buckner, F. S.; Gelb, M. H. *J. Med. Chem.* **2010**, *53*, 3887-3898.

40. Engel, J. C.; Ang, K. K.; Chen, S.; Arkin, M. R.; McKerrow, J. H.; Doyle, P. S. *Antimicrob. Agents. Chemother.* **2010**, *54*, 3326-3334.
41. Yang, P. Y.; Wang, M.; He, C. Y.; Yao, S. Q. *Chem. Commun.* **2012**, *48*, 835-837.
42. Rosenfeld, J.; Capdevielle, J.; Guillemot, J. C.; Ferrara, P. *Anal. Biochem.* **1992**, *203*, 173-179.

APPENDIX 1: Supporting information Chapter 1

Supplementary Table 1: Mass Spectrometry Results

Protein ID	Acc. #	Protein Name or Family	MW (kDa)	Unique peptides from iTRAQ	Spectra used for Quantitation	Relative Covalent Labeling						Unique peptides from in-gel Digest
						20 nM Hyp	SD (20 nM)	200 nM Hyp	SD (200 nM)	1 μ M Hyp	SD (1 μ M)	
Tb927.10.13780	71749044	<i>Tb</i> GSK3short	40.3	14	18	0.779	0.200	0.333	0.085	0.021	0.033	23
Tb927.11.12410 Tb927.11.12420	74025474 74025476	<i>Tb</i> CLK1 <i>Tb</i> CLK2	53.1 54.9	4	3	0.486	0.079	0.164	0.053	0.115	0.106	
Tb927.10.16030	71749466	<i>Tb</i> MAPK2	41.8	2	2	0.901	0.308	0.152	0.104	0.040	0.065	4
Tb927.10.12040	71748710	MAPK	47.3	10	8	0.853	0.182	0.701	0.143	0.335	0.084	
Tb927.8.3770	72392815	MAPK	46.9	8	7	1.148	0.446	0.854	0.095	0.499	0.082	2
Tb927.10.3900	71747166	CAMKL	83.1	3	1	0.815		1.006		1.062		
Tb927.9.12880	71745786	STE20	75.6	2	1	0.866		1.245		0.534		
Tb927.10.3230	71747040	MAPK	42.2	2	0							1
Tb927.11.16790	74026340	<i>Tb</i> ECK1	72.1	1	0							
Tb927.6.4220	72390491	<i>Tb</i> MAPK5	43.6	-	-							3
Tb927.10.12710	71748842	heat shock protein	90.8	32	23	1.015	0.189	1.177	0.199	0.969	0.263	1
Tb927.1.2390	115504281	beta tubulin	49.7	30	33	1.118	0.380	1.252	0.490	1.018	0.403	9

Protein ID	Acc. #	Protein Name or Family	MW (kDa)	Unique peptides from iTRAQ	Spectra used for Quantitation	Relative Covalent Labeling						Unique peptides from in-gel Digest
						20 nM Hyp	SD (20 nM)	200 nM Hyp	SD (200 nM)	1 μ M Hyp	SD (1 μ M)	
Tb927.3.4720	72387279	dynammin, putative	73.3	21	18	1.100	0.318	1.028	0.305	0.758	0.290	
Tb927.1.2400	115504283	alpha tubulin	49.8	17	18	1.104	0.212	1.207	0.296	1.137	0.305	7
Tb927.10.2110	71746820	elongation factor 1-alpha	49.1	16	27	1.149	0.308	1.117	0.352	1.220	0.326	
Tb927.10.14890	71749256	C-terminal kinesin KIFC1	90.8	14	10	1.062	0.216	1.137	0.274	1.312	0.361	
Tb927.9.12570	71745736	glycerolkinase	56.2	10	5	1.114	0.243	0.948	0.166	1.159	0.297	
Tb927.8.7100	72393479	acetyl-CoA carboxylase	242.9	10	4	1.234	0.314	1.061	0.269	1.273	0.248	
Tb927.6.4280	72390503	glyceraldehyde 3-phosphate dehydrogenase	39.0	9	7	1.000	0.194	0.976	0.215	1.046	0.253	2
Tb927.11.100	71754383	variant surface glycoprotein	51.0	9	3	0.952	0.140	0.842	0.033	1.133	0.183	1
Tb927.11.11330	74025264	heat shock protein 70	75.4	9	7	0.955	0.140	0.994	0.195	1.136	0.205	
Tb927.10.14140	71749110	pyruvate kinase	54.5	8	4	1.055	0.049	1.021	0.048	1.014	0.075	
Tb927.10.4570	71747296	elongation factor 2	94.3	7	4	0.866	0.324	0.940	0.159	0.987	0.200	
Tb927.3.1120	72386563	GTP-binding nuclear protein	24.4	5	4	0.966	0.117	0.935	0.035	0.985	0.114	
Tb927.10.2890	71746974	enolase	46.6	5	3	1.226	0.203	1.161	0.184	1.116	0.209	
Tb927.10.2010	71746802	hexokinase	51.4	5	3	0.802	0.059	0.811	0.017	0.756	0.011	
Tb927.7.4770	72391612	peptidyl-prolyl cis-trans isomerase	18.6	5	2	0.987	0.367	0.575	0.091	0.721	0.114	

Protein ID	Acc. #	Protein Name or Family	MW (kDa)	Unique peptides from iTRAQ	Spectra used for Quantitation	Relative Covalent Labeling						Unique peptides from in-gel Digest
						20 nM Hyp	SD (20 nM)	200 nM Hyp	SD (200 nM)	1 μ M Hyp	SD (1 μ M)	
Tb927.9.6290	71744742	arginine kinase	40.1	5	3	1.033	0.243	0.935	0.099	0.928	0.118	
Tb927.2.450	84043386	retrotransposon hot spot	98.0	5	1	5.098		0.877		1.371		
Tb927.10.5620	71747502	fructose-bisphosphate aldolase	41.1	4	3	0.864	0.187	0.884	0.246	0.903	0.223	5
Tb927.11.9590	71756217	S-adenosylhomocysteine hydrolase	48.4	4	0							
Tb927.11.16280	72388980	60S ribosomal protein L2	28.3	4	3	0.949	0.122	0.957	0.085	0.831	0.280	
Tb927.11.16130	74026208	nucleoside diphosphate kinase	16.8	3	1	0.873		0.750		0.907		
Tb927.11.3600	71755051	40s ribosomal protein S4	30.6	3	2	0.873	0.471	1.359	0.361	0.619	0.538	
Tb927.9.8880	71745150	actin A	41.9	3	3	1.130	0.108	1.242	0.098	1.330	0.137	
Tb927.10.6400	71747654	heat shock protein 60	59.5	3	1	0.675		0.860		0.736		
Tb927.6.3800	72390395	heat shock 70 kDa protein, precursor	71.5	3	3	0.897	0.119	0.831	0.169	1.007	0.277	
Tb927.9.11270	71745542	t-complex protein 1, eta subunit	61.5	3	1	1.015		0.788		1.002		
Tb927.10.10980	71748504	unnamed protein	80.7	3	2	1.486	0.718	1.171	0.197	1.266	0.814	
Tb927.10.14710	71749200	40S ribosomal protein S2	28.6	3	4	1.211	0.371	1.494	0.569	0.756	0.659	
Tb927.11.6300	71755579	40S ribosomal protein S5	21.4	3	2	1.182	0.395	1.015	0.057	1.048	0.005	

Protein ID	Acc. #	Protein Name or Family	MW (kDa)	Unique peptides from iTRAQ	Spectra used for Quantitation	Relative Covalent Labeling						Unique peptides from in-gel Digest
						20 nM Hyp	SD (20 nM)	200 nM Hyp	SD (200 nM)	1 μ M Hyp	SD (1 μ M)	
Tb927.11.13090	74025600	elongation factor 1 gamma	46.3	3	2	1.164	0.295	0.874	0.038	1.093	0.047	
Tb927.4.1790	72387810	ribosomal protein L3	54.4	3	0							
Tb927.10.1090	71746622	40S ribosomal protein S23	15.9	3	1	1.709		1.637		2.083		
Tb927.4.1300	72387712	hypothetical protein	42.0	2	2	1.304	0.059	1.129	0.198	1.241	0.044	1
Tb927.11.11370	74025270	activated protein kinase C receptor	34.7	2	0							
Tb927.11.7510	71755837	glucose-regulated protein 78	71.4	2	1	1.423		1.298		0.786		
Tb927.10.8440	71748046	hexose transporter	56.5	2	1	1.075		0.928		1.281		
Tb927.11.17000	74026380	hypothetical protein	110.1	2	0							
Tb927.10.560	71746518	40S ribosomal proteins S11	20.1	2	0							
Tb927.10.8940	71748126	hypothetical protein	45.4	2	0							
Tb927.11.900	71754537	isocitrate dehydrogenase	46.8	2	0							
Tb927.10.15120	71749292	40S ribosomal protein S13	17.3	2	2	0.922	0.112	1.103	0.419	0.653	0.038	
Tb927.9.11410	71745554	60S ribosomal protein L23	15.0	2	0							
Tb927.9.5840	71744678	tryparedoxin peroxidase	22.4	2	2	0.779	0.144	0.891	0.144	0.951	0.200	
Tb927.11.6140	71755547	40S ribosomal protein S15a	14.6	2	2	1.418	0.325	1.629	0.459	1.154	0.048	

Protein ID	Acc. #	Protein Name or Family	MW (kDa)	Unique peptides from iTRAQ	Spectra used for Quantitation	Relative Covalent Labeling						Unique peptides from in-gel Digest
						20 nM Hyp	SD (20 nM)	200 nM Hyp	SD (200 nM)	1 μ M Hyp	SD (1 μ M)	
Tb927.6.3650	72390377	ADP-ribosylation factor	20.1	2	1	0.921		0.697		0.736		
Tb927.8.8330	72393711	calpain	98.5	2	0							
Tb927.4.1860	72387824	ribosomal protein S19	18.8	2	0							
Tb927.10.14580	71749196	60S ribosomal protein L17	19.1	2	0							
Tb927.8.3150	72392691	t-complex protein 1 gamma subunit	60.8	2	0							
Tb927.6.2300	72390113	adenosine kinase	37.9	1	1	1.106		1.016		0.884		3
Tb927.10.2890	71746974	enolase	40.1									2
Tb927.11.3030	71754941	phosphoribosylpyrophosphate synthetase	49.2									1
Tb11.0390	72390641	40S ribosomal protein S14	15.5	1	1	0.669		0.607		0.743		
Tb927.10.6670	71747708	dynein light chain	14.1	1	0							
Tb927.9.2470	71744100	hypothetical protein	85.9	1	1	1.219		1.234		1.373		
Tb927.11.880	71754533	cyclophilin a	18.7	1	0							
Tb927.11.680	71754491	ribosomal protein L21E (60S)	18.0	1	2	0.785	0.217	1.552	0.310	0.554	0.003	
Tb927.9.10230	71745370	hypothetical protein	18.8	1	0							
Tb927.9.6070	71744718	40S ribosomal protein S3	30.4	1	1	0.706		0.725		0.774		

Protein ID	Acc. #	Protein Name or Family	MW (kDa)	Unique peptides from iTRAQ	Spectra used for Quantitation	Relative Covalent Labeling						Unique peptides from in-gel Digest
						20 nM Hyp	SD (20 nM)	200 nM Hyp	SD (200 nM)	1 μ M Hyp	SD (1 μ M)	
Tb927.7.3440	72391346	I/6 protein	27.8	1	0							
Tb927.10.3940	71747174	40S ribosomal protein S3a	29.4	1	2	0.952	0.184	1.000	0.181	1.013	0.236	
Tb927.7.6470	72391952	O-sialoglycoprotein endopeptidase	40.2	1	1	1.246		1.629		1.226		
Tb927.10.10360	71748404	microtubule associated protein	16.8	1	0							
Tb927.9.15150	71746208	60S ribosomal protein L5	34.6	1	1	0.775		0.782		0.553		
Tb927.8.6160	72393289	40S ribosomal protein S8	24.9	1	1	0.771		1.618		0.852		
Tb927.11.12630	74025518	hypothetical protein	57.1	1	0							
Tb927.10.11390	71748584	60S ribosomal protein L6	21.2	1	1	1.234		1.066		0.980		
Tb927.3.3760	72387087	tryparedoxin	15.9	1	0							
Tb927.3.1370	71746964	40S ribosomal protein S25	12.6	1	1	0.983		1.296		1.466		
Tb927.9.5690	71744668	60S acidic ribosomal protein	11.1	1	0							
Tb927.10.7340	71747826	40S ribosomal protein S24E	15.7	1	0							
Tb927.11.5440	71755409	malic enzyme	61.5	1	0							
Tb927.9.15190	71746212	ribosomal protein L15	26.4	1	0							

Protein ID	Acc. #	Protein Name or Family	MW (kDa)	Unique peptides from iTRAQ	Spectra used for Quantitation	Relative Covalent Labeling						Unique peptides from in-gel Digest
						20 nM Hyp	SD (20 nM)	200 nM Hyp	SD (200 nM)	1 μ M Hyp	SD (1 μ M)	
Tb927.11.3650	71755063	adenylosuccinate synthetase	66.7	1	0							
Tb927.3.3270	72386991	ATP-dependent phosphofructokinase	53.5	1	0							
Tb927.10.5350	71747450	dynein heavy chain	474.7	1	1	1.858		2.816		4.160		
Tb927.5.750	72388908	hypothetical protein	24.8	1	0							
Tb927.10.8230	71748004	protein disulfide isomerase	55.6	1	0							
Tb927.9.8420	71745040	QM-like protein	24.7	1	0							
Tb927.4.5200	72388488	nucleoporin (NUP54/57)	62.5	1	0							
Tb927.6.560	72389769	cathepsin B-like cysteine protease	37.2	1	0							
Tb927.8.2540	72392573	3-ketoacyl-CoA thiolase	46.2	1	0							
Tb927.11.10640	74025124	DNA-(apurinic or apyrimidinic site) lyase	63.8	1	0							
Tb927.11.6870	71755691	14-3-3 protein II	29.2	1	0							
Tb927.8.4580	72392977	hypothetical protein	58.1	1	0							

Supplementary Table 2: Summary of *T. brucei* CDXG kinases and RNAi results

Protein ID	Acc. #	Name	Kinase Group/Family	RNAi lethal	Knockdown (%)	Pubmed ID
Tb927.10.13780	71749044	<i>Tb</i> GSK3short	CMGC/GSK	Y	74	18644955
Tb927.10.16030	71749466	<i>Tb</i> MAPK2	CMGC/MAPK	N	75	12429824
Tb927.11.12410	74025474	<i>Tb</i> CLK1	CMGC/CLK	Y	78	
Tb927.11.12420	74025476	<i>Tb</i> CLK2	CMGC/CLK	N	76	
Tb927.10.12040	71748710	<i>Tb</i> MAPK11	CMGC/MAPK	N	70	
Tb927.8.3770	72392815	<i>Tb</i> MPK10	CMGC/MAPK	N	72	
Tb927.10.3900	71747166		CMGC/CAMKL	N	58	
Tb927.9.12880	71745786		STE/STE20	N	55	
Tb927.7.2420	72391142	<i>Tb</i> GSK3long	CMGC/GSK	N	65	18644955
Tb927.6.4220	72390491	<i>Tb</i> MAPK5	CMGC/MAPK	N	63	16835456
Tb927.6.1780	72390009	<i>Tb</i> MAPK12	CMGC/MAPK	N	68	17368580
Tb927.10.3230	71747040		CMGC/MAPK	N	69	

Protein ID	Acc. #	Name	Kinase Group/Family	RNAi lethal	Knockdown (%)	Pubmed ID
Tb927.2.4510	84043702	<i>TbCRK9</i>	CMGC/CDK	N	60	
Tb927.11.16790	74026340	<i>TbECK1</i>	CMGC/CDKL	N	79	15387824
Tb927.11.8170	71755953	<i>TbMPK8</i>	CMGC/CDKL	N	70	
Tb927.2.4200	84043666		CMGC/CLK	N	80	
Tb927.2.1820	84043450		CMGC/CAMKL	N	71	
Tb927.2.5230	84043810		CAMK	N	70	
Tb927.10.14300	71749142		STE/STE11	N	61	
Tb927.10.15880	71749438		Other	N	62	
Tb927.11.14500	74025880		Other	N	68	

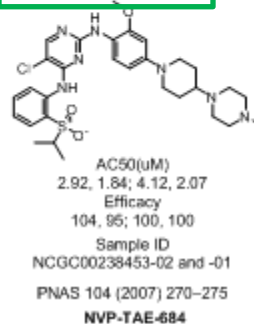
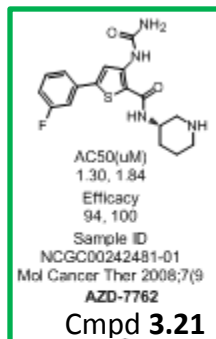
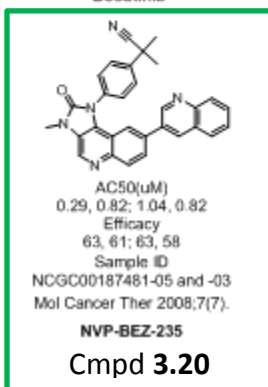
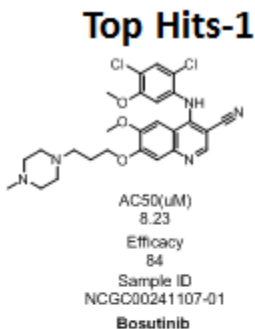
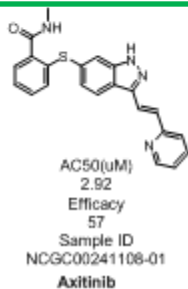
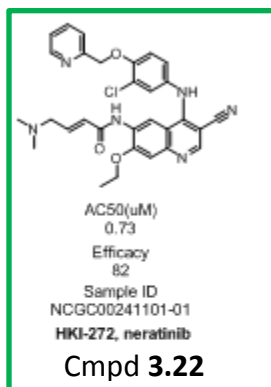
Supplementary Table 3: Primers used in this study (All primers are listed as 5' to 3').

Primer name	Sequence	Purpose
71749142RNAiF 71749142RNAiR	ccccaagcttaaccagctaataatccttggtg aaactcgagcagctcagtttgctgagaaac	RNAi for Tb927.10.14300
71747166RNAiF 71747166RNAiR	aaccaagctttgacacctcattcagatggg aaaaaactcgagttcaccacggttccac	RNAi for Tb927.10.3900
71749044RNAiF 71749044RNAiR	aaaagcttacttatacacagctcggggg caactcgaggcagctcattgaaaaggat	RNAi for Tb927.10.13780
72390009RNAiF 72390009RNAiR	ccccaagatttgagattcctgaacatcc aactcgagcagttaagtagcgcctgttcag	RNAi for Tb927.6.1780
71748710RNAiF 71748710RNAiR	aaaaagcttctgaagcaagtgctccgaac aacctcgagctgtacatcggattcctcgta	RNAi for Tb927.10.12040
72391142RNAiF 72391142RNAiR	aagcttcatctgagcgggaactctc ctcgagcagccacgggtgtaaaatct	RNAi for Tb927.7.2420
84043450RNAiF 84043450RNAiR	aagcttcggctcggctatacttctcag ctcgagaagatccgttgaaaccacg	RNAi for Tb927.2.1820
74025474RNAiF 74025474RNAiR	aagcttgctccagcaatacccctca ctcgagatactcttccgctccggt	RNAi for Tb927.11.12410
74025476RNAiF 74025476RNAiR	aagctttaactcatccaccattgcca ctcgagcttcccgaatgtaccctca	RNAi for Tb927.11.12420
71749438RNAiF 71749438RNAiR	aagcttgctcgtgtgtacattccct ctcgagcgcggtacggtaaagtatct	RNAi for Tb927.10.15880
74025880RNAiF 74025880RNAiR	aagcttcggcatttctacagtggtt ctcgagcacgaggcatttagagcaca	RNAi for Tb927.11.14500
72392815RNAiF 72392815RNAiR	aagctgtttctcagcagtgaggagc ctcgagttggatcctgcgtatcctc	RNAi for Tb927.8.3770

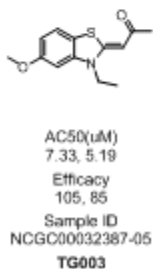
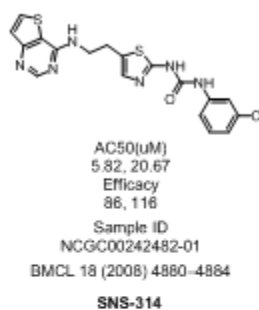
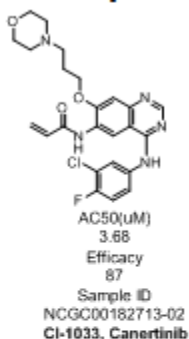
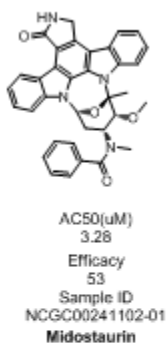
Primer name	Sequence	Purpose
84043810RNAiF 84043810RNAiR	aagcttcacgacgacattaaagcaga ctcgaggcagaaggcctgttactg	RNAi for Tb927.2.5230
71745786RNAiF 71745786RNAiR	aagctcaagatggcggttactgt ctcgagttgtcagcgttgaatctgc	RNAi for Tb927.9.12880
84043702RNAiF 84043702RNAiR	aagctttattaccatcgcaagggg ctcgagatcgaaaatgcggtcaataggt	RNAi for Tb927.2.4510
74026340RNAiF 74026340RNAiR	aaacaagctgttgcaactcgctggtatcg aaactcgagacgcccacgtggtattctta	RNAi for Tb927.11.16790
71755953RNAiF 71755953RNAiR	aggaagctttgtgtaccaaggaaggga tttctcgagtggcaacgatacaccgggta	RNAi for Tb927.11.8170
84043666RNAiF 84043666RNAiR	aaaaataaaagctgccacgcccgtgc aaactcgagacgactcctttccttctg	RNAi for Tb927.2.4200
71749466RNAiF 71749466RNAiR	accaagcttctacatcgctacacacc ccactcgagcttagcaaaatcgcatgttc	RNAi for Tb927.10.16030
71747040RNAiF 71747040RNAiR	ctttaagcttatccgtcccgtcaatgag aaactcgaggcgaaactgatcaaacatgg	RNAi for Tb927.10.3230
72390491RNAiF 72390491RNAiR	cccaagcttattggcgcaaatcgttat aacctcgagtgtgactaaatcaatcccct	RNAi for Tb927.6.4220
71749044pETf2 71749044pETr2	caccatgctgctcaacctaccgatgctgc ttacttctcagcagatactcccgtgctgc	Recombinant expression of <i>TbGSK3short</i>
74025474pETf 74025474pETr	caccatggatagcataagctctggtctaac ttacatgacaggtgtggggggagttaga	Recombinant expression of <i>TbCLK1</i>
pC9044f pC9044r	ggtacctgctcaacctaccgatgct ctcgagcttctcagcagatactccc	Endogenous tagging expression of <i>TbGSK3short</i>
pC5474f pC5474r	ctggtaccgatagcataagctctggtctaac gactcgagcatgacaggtgtggggggag	Endogenous tagging expression of <i>TbCLK1</i>

APPENDIX 2: Supporting information for Chapter 2

Supplemental Figure 1: NCGC HTS results 4-2011 (NCGC: Mindy Davis, Min Shen, Craig Thomas)

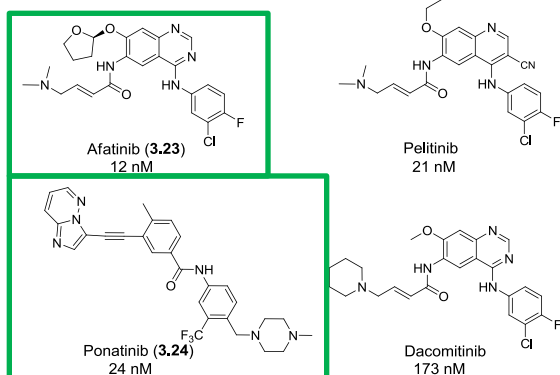


Top Hits-2



Supplemental Figure 2: NCGC HTS results 4-2012 (NCGC: Mindy Davis, Min Shen, Craig Thomas)

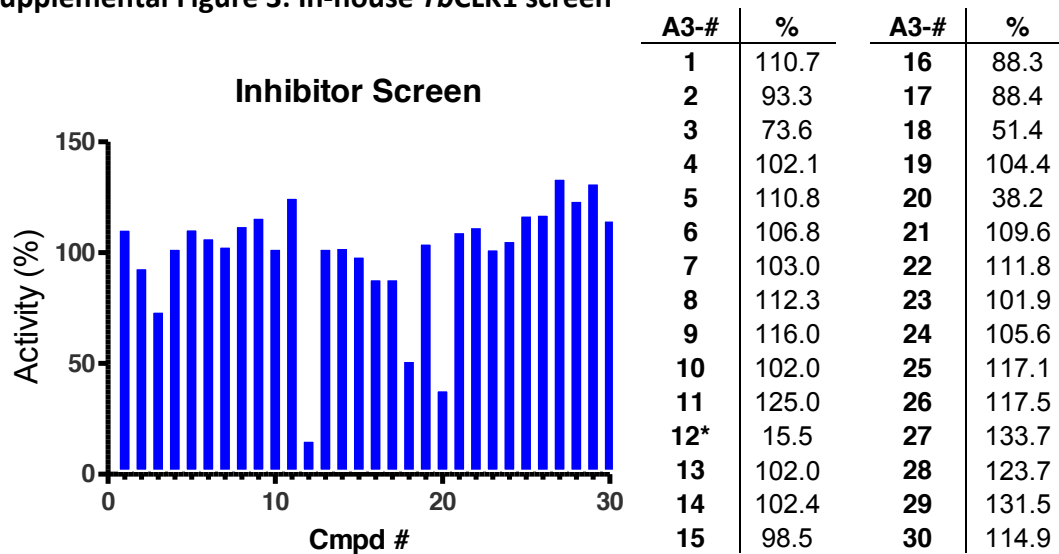
Sample ID	Sample Name	Curve Class2	AC50(uM)	Efficacy	
NCGC00185000-02	Afatinib	-1.1	0.0122	-89.0037	
NCGC00263152-01	Ponatinib	-1.1	0.0244	-78.0251	
NCGC00263185-01	Dacomitinib	-1.1	0.1729	-79.205	
NCGC00253463-01	null	-1.1	0.5468	-86.3665	
NCGC00263195-01	AV-412	-1.1	0.6135	-76.6971	
NCGC00263088-01	BAG-956	-1.1	0.6884	-97.285	
NCGC00263200-01	RO495	-1.1	0.8666	-82.9407	
NCGC00263176-01	BMS-794833	-1.1	1.7291	-92.5655	
NCGC00092289-01	null	-1.1	1.94	-76.849	
NCGC00263172-01	DCC-2036	-1.1	2.1768	-88.8924	
NCGC00263103-01	Pelitinib	-1.2	0.0218	-65.4426	
NCGC00263164-01	Cabozantinib	-1.2	0.5468	-70.2973	
NCGC00263144-01	WZ-4002	-1.2	0.6884	-70.2309	Cmpd 3.6
NCGC00263206-01	AMG-47a	-1.2	3.0748	-69.2165	
NCGC00167767-01	null	-2.1	2.7404	-104.404	
NCGC00263132-01	PF-477736	-2.2	4.3432	-81.7436	



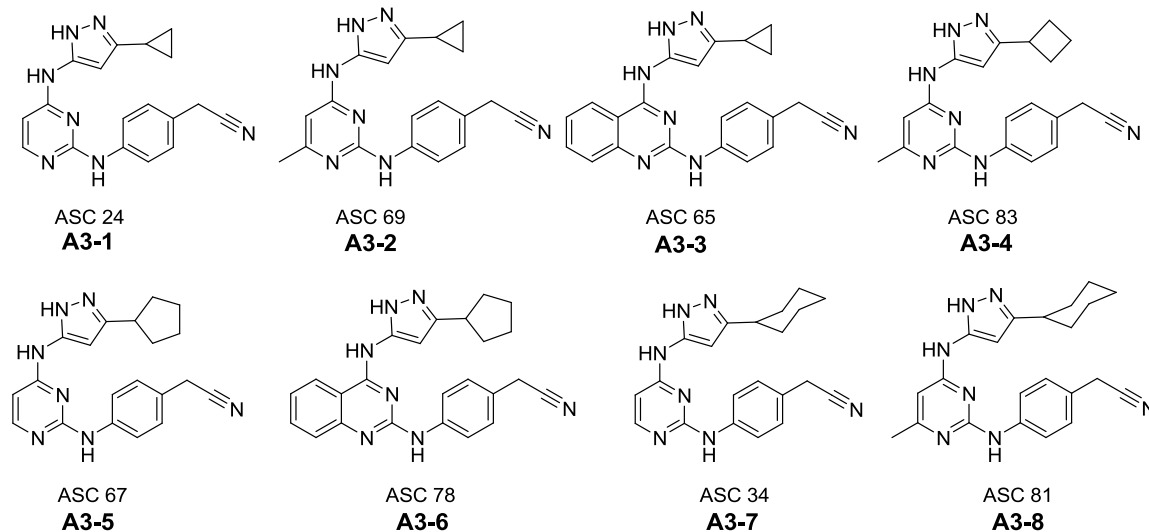
NCGC *Tb*CLK1 HTS screening:

Compound (23 nL; 7 or 24 pt dose-response + control) was transferred by pin tool to buffer (2 μ L; 0.01 M Tris pH 7.5, 10 mM MgCl₂, 0.02 M EGTA, 0.01% Tween 20, 1 mM DTT, 0.05% BSA) containing *Tb*CLK1 (1.35 μ M). After 15 minutes, ATP (15 μ M) and MBP (0.24 mg/ml) in buffer (1 μ L; 0.01 M Tris pH 7.5, 10 mM MgCl₂, 0.02 M EGTA, 0.01% Tween 20, 1 mM DTT, 0.05% BSA) was added to the solution. After 60 minutes, ADP-Glo Reagent (2.5 μ L; Promega) was added. After 40 minutes, Kinase Detection Reagent (5 μ L; Promega) was added. The reaction was allowed to incubate for 30 minutes. Luminescence was quantified using a ViewLux microplate uHTS reader (PerkinElmer).

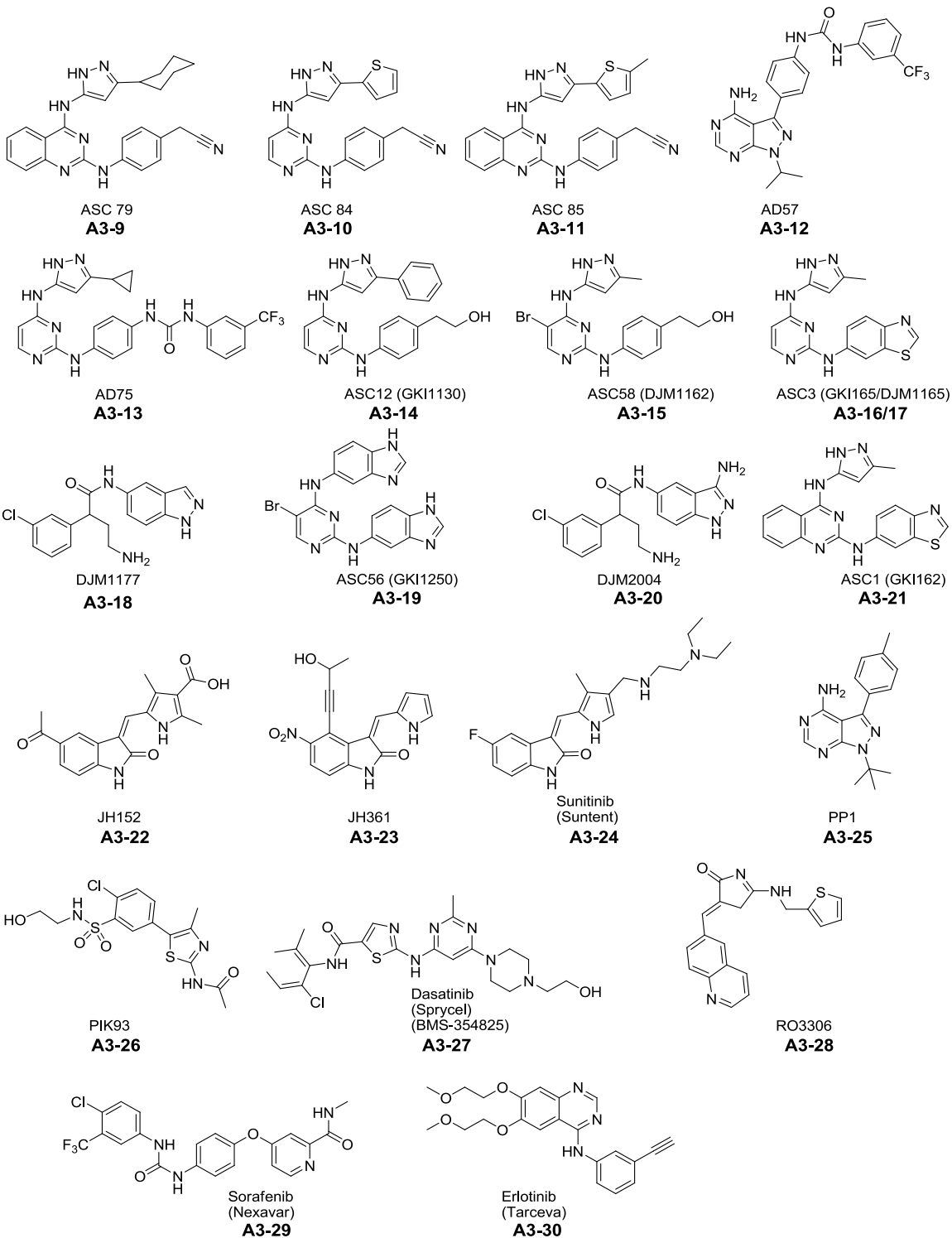
Supplemental Figure 3: In-house *Tb*CLK1 screen



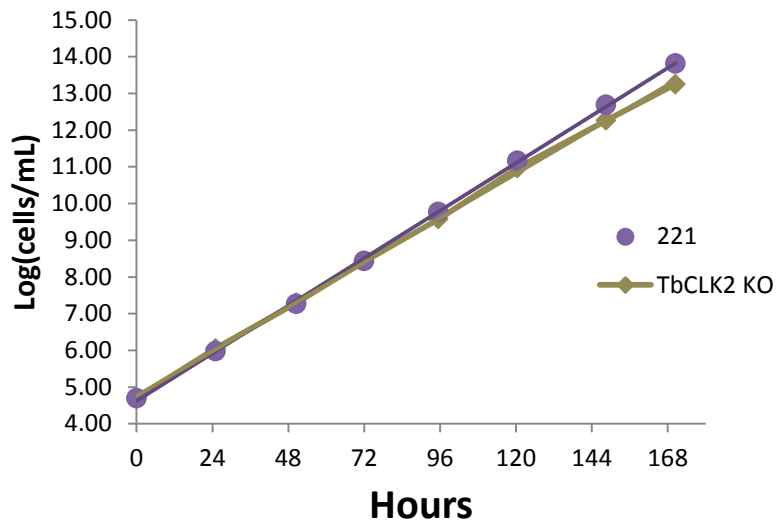
Kinases were incubated with compound at 1 μ M (n=1) and 100 μ M ATP for 30 min before initiating reactions with substrate and γ ³²P-ATP. Phosphotransfer was quantified and normalized to DMSO control. * Compound 12 (AD57 **3.25**) was run at 10 μ M.



Supplemental Figure 3 (cont.): In-house *Tb*CLK1 screen

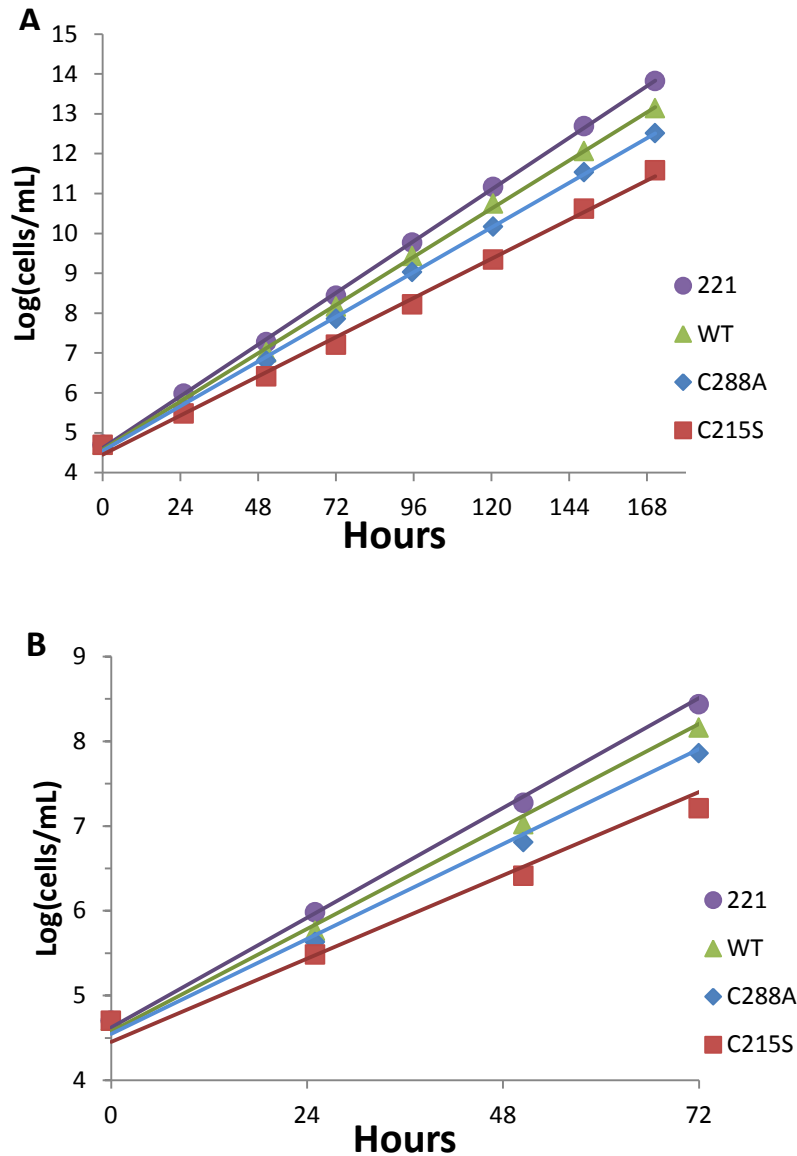


Supplemental Figure 4: Growth curves comparing unmodified cells and *TbCLK2* KO cells. (Nathan Gushwa)



Cells were seeded at the same density at day 0 in their respective media. Cell density was measured about every 24 hours by cell counting. Populations were diluted as needed to maintain linear growth ($< 5 \times 10^6$ cells/mL; doubling time: 8-12 hours). Dilutions were recorded and used to adjust the daily cell density numbers. (mean, $n=3$)

Supplemental Figure 5: Growth curves comparing *TbCLK1* modified cell lines. (Nathan Gushwa)



Cells were seeded at the same density at day 0 in their respective media. Cell density was measured about every 24 hours by cell counting. Populations were diluted as needed to maintain linear growth ($< 5 \times 10^6$ cells/mL; doubling time: 8-12 hours). Dilutions were recorded and used to adjust the daily cell density numbers. (mean, $n=3$) **(A)** Cell density was measured through 171 hours (>7 days). **(B)** Same data in (A), but focusing on the time between 0 and 72 hours. This time period is most relevant to our proliferation assays (24 and 48 hours). There is a difference in growth rates amongst the cell lines, but growth is linear through the study.

Supplemental Table 1: Invitrogen SelectScreen Kinase Profiling Results at 1 μ M AD57 (3.25) (Arvin Dar)

	AD57				
ABL1	89	FLT4 (VEGFR3)	98	PDGFRA D842V	94
ABL1 E255K	82	FRAP1 (mTOR)	33	PDGFRA T674I	91
ABL1 G250E	56	FRK (PTK5)	89	PDGFRA V561D	100
ABL1 T315I	88	FYN	72	PDGFRB (PDGFR beta)	91
ABL1 Y253F	72	GRK4	2	PDK1	12
ABL2 (Arg)	66	GRK5	-3	PDK1 Direct	27
ACVR1B (ALK4)	2	GRK6	9	PHKG1	9
ADRBK1 (GRK2)	10	GRK7	17	PHKG2	3
ADRBK2 (GRK3)	-1	GSG2 (Haspin)	21	PI4KA (PI4K alpha)	0
AKT1 (PKB alpha)	-10	GSK3A (GSK3 alpha)	20	PI4KB (PI4K beta)	6
AKT2 (PKB beta)	2	GSK3B (GSK3 beta)	51	PIK3C2A (PI3K-C2 alpha)	3
AKT3 (PKB gamma)	-4	HCK	90	PIK3C2B (PI3K-C2 beta)	1
ALK	35	HIPK1 (Myak)	76	PIK3C3 (hVPS34)	-2
AMPK A1/B1/G1	9	HIPK2	69	PIK3R1 (p110 alpha/p85)	3
AMPK A2/B1/G1	25	HIPK3 (YAK1)	63	PIK3R1 (p110 delta/p85)	16
AURKA (Aurora A)	87	HIPK4	100	PIK3CG (p110 gamma)	2
AURKB (Aurora B)	96	IGF1R	17	PIM1	6
AURKC (Aurora C)	77	IKBKB (IKK beta)	17	PIM2	-8
AXL	82	IKBKE (IKK epsilon)	4	PKN1 (PRK1)	-5
BLK	88	INSR	52	PLK1	3
BMX	92	INSRR (IRR)	84	PLK2	74
BRAF	67	IRAK1	87	PLK3	58
BRAF V599E	85	IRAK4	24	PRKACA (PKA)	-3
BRSK1 (SAD1)	6	ITK	7	PRKCA (PKC alpha)	7
BTK	76	JAK1	76	PRKCB1 (PKC beta)	9
CAMK1 (CaMK1)	85	JAK2	62	PRKCB2 (PKC beta II)	-3
CAMK1D (CaMK1 delta)	42	JAK2 JH1 JH2	43	PRKCD (PKC delta)	24
CAMK2A (CaMKII alpha)	-7	JAK2 JH1 JH2 V617F	52	PRKCE (PKC epsilon)	5
CAMK2B (CaMKII beta)	5	JAK3	18	PRKCG (PKC gamma)	4
CAMK2D (CaMKII delta)	8	KDR (VEGFR2)	98	PRKCH (PKC eta)	8
CAMK4 (CaMKIV)	72	KIT	50	PRKCI (PKC iota)	0
CDC42 BPA (MRCKA)	11	KIT T670I	44	PRKCN (PKD3)	8
CDC42 BPB (MRCKB)	8	LCK	84	PRKCQ (PKC theta)	8
CDK1/cyclin B	2	LRRK2	72	PRKCZ (PKC zeta)	7
CDK2/cyclin A	-4	LRRK2 G2019S	56	PRKD1 (PKC mu)	6
CDK5/p25	6	LTK (TYK1)	65	PRKD2 (PKD2)	23
CDK5/p35	-14	LYN A	89	PRKG1	-1
CDK7/cyclin H/MNAT1	-14	LYN B	84	PRKG2 (PKG2)	4
CDK9/cyclin T1	47	MAP2K1 (MEK1)	5	PRKX	1
CHEK1 (CHK1)	-12	MAP2K2 (MEK2)	8	PTK2 (FAK)	92
CHEK2 (CHK2)	-10	MAP2K6 (MKK6)	96	PTK2B (FAK2)	69
CHUK (IKK alpha)	75	MAP3K8 (COT)	7	PTK6 (Brk)	65
CLK1	46	MAP3K9 (MLK1)	88	(cRAF) Y340D Y341D	87
CLK2	73	MAP4K2 (GCK)	69	RET	101
CLK3	5	MAP4K4 (HGK)	15	RET V804L	100
CSF1R (FMS)	97	MAP4K5 (KHS1)	103	RET Y791F	101
CSK	86	MAPK1 (ERK2)	-4	ROCK1	81
CSNK1A1 (CK1 alpha 1)	-16	MAPK10 (JNK3)	0	ROCK2	92
CSNK1D (CK1 delta)	4	MAPK11 (p38 beta)	19	ROS1	26
CSNK1E (CK1 epsilon)	10	MAPK12 (p38 gamma)	77	RPS6KA1 (RSK1)	90
CSNK1G1 (CK1 gamma 1)	2	MAPK13 (p38 delta)	28	RPS6KA2 (RSK2)	89
CSNK1G2 (CK1 gamma 2)	3	MAPK14 (p38 alpha)	-8	RPS6KA3 (RSK2)	74
CSNK1G3 (CK1 gamma 3)	2	MAPK14 (p38 alpha) Direct	-1	RPS6KA4 (MSK2)	86
CSNK2A1 (CK2 alpha 1)	1	MAPK3 (ERK1)	6	RPS6KA5 (MSK1)	62
CSNK2A2 (CK2 alpha 2)	1	MAPK8 (JNK1)	15	RPS6KA6 (RSK4)	60
DAPK1	17	MAPK9 (JNK2)	45	RPS6KB1 (p70S6K)	93
DAPK3 (ZIPK)	40	MAPKAPK2	-2	SGK (SGK1)	5
DCAMKL2 (DCK2)	5	MAPKAPK3	1	SGK2	18
DNA-PK	22	MAPKAPK5 (PRAK)	9	SGKL (SGK3)	4
DYRK1A	-4	MARK1 (MARK)	-3	SNF1LK2	0
DYRK1B	-5	MARK2	-1	SPHK1	-4
DYRK3	81	MARK3	2	SPHK2	6
DYRK4	1	MARK4	-3	SRC	88
EEF2K	5	MATK (HYL)	22	SRC N1	89
EGFR (ErbB1)	11	MELK	65	SRMS (Srm)	35
EGFR (ErbB1) L858R	17	MERTK (cMER)	85	SRPK1	59
EGFR (ErbB1) L861Q	4	MET (cMet)	47	SRPK2	33
EGFR (ErbB1) T790M	62	MET M1250T	24	STK22B (TSSK2)	2
(ErbB1) T790M L858R	79	MINK1	-5	STK22D (TSSK1)	8
EPHA1	56	MKNK1 (MKN1)	20	STK23 (MSSK1)	66
EPHA2	97	MST1R (RON)	52	STK24 (MST3)	23
EPHA4	95	MST4	18	STK25 (YSK1)	13
EPHA5	94	MUSK	98	STK3 (MST2)	-5
EPHA8	99	MYLK2 (skMLCK)	4	STK4 (MST1)	24
EPHB1	89	NEK1	19	SYK	40
EPHB2	100	NEK2	-8	TAOK2 (TAO1)	45
EPHB3	85	NEK4	16	TBK1	3
EPHB4	88	NEK5	-1	TEK (Tie2)	98
ERBB2 (HER2)	24	NEK6	-1	TXK	15
ERBB4 (HER4)	14	NEK7	-3	TYK2	-1
FER	72	NEK9	42	TYRO3 (RSE)	59
FES (FPS)	42	NTRK1 (TRKA)	103	YES1	86
FGFR1	81	NTRK2 (TRKB)	101	ZAP70	6
FGFR2	94	NTRK3 (TRKC)	103		
FGFR3	64	NUAK1 (ARK5)	19		
FGFR3 K650E	82	PAK1	6		
FGFR4	59	PAK2 (PAK65)	9		
FGR	93	PAK3	3		
FLT1 (VEGFR1)	88	PAK4	-2		
FLT3	100	PAK6	0		
FLT3 D835Y	99	PAK7 (KIAA1264)	1		
		PASK	5		
		PDGFRA (PDGFR alpha)	99		

Supplemental Table 2: *Tb*CLK modified cell lines.

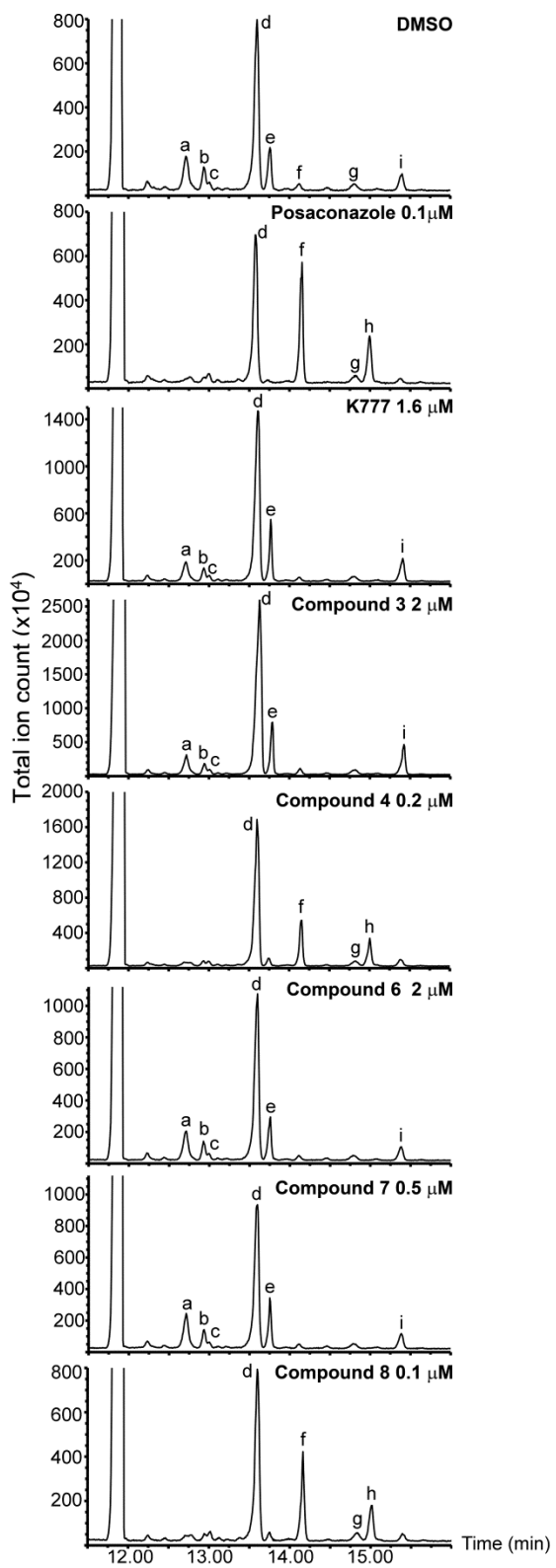
	<u>Genotype</u>	<u>Antibiotics</u>	<u>Notes</u>
221	TbCLK1 ^{WT} /TbCLK1 ^{WT}	none	parental strain, untransfected
WT	HA-TbCLK1 ^{WT} /deletion	hygromycin/G418	haploid, expressing HA-tagged <i>Tb</i> CLK1
C288A	HA-TbCLK1 ^{C288A} /deletion	hygromycin/G418	haploid, expressing HA-tagged <i>Tb</i> CLK1 C288A
C215A	HA-TbCLK1 ^{C215A} /deletion	hygromycin/G418	haploid, expressing HA-tagged <i>Tb</i> CLK1 C215S
C215S	HA-TbCLK1 ^{C215S} /deletion	hygromycin/G418	haploid, expressing HA-tagged <i>Tb</i> CLK1 C215S
<i>Tb</i>CLK2 KO		puromycin/blastidicin	both alleles of <i>Tb</i> CLK2 have been deleted

*Intermediate cell lines of **WT**, **C288A**, and **C215A/S** that lack the deletion have been used, but would be noted as such.

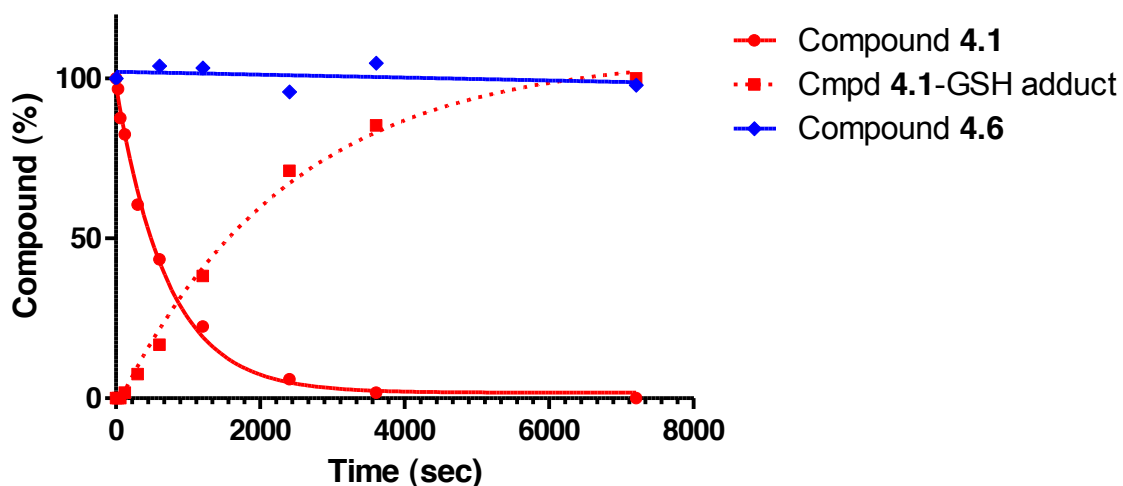
APPENDIX 3: Supporting information for Chapter 3

Supplemental Table 1. Rank ordering of experimentally and computationally determined binding affinities for compounds **4.5**, **4.4**, **4.13**, and **4.12**.

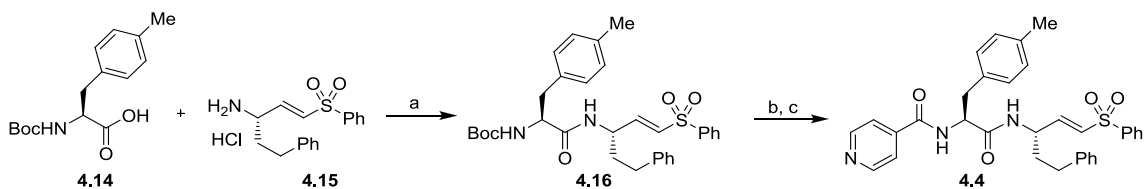
Compound	K _D (nM)	Docking Score (Glide XP)	MM/GBSA score
4.5	≤ 5	-10.8	-84.2
4.4	≤ 5	-10.6	-91.9
4.13	75 ± 26	-11.5	-77.9
4.12	615 ± 264	-9.71	-63.9



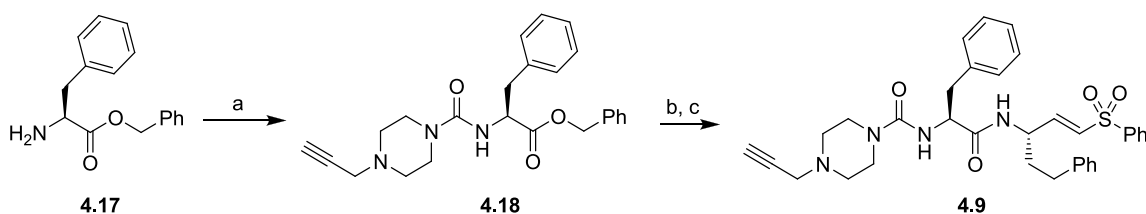
Supplemental Figure 1. GC/MS spectra showing sterol composition of intra-cellular *T. cruzi* parasites treated with indicated test compounds. Peaks are assigned as **a** - cholesta-7,24-dien-3 β -ol, [M]⁺⁺ = m/z 454; **b** - cholesta-8,24-dien-3 β -ol (zymosterol), [M]⁺⁺ = m/z 470; **c** - 24-methyl-7-en-cholesta-en-3 β -ol, [M]⁺⁺ = m/z 472; **d** - ergosta-7,24-diene-3 β -ol (episterol), [M]⁺⁺ = m/z 470; **e** - ergosta-8,24-diene-3 β -ol (fecosterol), [M]⁺⁺ = m/z 470; **f** - lanosterol, [M]⁺⁺ = m/z 498; **g** - 4-methylepisterol, [M]⁺⁺ = m/z 484; **h** - eburicol, [M]⁺⁺ = m/z 512; **i** - 24-ethyl-7,24(24')-en-cholesta-dien-3 β -ol, [M]⁺⁺ = m/z 484.



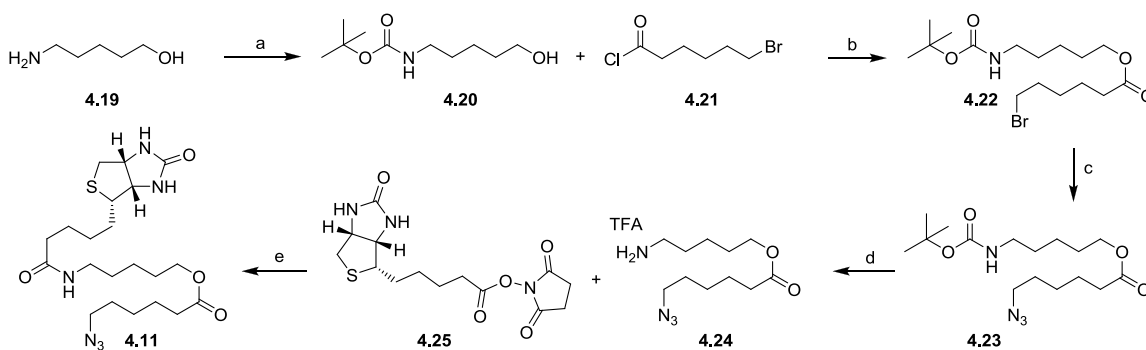
Supplemental Figure 2. Time-course of reaction between compounds **4.1** or **4.6** with 10 mM glutathione *in vitro*.



Supplemental Scheme A3.1. Synthesis of compound **4.4**. Conditions: a) HATU, DIEA, DMF; b) HCl, dioxane; c) Pyridyl acid, HATU, DIEA, DMF.

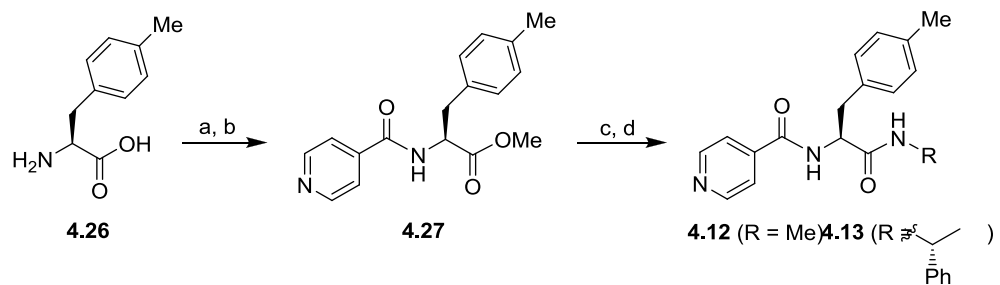


Supplemental Scheme A3.2. Synthesis of activity-based probe **4.9**. Conditions: a) triphosgene, CH_2Cl_2 , aq. NaHCO_3 , then *N*-propargylpiperazine; b) NaOH, MeOH; c) **4.15**, HATU, DMF, 11% overall.



Supplemental Scheme A3.3. Synthesis of cleavable biotin azide reagent **4.11**.

Conditions: a) di-tert-butyl dicarbonate, H_2O , NaOH, 88%; b) CH_2Cl_2 , Et_3N ; c) NaN_3 , DMF, 49% over two steps; d) TFA, CH_2Cl_2 ; e) DMF, Et_3N , 59% over two steps.



Supplemental Scheme A3.4. Synthesis of analogs **4.12** and **4.13**. Conditions: a) SOCl_2 , MeOH; b) isonicotinic acid, HATU, DIEA, DMF; c) LiOH, THF, water; d) R-NH₂, HATU, DIEA, DMF.

Experimental Procedures

Computational modeling

The binding of compounds **4.4**, **4.5**, **4.12**, and **4.13** to CYP51 were modeled using the induced fit docking software (Schrödinger Inc.). As previously described [1], these compounds were docked to a structural model that was developed based on a crystal structure of *Tc*CYP51 (PDB code 2WUZ) [2]. In brief, the protein and the compounds were first prepared for docking using the Protein Preparation Wizard and the Ligprep modules, respectively. The Fe³⁺ ion of the heme group was defined as a metal site during the docking calculations. The induced fit docking protocol was employed to address protein flexibility via the following procedures [3]. Binding models were first generated by rigid docking with a softened potential using the GLIDE standard precision (SP) scoring function [4]. Models with top docking scores were then subjected to binding site optimization in which amino acid residues within 5 Å from the docked compound were energy minimized in the presence of the ligand using PRIME [5]. The resulting low-energy complex structures that were within 30 kcal/mol from the lowest-energy complex were used in the subsequent re-docking calculations. Each compound was redocked to these “induced fit” protein models and was evaluated using the GLIDE extra precision (XP) scoring function. Finally, the predicted binding mode with the most favorable docking score was further refined using the PRIME MM-GBSA protocol by energy minimizing the ligand and the surrounding protein residues within 5 Å [6].

TcCYP51 Binding Assay

All K_D values were determined using the UV-vis binding assay described previously [7]. Although CYP51 concentrations of 0.5 μM are required in this assay, application of the Morrison equation [8] allows us to estimate ligand binding with a lower limit of $K_D \sim 5$ nM.

CYP Inhibition and Reversibility Studies

Mammalian CYP inhibition values and reversibility studies were conducted by WuXi AppTech (Tianjin, China 300457) according to standard protocols.

Cruzain Kinetics

Cruzain (0.1nM) in assay buffer (100mM sodium acetate pH 5.5, 5 mM DTT, 10 mM EDTA, and 0.01% Triton-X 100) was added to inhibitor dilutions in equal volume of 1 μM Z-Phe-Arg-AMC (Bachem) in the same buffer. Total assay volume is 20 μL in 384-well plate format or 200 μL in 96-well plate format. Progress curves were determined using a Flexstation 3 fluorescent plate reader (Molecular Devices) with excitation wavelength of 355 nm and the emission wavelength of 460 nm for 360 seconds at room temperature; for inhibitor concentrations ranging from 25 - 0.001 μM . Inhibitor dilutions that

produced exponential progress curves over a wide range of k_{obs} were used to determine kinetic parameters. Using Prism5 software (GraphPad), the value of k_{obs} , the rate constant for loss of enzymatic activity, was determined from an equation for pseudo first order dynamics ($P = ([P]_i/k_{obs})[1 - \exp(-k_{obs} * t)]$); where product formation = P, initial rate = $[P]_i$, time = t, and the first-order rate constant = k_{obs}). For inhibitors where k_{obs} varied hyperbolically with [I], nonlinear regression analysis was performed with Prism5 to determine k_{inact}/K_i using $k_{obs} = k_{inact}[I]/([I] + K_i*(1 + [S]/K_m))$. IC₅₀ values were determined using sigmoidal dose-response variable slope model.

Minimum Trypanocidal Concentration Assay

The minimal trypanocidal concentration assay is derived from the established protocol [9] for evaluating 'cidal' activity, except that each compound is evaluated at multiple concentrations (typically 7 or 10) ranging from 0 to 50 μ M. As in the original protocol, *T. cruzi* infected J744 macrophages are treated with test compound for 27 days (or until *T. cruzi* parasites appear). During the course of treatment, media containing test compound is replaced every 72 hours. After 27 days of treatment, the cultures are maintained under drug-free conditions until day 40. The MTC is taken as the lowest concentration of test compound that prevents reappearance of *T. cruzi* parasites as read at day 40.

Imaging-Based Screening Assay

Antiparasitic EC₉₀ values were determined using an adaption [7] of the high-content screening approach described previously [10]. The assay employed mouse C2C12 myoblasts (ATCC#CRL-1772) infected with *T. cruzi* CAI-72 [11] trypomastigotes.

Ergosterol Biosynthesis - GC/MS Methods

To grow intracellular amastigotes for lipid analysis, C2C12 mouse skeletal myoblasts (4 x 10⁶ cells) cultured in 150 cm² flasks as described elsewhere [7] were infected with 80 x 10⁶ trypomastigotes for 24 hours. The cultures were treated with test compounds 72 hours post-infection at concentrations ~2-4-fold below their MTC, so as to retain a viable parasite population for analysis. Posaconazole at 100 nM was used as a positive control. The cysteine protease cruzain inhibitor K777 [12] at 1.6 μM was used as a negative control to ensure that *T. cruzi* inhibition via other pathways did not affect composition of the membrane sterols. Cells were harvested 96 hours post-infection and after 24 hours of drug treatment.

Gas chromatography/mass spectrometry (GC-MS) was used for lipid analysis as described previously [7]. Our GC/MS-based assays measure total sterol abundances, and allow us to delineate the effects of individual compounds on the ergosterol biosynthesis pathway of *T. cruzi*. Total lipids were extracted by treating the dry cell

pellets with 2 mL chloroform for 24 hours. Polar lipids were removed by washing the organic phase thrice with 3 mL water. The organic phase was retained and subsequently dried on an evaporator using a stream of nitrogen gas. The sterol-containing residue was re-suspended in 2 mL of chloroform-methanol (9:1 ratio), washed thrice with 3 mL water, dried under N₂ gas and re-suspended in acetonitrile. The acetonitrile phase was dried under a stream of N₂ and then washed with three portions of water (3 X 3 mL). The extracted sterols were dissolved in hexane (50 µL) and derivatized with 75 µL *N,N*-bis(trimethylsilyl)-2,2,2-trifluoroacetamide (BSTFA) (Pierce) by incubating at 37 °C for 2 h. Sterols were analyzed by GC/MS as the respective trimethylsilyl (TMS) derivatives using an Agilent HP 6850 GC system coupled with an Agilent 5979 mass-selective detector (MSD) operating at 70 eV. The individual sterols were separated using a DB-5-MS column (30 m x 0.25 mm inner diameter, 0.25-µm film thickness) in which the temperature was held at 200 °C for 1 min, followed by an increase of 15 °C/min up to 300 °C and finally held at 300 °C for 20 min.

Cellular Labeling Assay and “click” chemistry: C2C12 cells (1.5×10^6 cells) were pre-treated with competitor compounds in culture media (1 mL, DMEM with 10% FBS) for 1 hour at 37 °C. Then additional competitor compound and **4.9** (2 µM) were added in culture media (1 mL), and the cells were incubated for an additional hour. The cells were then pelleted, washed with phosphate buffered saline pH 7.4 (2 x 1 mL), and lysed with 1% SDS in phosphate buffered saline (400 µL). The denatured lysates were sonicated to

reduce viscosity, centrifuged to pellet insoluble material, and normalized for total protein content. The normalized lysates (22 μL , 3 mg/mL) were mixed with TAMRA azide **4.10** (0.5 μL , 5 mM), TCEP (0.5 μL , 50 mM, pH \sim 7.0), TBTA ligand in 1:4 DMSO/*t*-butyl alcohol (1.5 μL , 1 mM), and CuSO_4 (0.5 μL , 50 mM). Reactions were incubated at room temperature for one hour and then resolved by SDS-PAGE. The resulting gel was scanned for fluorescence (Typhoon Imaging System, Molecular Dynamics), and coomassie stained.

Enrichment of 4.9-labeled proteins and mass spectrometric analysis. C2C12 cells were treated with probe **4.9**, lysed and denatured as described above. The labeled and denatured lysates were sonicated to reduce viscosity and centrifuged to give pellets of the insoluble material. The buffer was then exchanged by using a NAP-5 column (GE Healthcare) equilibrated with 1% SDS in phosphate buffered saline pH 7.4 (PBS). The samples were then normalized for protein content. The normalized lysates (1 mL, 2 mg/mL) were subjected to a “click” reaction with a cleaveable biotin azide **4.11** using the general conditions described above. After the “click” reaction, the proteins were precipitated using cold acetone (9 mL) and pelleted. The protein pellet was washed twice with cold acetone (10 mL) and solubilized with 1% SDS in 50 mM Tris pH 8.0 (200 μL). This solution was diluted with 1% Nonidet P-40 (NP-40) in PBS (300 μL). The buffer was then exchanged using a NAP-5 column (GE Healthcare) equilibrated with 1% NP-40, 0.1% SDS in PBS. Biotinylated proteins were immobilized at 4° C using Streptavidin-coupled Dynabeads (Invitrogen). The unbound material was washed from the

Dynabeads with 1% NP-40, 0.1% SDS in PBS (3 x 1 mL), 6M urea in PBS (2 x 1mL at 4° C), PBS (1 mL), 0.5% SDS in PBS (2 x 1 mL), and PBS (2 x 1 mL) in that order. Proteins conjugated to biotin through the base labile ester linkage were selectively eluted with aqueous sodium hydroxide (0.4 N). The supernatant was neutralized with aqueous hydrochloric acid (0.8 N), and then run on an SDS-PAGE gel. The gel was colloidal coomassie stained.

The competed band at ~35 kDa was cut, and the proteins contained in the bands of interest were digested in-gel with trypsin as described previously [13]. Using the protocol we described recently [14], the extracted digests were vacuum-evaporated and re-suspended in 10 µL of 0.1% formic acid in water. The digests were separated by nano-flow liquid chromatography using a 75 µm x 150 mm reverse phase 1.7 µm BEH 130 C18 column (Waters) at a flow rate of 350 nL/min in a NanoAcquity™ Ultra performance UPLC system (Waters). Solvent A was 0.1% formic acid in water and solvent B was 0.1% formic acid in acetonitrile. Following equilibration of the column in 5% solvent B, approximately one-half of each digest (5 µL) was injected, then the organic content of the mobile phase was increased linearly to 40% over 60 min, and then to 50% in 1 min. The liquid chromatography eluate was coupled to a hybrid linear ion trap-Orbitrap mass spectrometer (LTQ-XL, Thermo Scientific, San Jose, CA) equipped with a nanoelectrospray ion source. Spraying was from an uncoated 15 µm inner diameter spraying needle (New Objective, Woburn, MA). Peptides were analyzed in

positive ion mode and in information-dependent acquisition mode to automatically switch between MS and MS/MS acquisition. MS spectra were acquired in profile mode using the Orbitrap analyzer in the m/z range between 300 and 1800. For each MS spectrum, the 6 most intense multiple charged ions over a threshold of 1000 counts were selected to perform CID (collision induced dissociation) experiments. Product ions were analyzed on the linear ion trap in centroid mode. The CID collision energy was automatically set to 25%. A dynamic exclusion window of 0.5 Da was applied that prevented the same m/z from being selected for 60 seconds after its acquisition.

Mass spectrometry data analysis. Peak lists were generated using PAVA [15]. The peak lists were searched against the murine and human subset of the UniProtKB database as of July 6, 2011 (containing 182779 entries) using ProteinProspector version 5.8.0. Peptide tolerance in searches was 20 ppm for precursor and 0.6 Da for product ions, respectively. Cleavage specificity was selected to Trypsin. Peptides containing two mis-cleavages were allowed. Carbamidomethylation of cysteine, acetylation of the N-terminus of the protein, pyroglutamate formation from N-terminal glutamine, and oxidation of methionine were allowed as variable modifications.

The number of modification was limited to two per peptide. Hits were considered significant when two or more peptide sequences matched a protein entry and the Prospector score was above the significance level. A minimal ProteinProspector protein

score of 20, a peptide score of 15, a maximum expectation value of 0.05 and a minimal discriminate score threshold of 0.0 were used for initial identification criteria. For identifications based on one peptide sequence with high scores, the MS/MS spectrum was reinterpreted manually by matching all the observed fragment ions to a theoretical fragmentation obtained using MS Product (Protein Prospector).

Reactivity with Glutathione

Compounds **4.1** and **4.6** were incubated with reduced glutathione (10mM) in phosphate buffered saline (pH 7.4). The reactions were quenched with a solution of trifluoroacetic acid (2%) in methanol and then separated by reverse phase liquid chromatography. The amount of unreacted compound remaining was quantified by integrating the compound peak observed by evaporative light scattering. This value was normalized to an internal control of quinine, and then used to report the fraction of compound remaining.

Synthetic Procedures.

¹H NMR spectra were recorded on a Varian INOVA-400 400 MHz spectrometer. Chemical shifts are reported in δ units (ppm) relative to residual solvent peaks. Coupling constants (*J*) are reported in hertz (Hz). Compounds **4.1**, **4.2**, and **4.3** were prepared as described previously [21] and TAMRA azide **4.10** was prepared according to the reported procedure [16]. Compound **4.9** was prepared as described below and exhibited

spectral characteristics consistent with that reported previously [17]. Intermediates **4.15** was prepared as described previously [18]. All reagents and solvents were purchased from Aldrich Chemical or Acros Organics and used as received. Column chromatography was carried out using a Biotage SP1 flash chromatography system and silica gel cartridges from Biotage. Analytical TLC plates from EM Science (Silica Gel 60 F254) were employed for TLC analyses. Preparative HPLC purifications were performed using a Biotage Parallax Flex equipped with Waters Xbridge 19 x 50mm C18 5 μ M OBD columns.

***tert*-Butyl N-[(2S)-1-[[*(E,3S)*-1-(benzenesulfonyl)-5-phenylpent-1-en-3-yl]amino]-3-(4-methylphenyl)-1-oxopropan-2-yl]carbamate (16).**

(2S)-3-(4-methylphenyl)-2-[(2-methylpropan-2-yl)oxycarbonylamino]propanoic acid (**4.14**, 0.040 g, 0.143 mmol) and (*1E,3S*)-1-(benzenesulfonyl)-5-phenylpent-1-en-3-amine hydrochloride (**4.15**, 0.051 g, 0.150 mmol) were dissolved in 1 mL DMF. HATU (0.086 g, 0.225 mmol) was added followed by diisopropylethylamine (0.150 mL, 0.858 mmol). The reaction mixture was stirred at room temperature until complete as judged by LC-MS. The reaction mixture was diluted with 5 mL ethyl acetate, shaken with 5 mL 1N HCl, 5 mL 50% saturated NaHCO₃ then 5 mL brine. The organic layer was further dried with MgSO₄. Solvent was removed to afford the crude product (89 mg) that was used without further purification. ¹H NMR (400 MHz, CDCl₃) δ 1.26 - 1.54 (m, 9 H), 1.74 - 1.96 (m, 2 H), 2.28 (s, 3 H), 2.49 - 2.64 (m, 2 H), 2.91 - 3.01 (m, 2 H), 4.18- 4.28 (m, 1 H), 4.59 - 4.71 (m, 1 H), 4.97 (br s, 1 H), 6.11 - 6.22 (m, 2 H), 6.79 (dd, *J* = 15.11, 4.67 Hz, 1

H), 6.98 - 7.11 (m, 5 H), 7.14 - 7.20 (m, 1 H), 7.19 - 7.25 (m, 1 H), 7.48 - 7.54 (m, 2 H), 7.57 - 7.66 (m, 1 H), 7.81 - 7.85 (m, 1 H).

General procedure for preparing pyridyl analogs 4.3 and 4.4.

tert-Butyl N-[(2S)-1-[[*(E,3S)*-1-(benzenesulfonyl)-5-phenylpent-1-en-3-yl]amino]-3-(4-methylphenyl)-1-oxopropan-2-yl]carbamate (**4.16**) is treated with 4N HCl in dioxane (~100 molar equiv HCl) at room temperature until cleavage of the Boc group is complete, as judged by LC/MS analysis. The solvent is then removed and the residue re-suspended in DMF (~0.15 M). The requisite pyridine carboxylic acid (1.2 equiv), HATU (1.5 equiv) and diisopropylethylamine (6.0 equiv) are added and the reaction mixture stirred until the coupling is judged complete by LC/MS analysis, typically 12 hours. The reaction mixture is then diluted with ethyl acetate (at least 5x volume) and shaken with an equivalent volume of 50% saturated aqueous NaHCO₃. The organic layer is separated and then washed with saturated NaCl, and dried (MgSO₄), filtered, and concentrated to afford the crude product, which is purified using automated silica gel chromatography with a mobile phase of ethyl acetate/hexanes.

N-[(2S)-1-[[[(E,3S)-1-(benzenesulfonyl)-5-phenylpent-1-en-3-yl]amino]-3-(4-methylphenyl)-1-oxopropan-2-yl]pyridine-2-carboxamide trifluoroacetate salt (4.3).

Compound **4.3** was prepared according to the general procedure except that additional purification by HPLC was required. 10% overall yield from **4.14**. ^1H NMR (400 MHz, CDCl_3) δ 1.71 - 1.82 (m, 1 H), 1.82 - 1.92 (m, 1 H), 2.29 (s, 3 H), 2.48 - 2.57 (m, 2 H), 3.16 (dd, $J = 7.05, 2.11$ Hz, 2 H), 4.65 (dd, $J = 5.13, 3.48$ Hz, 1 H), 4.75 (q, $J = 7.33$ Hz, 1 H), 6.22 (dd, $J = 15.11, 1.74$ Hz, 1 H), 6.39 (d, $J = 7.87$ Hz, 1 H), 6.80 (dd, $J = 15.11, 4.85$ Hz, 1 H), 6.96 - 7.02 (m, 2 H), 7.02 - 7.09 (m, 2 H), 7.09 - 7.22 (m, 4 H), 7.48 (ddd, $J = 7.55, 4.81, 1.01$ Hz, 1 H), 7.52 - 7.60 (m, 2 H), 7.60 - 7.67 (m, 1H), 7.82 - 7.92 (m, 2 H), 8.14 (d, $J = 7.87$ Hz, 1 H), 8.53 - 8.66 (m, 2 H); ^{13}C NMR (100 MHz, CDCl_3) δ 21.1, 31.7, 35.5, 37.3, 49.3, 55.3, 122.7, 126.2, 126.8, 127.6, 128.3, 128.5, 129.1, 129.3, 129.5, 130.6, 133.2, 133.4, 136.9, 138.0, 140.2, 140.3, 145.5, 147.9, 148.6, 164.3, 170.4; MS: (m/z): $[\text{M}+\text{H}]^+ = 568.6$

N-[(2S)-1-[[[(E,3S)-1-(benzenesulfonyl)-5-phenylpent-1-en-3-yl]amino]-3-(4-methylphenyl)-1-oxopropan-2-yl]pyridine-4-carboxamide (4.4).

Compound **4.4** was prepared according to the general procedure in 45% overall yield from **4.14**. ^1H NMR (400 MHz, CDCl_3): 1.71 - 1.84 (m, 1 H), 1.84 - 1.97 (m, 1 H), 2.23 (s, 1 H), 2.34 (s, 2 H), 2.46 - 2.64 (m, 2 H), 2.77 - 2.86 (m, 1 H), 3.01 (dd, $J = 13.46, 8.88$ Hz, 1 H), 3.17 (dd, $J = 13.92, 5.86$ Hz, 1 H), 4.59 - 4.73 (m, 1 H), 5.75 (d, $J = 8.24$ Hz, 1 H), 6.19 (d, $J = 1.65$ Hz, 1 H), 6.15 (d, $J = 1.47$ Hz, 1 H), 6.77 (d, $J = 5.49$ Hz, 1 H), 6.73 (d, $J = 5.31$

Hz, 1 H), 6.87 (d, $J = 7.33$ Hz, 1 H), 6.99 - 7.29 (m, 7 H), 7.43 - 7.72 (m, 3 H), 7.72 - 8.00 (m, 2 H), 8.62 - 8.84 (m, 2 H); MS: (m/z): $[M+H]^+ = 568.6$

General procedure for reduction of 4.1, 4.3, and 4.4 to afford analogs 4.6, 4.7, and 4.8.

The vinylsulfone starting material was dissolved in ethanol (0.015 mM concentration) and one weight equivalent of 10% Pd on C was added. The flask was evacuated and filled with hydrogen three times, and the reaction mixture maintained under a hydrogen atmosphere (1 atm) until the vinyl group was fully reduced, but before over-reduction of the pyridyl ring began to occur. Typically this required ~ 6 hrs. The hydrogen balloon was then removed and the reaction mixture filtered through Celite. The filtrate was concentrated and the crude product purified using automated silica gel chromatography with a mobile phase of ethyl acetate/hexanes.

N-[(2S)-1-[[[(3S)-1-(benzenesulfonyl)-5-phenylpentan-3-yl]amino]-1-oxo-3-phenylpropan-2-yl]-4-methylpiperazine-1-carboxamide (4.6).

Compound **4.6** was prepared from **4.1** according to the general procedure, except that additional purification by preparative HPLC was required. ^1H NMR (400 MHz, CDCl_3): 1.53 - 1.83 (m, 4 H), 2.35 - 2.61 (m, 2 H), 2.61 - 2.85 (m, 3 H), 2.86 - 3.03 (m, 3 H), 3.09 (dd, $J = 13.73, 6.96$ Hz, 1 H), 3.28 (br s, 2 H), 3.45 (br s, 1 H), 3.75 - 3.91 (m, 1 H), 3.96 (br s, 1 H), 4.05 (br s, 2 H), 4.29 - 4.46 (m, 1 H), 6.27 (d, $J = 6.59$ Hz, 1 H), 6.90 (d, $J = 8.97$ Hz,

1 H), 7.04 (d, $J = 6.96$ Hz, 2 H), 7.08 - 7.35 (m, 7 H), 7.53 (t, $J = 7.78$ Hz, 2 H), 7.64 (t, $J = 7.42$ Hz, 1 H), 7.72 - 7.91 (m, 2 H); MS: (m/z): $[M+H]^+ = 577.2$

N-[(2S)-1-[[[(3S)-1-(benzenesulfonyl)-5-phenylpentan-3-yl]amino]-3-(4-methylphenyl)-1-oxopropan-2-yl]pyridine-2-carboxamide (4.7).

Compound **4.7** was prepared from **4.3** according to the general procedure. Yield: 20%
 ^1H NMR (400 MHz, CDCl_3): 1.55 - 1.73 (m, 3 H), 1.82 - 1.95 (m, 1 H), 2.19 - 2.35 (m, 3 H), 2.37 - 2.51 (m, 2 H), 2.84 - 3.04 (m, 2H), 3.04 - 3.24 (m, 2 H), 3.88 (tq, $J = 9.21, 4.36$ Hz, 1 H), 4.59 - 4.74 (m, 1 H), 5.96 (d, $J = 9.16$ Hz, 1 H), 6.93 - 7.00 (m, 2 H), 7.03 - 7.22 (m, 6 H), 7.44 (ddd, $J = 7.55, 4.81, 1.19$ Hz, 1 H), 7.51 - 7.59 (m, 2 H), 7.60 - 7.68 (m, 1 H), 7.80 - 7.89 (m, 2 H), 8.13 (dt, $J = 7.83, 1.03$ Hz, 1 H), 8.47 - 8.59 (m, 2 H); MS: (m/z): $[M+H]^+ = 570.2$

N-[(2S)-1-[[[(3S)-1-(benzenesulfonyl)-5-phenylpentan-3-yl]amino]-3-(4-methylphenyl)-1-oxopropan-2-yl]pyridine-4-carboxamide (4.8).

Compound **4.8** was prepared from **4.4** according to the general procedure. Yield: 26%
 ^1H NMR (400 MHz, CDCl_3): 8.72 (br s, 2H), 7.84 - 7.90 (m, 2H), 7.65 - 7.71 (m, 1H), 7.57 - 7.61 (m, 2H), 7.52 - 7.57 (m, 2H), 7.16 - 7.23 (m, 3H), 7.11 - 7.16 (m, 4H), 6.99 - 7.03 (m, 2H), 6.96 (d, $J = 7.33$ Hz, 1H), 5.82 (d, $J = 8.97$ Hz, 1H), 4.70 (dt, $J = 6.23, 7.87$ Hz, 1H), 3.90 (td, $J = 4.56, 9.02$ Hz, 1H), 3.04 - 3.19 (m, 2H), 2.85 - 3.02 (m, 2H), 2.42 - 2.55 (m,

2H), 2.34 (s, 3H), 1.78 - 2.00 (m, 1H), 1.54 - 1.76 (m, 3H); ¹³C NMR (100 MHz, CDCl₃) 170.4, 165.4, 150.6, 140.7, 140.6, 139.0, 137.3, 133.8, 132.9, 129.8, 129.3, 129.0, 128.5, 128.2, 128.0, 126.2, 120.8, 55.4, 52.9, 48.3, 37.8, 36.6, 32.0, 28.0, 21.1 MS: (*m/z*): [M+H]⁺ = 570.2

Benzyl (2S)-3-phenyl-2-[(4-prop-2-ynylpiperazine-1-carbonyl)amino]propanoate (4.18).

Benzyl (2S)-2-amino-3-phenylpropanoate hydrochloride (**4.17**, 2.0 g, 6.85 mmol, 1.0 equiv) was dissolved in methylene chloride (70 mL), and the solution was cooled in an ice/water bath. A saturated aqueous solution of sodium bicarbonate (70 mL) was then added, followed by a solution of triphosgene (0.81 g, 2.74 mmol, 0.4 equiv) in methylene chloride (15 mL) was added slowly. The reaction was stirred for 1 hour and then quenched by addition of excess benzylamine. The aqueous layer was extracted with methylene chloride (2 x 50 mL) and the combined organic layers were then washed with 1M hydrochloric acid (2 x 10 mL). Hexane was added to precipitate the desired isocyanate intermediate (1.7 g, 89% crude yield) which contained a small amount of starting material. A portion of this material was used without further purification in the next reaction.

A flask was charged with 4-(prop-2-ynyl)piperazine dihydrochloride (120 mg, 0.61 mmol, 1.0 equiv) and dissolved in a mixture of tetrahydrofuran (1 mL), dimethylformamide (0.4 mL), and diisopropylethylamine (0.42 mL, 2.44 mmol, 4.0 equiv). In another flask, the crude benzyl (2S)-2-isocyanato-3-phenylpropanoate (171 mg, 0.061 mmol, 1.0 equiv)

was dissolved in tetrahydrofuran (1 mL) and chilled in an ice/water bath. The 4-(prop-2-ynyl)piperazine solution was then added dropwise to the solution of isocyanate. After 1 hour, the reaction was diluted with ethyl acetate (10 mL) and washed with water (5 x 5 mL) and brine (5 mL). The organic layer was then dried over sodium sulfate, filtered, and concentrated. The crude residue was subjected to automated flash chromatography using a solvent gradient of methanol in dichloromethane (0-4%). The relevant fractions were collected and concentrated to afford benzyl (2S)-3-phenyl-2-[(4-prop-2-ynylpiperazine-1-carbonyl)amino]propanoate (**4.18**, 195 mg, 79% yield, ~85% pure) as a clear gum. This material was used in the next step without further purification. MS: 406.15 (M+H).

N-[(2S)-1-[[[(E,3S)-1-(benzenesulfonyl)-5-phenylpent-1-en-3-yl]amino]-1-oxo-3-phenylpropan-2-yl]-4-prop-2-ynylpiperazine-1-carboxamide (4.9).

A flask was charged with benzyl (2S)-3-phenyl-2-[(4-prop-2-ynylpiperazine-1-carbonyl)amino]propanoate (**4.18**, 90 mg, 0.22 mmol, 1.0 equiv) and methanol (2 mL) was added. Aqueous sodium hydroxide (2M, 0.24 mL) was added, and the resulting mixture was stirred overnight and then concentrated to dryness. The resulting crude (2S)-3-phenyl-2-[(4-prop-2-ynylpiperazine-1-carbonyl)amino]propanoic acid (MS: 316.04 (M+H)) was dissolved in dimethylformamide (1 mL) and treated with (E,3S)-1-(benzenesulfonyl)-5-phenylpent-1-en-3-amine (75 mg, 0.22 mmol, 1.0 equiv) and diisopropylethylamine (118 μ L, 0.67 mmol, 3.0 equiv). The mixture was then cooled in a

dry ice-acetone bath, and HATU (92.8 mg, 0.24 mmol, 1.1 equiv) was added in one portion. The reaction was allowed to slowly warm to -40°C over 3 hours and was then diluted in ethyl acetate (20 mL) and washed with water (5 x 5 mL) and brine (5 mL). The organic layer was dried over sodium sulfate, filtered and concentrated. The crude material was then purified by reverse phase HPLC, affording 17 mg of compound **4.2** as a white solid (11% over three steps). ¹H NMR (400MHz, CDCl₃): 1.79 (m, 1H), 1.89 (m, 1H), 2.50 (t, *J* = 2.1 Hz, 1H), 2.52-2.62 (m, 2H), 2.92-3.07 (m, 6H), 3.60 (m, 4H), 3.72 (m, 2H), 4.37 (q, *J* = 7.7Hz, 1H), 4.61 (m, 1H), 5.12 (m, 1H), 5.98 (m, 1H), 6.07 (dd, *J* = 15.4, 1.2 Hz, 1H), 6.75 (dd, *J* = 15.4, 5.3 Hz, 1H), 7.00-7.90 (m, 15H). LCMS: 599.17 (M+H); MW: 598.75).

***Tert*-butyl 5-hydroxypentylcarbamate (4.20)**: 5-amino-1-pentanol (**4.19**, 500 mg, 4.85 mmol) and di-*tert*-butyl dicarbonate (881 mg, 4.04 mmol) were dissolved in water (5 mL). The solution was adjusted to pH 12 with 1N NaOH (aq) and stirred for 20 min. The reaction was extracted three times with ethyl acetate. The organic layers were combined, washed with brine, dried over sodium sulfate, and concentrated to afford the product as a clear oil (719 mg, 88% yield). ¹H NMR (400MHz, CDCl₃): 1.30-1.70 (m, 15H), 3.12 (q, *J* = 6.5 Hz, 2H), 3.63 (m, 2H), 4.53 (b s, 1H).

5-(tert-butoxycarbonylamino)pentyl 6-bromohexanoate (4.22): *Tert*-butyl 5-hydroxypentylcarbamate (**4.20**, 696 mg, 3.42 mmol) and triethylamine (416 mg, 4.11 mmol) were mixed in methylene chloride (5 mL) and cooled in an ice-water bath. The 6-bromohexanoyl chloride (**4.21**, 1.464 g, 6.06 mmol) in methylene chloride (5 mL) was added slowly to the reaction. After 20 min, water (5 mL) was added, and the aqueous layer was acidified with 1N HCl (aq). The organic layer was separated, washed with saturated NaHCO₃ (aq) and brine, dried over sodium sulfate, and concentrated under reduced pressure. A clear oil containing a mixture (2:1) of the desired product and the starting acid chloride was obtained after silica flash chromatography using a gradient of ethyl acetate up to 25% (v/v) in hexanes. The crude material was used without further purification. ¹H NMR (400MHz, CDCl₃): 1.30-1.55 (m, 15H), 1.59-1.69 (m, 4H), 1.82-1.92 (m, 2H), 2.31 (t, *J* = 7.4 Hz, 2H), 3.11(q, *J* = 6.6 Hz, 2H), 3.40 (t, *J* = 6.8 Hz, 2H), 4.05 (t, *J* = 6.7 Hz, 2H), 4.51 (br s, 1H).

5-(tert-butoxycarbonylamino)pentyl 6-azidohexanoate (4.23): Crude 5-(*tert*-butoxycarbonylamino)pentyl 6-bromohexanoate (**4.22**, 220mg) and sodium azide (301 mg, 4.63 mmol) were dissolved in dimethylformamide (10 mL). The vented mixture was stirred at 50°C behind a blast shield for 48 hrs. The reaction was diluted in ethyl acetate (80 mL) and washed extensively with water. The organic layer was washed with brine, dried over sodium sulfate, and concentrated. Silica flash chromatography using a gradient of ethyl acetate up to 25% (v/v) in hexanes yielded the desired product as a

yellow oil, (110 mg, 49% yield over 2 steps). ^1H NMR (400MHz, CDCl_3): 1.30-1.55 (m, 15H), 1.55-1.70 (m, 6H), 2.30 (t, $J=7.4$ Hz, 2H), 3.10 (q, $J=6.5$ Hz, 2H), 3.26 (t, $J=6.9$ Hz, 2H), 4.05 (t, $J=6.7$ Hz, 2H), 4.50 (br.s, 1H).

5-aminopentyl 6-azidohexanoate (4.24): 5-(*tert*-butoxycarbonylamino)pentyl 6-azidohexanoate (**4.23**, 50mg, 0.146 mmol) was dissolved in methylene chloride (0.6 mL) and cooled in an ice bath. Trifluoroacetic acid (0.4 mL) was added drop-wise then stirred for 20 min. Volatiles were removed under reduced pressure to afford the trifluoroacetic acid salt in quantitative yield as a yellow oil. ^1H NMR (400MHz, CDCl_3): 1.34-1.47 (m, 4H), 1.55-1.75 (m, 8H), 2.31 (t, $J = 7.5$ Hz, 2H), 2.96 (m, 2H), 3.26 (t, $J = 6.9$ Hz, 2H), 4.05 (t, $J = 6.6$ Hz, 2H), 7.80(m, 3H).

Base cleavable biotin azide (4.11): 5-aminopentyl 6-azidohexanoate (**4.24**, TFA salt, 31.3 mg, 0.088 mmol), biotin-NHS ester (**4.25**, 20 mg, 0.059 mmol), triethylamine (17.8 mg, 0.175 mmol) were dissolved in dimethylformamide (0.8 mL) and stirred at room temperature for 48 hrs. The reaction was diluted into ethyl acetate (20 mL), washed three times with 0.5N HCl (aq), once with brine, dried over magnesium sulfate, and concentrated. Silica flash chromatography using a stepwise 0.5% increments methanol to 10% (v/v) in methylene chloride yielded the desired product as a white solid (16.2 mg, 59% yield). ^1H NMR (400MHz, CDCl_3): 1.29-1.79 (m, 16H), 2.18 (t, $J = 7.4$ Hz, 2H), 2.30 (t, $J = 7.4$ Hz, 2H), 2.72 (d, $J = 12.8$ Hz, 1H), 2.89 (dd, $J = 4.8, 12.7$ Hz, 1H), 3.13 (m, 2H), 3.21

(q, $J = 6.5$ Hz, 2H), 3.26 (t, $J = 7.4$ Hz, 2H), 4.05 (t, $J = 6.7$ Hz, 2H), 4.30 (dd, $J = 7.4, 4.9$ Hz, 1H), 4.50 (dd, $J = 7.4, 5.0$ Hz, 1H), 5.71 (br s, 1H), 6.09 (m, 1H), 6.39 (s, 1H). ^{13}C NMR (400MHz, CDCl_3): 23.28, 24.49, 25.64, 26.22, 28.07, 28.23, 28.31, 28.53, 29.21, 34.06, 36.00, 39.28, 40.47, 51.21, 55.60, 60.29, 61.89, 64.14, 173.14, 173.57.

Methyl (2S)-3-(4-methylphenyl)-2-(pyridine-4-carbonylamino)propanoate (4.27).

To a flask containing anhydrous methanol (5.0 mL) and cooled in an icebath was added thionyl chloride (0.71 mL, 9.77 mmol, 3.5 equiv). After stirring for 5 minutes, (2S)-2-amino-3-(4-methylphenyl)propanoic acid (**4.26**, 0.50 g, 2.79 mmol) was added and the reaction mixture was warmed to room temperature with stirring for 64 hours. The solvent was evaporated and the residue was triturated twice with ether to afford methyl (2S)-2-amino-3-(4-methylphenyl)propanoate hydrochloride (0.61 g, 95%).

^1H NMR (400 MHz, D_2O): 2.31 (s, 3 H), 3.19 (dd, $J = 14.47, 7.33$ Hz, 1 H), 3.28 (dd, $J = 14.56, 5.95$ Hz, 1 H), 3.82 (s, 3 H), 4.39 (dd, $J = 7.33, 6.04$ Hz, 1 H), 7.11 - 7.21 (m, 2 H), 7.21 - 7.30 (m, 2 H).

A solution of (2S)-2-amino-3-(4-methylphenyl)propanoate hydrochloride (0.61 g, 2.67 mmol) in DMF (10 mL) was treated with isonicotinic acid (0.39 g, 3.20 mmol), HATU (1.52 g, 4.01 mmol) and diisopropylethylamine (2.32 mL, 13.4 mmol) and the reaction mixture stirred overnight. The reaction mixture was then poured into ethyl acetate and washed with 50% saturated NaHCO_3 , saturated NaCl, and dried (MgSO_4), filtered, and

concentrated. The crude product was purified by automated silica gel chromatography to afford **4.27** (0.36 g, 45% overall) ^1H NMR (400 MHz, CDCl_3): 2.24 - 2.39 (m, 3 H), 3.21 (qd, $J = 13.83, 5.40$ Hz, 2 H), 3.73 - 3.83 (m, 3 H), 4.96 - 5.12 (m, 1 H), 6.61 (d, $J = 6.41$ Hz, 1 H), 6.98 (d, $J = 7.69$ Hz, 2 H), 7.09 (d, $J = 7.87$ Hz, 2 H), 7.49 - 7.63 (m, 2 H), 8.65 - 8.80 (m, 2 H).

General procedure for synthesis of 4.12 and 4.13 from 4.27.

Intermediate **4.27** (0.36 g, 1.22 mmol) is dissolved in 5 mL 4:1 THF:water and then lithium hydroxide (0.117 g, 4.86 mmol) is added. The reaction mixture is stirred at room temperature and usually forms a suspension. When the hydrolysis reaction is judged complete by LC/MS, the reaction mixture is adjusted to pH ~ 4 by slow addition of ~ 0.8 mL 2N HCl. The solvent is then removed in vacuo and the crude lithium carboxylate intermediate used in subsequent coupling reactions without further purification.

A portion of the lithium carboxylate intermediate is dissolved in DMF (0.15 M) and the requisite amine (2.2 equiv), HATU (2.0 equiv), and diisopropylethylamine (5.0 equiv) are added. The reaction mixture is stirred at room temperature until the reaction is judged complete by LC/MS (typically 2 hrs). The reaction mixture is then diluted with ethyl acetate (at least 5x volume) and shaken with an equivalent volume of 50% saturated NaHCO_3 . The organic layer is then separated, washed with saturated NaCl, dried (MgSO_4), filtered, and concentrated. The crude product is then purified by automated

silica gel chromatography (ethyl acetate/hexanes) and/or preparative HPLC to afford the final product.

N-[(2S)-1-(methylamino)-3-(4-methylphenyl)-1-oxopropan-2-yl]pyridine-4-carboxamide (4.12).

Prepared according to the general procedure, employing methylamine in the coupling reaction. Yield: 33% ¹H NMR (400 MHz, CDCl₃): 2.15 - 2.33 (m, 3 H), 2.60 - 2.77 (m, 3 H), 2.97 - 3.12 (m, 1 H), 3.12 - 3.24 (m, 1 H), 4.78 (td, *J* = 7.83, 6.50 Hz, 1 H), 6.15 (d, *J* = 4.58 Hz, 1 H), 6.93 - 7.17 (m, 3 H), 7.46 - 7.63 (m, 3 H), 8.53 - 8.72 (m, 2 H); MS: (*m/z*): [M+H]⁺ = 298.1

N-[(2S)-3-(4-methylphenyl)-1-oxo-1-[(1R)-1-phenylethyl]amino]propan-2-yl]pyridine-4-carboxamide (4.13).

Prepared according to the general procedure, employing (*R*)-phenethylamine in the coupling reaction. Yield: 43% ¹H NMR (400 MHz, CDCl₃): 1.29 (d, *J* = 6.96 Hz, 2 H), 2.36 (s, 2 H), 3.08 (dd, *J* = 13.37, 8.97 Hz, 1 H), 3.27 (dd, *J* = 13.37, 6.04 Hz, 1 H), 4.69 - 4.87 (m, 1 H), 4.95 (quin, *J* = 6.96 Hz, 1 H), 6.15 (d, *J* = 7.33 Hz, 2 H), 6.25 (br. s., 2 H), 7.08 - 7.38 (m, 7 H), 7.85 (br s, 2 H), 8.15 (d, *J* = 7.14 Hz, 1 H), 8.63 (br s, 2 H); MS: (*m/z*): [M+H]⁺ = 388.2

References

1. Doyle, P. S.; Chen, C. K.; Johnston, J. B.; Hopkins, S. D.; Leung, S. S.; Jacobson, M. P.; Engel, J. C.; McKerrow, J. H.; Podust, L. M. *Antimicrob. Agents. Chemother.* **2010**, *54*, 2480-2488.
2. Chen, C. K.; Leung, S. S.; Guilbert, C.; Jacobson, M. P.; McKerrow, J. H.; Podust, L. M. *PLoS Negl. Trop. Dis.* **2010**, *4*, e651.
3. Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. *J. Med. Chem.* **2006**, *49*, 534-553.
4. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739-1749.
5. Zhu, K.; Shirts, M. R.; Friesner, R. A.; Jacobson, M. P. *J. Chem. Theory. Comput.* **2007**, *3*, 640-648.
6. Huang, N.; Kalyanaraman, C.; Irwin, J. J.; Jacobson, M. P. *J. Chem. Inf. Model.* **2006**, *46*, 243-253.
7. Gunatilleke, S. S.; Calvet, C. M.; Johnston, J. B.; Chen, C. K.; Erenburg, G.; Gut, J.; Engel, J. C.; Ang, K. K.; Mulvaney, J.; Chen, S.; Arkin, M. R.; McKerrow, J. H.; Podust, L. M. *PLoS Negl. Trop. Dis.* **2012**, *6*, e1736.
8. Morrison, J. F. *Biochim. Biophys. Acta* **1969**, *185*, 269-286.
9. Engel, J. C.; Doyle, P. S.; Hsieh, I.; McKerrow, J. H. *J. Exp. Med.* **1998**, *188*, 725-734.
10. Engel, J. C.; Ang, K. K.; Chen, S.; Arkin, M. R.; McKerrow, J. H.; Doyle, P. S. *Antimicrob. Agents. Chemother.* **2010**, *54*, 3326-3334.

- 11 Engel, J. C.; Doyle, P. S.; Dvorak, J. A. *J. Protozool.* **1985**, *32*, 80-83.
- 12 McKerrow, J. H.; Doyle, P. S.; Engel, J. C.; Podust, L. M.; Robertson, S. A.; Ferreira, R.; Saxton, T.; Arkin, M.; Kerr, I. D.; Brinen, L. S.; Craik, C. S. *Mem. I. Oswaldo Cruz.* **2009**, *104*, 263-269.
- 13 Rosenfeld, J.; Capdevielle, J.; Guillemot, J. C.; Ferrara, P. *Anal. Biochem.* **1992**, *203*, 173-179.
- 14 Bhattacharya, M.; Su, G.; Su, X.; Oses-Prieto, J. A.; Li, J. T.; Huang, X.; Hernandez, H.; Atakilit, A.; Burlingame, A. L.; Matthay, M. A.; Sheppard, D. *Am J. Physiol. Lung. Cell. Mol. Physiol.* **2012**, *303*, L12-L19.
- 15 Guan, S.; Price, J. C.; Prusiner, S. B.; Ghaemmaghami, S.; Burlingame, A. L. *Mol. Cell. Proteomics.* **2011**, *10*.
- 16 Speers, A. E.; Cravatt, B. F. *Chem. Biol.* **2004**, *11*, 535-546.
- 17 Yang, P.-Y.; Wang, M.; He, C. Y.; Yao, S. Q. *Chem. Commun.* **2012**, *48*, 835-837.
- 18 Somoza, J. R.; Zhan, H.; Bowman, K. K.; Yu, L.; Mortara, K. D.; Palmer, J. T.; Clark, J. M.; McGrath, M. E. *Biochemistry* **2000**, *39*, 12543-12551.

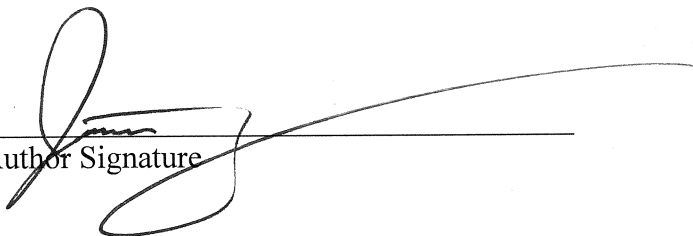
Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature

A handwritten signature in black ink, written over a horizontal line. The signature is stylized and appears to be a first name followed by a last name.

Date

3/27/13