

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

Drug Target Discovery Using Designer Drug Sensitive Yeast

### Permalink

<https://escholarship.org/uc/item/42b8231t>

### Author

Goldgof, Gregory Mark

### Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Drug Target Discovery Using Designer Drug Sensitive Yeast

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Biomedical Sciences

by

Gregory Mark Goldgof

Committee in charge:

Professor Elizabeth Winzeler, Chair  
Professor James Andrew McCammon  
Professor Victor Nizet  
Professor Scott Rifkin  
Professor Deborah Spector

2017

Copyright

Gregory Mark Goldgof, 2017

All rights reserved.

The Dissertation of Gregory Mark Goldgof is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

---

---

---

---

---

Chair

University of California, San Diego

2017

## TABLE OF CONTENTS

Signature Page.....	iii
Table of Contents .....	iv
List of Figures .....	vi
List of Tables .....	viii
Acknowledgements .....	ix
Vita .....	xii
Abstract of the Dissertation .....	xiv
1. Introduction.....	1
History of Antimicrobial Drug Discovery.....	1
History of Antimicrobial Mechanism of Action Studies.....	3
Modern Antimicrobial Drug Discovery.....	7
Malaria: The Disease .....	10
<i>Plasmodium</i> Life Cycle .....	14
Current Antimalarial Drug Discovery.....	17
Modern Approaches to Drug Target Identification .....	21
2. Method Validation Using Known Drug Target Combinations .....	36
3. The Spiroindolone Antimalarials Target the P-Type ATPase Pma1 in <i>S. cerevisiae</i> and PfAtp4 in <i>P. falciparum</i> .....	44
Abstract.....	44
Introduction .....	45
Results .....	47
Discussion.....	59
4. The Anti-trypanosomal MMV001239 Inhibits Ergosterol Biosynthesis by Inhibiting <i>T. cruzi</i> Cyp51p .....	63
Abstract.....	63
Introduction .....	64
Results .....	73
Identifying a Compound for Proof-of-Concept .....	73
Identification and Characterization of MMV001239-resistant <i>S. cerevisiae</i> .....	76
MMV001239 Binds to <i>T. cruzi</i> CYP51 .....	81
Characterizing the Binding Site through X-ray Crystallography and Molecular Docking .....	90
Discussion.....	95
5. The Phenyl-amino-methyl-quinolinols (PAMQs) Inhibit Cyclic AMP Signaling in <i>P. falciparum</i> .....	99

Abstract.....	99
Introduction .....	100
Results .....	105
Directed Evolution.....	105
Whole Genome Sequencing Results .....	106
PAMQs Inhibit Intracellular cAMP Signalling .....	112
Evolutionary Sequence Analysis and Structural Biology .....	118
6. Methods.....	125
7. Conclusions and Future Studies .....	143
8. References .....	148

## LIST OF FIGURES

Figure 1.1 An improved method for directed evolution target identification .....	16
Figure 1.2 Inhibition of ABC <sub>16</sub> -Monster vs. wild-type yeast .....	30
Figure 1.3 An Improved method for directed evolution target identification .....	32
Figure 2.1 Cycloheximide resistance-conferring mutations found in Rpl28p .....	38
Figure 2.2 Etoposide-conferring mutations found in Top2p .....	42
Figure 3.1 KAE609, an experimental spiroindolone antimalarial known to clear malaria parasites from infected adults .....	46
Figure 3.2 Inhibition of wild-type and ABC <sub>16</sub> -Monster <i>S. cerevisiae</i> , as judged by the optical density at 600 nm after 18 hours of incubation .....	48
Figure 3.3 KAE609 directed evolution .....	49
Figure 3.4 Alignment of ScPma1p to PfAtp4p .....	50
Figure 3.5 Validation of KAE609 activity on Pma1p .....	52
Figure 3.6 Illustrations of the ScPMA1 homology model and KAE609 docked pose .....	56
Figure 3.7 KAE609 halogenation .....	58
Figure 4.1 Target identification via directed evolution .....	72
Figure 4.2 MMV Malaria Box drug sensitivity screens .....	74
Figure 4.3 The chemical structure of MMV001239 .....	74
Figure 4.4 MMV001239 activity on <i>T. cruzi</i> intracellular amastigote .....	75
Figure 4.5 MMV001239 dose-response curve .....	76
Figure 4.6 IC <sub>50</sub> changes in MMV001239 resistant strains .....	78

Figure 4.7 The ergosterol biosynthetic pathway .....	82
Figure 4.8 14-alpha demethylase sequence alignment.....	86
Figure 4.9 TcCyp51p absorbance difference spectra .....	88
Figure 4.10 Determining binding potency with absorbance difference spectra .....	89
Figure 4.11 Image of crystals used for X-RAY crystallography .....	90
Figure 4.12 Electron density diagram for MMV001239 in TcCyp51 .....	91
Figure 4.13 Determining the binding pose of MMV001239 using computational docking .....	92
Figure 4.14 Comparing the binding poses of a range of Cyp51p/Erg11p inhibitors .....	94
Figure 5.1 Structure of PAMQs .....	103
Figure 5.2 The cyclic AMP second messenger system in <i>S. cerevisiae</i> .....	110
Figure 5.3 The effect of PAMQs on cAMP levels in <i>S. cerevisiae</i> .....	113
Figure 5.4 The effect of PAMQs on cAMP levels in <i>P. falciparum</i> .....	115
Figure 5.5 The effect of PAMQs on cAMP levels in human red blood cells .....	117
Figure 5.6 The effect of PAMQs on cAMP levels in human liver cells .....	117
Figure 5.7 Sequence alignments of class III adenylyl cyclases .....	119
Figure 5.8 Phylogenetic tree of class III adenylyl cyclases .....	120
Figure 5.9 Structural biology of adenylyl cyclase monomer .....	121
Figure 5.10 Structural biology of adenylyl cyclase homodimer .....	122



## LIST OF TABLES

Table 1.1 Antimalarial Drug Resistance .....	13
Table 1.2 Comparing methods of drug target identification (part 1) .....	24
Table 1.3 Comparing methods of drug target identification (part 2) .....	25
Table 1.4 The advantages and disadvantages of performing directed evolution experiments in three selected organisms .....	28
Table 1.5 Simple rules for prioritizing potential targets .....	34
Table 2.1 Summary of directed evolution and sequencing .....	40
Table 2.2 Summary of genes with missense mutations for etoposide selections.....	41
Table 4.1 The ABC16-monster strain as a model for <i>T. cruzi</i> . .....	68
Table 4.2 Summary of directed evolution protocol .....	77
Table 4.3 Summary of directed evolution sequencing results .....	79
Table 4.4 Sequencing statistics.....	80
Table 4.5 Chemical structures of varied Cyp51p/Erg11p inhibitors .....	84
Table 5.1 Summary of PAMQ activity against varied pathogens .....	104
Table 5.2 A summary PAMQ selections.....	106
Table 5.3 Summary of directed evolution sequencing results .....	107
Table 5.4 Summary of genes with amino acid changes in drug resistant clones .....	107
Table 5.5 Sequencing statistics.....	108
Table 6.1 X-ray data collection and refinement statistics .....	139

## ACKNOWLEDGEMENTS

I would first like to thank my mentor Professor Elizabeth Winzeler for the opportunity to work in her lab, along the other members of the Winzeler Lab. In addition, I would like to acknowledge Professor Yo Suzuki and Professor Larisa Podust for their contributions to this work. I also thank Paula Maguina, who keeps the lab running and is always patient with my questions.

I would like to thank my senior colleagues whose wisdom and companionship have made my PhD possible. Specifically, Dr. Sabine Otilie, Dr. Jacob Durant and Dr. Greg LaMonte have been invaluable to me. My gratitude also goes to my technician and friend Eddie Vigil, with whom I have enjoyed working for the past few years, as well as the undergraduate students who I had the pleasure to work with including Rebecca Stanhope, Maximo Prescott, Jennifer Yang, Jake Schenken, and Felicia Gunawan.

My appreciation also goes to Dr. Omar Vandal and his colleagues at the Bill and Melinda Gates Foundation for their financial support and intellectual guidance over the past four years. I would also like to thank Dr. Paul Insel and Mary Alice Kiisel from the UC San Diego Medical Scientist Training Program and Prof. Deborah Spector, Leanne Nordeman and Gina Butcher from the Biomedical Sciences PhD Program.

Finally, I thank my family, without whom none of this would be possible. My parents and grandparents immigrated to this country in search of freedom and opportunity. My accomplishments would not have been possible if not for

their sacrifices. I gratitude also goes to my two supportive sisters for always being there.

Chapter 3, in part, has been submitted for publication as it may appear in the journal *Scientific Reports*, 2016, Goldgof GM, Jacob D. Durrant, Sabine Otilie, Edgar Vigil, Kenneth E. Allen, Felicia Gunawan, Maxim Kostylev, Kiersten A. Henderson, Jennifer Yang, Jake Schenken, Gregory M. LaMonte, Micah J. Manary, Ayako Murao, Marie Nachon, Rebecca Stanhope, Maximo Prescott, Case W. McNamara, Carolyn W. Slayman, Rommie E. Amaro, Yo Suzuki and Elizabeth A. Winzeler under the title, “Comparative chemical genomics reveal the spiroindolone antimalarial KAE609 (Cipargamin) is a P- type ATPase inhibitor.” The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, has been submitted for publication and may appear in the journal *ACS Chemical Biology*, 2016, Goldgof GM, Sabine Otilie, Claudia Calvet, Gareth K Jennings, Gregory M. LaMonte, Jake Schenken, Edgar Vigil, Priyanka Kumar, Laura-Isobel McCall, Eduardo Soares Constantino Lopes, Felicia Gunawan, Jennifer Yang, Yo Suzuki, Jair Lage De Siqueira Neto, James H. McKerrow, Rommie E. Amaro, Larissa M. Podust, Jacob D. Durrant, and Elizabeth Winzeler. The dissertation author was the primary investigator and author of this paper.

Chapter 5 is being prepared for submission for publication. Co-authors include Sabine Otilie, Edgar Vigil, Nelissa Figuero, Felicia Gunawan, Jennifer Yang, Jake Schenken, Gregory M. LaMonte, Yo Suzuki and Elizabeth A.

Winzeler. The dissertation author was the primary investigator and author of this paper.

## VITA

- 2009 Bachelor of Science, Computer Science and Biological Sciences,  
Stanford University
- 2009 Masters of Science, Bioengineering, Stanford University
- 2017 Doctor of Philosophy, Biomedical Sciences, University of California,  
San Diego
- 2018 (Expected) Medical Doctor, University of California, San Diego

## PUBLICATIONS

1. **Goldgof GM**, Otilie S, et al. Rapid Chagas Disease Drug Target Discovery Using Directed Evolution in Drug-Sensitive Yeast. *ACS Chemical Biology*. 2017 Feb 17;12(2):422-434.
2. Van Voorhis WC,... **Goldgof GM**, ...Willis PA. Open Source Drug Discovery with the Malaria Box Compound Collection for Neglected Diseases and Beyond. *PLoS Pathogens*. 2016 Jul 28; 12(7):e1005763.
3. Corey VC, ...**Goldgof GM**, ...Winzeler EA. A broad analysis of resistance development in the malaria parasite. *Nature Communication*. 2016 Jun 15; 7:11901.
4. **Goldgof GM**, Durrant JD, Otilie S, Vigil E, et al. Comparative chemical genomics reveal the spiroindolone antimalarial KAE609 (Cipargamin) is a P-type ATPase inhibitor. *Scientific Reports*. 2016 Jun 13;6:27806.
5. Suzuki Y, ...**Goldgof GM**, ...Venter JC. Bacterial genome reduction using the progressive clustering of deletions via yeast sexual cycling. *Genome Res*. 2015 Mar;25(3):435-44.
6. Karas, BJ, Jablanovic, J., Sun, L, Ma, L, **Goldgof, GM.**, Stam, J, et al. (2013). Direct transfer of whole genomes from bacteria to yeast. *Nature Methods*. 2013 doi:10.1038/nmeth.2433

7. Bertin, M. J., Demirkiran, O., Navarro, G., Moss, N. A., Lee, J., **Goldgof, G. M.**, et al. (2015). Kalkipyronone B, a marine cyanobacterial  $\gamma$ -pyrone possessing cytotoxic and anti-fungal activities. *Phytochemistry*.

#### FIELDS OF STUDY

Major Field: Biomedical Sciences, Chemical Biology, Drug Discovery

Studies in Biomedical Sciences  
Professor Elizabeth Winzeler

ABSTRACT OF THE DISSERTATION

Drug Target Discovery Using Designer Drug Sensitive Yeast

by

Gregory Mark Goldgof

Doctor of Philosophy in Biomedical Sciences

University of California, San Diego, 2017

Professor Elizabeth Winzeler, Chair

Determining the protein target(s) and mechanism(s) of drug candidates found in phenotypic screens is critical to subsequent structure-activity-based development and optimization, but existing methods for target identification are limited. Here we present a method that applies directed evolution to a genetically

engineered, drug sensitive *Saccharomyces cerevisiae* strain. Whole genome sequencing of yeast clones that have evolved drug resistance, in concert with *in vitro* cell free assays and computer modeling, can be a useful tool for target identification and binding site characterization.

To demonstrate the ease and utility of this method, we applied it to the identification of the molecular target and binding site of a range of cytotoxic molecular compounds with activity against eukaryotic pathogens and human cancers. These studies include known drug target combinations, as well as application to experimental compounds with unknown drug targets. As proof of concept, the method correctly identified the precise binding pocket of the protein synthesis inhibitor, cycloheximide, as the ribosomal protein Rpl28. We also correctly identified topoisomerase II inhibitor as the target of the human cancer chemotherapeutic, etoposide.

We next used the method to identify novel drug target combinations, which were then validated using a combination of genetic, biochemical, structural and chemical structure activity relationships (SAR)-based assays. We identified a p-type ATPase, ScPma1, as the target of the spiroindolone antimalarials, of which KAE609 is currently in stage 2b clinical trials. We determined that the pre-clinical phenyl-amino-methyl-quinolinols (PAMQ) antimalarials inhibit the cyclic AMP signaling pathway, a mechanism of action that is different from existing commercial antimicrobials. We also demonstrated that MMV001239, a compound with antitrypanosomal activity, targets ScErg11, the yeast homolog of the T, cruzi Cyp51p, and a protein crucial for ergosterol biosynthesis. Taken



together, our approach expands on the number of tools available for analyzing compound-target interactions and can be applied to studies of other eukaryotic antimicrobials and chemotherapeutics.

# 1. Introduction

## History of Antimicrobial Drug Discovery

Humans have been using plants with antimicrobial activity to treat infections for thousands of years, with written records dating back to ancient Mesopotamia (Aboelsoud, 2010). For example, Chinese healers have used qinghao, also known as sweet wormwood or *Artemisia annua*, for centuries, with written records dating back to the 284 AD (Tu, 2011). Today, the antimalarial, artemisinin, is derived from this plant (Tu, 2011). In pre-Colombian Peru, the Quechua drank the predecessor of modern-day tonic water by combining the bark of the cinchona tree with sweetened water (Achan et al., 2011). This antimalarial drink contains the antimalarial quinine, which became the mainstay of European treatment for malaria from the 17<sup>th</sup> century until after World War II (Achan et al., 2011). Controversy still surrounds the precise mechanism of action of both of these antimalarials.

Modern, industrially produced, antimicrobials date back to the discovery of penicillin in London in 1928 by Scottish scientist Alexander Flemming at St. Mary's Hospital (Flemming, 1929) (ACS, 1999). Flemming discovered that a petri dish containing colonies of the bacteria *Staphylococcus*, contained a halo of area where no bacteria grew around a growing section of mold that had contaminated his experiment (Flemming, 1929) (ACS, 1999). He also discovered that the mold secreted a chemical that could kill a wide diversity of harmful bacteria, including *Streptococcus*, *Meningococcus* and *Diphtheria* (Flemming, 1929) (ACS, 1999).

However, the difficulty of isolating the active chemical, penicillin, from living fungus, meant that the drug would not become commercially available until 1943 (ACS, 1999).

In the meantime, during the 1930s, German scientist Gerhard Domagk and colleagues were researching the antimicrobial activity of dyes at IG Farben (now Bayer AG), an industrial chemical company. They discovered Prontosil, the first commercially available antibiotic drug (Rubin, 2007). It was soon discovered that Prontosil was a prodrug that required conversion in the body to sulfanilamide that had been used in the dye making industry since 1906 and that was no longer under patent (Otten, 1986). As a result, hundreds of manufacturers began to produce these sulfa antibiotics with widespread medical use. Alternatives would not become available until the introduction of penicillin and streptomycin in 1943, as well as chloramphenicol and tetracycline, later that decade (Palumbi, 2001).

The work by Flemming and others led to a new strategy for antimicrobial discovery: to purify natural products from microbe-derived cytotoxic extracts. This approach created a boom of discovery that began to produce new treatments for patients starting in the 1940s (Rubin, 2007). Other microbe-derived, natural product antibiotics include streptomycin in 1943, the first drug to cure tuberculosis, chloramphenicol (1947), tetracycline (1948), erythromycin (1952), vancomycin (1956), cephalosporins (1960s), and fluoroquinolones (1960s) (Rubin, 2007). They also include the antifungals nystatin (1954) and amphotericin B (1955) (Ng et al., 2003) (Dutcher, 1968) (Ostrosky-Zeichner et al., 2010).

### History of Antimicrobial Mechanism of Action Studies

While we have the first half of the 20<sup>th</sup> century to thank for many of the antimicrobial drug classes that we currently use today, determination of the mechanism of action for these compounds did not occur until decades later. For some antimicrobials currently in use, such as the antimalarials primaquine and artemisinin, scientists remain unsure about how they inhibit the growth of the parasite. However, starting in the 1950s and 1960s, scientists began to unravel the mechanism of action of some important antimicrobial drug classes.

During this time, the major approach was to determine which microbial functions the compound inhibited, and then try figure out where in the pathway the compound was acting. As early as the 1940s, scientists had reported that treatment with penicillin caused bacteria to display unusual cell shapes suggestive of a defect in the cell wall (Tipper and Strominger, 1965). Other scientists later observed that *E. coli* treated with penicillin in hypertonic sucrose baths produced spheroblasts. Spheroblasts are bacteria lacking cell wall that are held together only by their plasma membrane (Weibull, 1953).

The next major piece of evidence that advanced the hypothesis that penicillin kills bacteria by disrupting the cell wall was published in the 1950s. James Park at Walter Reed and Jack Strominger at Washington University of St. Louis demonstrated that penicillin inhibited cell wall synthesis was the discovery that *S. aureus* treated with the penicillin accumulates a uridine nucleotide that is structurally similar to the cell wall (Park and Strominger, 1957) (Strominger,

1957). By the 1960s, the steps in cell wall synthesis were mapped out, and the target of penicillin was discovered to be a transpeptidase that cross-links the peptidoglycan chains of the cell wall (Tipper and Strominger, 1965). This protein was creatively named, the penicillin-binding protein (PBP).

A similar approach was used during the 1970s to elucidate the mechanism of action of the azole antifungals, one of the most important classes of antifungal treatment to this day. The first report of an azole compound with activity against fungi was benzimidazole in 1943 (Ostrosky-Zeichner et al., 2010). However, azole class compounds did not enter the market until the 1980s, when the number of fungal infections skyrocketed as a result of the AIDS crisis. In fact, the fungus *Pneumocystis pneumonia* (PCP) accounted for 70% of the first 400 recorded AIDS deaths (Armstrong-James et al., 2014).

In the early 1970s, it was observed that treatment with azole antifungals could cause yeast to leak their intracellular contents, suggesting that they interfered with the plasma membrane in some way (Iwata et al., 1973). A few years later, another group discovered that yeast cannot grow in the presence of azole antifungals without ergosterol (Nes et al., 1978). Additionally, azole treatment induced the cellular accumulation of intermediates in the ergosterol biosynthetic pathway (Nes et al., 1978). One of these intermediates that accumulated was lanosterol, which led many groups to scientists believe that the 14 $\alpha$ -demethylase, the enzyme that acts on lanosterol, is the target (van den Bossche et al., 1978) (Sud and Feingold, 1981).

Another classic approach to drug target identification was to look at differences between drug resistant and drug sensitive strains. Resistance-based approaches were used to identify the targets of streptomycin, the first aminoglycoside antibiotic, and nalidixic acid, the precursor to the fluoroquinolone antibiotics. Streptomycin was isolated from the actinobacterium *Streptomyces griseus* in 1943 at Rutgers University under funding from Merck in the laboratory of Professor Selman Abraham Waksman. This laboratory was also responsible for the discovery of several other important antibiotics including the aminoglycoside, neomycin (1949) and the polypeptide, dactinomycin (1940). Later clinical trials demonstrated that streptomycin was the first antibiotic that could cure tuberculosis (Comroe, 1978), securing its place in medical history.

Like many other early antibiotics, its mechanism of action remained unknown until the 1960s. In the early 1960s, several groups helped to discover that streptomycin inhibits protein synthesis, both in living cells and in cell free *in vitro* cultures (Cox et al., 1964). Furthermore, in cell free *in vitro* cultures, ribosomes isolated from streptomycin resistant bacteria were streptomycin resistant, whereas ribosomes isolated from streptomycin susceptible bacteria were streptomycin sensitive (Cox et al., 1964). Two articles in *PNAS* in 1964, supported narrowing down the target to 30s subunit of the 70s prokaryotic ribosome (Cox et al., 1964, Davies et al., 1964). In the late 1960s the mechanism of action was further narrowed down to the initiation of the initiation step of protein synthesis, rather than elongation or termination (Luzzato et al., 1968).

Naladixic acid was introduced in 1962 to treat urinary tract infections, and is the first antibiotic in the fluoroquinolone class. Interestingly, this compound was originally discovered as a contaminant in the production of the antimalarial chloroquine. By 1969, the target of this compound was identified as DNA gyrase A, using genetic mapping studies on drug resistant *E. coli* (Hane and Wood, 1969). In the 1980, using this same method, DNA gyrase B was identified as an additional target gene in *E. coli* (Nakamura et al., 1989). Bacterial DNA gyrases belong to a class of enzymes called type II topoisomerases that regulate the winding and unwinding of DNA. In 1990, an additional type II topoisomerases, named DNA topoisomerase IV, was discovered (Kato et al., 1990), and soon after, determined to be an additional target of the fluoroquinolones (Ferrero et al., 1994).

In summary, during the golden age of antimicrobial discovery, the methods used to identify the mechanism of action of these compounds usually took well over a decade to come to an answer, even under intense scrutiny by many competing and collaborating research laboratories. As a result, mechanism of action studies post-dated the actual drug development process. One common approach was to determine the differences between resistant and non-resistant strains, using methods such as genetic mapping. The other main approach was to determine which cellular processes were disrupted, and then to narrow down the pathway and the specific enzyme.

A principal weakness of the pathway-based approach was that antimicrobials with specific targets often have affects on many cellular processes.

For example, by inhibiting protein translation, streptomycin causes dysfunction throughout many cellular pathways, since they all rely on the synthesis of proteins for their functioning. As a result, research in the late 50s went in many directions, with some noting that streptomycin treated cells exhibited permeability defects similar to a detergent, suggesting a membrane biogenesis defect. Meanwhile others noted that increased excretion of nucleotide precursors suggested a defect in DNA synthesis (Anand and Davis, 1960). This controversy was only settled by examining the differences between resistant and sensitive bacteria. I hope that this dissertation demonstrates that this fundamental approach to studying how antimicrobial drugs function is as important today as it was over half a century ago, especially when paired with the myriad pathway-based assays that have been developed.

### **Modern Antimicrobial Drug Discovery**

Modern drug discovery is characterized by the screening of large libraries of compounds, typically in the millions (Payne et al., 2007). Over the past few decades, advances in cell and molecular biology, liquid-handling, automation, image analysis, chemical library, and other areas of science and technology, have driven improvements in the ability to perform drug screens, in terms of miniaturization, reduced cost, and higher precision (Flannery et al., 2013). For example, the ability to take photos of every well in a high-density microtiter plate, combined with automated image analysis, has led to an explosion in the types of phenotypes that can be rapidly assayed. This approach has been named 'high-



content screening,' because of the numerous variables that can be rapidly analyzed using automated photography and image analysis.

Typically, high throughput drug screens are separated into two main categories: phenotypic vs. target-based. In typical antimicrobial phenotypic screens, a library of compounds is tested for the ability to inhibit the growth or kill the pathogenic microbe. The strength of this approach is that, from the onset, any compounds discovered in this way are known to be able to enter the cell and evade degradation or deactivation once within the organism. The major weakness of this approach is that it does not provide information about the mechanism of action (MoA) of the compound (Payne et al., 2007) (Swinney and Anthony, 2011). Most first in class antimicrobials were developed through low-throughput phenotypic assays, such as the observation by Ancient Chinese physicians that sweet wormwood can reduce a child's fever, or the observation by Alexander Fleming that mold can produce chemicals that kill bacteria.

Target-based screens, on the other hand, involve screening the compound library against recombinant proteins that are believed to be essential for the microbe's survival. Using this approach, drug developers can get an excellent understanding of the structure-activity-relationships (SARs) between the chemical class and its target. The major disadvantage of this approach is that, *in cellulo*, the compounds may not be able to enter the cell and get to the desired target (Payne et al., 2007) (Swinney and Anthony, 2011).

Knowledge of the structure of the target of the compound, along with SARs, is the key tool that medicinal chemists use when a compound lead is

optimized (Payne et al., 2007). Typical compound optimization programs at pharmaceutical companies will make hundreds of derivatives of the lead compounds in order to improve pharmacodynamics (how the compound binds to its molecular target) and pharmacokinetics (how the compound is able to enter the human body and get to the desired tissue). However, knowledge of the target-binding interaction, while helpful, is not necessary for compound development, since SAR can be inferred from phenotypic assays (Payne et al., 2007). (Swinney and Anthony, 2011).

Since target-based screening requires that scientists already know what protein in the cell we want to target, it has a hard time at discovering drugs that work through novel mechanisms of action. Between 1999 and 2008, only 11 new molecular entities (NME) were approved by the FDA to treat infectious diseases. Of these, seven were discovered through phenotypic screens, three were discovered through target-based screens, and one was a biologic drug that modulated the immune system (Swinney and Anthony, 2011).

One thing to keep in mind is that most phenotypic screens are looking for compounds with new mechanisms of action, so one can counter-screen potential lead molecules against target-based assays or pathogens with known drug resistance mechanisms. On the other hand, the purpose of many target-based screens is to discover new chemical structures or modifications on an existing drug that improve some aspect of a clinically approved drug, such as widening the range of organisms that it can kill. For this reason, a majority of drugs from target-based screens that come to market, not because they have a new

mechanism of action, but because they exhibit improved or different features from the original drug (Swinney and Anthony, 2011).

### **Malaria: The Disease**

Malaria describes a group of mosquito-transmitted human diseases caused by protists of the genus *Plasmodium*. The World Health Organization estimates that 3.2 billion people currently live at risk for malaria. In 2015, there were over 200 million estimated cases, causing nearly half a million deaths (WHO, 2015). Although malaria is present in most tropical and sub-tropical regions, in 2015, 88% of cases and 90% of deaths were in sub-Saharan Africa.

Malaria in humans is caused by five species of the genus *Plasmodium*, with *P. falciparum* and *P. vivax* responsible for majority of cases. *P. falciparum*, which is responsible for the majority of deaths, is most prevalent in Africa, where the preponderance of the deaths occur in children under the age of five years. Another particularly vulnerable group is pregnant women. There are around 400 different species of *Anopheles* mosquito species believed to transmit the parasite, although around thirty are responsible for most disease (WHO, 2015).

The symptoms of malaria begin 7-15 days after infection by the parasite-carrying mosquito. The first symptoms include fever, chills, headache and vomiting. These symptoms typically re-occur every 2-3 days, as the parasite undergoes its asexual blood stage reproductive cycle. These symptoms can be mild and therefore hard to distinguish from other infections, making diagnosis difficult, especially in resource-limited settings. Diagnosis requires microscopic

examination of the patient's blood for infected red blood cells. If not treated within 24 hours, *P. falciparum* malaria can progress to severe illness, especially in those who have not developed partial immunity by previously having the disease, such as young children and travellers (Braunwald, 2001).

The symptoms of severe malaria include anemia, respiratory distress, multi-organ involvement and cerebral malaria (Braunwald, 2001). Infected red blood cells adhere to the endothelium of small blood vessels, leading to microinfarctions and microhemorrhages throughout the body (Braunwald, 2001). The manifestations of these infarctions in the central nervous system are collectively called cerebral malaria. These microinfarctions and microhemorrhages lead to swelling in the brain that may cause seizures, coma or death (Seydel et al., 2015). The respiratory distress seen in severe anemia may be the result of a combination of metabolic acidosis, severe anemia, and damage to the lungs. Other complications include renal failure, hepatomegaly, splenomegaly and retinal damage. In pregnant women, severe malaria can cause acute respiratory distress syndrome occurring in up to 29% of patients and is associated with stillbirths, infant mortality, and low birth weight (Braunwald, 2001).

A range of strategies is currently employed to reduce malaria prevalence. Since the mosquitos that transmit the disease require standing bodies of water, economic development indirectly reduces infection rates as marshes and wetlands are paved over, a possible benefit when you 'pave paradise to put up a parking lot,' to quote Joni Mitchell. In terms of active human interventions, the

main strategies include vector control through the use of insecticides and mosquito nets in the home, chemoprevention using antimalarials and rapid treatment to prevent death and reduce transmission. Chemoprevention in pregnant women has been a focus of recent public health efforts in sub-Saharan Africa, since infection can cause maternal anemia, placental infection and low birth weight. As a result of improved development and public health strategies, new cases of malaria have fallen by 37% globally since the year 2000, with the death rate declining among the most vulnerable groups by 60% (WHO, 2015).

Drug resistance in malaria has emerged as a top challenge for the treatment and prevention of malaria, as well as eradication efforts. Resistance has been documented against all anti-malarial drugs dating back to chloroquine-resistance a half a century ago in both South America and South East Asia (WHO, 2014) (WHO, 2015). In part, due to the massive deployment of antimalarials globally as well as better drug resistance monitoring, drug resistance is emerging more rapidly in recent years. Table 1.1 summarizes the timing of drug resistance for major classes of anti-malarials.

**Table 1.1 Antimalarial Drug Resistance.** A summary of the year of introduction and the year of first reports of drug resistance for antimalarial drugs (Bjorkman and Phillips-Howard, 1990, Wernsdorfer and Payne, 1991, Wongsrichanalai et al., 2001, Engel and Straus, 2002, McClure and Day, 2014) (Carrara et al., 2013) (Leang et al., 2013b) (Leang et al., 2013a) (Ashley et al., 2014) (WHO, 2014).

<b>Antimalarial</b>	<b>Introduced</b>	<b>Reported Drug Resistance</b>	<b>Difference in Years</b>
Quinine	1632	1910	278
Chloroquine	1945	1957	12
Proguanil	1948	1949	1
Sulfadoxine-pyrimethamine	1967	1967	0
Mefloquine	1977	1982	5
Halofantrine	1988	1992	4
Atovaquone	1996	1996	0
Atovaquone-Proguanil	1997	2001	4
Artemisinin	1998	2008	10

### ***Plasmodium* Life Cycle**

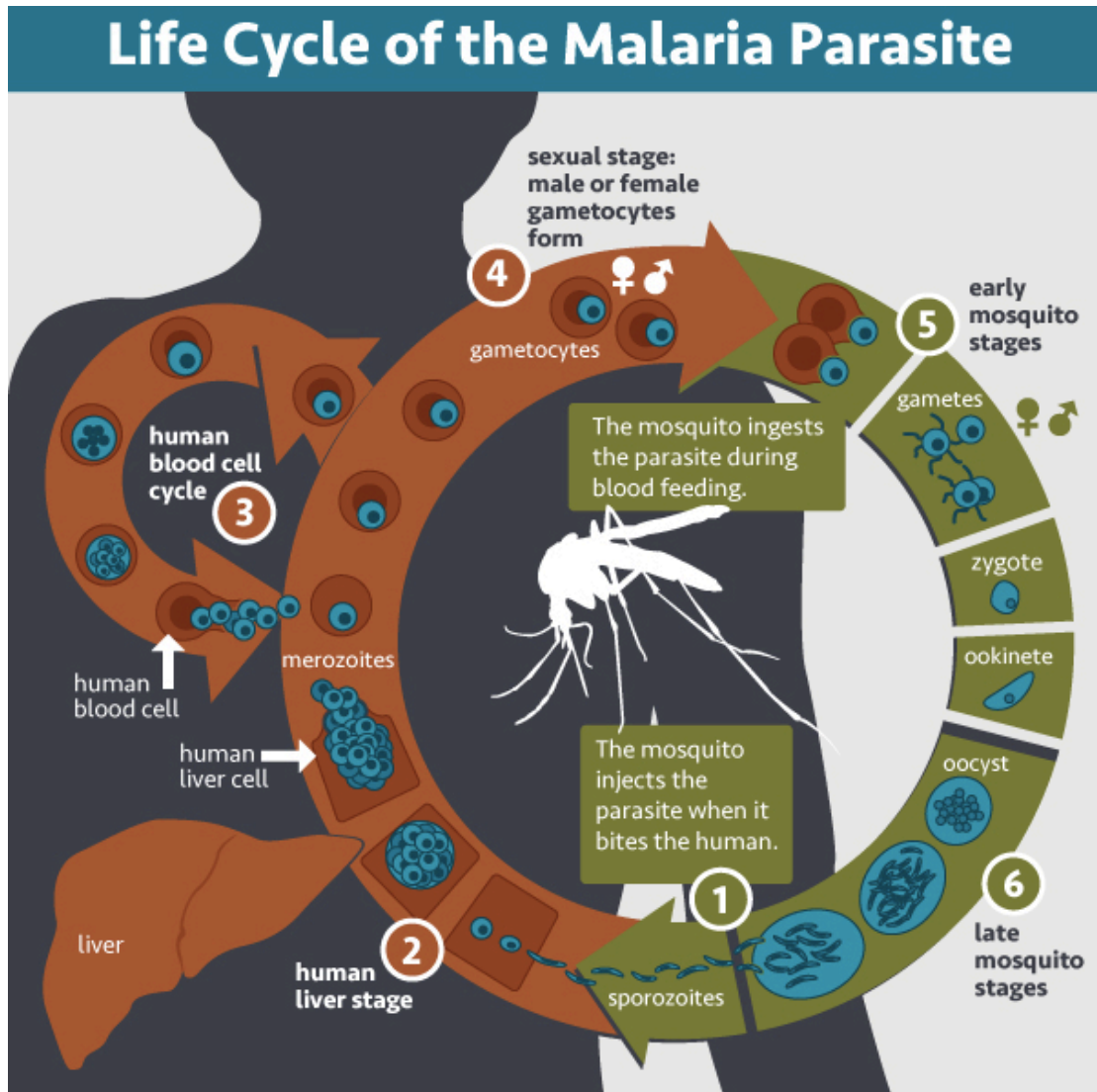
To understand how antimalarials work, it is first important to understand the life cycle of the malaria parasites (Fig 1.1). Here we will focus on the life cycle of *P. falciparum* and *P. vivax*. When an infected female *Anopheles* mosquito takes a blood meal from a human, *Plasmodium* sporozoites (a word roughly meaning 'animal seed' in Greek) are injected from the mosquito's salivary glands, into the blood stream of the human. These sporozoites then migrate to the liver and infect or seed individual hepatocytes. Within a hepatocyte, a single sporozoite will multiply asexually into thousands of merozoites (meaning 'animal piece'). This stage of the life cycle is commonly called the liver stage.

After 10-14 days, the merozoites will burst the hepatocyte and enter the blood stage and infect human erythrocytes (red blood cells). It is at this moment that the symptoms of malaria will begin and the human host will become both symptomatic and infectious. Asexual stage parasites may follow two paths. Most commonly, they divide into 8-32 merozoites, burst the red blood cell and continue to reproduce asexually. Most antimalarial drugs target this stage of the disease. This stage of the disease is also responsible for malaria's morbidity and mortality. Alternatively, they may enter the sexual development pathway and become gametocytes. Gametocytes (meaning 'marriage cells') are the only sexually reproducing cell types within the malaria life cycle and come in both male and female varieties.

Through an unknown mechanism, parasites in some infected red blood cells, instead of reproducing asexually, will differentiate into a single male or

female gametocyte. This process is often referred to as sexual stage of the life cycle. These gametocytes may then be taken up from the blood by a female *Anopheles* mosquito during a blood meal. The male and female gametocytes will mate in the midgut of the mosquito, forming a zygote ('yoked animals') and then a motile ookinete ('moving egg'). The ookinete will dig through the wall of the mosquito's midgut and then attach itself to the outer lining to form an oocyst ('sac of eggs'). Over the next 7-10 days, the oocyst will divide into thousands of infective sporozoites. When the oocyst bursts, the sporozoites will then migrate to the mosquito's salivary glands, and a new cycle of parasite transmission from mosquito to human can begin.





**Figure 1.1** The life cycle of the malaria parasite. This figure describes how the parasite propagates itself, both within the mosquito and human portions of the life cycle. Figure made publicly available courtesy of the NIAID (NIAID, 2016).

In malaria caused *P. falciparum*, once the blood stage begins, there are not dormant infected hepatocytes that can re-infect the individual. In contrast, in infections caused by *P. vivax*, dormant forms called hypnozoites ('sleeping animals') can lay dormant in the liver for many years. Through an unknown mechanism, they may awaken and reinfect the host's blood stream at any time. This presents a particular challenge to efforts to eliminate malaria, because most antimalarial drugs are unable to kill these sleeping hypnozoites.

Only the 8-aminoquinolones, such as primaquine, are known to kill hypnozoites (Braunwald, 2001). These drugs unfortunately can cause life-threatening side effects in patients with a genetic condition called G6PD deficiency (Braunwald, 2001). Over 400 million people worldwide have G6PD deficiency, especially those whose ethnic background comes from malaria endemic regions, leading to the hypothesis that this trait has historically provided some protection against death from malaria, undergoing selection in these populations (Cappellini and Fiorelli, 2008). This unfortunate fact results makes primaquine a poor candidate for eradication programs, which would require treating large populations of people not currently experiencing any life-threatening illness.

### **Current Antimalarial Drug Discovery**

Since resistance has been reported against all known antimalarials, a renewed effort to identify next-generation antimalarials has been underway for the past decade (Flannery et al., 2013). In order to be useful against multi-drug

resistant parasites, these new compounds need to have novel mechanisms of action. In addition, renewed interest in global malaria eradication will require additional antimalarials with activity against multiple stages of the parasite life cycle (Tanner and de Savigny, 2008).

As a result, high-throughput phenotypic screens have been performed in partnerships between academia and the pharmaceutical industry against a range of human stages of the *Plasmodium* life cycle, including the asexual blood stage, the liver stage, and the gametocyte/sexual stage. In 2008 and 2010, the results of three large-scale screens for novel inhibitors of the asexual blood stage were reported, revealing thousands of new compounds with antimalarial activity. These sets included the Novartis-GNF Malaria Box, which contains 6,000 hits from 1.7 million screened compounds (Plouffe et al., 2008), the GlaxoSmithKline TCAMS set, which contains 13,000 hits from 2 million screened compounds (Gamo et al., 2010), and the St. Jude's Children's hospital set of 1,300 hits from 300,000 screened chemicals (Guiguemde et al., 2010).

As a result, by 2010, scientists had reported thousands of compounds with antimalarial activity, but unknown function. To help identify their mechanism of action, the Medicines for Malaria Venture (MMV) compiled 400 of the most interesting chemotypes and assembled them into an open-access library, called the MMV Malaria Box, that was sent to hundreds of research groups around the world (Spangenberg et al., 2013). Chapters 4 and 5 discuss the molecular mechanism of some of the compounds from this chemical library.

In addition, the most promising hits from each of the original screens have been further developed and taken into clinical trials. The first of these compounds were the spiroindolones, which were discovered as part of the Novartis screen (Rottmann et al., 2010a) (Yeung et al., 2010a). The original hit from the screen was optimized for potency and bioavailability, leading to the molecule KAE609, also known as cipargamin (formerly named NITD609).

In 2014, a successful Phase II trial in Thailand was reported in the *New England Journal of Medicine*. In this trial, investigators found that KAE609 has a median parasite clearance half-life of 0.9 hours for *P. falciparum* and 0.95 hours for *P-vivax*, making it faster acting than any existing antimalarial. Speed of parasite clearance from the blood is correlated with the rate at which symptoms are reduced, meaning that faster acting drugs are likely to cure patients more quickly and require shorter treatment duration. In addition, since KAE609 shows in vitro activity against gametocytes, it is hoped that once approved this compound will reduce transmission of the disease. Since there were no major side effects reported, more trials are under way (White et al., 2014).

If approved, this compound would be the first antimalarial with a new mechanism of action since artemisinin, around twenty years prior. In addition, because of its rapid parasite clearance rate, additional Phase II trials are under way to assess KAE609 as a single-dose therapy, as well as to determine if it has the ability to inhibit transmission after treatment through activity against sexual stage gametocytes (ClinicalTrials.gov, 2016a, c). The mechanism of action of KAE609 is discussed in Chapter 3.

Concurrently, phenotypic screening has continued against the liver and gametocyte stages of the *Plasmodium* life cycle. Compounds that are active against the liver stage should have prophylactic activity, because they kill the parasite before it enters the blood stage, which is responsible for the clinical manifestations of the disease. Compounds that are active against the gametocyte stage are often referred to as ‘transmission-blocking,’ because they kill the infectious forms of the parasite. Compounds that only kill asexual blood stage parasites, such as current first line artemisinin-based therapies, leave the infected person contagious for up to 55 days (Bousema et al., 2010).

In liver-based screens, a human hepatocyte cell line is infected with *P. yoelii*, a species of *Plasmodium* that infects mice. *P. yoelii* sporozoites are harvested manually using microdissection microscopes from the salivary glands of mosquitos that are fed infected blood. In 2011, 5,697 compounds with activity against asexual blood stage *Plasmodium* were tested for their ability to inhibit growth of the liver stage in this assay. Of these, 275 had  $IC_{50}s < 10\mu M$ , 229 had  $IC_{50}s < 1\mu M$  and 86 had  $IC_{50}s$  less than 100nM (Meister et al., 2011). This study led to the development the imidazolopiperazine (IP) KAF156, which is currently in a Phase II clinical trial (ClinicalTrials.gov, 2016b). KAF156 is known as a dual stage inhibitor, since it is believed to acts on both the liver and blood stages (Kuhlen et al., 2014).

More recently, efforts to identify compounds with transmission-blocking activity through phenotypic screens against *Plasmodium* gametocytes have been underway (Lucantoni et al., 2015) (Plouffe et al., 2016). These assays are

performed using *P. falciparum* in human red blood cells. Asexual parasite inhibitors are used to isolate infected by gametocytes. These cells are then imaged at each substage of gametocyte development, in order to determine the efficacy of the compound against the parasite at each of these substages.

Although previous studies focused on compounds with activity against blood stage asexual parasites, combined with liver stage or gametocytocidal activity, screens are currently underway to identify compounds that have activity against just the liver or gametocytes stages. This interest comes from the observation that the majority of parasite replication in humans occurs during the asexual blood stage. Therefore, compounds that act on the asexual blood stage are more likely to select for drug resistance than compounds that act on the other stages (Burrows et al., 2013). While liver stage only compounds would not be able to treat active disease, since the disease's clinical manifestation are the result of the asexual blood stage, liver stage only compounds would be superior for chemoprophylaxis if they enabled large populations to be protected without the risk of drug resistance. Similarly, gametocyte-specific compounds could be employed to block malaria transmission in a population without a large risk of drug resistance, due to the relatively small number of circulating gametocytes in each infected human.

### **Modern Approaches to Drug Target Identification**

Phenotypic screening is a powerful tool for discovering antiinfective and antiproliferative compounds (Wagner, 2016). Phenotypic cellular screens are

efficient because they effectively multiplex many critical targets, identifying cytotoxic inhibitors of DNA replication, protein synthesis, secretion, transcription, or other cellular processes via a single, miniaturized assay. Although not necessarily critical for compound development, discovering a specific protein target can inform subsequent drug development, enabling both structure-guided drug design and a better understanding of off-target activity (Payne et al., 2007).

Characterizing drug candidate mechanisms and protein targets is critical to structure-activity based development and optimization. However, phenotypic screens rarely provide this information, requiring follow-up studies. Here we present a novel genomics approach involving directed evolution in a genetically engineered, drug-sensitive *Saccharomyces cerevisiae* strain that greatly accelerates drug target identification.

Traditional methods have relied on assays of macromolecular synthesis (DNA, RNA, protein, cell wall and membrane components). Over time, these assays have been expanded to many pathogens and miniaturized to high-throughput microtiter format (Payne et al., 2007). In addition, genomic approaches have been employed, including gene over, under expression and resistant mutant analysis (Payne et al., 2004). One such genomic approach is the haploinsufficiency screen, which can produce an overwhelming amount of difficult-to-interpret data (e.g., at least 2,452 gene deletions confer resistance or sensitivity to cycloheximide, according to the *Saccharomyces* Genome Database (Cherry et al., 2012)). Proteomics approaches, which detect compound-binding proteins from a cell lysate, are difficult and expensive to perform, and can

also produce high false-positive rates (Burdine and Kodadek, 2004), as do computational approaches such as proteome-wide virtual screening (Durrant et al., 2010). Tables 1.2 and 1.3 compare and contrasts some modern approaches to drug target identification. It is important to note that all methods for target identification require validation through other approaches to prove causality.



**Table 1.2 Comparing methods of drug target identification (part 1).** This table compares and contrasts two methods of genome-wide drug target identification, with directed evolution, the primary method employed in this thesis. (Pierce et al., 2006, Pierce et al., 2007).

	Haploinsufficiency and Deletion Library Screens (HipHop)	Cell Lysate Binding Assays	Directed Evolution
Description	Compound is added to pooled or arrayed libraries of haploid knockouts (for non-essential genes) and single allele diploid knockouts for essential genes.	Chemical is labeled with a tag (such as biotin). Chemical is exposed to cell lysate. Proteins that bind chemical pulled out using tag and identified using mass spectrometry.	Cells are grown under drug pressure. Once resistance is observed, a resistant clone is sequenced.
Resolution	This method can provide simultaneous data on how gene dosage affects phenotype for all genes in the genome. However, it cannot provide the protein structural information that comes from single nucleotide mutation that affect individual amino acids.	Single gene. Structural information may be determined through additional assays of the chemical-protein complex.	Single nucleotide resolution that allows you to see which parts of the protein are interacting with the compound immediately using bioinformatics if a structural model is available.
Sensitivity	Unknown. If gene dosage does not affect expression levels due to regulatory feedback, then no change would be expected.	Unknown. Lysate preparation may change conformation of many proteins.	Unknown. True target may be hidden if total resistance achieved through indirect resistance locus.
Specificity	Low. In a typical screen, many genes are implicated, complicating interpretation.	Low. In a typical screen, many proteins are implicated, complicating interpretation.	High specificity since very few mutations are observed per experiment.

**Table 1.3 Comparing methods of drug target identification (part 2).** A continuation of table 1.2, comparing and contrasting contrasts two methods of genome-wide drug target identification, with directed evolution, the primary method employed in this thesis. (Pierce et al., 2006, Pierce et al., 2007)

	Haploinsufficiency and Deletion Library Screens (HipHop)	Cell Lysate Binding Assays	Directed Evolution
Major Source of Bias	Unclear what proportion of essential genes exhibit haploinsufficiency phenotypes.	Bias towards stickier genes. This requires control screens against compounds with known target.	Bias towards mechanisms of pleiotropic drug resistance. Deleting genes that confer non-specific drug resistance, since they are typically non-essential, can mitigate this bias.
Difficulty	Medium to high. Methods using arrayed libraries either require specialized robotics or a high amount of labor. Pooled methods are faster and easier, but require ability to sequence barcodes.	High. Requires knowledge of structure activity relationship, so that you know which part of the compound can be tagged. Requires medicinal chemistry capacity to tag the compound.	Low. Does not require specialized equipment, sequencing methods or analysis.
Biggest Pro	Data from all genes is provided.	Can identify multiple binding partners	Fast and easy. High specificity and single nucleotide resolution.
Biggest Con	Low specificity making data interpretation challenging. May require specialized equipment or analysis pipeline.	Very difficult to do. Requires knowledge of structure activity relationship, so that you know which part of the compound can be tagged. Requires ability to tag compound, which requires a chemist.	Does not work for compounds that are not cytotoxic to organism. True target may be hidden if total resistance achieved through indirect resistance locus.

Directed evolution, or related methods of identifying changes in drug resistant organisms, has been used for many years to discover targets in bacteria. For example, the mechanism of action of the fluoroquinolones was identified in the late 1960s through genetic mapping of drug resistant bacteria (Hane and Wood, 1969). Similarly, the mechanism of action of streptomycin was determined based on the observation that ribosomes from streptomycin resistant bacteria, continued to be streptomycin resistant in *in vitro* protein translation cultures, whereas ribosomes from streptomycin sensitive bacteria, were also sensitive in cell free translation assays (Cox et al., 1964) (Davies et al., 1964).

In the past, the use of genetic-based approaches has been dependent on the availability of complete genetic maps or genome sequences, as well as rapid methods for mapping the resistance locus to a given chromosomal position. The advent of whole-genome tiling arrays and whole-genome sequencing has made the method more practical in single-celled organisms such as *Mycotuberculosis bacilli* (Aye et al., 2016, Sandhu and Akhter, 2016), *Plasmodia* parasites (Baragana et al., 2015), and trypanosoma (Khare et al., 2015).

Unlike other essential gene products that are often described as possible drug targets, drug targets discovered through directed evolution are known from the outset to be druggable. The method has been successfully used to identify the targets of several antimalarial compounds initially found through phenotypic screens, as well as to identify genes contributing to the resistome (Hoepfner et al., 2012b) (McNamara et al., 2013) (Rottmann et al., 2010a). Druggable targets discovered through target identification may later be used in protein-based drug

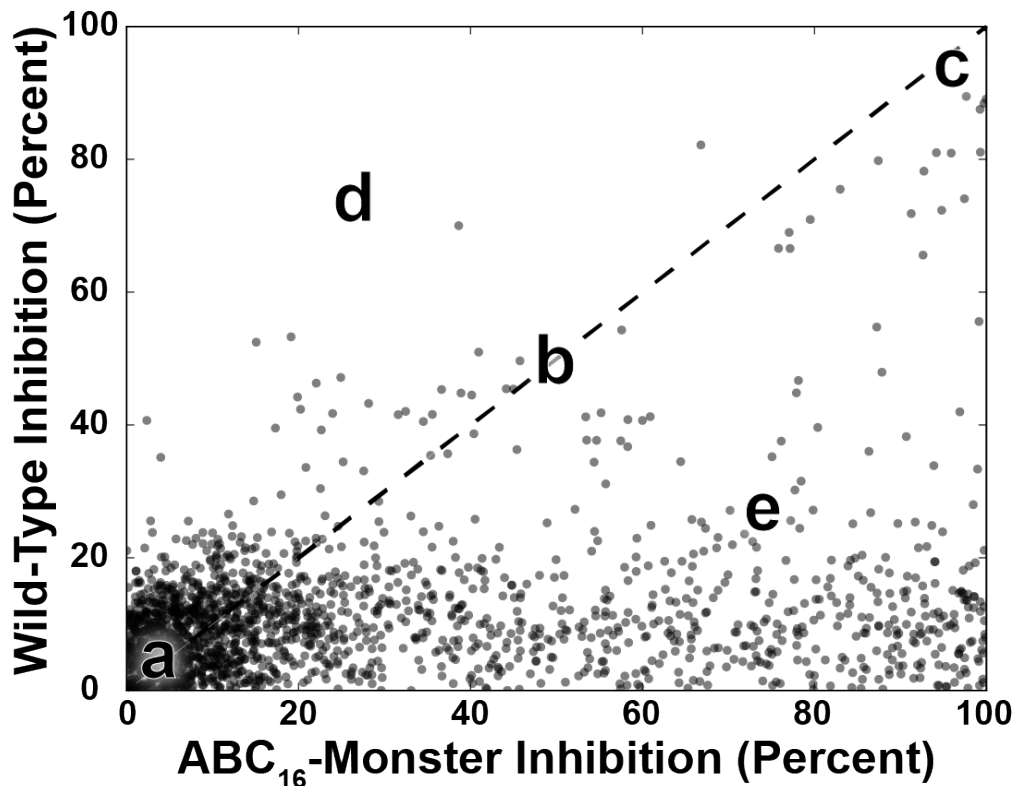
screens to identify improved inhibitors. Their 3D structures may also be used for structure-guided chemical optimization during the drug development process.

Directed evolution followed by whole genome sequencing has its limitations. *In vitro* evolution followed by genome sequencing is very challenging in organisms that are multicellular, sexual, or difficult to culture. Many organisms also have lengthy cell cycles, large genomes that are prohibitively difficult to sequence, and multidrug efflux pumps that render them insensitive to many compounds of interest. Furthermore, resistance in multiploid organisms often requires simultaneous mutations in all alleles. Directed evolution in evolutionarily related surrogate species, such as the haploid yeast *S. cerevisiae*, an organism with a small, well-characterized genome, has many advantages that are summarized in Table 1.4. One major drawback of using most strains of the yeast *S. cerevisiae* is the fact that drug resistance often evolves through mutations in efflux-pump genes rather than the true drug target.

**Table 1.4 The advantages and disadvantages of performing directed evolution experiments in three selected organisms.** This table compares and contrasts different cell lines for their potential use in drug target identification using directed evolution and whole genome sequencing. (Gardner et al., 2002) (van Staveren et al., 2009) (Cherry et al., 1997a).

	<i>Plasmodium falciparum</i>	Human Cell Lines	<i>Saccharomyces cerevisiae</i>
<b>Genome Consideration</b>			
Genome size (larger genomes are more expensive to sequence)	25MB	3,000MB	12.5MB
Ploidy during selection (haploid is ideal for directed evolution interpretation)	Haploid	Aneuploid	Haploid
Number of genes	5.3k	30k	5.8k
Percentage coding (the lower the percent coding, the higher the probability that mutations will arise that modify gene regulation rather than target structure)	50%	<2%	70%
Genome annotation	Poor	Intermediate	High
<b>Directed evolution Considerations</b>			
ABC-Transporters	17	49 plus both genetic and transcriptional amplification	30 (14 for ABC <sub>16</sub> -Monster)
Compound requirements	High due to media changes	High due to media changes	Low, because no media changes are required. Potentially high due to drug export.
Labor intensity	High due to daily microscopic evaluation	High due to daily microscopic evaluation	Low, since visual inspection is sufficient to evaluate growth.
Dividing time	8 hours	20-96 hours	2 hours
Time for drug selection	Months	Months to years	Days to weeks+

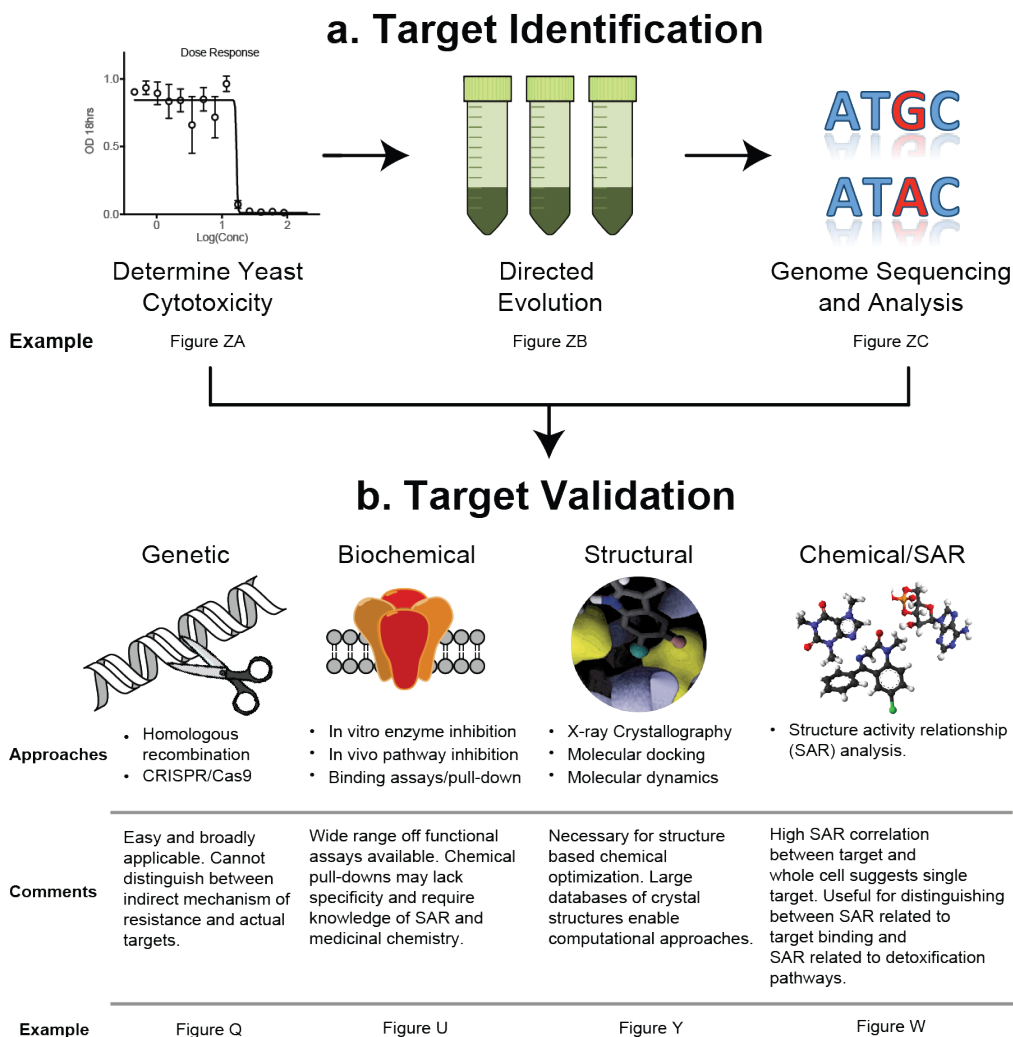
To address these challenges, we have developed an improved method for directed evolution target identification. We use a recently engineered *S. cerevisiae* strain, the ABC<sub>16</sub>-Monster, that had sixteen efflux-pump genes deleted and is broadly sensitive to cytotoxic compounds, and is, therefore, well suited to directed evolution experiments for compounds that are effective against other pathogenic organisms. It has been shown that the number of compounds used in clinical trials that are effective against the engineered strain is two-fold greater than that of the wild type (Suzuki et al., 2011). To verify enhanced ABC<sub>16</sub>-Monster susceptibility to drug-like molecules typical of those identified in phenotypic screens, we have also confirmed that ABC<sub>16</sub>-Monster has increased sensitivity to the 3,835 known antimalarials of the Novartis malaria box. The average inhibition of the ABC<sub>16</sub>-Monster and wild-type strains was 17.7% and 9.3%, respectively ( $p = 3.2e-45$  per a two-tailed t-test) (Fig 1.2).



**Figure 1.2 Inhibition of ABC<sub>16</sub>-Monster vs. wild-type yeast.** We screened 3,835 antimalarial compounds from the Novartis's collection of blood stage inhibitors. These compounds are effective against wild type (3D7) and a chloroquine-resistant mutant (W2) of *P. falciparum*, but not against human cells (Plouffe et al., 2008). Most compounds in this set are commercially available. The assay was done in duplicate at a compound concentration of 12.5 $\mu$ M for 16 hours at 30 deg. C. Cell growth was quantified using resazurin (Hoepfner et al., 2012a), with the reducing power of cells as a proxy for cell growth. The advantage of the ABC16 monster strain was striking. 1,672 compounds out of 3,835 inhibited the ABC16 monster growth by 75% or more, while only 54 compounds inhibited a wild-type strain to this range. Each point represents a single known antimalarial compound. **(a)** Compounds that were ineffective against both yeast strains. **(b)** Compounds that had similar potency in both strains, suggesting that they are not subject to ABC export. **(c)** Compounds arguably so potent that they could overcome any export. **(d)** Compounds that were paradoxically more effective in wild-type yeast (few). **(e)** Compounds that were more effective in ABC<sub>16</sub>-Monster, suggesting that they are exported in the wild-type strain.

Our method involves three steps outlined in Figure 1.3a. First, a drug candidate of unknown mechanism is tested for cytotoxicity against *the S. cerevisiae* ABC<sub>16</sub>-Monster strain. After verifying compound cytotoxicity, directed evolution is used to select for mutant cells with resistance to the cytotoxic compound of interest. Since *S. cerevisiae* doubles every two hours, drug resistance can be selected for in days rather than the months required in many other organisms such as asexual blood stage *P. falciparum*.





**Figure 1.3 An improved method for directed evolution target identification.** (a) Target identification via directed evolution and whole-genome sequencing, using a drug-sensitive yeast strain. (b) Target validation via genetic, biochemical, structural, and chemical/SAR follow-up experiments.

Third, the whole genomes of any resistant clones are sequenced.

Because the yeast genome is small (12.5 MB), sequencing is cost effective, and the number of mutations in resistant lineages is limited. The correct target can frequently be identified by disregarding mutated genes that are not essential or which are otherwise associated with multidrug resistance, such as certain well characterized transcription factors or efflux pumps. Mutated genes that confer resistance in multiple independent lineages are particularly attractive as potential targets. Rules for prioritizing potential drug targets from whole genome data are described in Table 1.3. Any ambiguity that persists after comparative sequencing can be resolved using a number of target-validation strategies (Fig. 1.3b), including genetic assays such as CRISPR/Cas9, enzymatic assays, structural analysis such as x-ray crystallography or homology modeling, and chemical structural activity relationship (SAR) analysis. It's important to note, genetic validation such as inserting discovered mutations into a clean genetic background using CRISPR/Cas9 can be used to confirm the importance of an individual mutation to the resistance phenotype, but does not prove that the mutation is in the target. Building a cohesive argument using a combination of the strategies is necessary to prove a Drug target interaction.

**Table 1.5 Simple rules for prioritizing potential targets.** These rules are examples of considerations use when determining whether mutated genes in chemogenomic studies are drug targets or resistance genes that do not directly bind the drug.

The gene that encodes the potential target is likely to be essential.
Resistance-conferring mutations in different lineages will arise more frequently in the gene associated with the true target than in other genes.
Mutations in genes associated with multidrug resistance (e.g., transcription factors, ABC-transporters, and MFS transporters) likely confer resistance through mechanisms unrelated to direct small-molecule binding to the primary target.
Targets that are known or suspected to be druggable, or that have homologous proteins in related organisms that are druggable (e.g., <i>pf</i> ATP4p), are more likely to be the true target.

## **Acknowledgements**

Figure 1.1 is a public domain photo from the NIAID's website:

<http://www.niaid.nih.gov/topics/malaria/pages/lifecycle.aspx>. The experiments in Figure 1.2 were performed at the Genomics Institute of the Novartis Research Foundation by Dr. Yo Suzuki of the J. Craig Venter Institute and Dr. Case McNamara, formerly of the Genomics Institute of the Novartis Research Foundation. Analysis was performed by Gregory M. Goldgof and Dr. Jacob D. Durrant.

## 2. Method Validation Using Known Drug Target Combinations

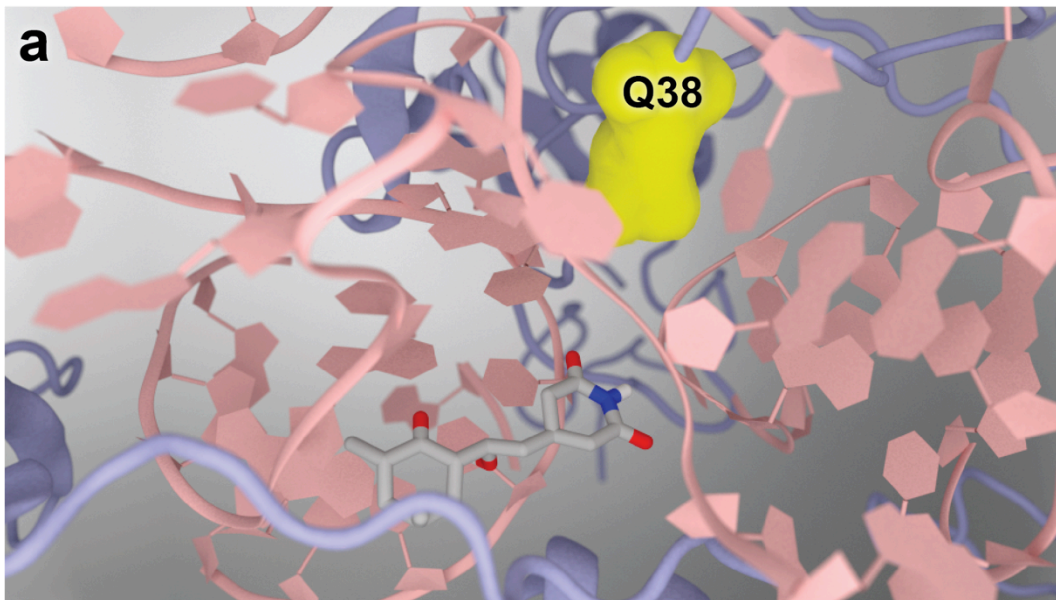
To test the validity of our approach, we first decided to evolve resistance to compounds with known drug targets. Evolving resistance to known drugs may also lead new insights into the resistome for those compounds. Since the focus of this dissertation is identification of novel Drug target interactions, description of this work will remain brief.

The first compound we evolved resistance to was the eukaryotic protein translation inhibitor, cycloheximide. This compound has activity against most eukaryotes, including fungi, *Plasmodium* and mammalian cells. In yeast, It acts through inhibition of the ribosomal protein Rpl28. Rpl28p is a component of the 60S subunit (the large subunit) of the ribosome and is believed to have peptidyl transferase activity (Mager et al., 1997).

ABC<sub>16</sub>-Monster cells were passaged six times in the presence of increasing cycloheximide concentrations, ultimately producing three resistant mutants that grew at ~40x the parent strain IC<sub>50</sub> concentration. One of the cycloheximide-resistant strains had mutations in two transcription factors: *FKH2* and *YAP1*. How these mutations confer resistance to cycloheximide is unclear, however, many transcription factors regulate the cellular response to stress and can confer drug resistance through indirect means such as drug export, sequestration and metabolism (Balzi and Goffeau, 1995, Onda et al., 2004). In fact, an ortholog of *YAP1* in *Candida* has previously been show to mediate multidrug resistance by up-regulating expression of a major facilitator superfamily transporter, a type of transporter not deleted in the ABC<sub>16</sub>-Monster (Chen et al.,

2007). Whole-genome sequencing revealed no additional missense or nonsense mutations.

In contrast, two of the strains had distinct missense mutations at the locus encoding the Q38 amino acid of the *RPL28* gene (Q38E and Q38L). *RPL28* encodes ribosomal 60S subunit protein L28, the known cycloheximide target. Furthermore, the Q38 residue maps onto the known cycloheximide-binding site, and others have similarly shown that single-amino acid variations at that residue confer resistance (Kaufer et al., 1983, Huang et al., 2013). Our technique therefore correctly identified not only the cycloheximide target, but also the cycloheximide-binding site, in two of three resistant mutants (see **Fig. 2.1**).



**Figure 2.1. Cycloheximide resistance-conferring mutations found in Rpl28p.** Two distinct amino-acid changes (Q38E and Q38L, shown in yellow) evolved in the cycloheximide-binding site of the ribosomal 60S subunit protein L28. Nucleic acid and protein strands are shown in pink and blue, respectively. Cycloheximide is shown in licorice. Portions of the structure have been removed to facilitate visualization. The illustration was generated using the 4U3U protein structure (de Loubresse et al., 2014).

The experience with cycloheximide foreshadowed both the impressive strengths of the directed evolution approach and its major challenge. The strength of directed evolution is the ability to not only identify the target of a chemical compound, but also the specific amino acid(s) that interact with the chemical at its binding site. On top of this, it can also be used to explore the types of changes at these sites that confer resistance, which can improve our understanding of the nature of the binding interaction.

On the other hand, the experiment with cycloheximide also underscores one of the risks of this approach. If resistance occurs as a result of mutations in genes that are not the direct target but confer a growth advantage to the cell through indirect resistance genes, such as the transcription factors that were mutated in one of the lineages, it may be challenging to discover the true target of the compound. Mutations in indirect resistance genes may confer sufficient resistance that further selection is impossible.

As a second proof of concept, we next used the technique to identify the known target of etoposide (Montecucco et al., 2015), a compound that kills cancer cells by forming a complex with the enzyme topoisomerase II and DNA. Topoisomerase II unwinds DNA double strands. By binding to the DNA-enzyme complex etoposide prevents the re-ligation of DNA strands, and thereby causes DNA strands to break. ABC<sub>16</sub>-Monster cells were passaged three times in the presence of increasing etoposide concentrations, producing five resistant mutants that grew at about 20-fold the parental strain IC<sub>50</sub> concentration.



**Table 2.1: Summary of directed evolution and sequencing.** This table summarizes the sequencing data from the five drug resistant strains evolved under selection by etoposide. It is useful to note how few missense mutations occurs per strain.

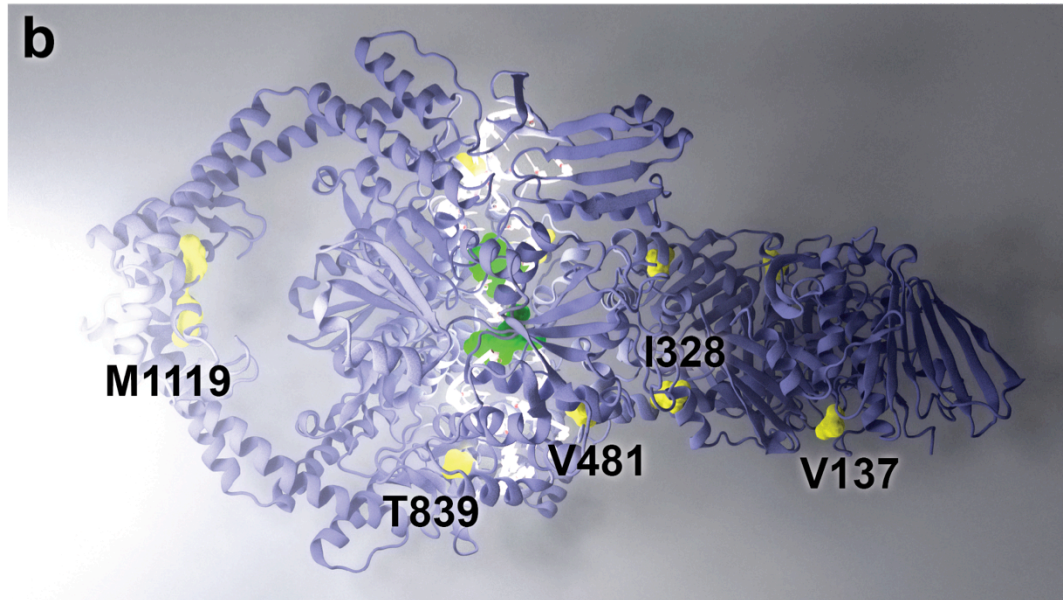
	R1	R2	R3	R4	R5
Total SNVs	3	3	5	6	7
Intergenic	2	1	1	2	2
Silent SNVs	0	0	0	1	1
Missense	1	2	4	3	4
Insertions and Deletions	None	None	None	None	None

Table 2.1 summarizes the sequencing results and Table 2.2 details the genes that had missense mutations in them encoding amino acid changes. One of the resistant strains had a single mutation in the non-essential gene *IRA2*, an inhibitory regulator of the RAS-cAMP pathway (Tanaka et al., 1989). The remaining four strains each had distinct missense mutations in *TOP2* (V137G, I328N, M1119R, T839N, and V481F). *TOP2* encodes DNA topoisomerase II, the known etoposide target, once again demonstrating the utility of our method.

**Table 2.2 Summary of genes with missense mutations for etoposide selections.** Mutations in *TOP2*, the known target of etoposide, are in bold. Few missense mutations occur per resistant strain. One gene, *TOP2*, is mutated in four out of five strains. *TOP2* is an essential gene that is not hypervariable or previously associated with pleiotropic drug resistance. The combination of these facts strongly suggests that *TOP2* is the target of the drug etoposide.

Lineage	Gene of Mutated Protein:Amino Acid Change
R1	<i>IRA2</i> :T1551I
R2	<b><i>TOP2</i>:V137G,</b> <b><i>TOP2</i>:I328N</b>
R3	<b><i>TOP2</i>:M119R,</b> <i>SHG1</i> :D105Y, <i>RFU1</i> :A98S, <i>AZF1</i> :G307D)
R4	<b><i>TOP2</i>:T398N,</b> <i>RFU1</i> :A98S, <i>AZF1</i> :G307D
R5	<b><i>TOP2</i>:V481F,</b> <i>TDA9</i> :N420K, <i>SPO71</i> :A421P, <i>UPC2</i> :M892I

Unlike in the case of cycloheximide, the mutated amino acids in Top2p did not cluster at one site or at the site of action of the compound (Fig 2.2). The method by which they confer resistance in this case is less clear than in the case with the mutations in *RPL28*. Perhaps these mutations affect the function of the protein or perhaps they indirectly affect the shape of the protein in a way that reduces its affinity for etoposide.



**Figure 2.2. Etoposide-conferring mutations found in Top2p.** Five distinct amino-acid changes (V137G, I328N, M1119R, T839N, and V481F, shown in yellow) evolved in DNA topoisomerase II. Nucleic acid and protein strands are shown in glowing white and diffuse blue, respectively. The crystallographic positions of etoposide (in green) and the nucleic acid were taken from the 3QX3 structure (Wu et al., 2011); the structure of the yeast protein dimer was taken from 4GFH (Schmidt et al., 2012).

## Acknowledgments

Gregory Mark Goldgof (G.M.G.) designed the experiments and wrote the manuscript. Directed evolution and IC50 experiments were performed by G.M.G., Jennifer Yang (J.Y.), Edger Vigil (E.V.), (J.Y.), Jack Schenken (J.S.), Rebecca Stanhope (R.S.), and Maximo Prescott (M.P.) Whole-genome sequencing was performed by G.M.G. and the UCSD Institute for Genomic Medicine Core Facility. Whole genome sequence analysis was performed by G.M.G., Felicia Gunawan (F.G.), and Micah J. Manary (M.J.M). Renderings of mutations in figures were produced by Jacob D. Durant (J.D.D.).

### 3. The Spiroindolone Antimalarials Target the P-Type ATPase Pma1 in *S. cerevisiae* and PfAtp4 in *P. falciparum*

#### Abstract

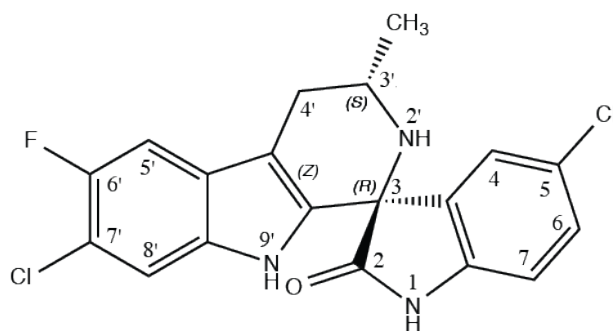
*In vitro* evolution and whole-genome sequencing are increasingly used to discover the target(s)/mechanism(s) of antimicrobial compounds found in high-throughput phenotypic screens, but it is challenging to subsequently determine whether or not resistance-conferring mutations have occurred in the gene encoding the compound's biochemical target or a resistance protein (e.g., a drug pump). This determination is especially difficult if the gene is poorly characterized and/or genetic studies in the targeted organism are difficult. Here we show that, despite differences in compound potency, the mechanism of action of the spiroindolone antimalarial KAE609 is conserved across fungal and apicomplexan eukaryotic phyla, allowing the genetically tractable model organism *S. cerevisiae* to be used as a surrogate.

Specifically, we show that *S. cerevisiae*, like *Plasmodium falciparum*, selects for mutations in a gene encoding a P-type ATPase (*ScPMA1*) after exposure to KAE609. Using an *in vitro* cell-free assay, we also demonstrate that KAE609 directly inhibits ScPMA1p activity, suggesting that PfATP4p is the direct KAE609 target rather than a drug pump. We also show that treatment with KAE609 causes an 80% increase in cytoplasmic hydrogen ion concentration. Computer docking into a ScPma1p homology model identifies a binding mode

that supports *in vitro* experimental structure-activity relationships. Our data demonstrate that drug-sensitive yeast models can be used to study compound-target interactions, providing a rapid method for characterizing the mechanisms of action of antimicrobial compounds.

### Introduction

The spiroindolones, a novel class of orally bioavailable antimalarials discovered in a phenotypic whole-cell screen, have been shown in Phase II clinical trials to rapidly clear parasites from adults with uncomplicated *P. vivax* and *P. falciparum* malaria (White et al., 2014). KAE609, also known as cipargamin or NITD609 (Fig. 3.1), a representative compound, works twice as fast in patients as the current gold standard treatment, artemisinin. KAE609 possesses both the potency (~550 pM against asexual blood stage *P. falciparum* (Rottmann et al., 2010b)) and favorable pharmacokinetics (elimination half-life of ~24 hours in humans (Leong et al., 2014)) needed for a single-dose cure, a feature that could help slow the onset of parasite resistance and that is not shared by existing, approved antimalarial drugs. KAE609 is also unique in its ability to block transmission to mosquitoes (van Pelt-Koops et al., 2012).



**Figure 3.1 KAE609, an experimental spiroindolone antimalarial.** The spiroindolones have been shown in Phase II clinical trials to rapidly clear parasites from adults with uncomplicated *P. vivax* and *P. falciparum* malaria (White et al., 2014) (White et al., 2014). KAE609 possesses both the potency (~550 pM against asexual blood stage *P. falciparum*) (Rottmann et al., 2010b) and favorable pharmacokinetics (elimination half-life of ~24 hours in humans) (Leong et al., 2014) needed for single-dose cure. Spiroindolone is a shorthand name for spiro-tetrahydro  $\beta$ -carbolines, which are composed of a  $\beta$ -carboline attached to an indolone group through a single shared carbon bound to three carbons and a nitrogen. Spiro compounds are those that create spirals as a result of two ring systems that are bound together by a single atom.

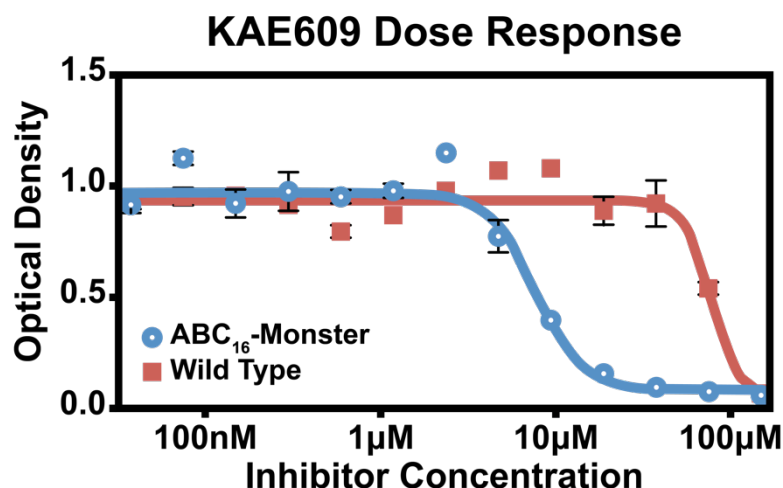
Despite promising activity in both cellular and organismal assays, the spiroindolones act by a relatively uncharacterized mechanism. Directed evolution experiments in parasites have shown that resistance is conferred by mutations in the gene encoding the parasite plasma membrane P-type ATPase, *PfATP4* (Rottmann et al., 2010b). Biophysical studies have shown that parasites treated with KAE609 are not only unable to extrude intracellular sodium, but also exhibit changes in intracellular pH. On the other hand, *PfATP4* mutations also confer resistance to a variety of unrelated chemical scaffolds with antimalarial activity, suggesting that *PfATP4* may be a multidrug resistance gene instead of the true spiroindolone target (Jimenez-Diaz et al., 2014b, Lehane et al., 2014, Vaidya et al., 2014, Flannery et al., 2015).

Despite its importance as a possible novel malaria drug target, attempts to overexpress *PfATP4* recombinantly have been unsuccessful, potentially because overexpression of this transporter appears to be toxic to some cells. There are also substantial hurdles associated with characterizing *PfATP4* function in parasites. These include the recombinogenic 89% AT-rich genome of the parasite, the parasite's slow growth rate, its low biomass, its intracellular replication, and complications related to how it is cultured. Given the limitations of direct work in malaria parasites, we investigated spiroindolone activity against baker's yeast (*Saccharomyces cerevisiae*), a more genetically tractable model system.

## Results

Using yeast proliferation as a readout, we found the half maximal inhibitory concentration ( $IC_{50}$ ) of KAE609 against a wild-type strain (SY025) to be prohibitively high for drug-selection studies ( $IC_{50} = 89.4 \pm 18.1 \mu\text{M}$ , 9 observations). Reasoning that the yeast cells might be expelling KAE609 via drug efflux pumps, we next tested a strain that lacks 16 genes encoding ATP-binding cassette (ABC) transporters, termed "ABC<sub>16</sub>-Monster" (Suzuki et al., 2011). As predicted, KAE609 was more potent against ABC<sub>16</sub>-Monster ( $IC_{50} = 6.09 \pm 0.746 \mu\text{M}$ ), suggesting that this unique yeast strain could be a useful surrogate for malaria parasites (Fig. 3.2).



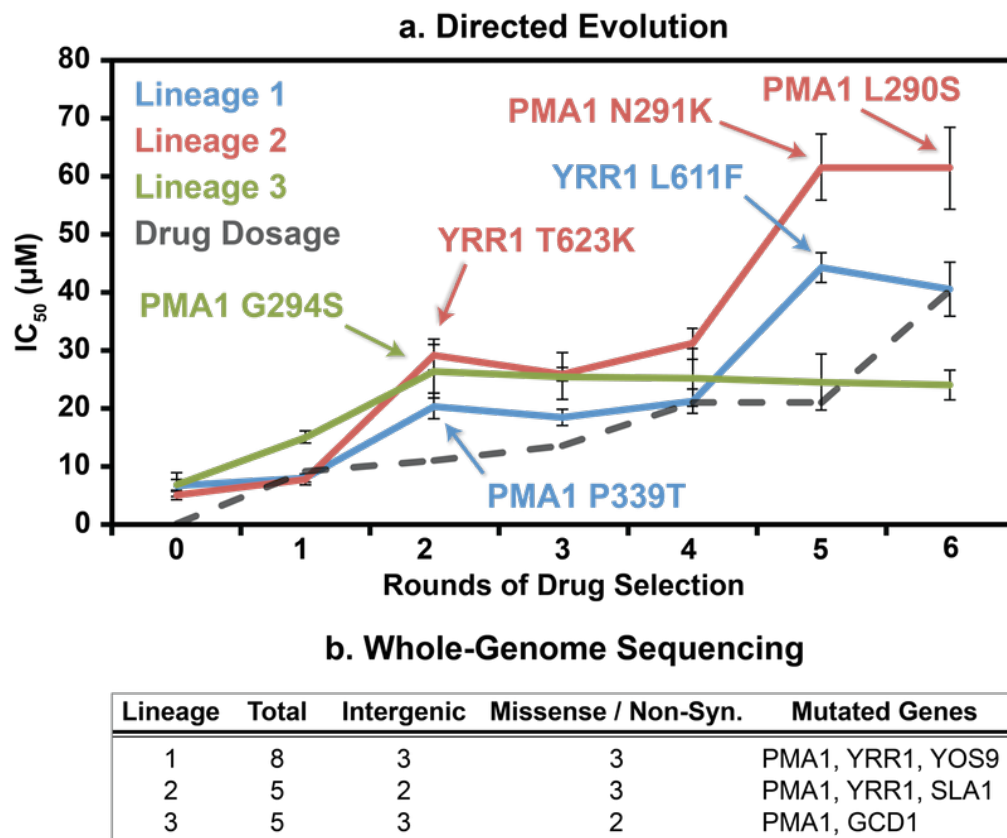


**Figure 3.2 Inhibition of wild-type and ABC<sub>16</sub>-Monster *S. cerevisiae*.** Growth was evaluated using optical density at 600 nm after 18 hours of incubation.

Using the same method that previously identified *PfATP4* as a KAE609 resistance gene (Rottmann et al., 2010b), we next sought to investigate whether the KAE609 mechanism of action is conserved in yeast. ABC<sub>16</sub>-Monster cells were exposed to increasing KAE609 concentrations in three clonal cultures. In all three cultures, compound resistance emerged after two rounds of selection, with new IC<sub>50</sub> values of  $20.4 \pm 2.24$ ,  $29.1 \pm 2.64$ , and  $26.4 \pm 4.64$  µM, respectively (Fig 3.3a). After an additional three rounds of selection, two of the cultures developed additional resistance ( $40.5 \pm 4.74$  and  $61.5 \pm 7.13$  µM). To determine the genetic basis of this *in vitro* resistance, we prepared genomic DNA from the terminal selection flasks. Samples were fragmented, labeled, and sequenced with >40-fold coverage. The sequences were then compared to the sequence of the parental clone.

Sequencing revealed 5–8 single nucleotide variants (SNVs) in each line and no additional copy number variants (CNVs) beyond the 16 ABC<sub>16</sub>-transporter

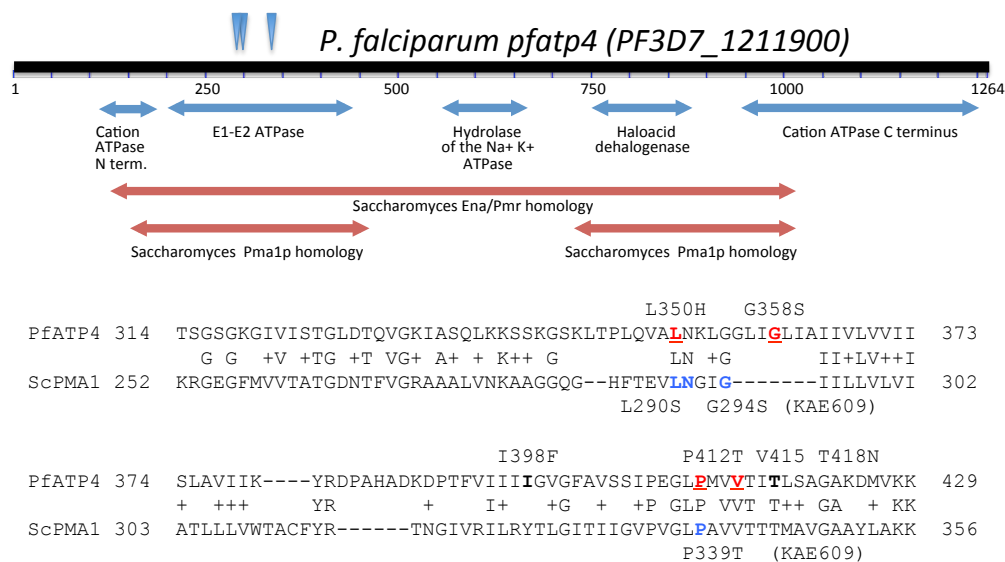
deletions and selection-marker insertions characteristic of the strain. Among the SNVs, there were 2–3 missense mutations in protein-coding genes per clone (Fig. 3.3b). *ScPMA1* was also the only gene mutated in all three clones. *ScPMA1* is an essential gene that encodes a P-type ATPase responsible for maintaining hydrogen-ion homeostasis across the plasma membrane in yeast (Serrano et al., 1986) and was the only essential gene among those identified.



**Figure 3.3 KAE609 directed evolution.** (a) Increasing ABC<sub>16</sub>-Monster resistance developed over multiple rounds of drug selection. Resistance mutations are highlighted. (b) Whole-genome sequencing after six rounds of selection

The mutations that were identified in *ScPMA1* are located in the E1-E2 ATPase domain in a region that is homologous to *PfATP4*. Interestingly, the

mutations that were identified in yeast were near or at the same relative position as mutations that had previously been identified after exposing malaria parasites to spiroindolones (Rottmann et al., 2010b) or dihydroisoquinolines (Jimenez-Diaz et al., 2014a). For example, the P412T mutation that confers resistance to the dihydroisoquinoline (+)-SJ733 aligns with the same amino acid change (P339T) that confers resistance to KAE609 in *S. cerevisiae* (Fig 3.4).



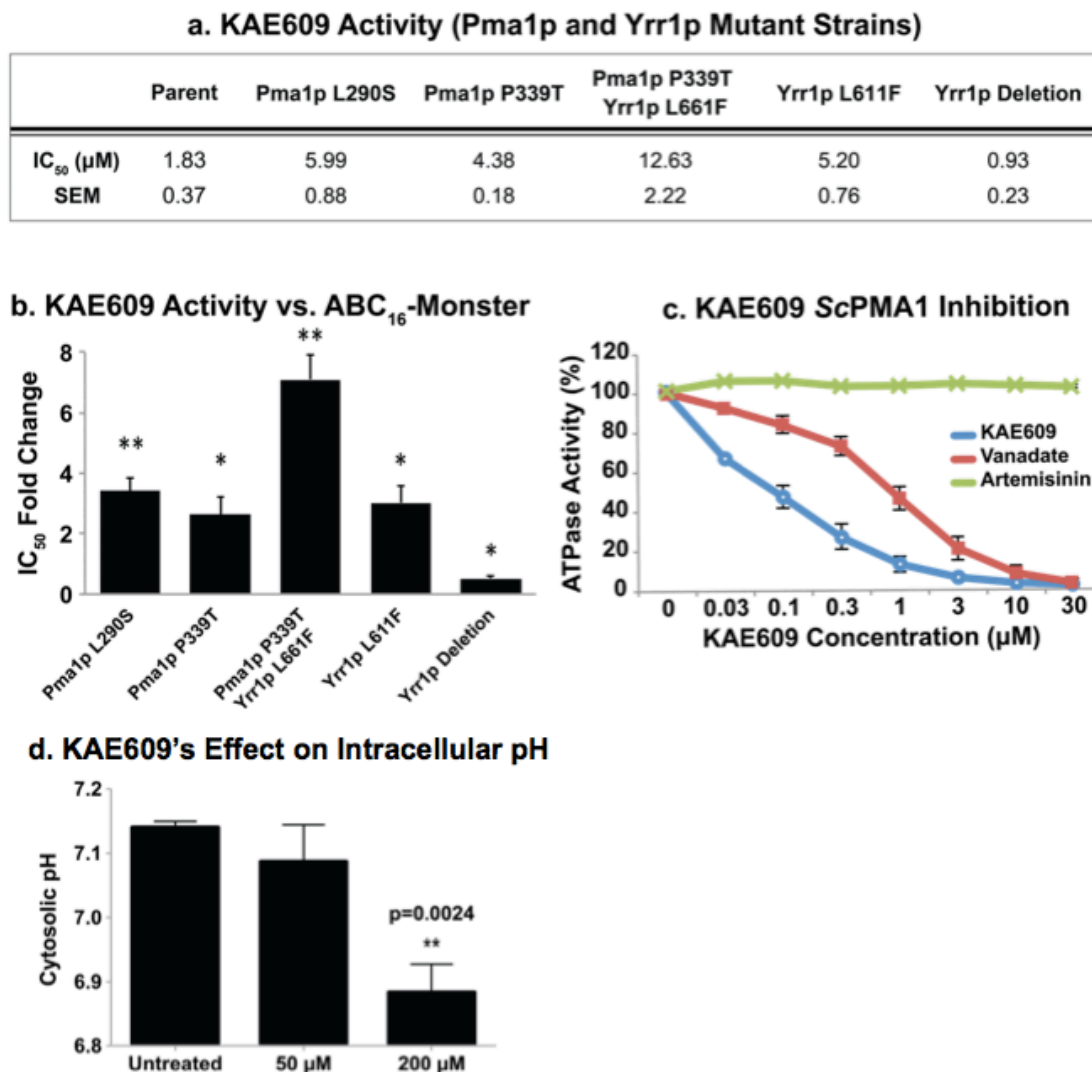
**Figure 3.4 Alignment of ScPma1p to PfAtp4p.** PSI-BLAST was used to align ScPma1p and PfAtp4p. The mutations that confer resistance to KAE609 in *S. cerevisiae* are shown in blue. In red are mutations that are associated with resistance to dihydroisoquinolines (L350, G358, P412T, V415D) in *P. falciparum*. In black are mutations that are associated with resistance to KAE609 in *P. falciparum* (I398F, T418N). It is notable that the P412T mutation that confers resistance to the dihydroisoquinoline (+)-SJ733 aligns with the same amino acid change (P339T) that confers resistance to KAE609 in *S. cerevisiae*.

Additionally, the transcription factor *ScYRR1* was mutated in two lineages. Sanger sequencing at each round of selection for *ScPMA1* and *ScYRR1* was used to determine when each mutation arose in its respective lineage. This same sequencing also identified an additional clone in Lineage 2 with its own distinct

*ScPMA1* mutation. Mutations in *ScPMA1* and *ScYRR1* each correlate with increased KAE609 resistance (Fig 3.3a).

Given that ScYrr1p is a transcription factor known to modulate a multidrug resistance phenotype, it is unlikely to be the KAE609 target. To support this hypothesis, we performed 103 additional directed evolution experiments in ABC<sub>16</sub>-monster against 26 different compounds with blood stage *P. falciparum* activity. None of the 103 genomes sequenced had *ScPMA1* mutations. However, 22 clones resistant to six unrelated compounds also had *ScYRR1* mutations. These findings suggest that *ScPMA1* is the spiroindolone target, and *ScYRR1* is a more general resistance gene.

Based on the initial evidence from directed evolution, we opted to further validate ScPma1p using genetic, structural, biochemical, and chemical/SAR-based methods. Genetic validation using the CRISPR/Cas system confirmed that mutations in *ScPMA1* and *ScYRR1* cause KAE609 resistance and that they have a multiplicative effect, as observed in the directed evolution experiments. However, ScYrr1p does not appear to be the primary KAE609 target. When we deleted *ScYRR1*, the resulting strain was viable, demonstrating that *ScYRR1* is not essential. Furthermore, KAE609 had double the potency against the deletion mutant, further suggesting that ScYrr1p confers resistance through an indirect mechanism (Fig 3.5a and Fig3.5b).



**Figure 3.5 Validation of KAE609 activity on Pma1p.** a) and b) The CRISPR/Cas system was used to introduce the same mutations observed in the directed evolution experiments into a clean genetic background. Both the Pma1p and Yrr1p mutations again conferred KAE609 resistance. ScYrr1 deletion doubled KAE609 potency, consistent with the hypothesis that ScYrr1 is an indirect resistance gene and not the primary KAE609 target. c) KAE609 inhibition of ScPMA1p in the vesicle-based assay. Inhibition is more potent than both inorganic orthovanadate, a non-specific ATPase inhibitor, and artemisinin, an unrelated antimalarial. d) The effect of KAE609 on intracellular pH was measured using pH-sensitive GFP. At 200μM intracellular hydrogen ion concentration increased by around 80%.

To determine whether KAE609 directly inhibits ScPma1p, we tested for ATPase inhibition in ScPma1p-coated vesicles harvested from *S. cerevisiae* cells with a temperature-dependent defect in secretory-vesicle/plasma-membrane fusion (Ambesi et al., 1997). In these experiments, KAE609 had an  $IC_{50}$  of 81.1 nM (SE = 1.19 nM), ten times more potent than inorganic orthovanadate, the non-specific ATPase inhibitor used as a positive control ( $IC_{50}$  = 842 nM, SE = 1.20 nM) (Fig 3.5c). Artemisinin, an unrelated antimalarial that served as a negative control, showed no activity. KAE609 inhibition of ScPma1p in a cell-free context is inconsistent with the hypothesis that ScPma1p modulates resistance by pumping the compound out of the cell, instead suggesting that KAE609 inhibits through direct binding.

As further validation, intracellular pH was measured using a strain of *S. cerevisiae* expressing the pH-sensitive green fluorescent protein. Since these cells are not ABC-transporter multi-knockouts, higher KAE609 dosages needed to be used. After 3 hours of treatment with 200  $\mu$ M KAE609, the cytoplasmic pH dropped from  $7.14 \pm 0.008$  to  $6.88 \pm 0.036$ . This equates to an 80.6% increase in the cytoplasmic hydrogen ion concentration ( $p = 0.0024$ ) (Fig. 3.5d). This finding is consistent with the model that KAE609 inhibits ScPma1p from pumping hydrogen ions from the cytoplasm to the extracellular space.

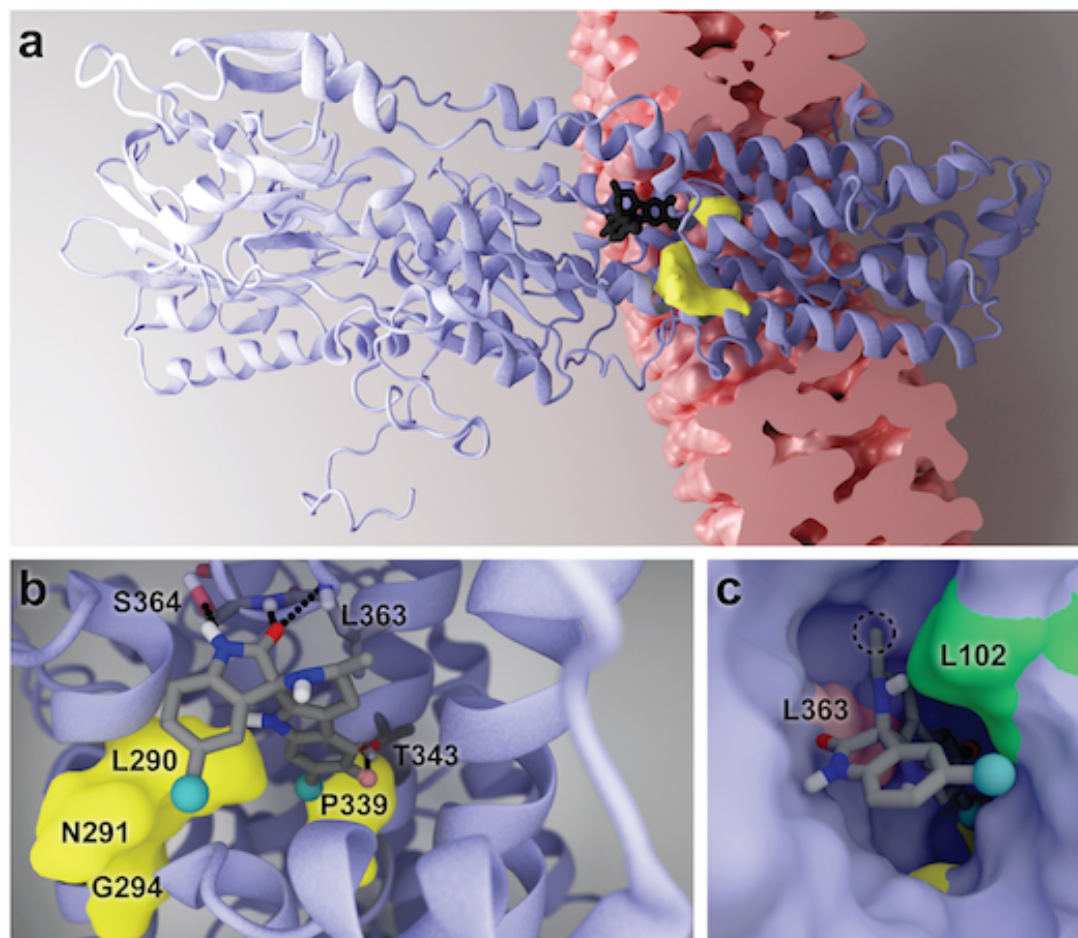
The KAE609 binding site and pose have not been previously characterized. We created a homology model of wild-type ScPma1p (UniProt ID: P05030) in the E2 (cation-free) state and mapped the four identified mutations (Leu290Ser, Pro339Thr, Gly294Ser, and Asn291Lys) onto the protein. The

altered amino acids line a well-defined, cytoplasm-accessible pocket within the membrane-spanning domain that is large enough to accommodate a small molecule (Fig. 3.6a and 3.6b). The computer-docking program Glide XP (Halgren et al., 2004, Friesner et al., 2006, Repasky et al., 2007) was next used to position KAE609 within the predicted wild-type pocket, with minimal manual adjustments. The docked pose placed the tricyclic tetrahydropyridoindole (THPI) moiety of the ligand between two residues that were mutated during directed evolution (Leu290 and Pro339) (Fig 3.7a), providing a plausible explanation for why these residues are so critical for KAE609 binding.

The docked pose suggests that receptor-ligand interactions are governed predominantly by high shape complementarity and hydrophobic contacts, provided by residues such as Leu102 (Fig. 3.6C). That both the Leu290Ser and Pro339Thr mutations substituted nonpolar with polar amino acids further supports the hypothesis that hydrophobicity plays a critical role. KAE609 may also form multiple hydrogen bonds with Ser364 and Leu363. In support of this, the Pro339Thr corresponds to a Pro415Thr PfATP4 mutation that confers resistance to the dihydroisoquinolines, a compound class that gives the same resistance profile as spiroindolones. The backbone amino groups of both these residues are predicted to form hydrogen bonds with the Ser364 ketone oxygen atom at the 2 position. Depending on the configuration of the Ser364 side chain, the Ser364 hydroxyl group may function as a hydrogen-bond acceptor bonded to the Ser364 nitrogen at the 1 position (Fig. 3.6b), or as another hydrogen-bond donor to the ketone oxygen atom at the 2 position. A hydrogen bond may also

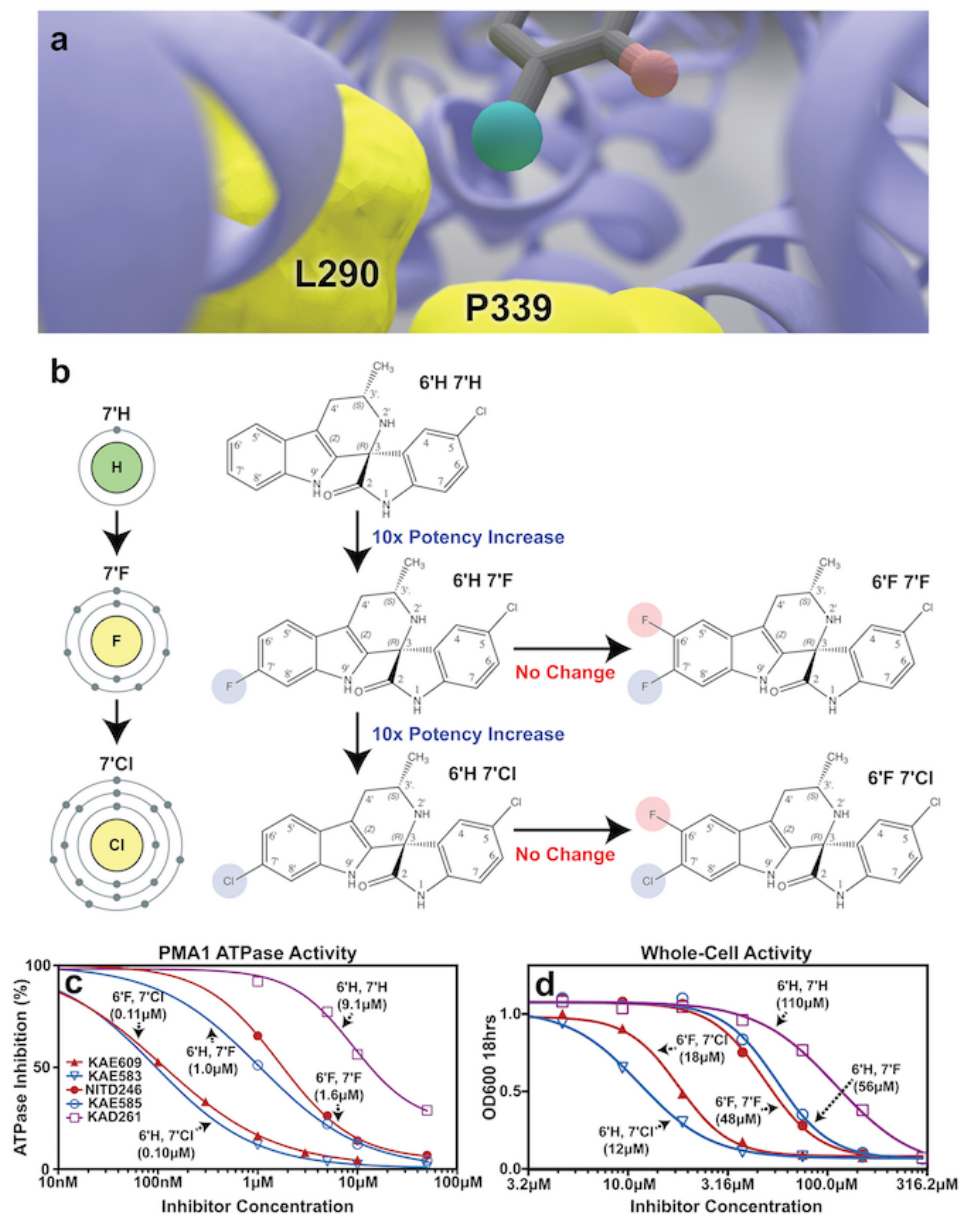
form between Thr343 and the KAE609 fluorine atom at the 6' position. H-F hydrogen bonds have been known to contribute to small-molecule binding in some contexts(Hao et al., 2012), though some argue that they are rare and typically weak (Fig. 3.6) (Dunitz and Taylor, 1997).





**Figure 3.6 Illustrations of the ScPMA1 homology model and KAE609 docked pose.** (a) The docked ligand is shown in black. Amino acids associated with evolved resistance are shown in yellow. The lipid-bilayer location was predicted using PDBID 1MHS(Kuhlbrandt et al., 2002), CHARMM-GUI(Jo et al., 2009), and the OPM database(Lomize et al., 2006). (b) The ligand with predicted hydrogen-bond partners (Ser364, Thr343, and Leu363). (c) The receptor in surface representation. Leu102 (green) is predicted to form hydrophobic contacts with KAE609. The position of Leu363 (pink) explains why the chirality at the 3' carbon atom is critical to potency; if inverted, the attached methyl group (circled) would clash sterically with Leu363

To verify the hypothesis that a hydrophobic interaction at the 7' position is key to potency, we tested the activity of KAE609 derivatives with substitutions at both the 6' and 7' positions (Fig. 3.7b). In the cell-free ScPma1p assay, substituting fluorine for the chlorine at position 7' results in a 10-fold decrease in potency (Fig. 3.7c). The removal of both the 6' and 7' halides causes a 100-fold decrease in potency. As the presence of fluorine at the 6' position has little effect on potency, the removal of the 7' halide appears to be exclusively responsible for the 100-fold decrease. Similar structure-activity relationships (SAR) are also observed in the whole-cell assay (Fig. 3.7d). The high degree of correlation between these two assays ( $r = 0.93$ ) suggests that the cytotoxic effect of KAE609 is primarily due to direct ScPma1p inhibition.



**Figure 3.7 KAE609 halogenation** (a) An illustration of the predicted binding pose, which positions the 7' chlorine atom (green) between two residues that were altered during directed evolution (L290S and P339T). Unlike the polar mutant residues, the nonpolar wild-type residues may stabilize the interaction with the 7' chlorine atom. (b) In yeast, halogenation at the 7' position (blue) has a substantial impact on potency, but halogenation at the 6' position (red) has little impact. Note that KAE609 is the molecule designated 6'F 7'Cl. (c) The blue, red, and purple curves indicate 7', 6'/7', and no halogenation. In the cell-free assay, 7' chlorination yields ~10- and ~100-fold greater potency than fluorination and no halogenation, respectively. (d) In the whole-cell assay, potency is increased ~3 and ~7 fold.

## Discussion

An open question is whether or not *Pf*ATP4p and *Sc*Pma1p have the same function. The function of *Pf*ATP4p in malaria parasites remains uncertain, but experiments have shown that it may affect both sodium and proton homeostasis (Spillman et al., 2013). In contrast, *Sc*PMA1p is only involved in maintaining proton homeostasis. Although *Sc*Pma1p and *Pf*Atp4p are highly homologous ( $P(n) = 4.60E-38$  per BLAST-P), *Pf*ATP4p and *Sc*Pma1p share the E1-E2 ATPase domain, but neither the Cation ATPase C terminal region (pfam00689) nor the hydrolase domain. *Pf*ATP4p also shows higher homology to the *Sc*Pmr1p and *Sc*Ena1p, *Sc*Ena2p and *Sc*Ena5p. However, these proteins are mostly (with the exception of *Sc*Ena5p) nonessential in *Saccharomyces* and thus they would not be likely targets for KAE609. Given that parasites lose the ability to extrude sodium after KAE609 treatment (Spillman and Kirk, 2015), it seems likely that the two proteins do not have the same cellular function.

On the other hand, the *Pf*Atp4p binding pose is likely to be similar to the *Sc*Pma1p pose presented here, especially as the yeast mutations are in a highly conserved region of the protein that is shared with *P. falciparum*. Both the 6' and 7' halides were originally added to reduce CYP2C9-mediated hepatic clearance, and the discovery that these modifications also increase potency up to 45-fold in whole-cell malarial assays was fortuitous (Yeung et al., 2010b). The current work suggests a structure-based explanation for the enhanced antimalarial activity of the halogenated compounds. On the other hand, further study of the malarial protein is warranted. There are some differences between the two proteins, as

reflected in subtle differences in the ABC<sub>16</sub>-monster and malarial whole-cell SAR. For example, 6' halogenation has little effect on ABC<sub>16</sub>-monster inhibition. In contrast, it increases potency against whole-cell *P. falciparum* (Yeung et al., 2010b). Additionally, 7' chlorination reduces malarial potency versus 7' fluorination, but the opposite is true in yeast (Yeung et al., 2010b).

Although our data strongly suggest that PfATP4 is the direct target of both the spiroindolones and by analogy, the dihydroisoquinolines, it is not clear that PfATP4 is the direct target of the aminopyrazoles (Flannery et al., 2015), or pyrazole amides (Vaidya et al., 2014) for example. PfATP4-resistance conferring or sensitizing mutations are located well away from the sites identified in yeast. It is possible that these compounds bind to another pocket, or that PfATP4 confers resistance through another mechanism. In addition, the aminopyrazole has only weak activity against yeast. Despite these questions, this work supports the hypothesis that PfATP4 is the direct KAE609 target in malaria and provides an additional method for characterizing P-type ATPase inhibitors. The data also suggest that ScPma1p could be an attractive target for novel antifungals (Perlin et al., 1997, Seto-Young et al., 1997, Monk et al., 2005).

## Acknowledgments

Gregory Mark Goldgof (G.M.G.), Dr. Sabine Otilie (S.O.), Dr. Jacob D. Durrant (J.D.D.), and Dr. Elizabeth A. Winzeler (E.A.W.) designed the experiments and wrote the manuscript. Preliminary feasibility studies were performed by Dr. Yo Suzuki (Y.S.) and Dr. Case W. McNamara (C.W.M.) Directed evolution experiments were performed by G.M.G. IC50 experiments were performed by G.M.G., S.O., Edger Vigil (E.V.), Jennifer Yang (J.Y.), Jack Schenken (J.S.), Rebecca Stanhope (R.S.), and Maximo Prescott (M.P.) Whole-genome sequencing was performed by the UCSD Institute for Genomic Medicine Core Facility. Whole genome sequence analysis was performed by G.M.G., Felicia Gunawan (F.G.), and Micah J. Manary (M.J.M). Homology bioinformatics studies were designed and performed by G.M.G. PCR studies were designed and performed by G.M.G, Gregory M. LaMonte (G.M.L.), E.V., and J.S. CRISPR mutant strains were designed and engineered by M.K., Y.S., Ayako Murao (A.M.), Marie Nachon (M.N.), and G.L. Intracellular pH studies were designed and performed by Kiersten Henderson (K.H.). Cell-free ATPase assays were designed and performed by Kenneth E. Allen (K.E.A.) and Carolyn W. Slayman (C.W.S.). Computational docking studies were designed and performed by G.M.G, J.D.D. and Dr. Rommie E. Amaro (R.E.A.).

Chapter 3, in part, has been submitted for publication as it may appear in the journal *Scientific Reports*, 2016, Goldgof GM, Jacob D. Durrant, Sabine Otilie, Edgar Vigil, Kenneth E. Allen, Felicia Gunawan, Maxim Kostylev, Kiersten A. Henderson, Jennifer Yang, Jake Schenken, Gregory M. LaMonte, Micah J.

Manary, Ayako Murao, Marie Nachon, Rebecca Stanhope, Maximo Prescott, Case W. McNamara, Carolyn W. Slayman, Rommie E. Amaro, Yo Suzuki and Elizabeth A. Winzeler under the title, "Comparative chemical genomics reveal the spiroindolone antimalarial KAE609 (Cipargamin) is a P- type ATPase inhibitor." The dissertation author was the primary investigator and author of this paper.

## 4. The Anti-Trypanosomal MMV001239 Inhibits Ergosterol Biosynthesis by Inhibiting *T. cruzi* Cyp51p

### Abstract

Recent advances in cell-based, high-throughput phenotypic screening have identified new chemical compounds that are active against many eukaryotic pathogens including kinetoplastid parasites. Identifying the molecular targets and binding modes of these compounds is critical for future drug development. In particular, subsequent structure-based chemical optimization and target-based screening require a detailed understanding of the binding event.

In this paper, we use a generalizable method for identifying the targets of antiparasitic compounds using directed evolution and whole genome sequencing in a drug-sensitive yeast strain. Using this approach, we identified lanosterol-14-alpha-demethylase (*TcCyp51p*) as the target of MMV001239, an open-access Medicines for Malaria Venture Malaria Box compound with activity against *Trypanosoma cruzi*. Direct drug-protein binding was confirmed through spectrophotometric binding assays and x-ray crystallography, revealing a binding site shared with other experimental anti-trypanosomal compounds targeting CYP51.



## Introduction

Phenotypic screens that test large chemical libraries for the ability to kill the organisms that cause infectious disease (Payne et al., 2007) (Swinney and Anthony, 2011). Inspired by success against other pathogens, researchers have recently used a similar approach to identify next-generation anti-trypanosomal drugs to treat Chagas' disease (Kaiser et al., 2015a, Pena et al., 2015) (Kaiser et al., 2015b, Llubra Montesino et al., 2015). Chagas' disease, the deadliest parasitic disease endemic to Latin America, is caused by the unicellular parasite *Trypanosoma cruzi*, which infects 5-10 million people worldwide, including 300,000-400,000 living in non-endemic regions like North America, Europe, and Australia. Global costs associated with the disease are estimated to be between 7 and 19 billion US dollars per year, with more than 10% paid by the USA and Canada (Lee et al., 2013).

Chagas' disease has two phases: the initial phase, which may be asymptomatic or acute, and the chronic phase. Initially, infection can last a few weeks as parasites multiply in tissues and organs (Bern, 2015). The symptoms are usually mild, but occasionally life-threatening myocarditis or meningoencephalitis can occur, leading to death in about ten percent of infected individuals. Ten to thirty percent of infected survivors subsequently develop chronic Chagas' disease, with symptoms that only develop after years or decades. The disease can cause tissue and organ damage, with potentially lethal complications from cardiomyopathy (Rassi et al., 2010).

Currently, only two drugs (benznidazole and nifurtimox) are available, but both have adverse drug reactions and variable efficacy. Furthermore, the recent BENEFIT (Benznidazole Evaluation for Interrupting Trypanosomiasis) clinical trial demonstrated that benznidazole is ineffective in symptomatic chronic stage patients (Marin-Neto et al., 2008). Safer, more effective drugs are needed to treat both the chronic stage of the disease and the acute stage in particularly vulnerable patients such as children and pregnant women (Altcheh et al., 2011).

To identify the next generation of compounds, many chemical libraries are currently being screened in both phenotypic (cell-based) and target (protein-based) screens, including libraries of active compounds known to be active against other parasitic diseases (Kaiser et al., 2015a, Pena et al., 2015) (Kaiser et al., 2015b, Llurba Montesino et al., 2015). One such library is the open-access Medicines for Malaria Venture (MMV) Malaria Box. The MMV Malaria box is a set of 400 compounds, 200 of which are drug like and 200 of which are intended to be tool compounds for discovering novel drug target pathways (Spangenberg et al., 2013). This library has been made freely available to research groups around the world, both to identify novel drug targets for malaria, but also to counter screen against other pathogens in the hopes of translating some of these drug classes and target pathways into neglected tropical diseases, such as Chagas' disease. Recently, 226/400 (56.5%) of the compounds in the open-access Medicines for Malaria Venture Malaria Box were shown to have activity against *T. cruzi*, with 25 (6.25%) having sub-micromolar activity (Kaiser et al., 2015a).

Given the ever increasing number of anti-trypanosomal compounds identified in phenotypic screens (Sykes et al., 2012) (De Rycker et al., 2012) (Merritt et al., 2014), a rapid method for identifying their molecular targets and binding sites is urgently needed. One approach to target identification is gene expression knockdown through RNA interference. Although RNAi can be easily used for target validation in *T. brucei* this method has not been successful in *T. cruzi* (DaRocha et al., 2004) (Elias and Faria, 2009). Another method to target identification is in vitro directed evolution, which often produces resistance-conferring mutations in Drug target genes that can be identified via subsequent whole genome sequencing. This approach has been highly successful in *P. falciparum* (Hoepfner et al., 2012b) (McNamara et al., 2013) (Rottmann et al., 2010a).

However, the advantages and limitations of directed evolution are organism dependent. Many organisms are difficult to culture, have lengthy cell cycles, possess large diploid genomes, and/or express multidrug efflux pumps that complicate resistance selection by conferring drug insensitivity generally (Zheng et al., 2004). Directed evolution in the surrogate species *S. cerevisiae*, which has a well-characterized, small haploid genome, is often effective. But yeast efflux pumps effectively expel many cytotoxic compounds, and resistance often evolves through mutations in these efflux-pump genes rather than through mutations in the true drug target.

Recently, directed evolution in *T. cruzi*, in combination with a yeast haploinsufficiency screen, identified the target of the experimental small molecule

inhibitor GNF7686 as cytochrome *b* (Khare et al., 2015). However, the approach took eleven months to evolve a single drug resistant culture with a mutation in cytochrome *b*. In addition, that strain had mutations that resulted in amino acid changes in two additional proteins. Had the mutation not been in a gene that has already been extensively studied as a drug target in *Plasmodium*, interpretation of this result may have been more challenging. Furthermore, while this approach has proven itself effective, given the large number of novel small molecule inhibitors under development for Chagas' disease, a more rapid approach is badly needed.

In the current work, we use an approach that relies on a synthetic drug-sensitive yeast strain (*S. cerevisiae*) that lacks 16 multidrug (ABC-transporter) export pumps (Suzuki et al., 2011). The advantages of using this so-called "ABC<sub>16</sub>-monster" strain as a directed evolution surrogate for more complex organisms such as *T. cruzi* are summarized in Table 4.1. In summary, the ABC<sub>16</sub>-Monster strain provides all of the advantages of working with *S. cerevisiae*, while eliminating a couple key drawbacks.

**Table 4.1 The ABC<sub>16</sub>-monster strain as a model for *T. cruzi*.** This table compares and contrasts options for directed evolution models for compounds that target *T. cruzi*. (Cherry et al., 1997b) (El-Sayed et al., 2005) (Andersson, 2011).

	<i>T. cruzi</i>	<i>S. cerevisiae</i>	ABC <sub>16</sub> -Monster
<b>Genomic Considerations</b>			
<b>Genome size</b> (smaller genomes cost less to sequence)	~60 Mb	~12.5 Mb	~12.5 Mb
<b>Number of Genes</b>	~12k (8k core genes)	~5.8k	~5.8k
<b>Percentage Coding</b>	59%	70.50%	70.50%
<b>Ploidy during selection</b>	Aneuploid	Haploid	Haploid
<b>Number of Chromosomes</b>	64-82 (although uncertainty remains)	16	16
<b>Genome Complexity</b>	50% of the genome consists of repeated sequences from retrotransposons or large antigen surface protein families	Low	Low
<b>Genome Characterization</b>	Poor	Extensive	Extensive
<b>Experimental Considerations</b>			
<b>ABC-Transporters</b>	30	30	14
<b>Ability to evolve resistance through regulation of drug export</b>	High	High	Low
<b>Labor Intensity</b>	High due to daily microscopic evaluation and media changes	Low	Low
<b>Time for Drug Selection</b>	11 months	< 2 weeks	< 2 weeks
<b>Compound requirements</b>	High due to media changes	Low	Lower
<b>Dividing Time</b>	22-200 h depending on strain and culture conditions	2 hours	2 hours
<b>Methods of Target Validation</b>	Limited	Many	Many

The advantages of yeast over *T. cruzi* for in vitro directed evolution are multifold, as *S. cerevisiae* has most of the desirable qualities for such an experimental approach. First, the dividing time is around 2 hours, in comparison to the 22-200 for *T. cruzi*. Thus, generation of drug resistant mutants is more likely to take one week, rather than one year. In addition, yeast is safe and easier to work with, making this technique available to many more research groups. Laboratory yeast strains can grow as haploids, meaning that a mutation need only be developed in a single allele to be effective, whereas *T. cruzi* is aneuploid, potentially complicating both the generation of resistance mutants and sequencing interpretation.

In terms of whole genome sequencing, *S. cerevisiae* also has a number of advantages over *T. cruzi*. First, the size of the yeast genome is approximately 12.5Mb, versus 60 Mb for *T. cruzi*, meaning that the cost of sequencing a drug resistant genome is approximately 5x less in yeast. The difference in the genome size, however, has little to do with the number of potential drug target genes, as *S. cerevisiae* has approximately 6,000 genes, compared to the 8,000 core *T. cruzi* genes. Approximately, 50% of the *T. cruzi* genome is comprised of repeated sequences from retrotransposons and large antigen surface protein families. While mutations in both of these are likely to arise during a year of selection, they are unlikely to contribute to our discovery of novel drug targets. *S. cerevisiae*, on the other hand, has relatively simple genome architecture, with over 70% of the genome encoding proteins, making it more likely that mutations

that are identified will be in the target gene, rather than intergenic or intronic regulatory sequences that may be difficult to interpret.

Finally, data interpretation in *S. cerevisiae*, is much simpler than many other organisms, since its genome has been fully sequenced and is better characterized than any other eukaryotic microbe. On the other hand, much of the *T. cruzi* genome is poorly characterized and only beginning to be studied. *S. cerevisiae*'s well characterized genome means that when mutations are identified in genes that are not already validated drug targets, scientists can rapidly determine using existing databases what the function of these genes are, if they are in related pathways, if they interact with each other physically or genetically, or if the genes are associated with drug resistance to other compounds.

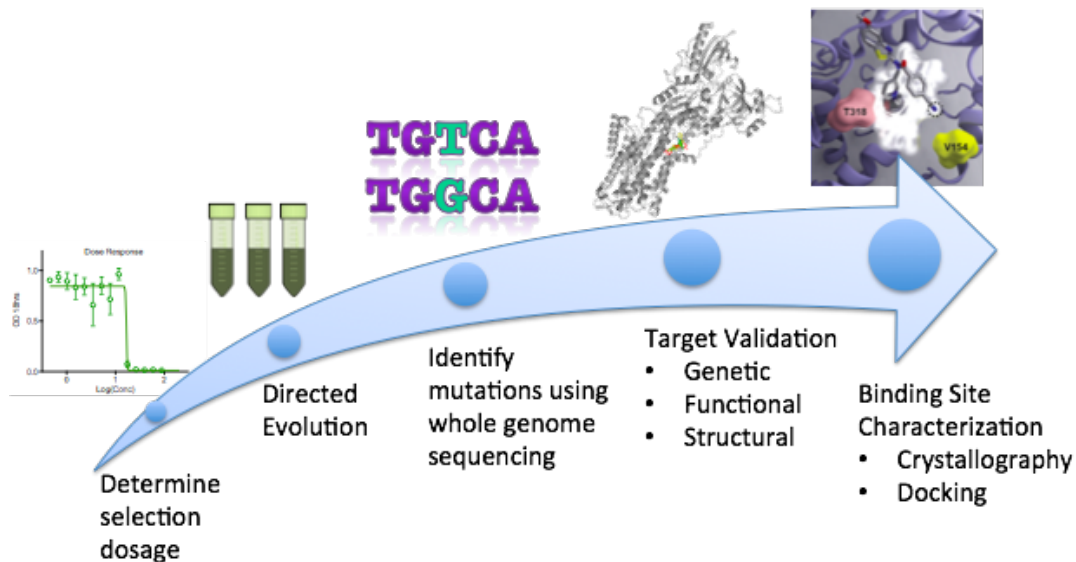
There are several drawbacks to using *S. cerevisiae* as a model for *T. cruzi*. The first is that there may be drug targets or pathways that are not sufficiently conserved between the two organisms. Such compounds are unlikely to have any cytotoxic effect against yeast, and therefore cannot be investigated using our method. Another major challenge is that many compounds with activity against *T. cruzi* are not active against *S. cerevisiae* due to its many methods of drug detoxification and multidrug resistance (MDR). One of the most important classes of MDR genes is the ABC-Transporter, members of which pump xenobiotics out of the cell and into metabolic compartments for sequestration and degradation. These pumps not only make directed evolution impossible for many compounds by preventing them from killing the cell, but also increase the

likelihood that resistance will emerge through mutations in the pumps themselves or their regulators.

The ABC<sub>16</sub>-Monster strain is a strain of drug-sensitive *S. cerevisiae*, in which the 16 ABC-transporters, including those most associated with drug resistance have been genetically deleted. As a result, this yeast strain is sensitive to roughly twice as many compounds, when it was tested against a collection of clinically approved drugs. At the same time, it has the same dividing time as other commonly used laboratory yeast strains (2 hours) and maintains all of the benefits of *S. cerevisiae* for target identification using directed evolution.

We use directed evolution in this strain (as outlined in Fig. 4.1) to identify lanosterol-14- $\alpha$ -demethylase (Cyp51) as the target of MMV001239, a member of the MMV Malaria Box with potent activity against *T. cruzi*. We show that resistance-conferring mutations in yeast can be used to predict the pathway targeted in more complex organisms and that whole genome sequencing can identify the direct molecular target. Delineation of the Drug target binding pose, which is conserved between fungi and *Trypanosoma* in the case of *ScErg11/TcCyp51*, underscores both the transferability of Drug target identification across taxonomic kingdoms as well as the general utility of our method.



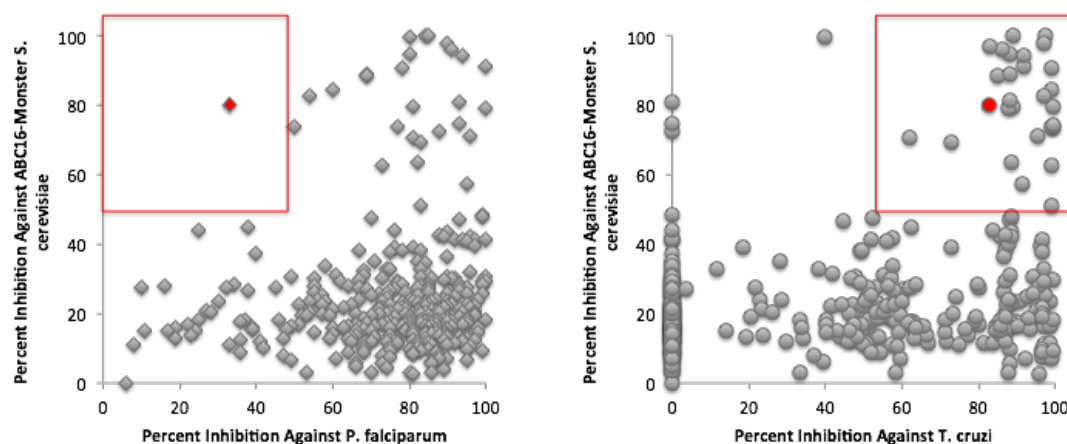


**Figure 4.1 Target identification via directed evolution.** First, determine the dose-response curve for your compound against the ABC<sub>16</sub>-Monster parental strain. Then, perform directed evolution using drug selection at concentrations around those that fully inhibit growth. After selection, clones are isolated to determine if they are drug resistant by comparing their dose-response curve to the parent strain. Both the parent and resistant lines are then sequenced to identify the mutation(s) that confer resistance. Once a potential target is identified, it can be validated using a variety of approaches. After that, the binding site should be characterized so that future structure based drug optimization is possible.

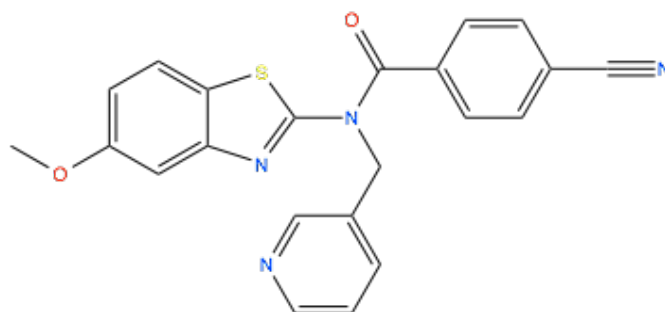
## Results

### Identifying a compound for proof-of-concept

To demonstrate that the ABC<sub>16</sub>-Monster yeast strain can be used as a surrogate for drug target screening and identification in *T. cruzi*, we first sought as a proof-of-concept to identify a model compound that is effective against both organisms. We tested the compounds of the Medicine For Malaria Venture (MMV) Malaria Box with known activities against *T. cruzi* for cytotoxicity against the yeast ABC<sub>16</sub>-Monster strain at 100  $\mu$ M (Fig. 4.2a). MMV001239 (Fig. 4.2b) was the only compound with activity against both the ABC<sub>16</sub>-Monster strain ( $IC_{50} = 2.20 \mu\text{M} \pm 0.23$ ) (Fig. 4.5) and *T. cruzi* ( $IC_{50} = 2.01 \mu\text{M} \pm 1.10$ ) (Fig. 4.4), but without minimal activity against *P. falciparum* ( $IC_{50} > 15 \mu\text{M}$ ) (Guiguemde et al., 2010) (Kaiser et al., 2015a).

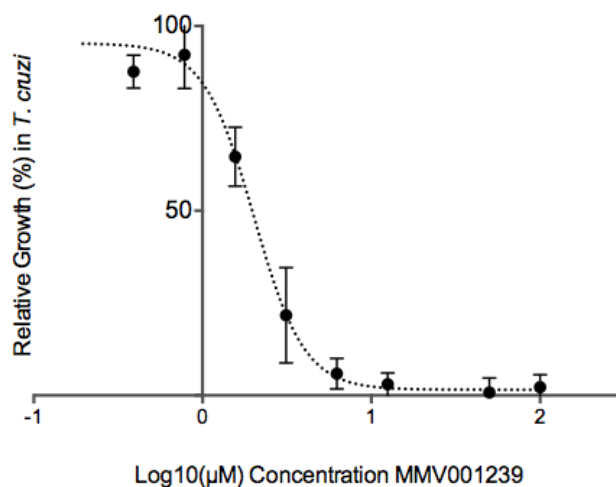


**Figure 4.2 MMV Malaria Box drug sensitivity screens.** Parts a and b plot the activity of the 400 Compound Open Source MMV Malaria Box. A plots blood stage *Plasmodium falciparum* (Guiguemde et al., 2010) against ABC<sub>16</sub>-Monster. The red box indicates compounds that are potent against ABC<sub>16</sub>-Monster and not *P. falciparum*. B plots *T. cruzi* intracellular amastigotes (Kaiser et al., 2015a) against ABC<sub>16</sub>-Monster. The red box indicates compounds that are potent against both ABC<sub>16</sub>-Monster and *T. cruzi*. MMV001239 (in red in both a and b) is the only compound that potently inhibits ABC<sub>16</sub>-Monster and *T. cruzi*, with minimal activity against blood stage *P. falciparum*.

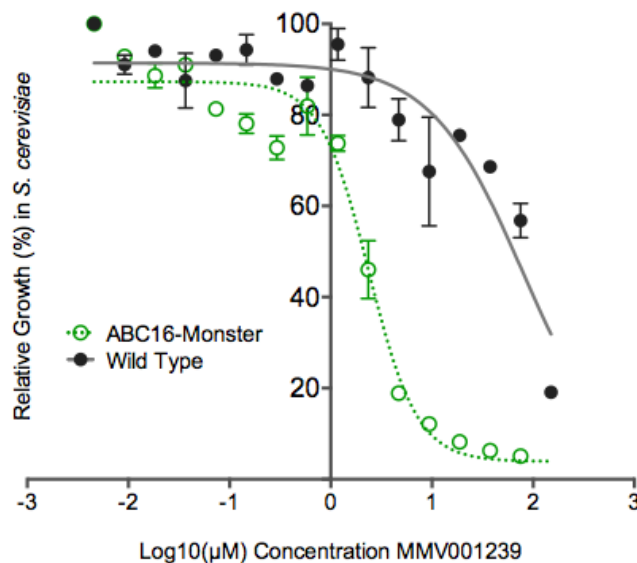


**Figure 4.3 The chemical structure of MMV001239.** MMV001239 is a member of the MMV Malaria Box, originally identified in the St. Jude's Children's hospital set of 1,300 hits from 300,000 screened chemicals (Guiguemde et al., 2010). Later studies showed that the compound only weakly inhibits *Plasmodium*, but has potent activity against *Trypanosoma cruzi*. The compound is composed of a pyridine-methyl-benzamide functional group on the bottom, a cyanobenzoyl functional group on the right, and methoxybenzothiazol functional group on the left. The full name of this compound is 4-cyano-N-(5-methoxybenzo[d]thiazol-2-yl)-N-(pyridin-3-ylmethyl)benzamide.

Its weak inhibition of *P. falciparum* and potent activity against *T. cruzi*, made it an attractive candidate for target identification using *S. cerevisiae*, since directed evolution in *Plasmodium* seemed both unlikely to succeed and of little interest to that research community. In addition, MMV001239 was only weakly active against wild-type *S. cerevisiae* ( $IC_{50} = 246.33 \mu\text{M} \pm 29.24$ ) (Fig. 3) and full wild-type growth inhibition was not achieved even at the highest concentration tested ( $300 \mu\text{M}$ ), implying that MMV001239 would make a poor candidate for directed evolution and target identification in a wild-type yeast strain. In contrast, the drug-sensitive ABC<sub>16</sub>-Monster strain is well suited to just this task.



**Figure 4.4 MMV001239 activity on *T. cruzi* intracellular amastigotes.** C2C12 myocytes were infected at a 15:1 parasite to host cell ratio with *T. cruzi* trypomastigotes and treated with MMV001239 as described in the methods. Compound activity was assessed after 72 h of treatment by determining the number of amastigotes per total host cells, normalizing to vehicle control and to positive control (uninfected cells).



**Figure 4.5 MMV001239 dose-response curve.** Data above are means and standard error for duplicates from a representative 18 hour, 16-point dose response experiment as described in the methods. This experiment was performed four times to determine average IC50 concentrations of  $2.20 \mu\text{M} \pm 0.23$  for the ABC<sub>16</sub>-Monster strain (○) and  $246.33 \mu\text{M} \pm 29.24$  for wild type *S. cerevisiae* (●).

#### Identification and characterization of MMV001239-resistant *S. cerevisiae*

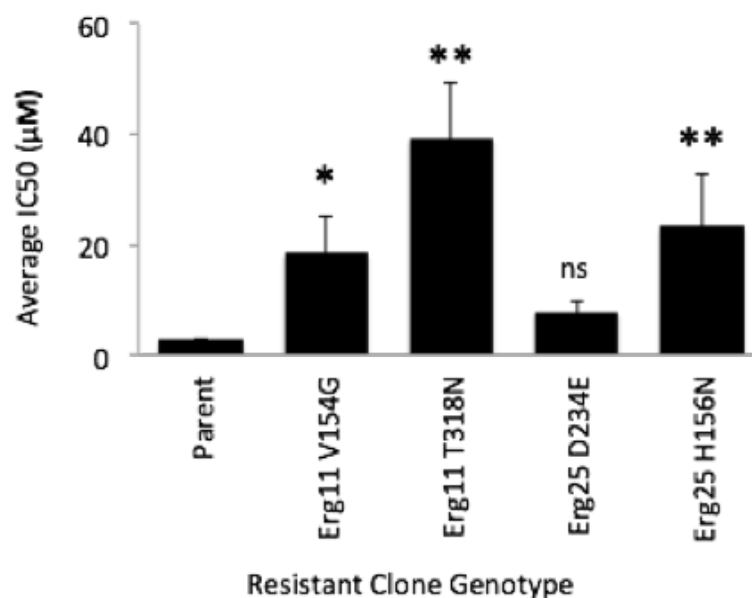
Directed evolution was performed using the ABC<sub>16</sub>-Monster strain in 20mL culture tubes in rich media (YPD) incubated at 30 deg. C with shaking. Lineage 1 was inoculated at an initial concentration of  $7 \mu\text{M}$  and grew to saturation after 2 days, after which it was passaged to  $50 \mu\text{M}$  and grew to saturation after 7 additional days. The resulting clone demonstrated 40-fold resistance relative to the parent.

**Table 4.2 Summary of directed evolution protocol.** This table summarizes the number of selections, their length of time, and the dosages used, to generate four mutants with drug resistance to MMV001239.

	<b>Mutant 1</b>	<b>Mutant 2</b>	<b>Mutant 3</b>	<b>Mutant 4</b>
Rounds of Selection	2	1	1	1
Total Days of Selection	9 days total: 2 days at 7 $\mu$ M, 7 days at 50 $\mu$ M	3 days	8 days	7 days
Selection Concentration	7 $\mu$ , 50 $\mu$ M	14 $\mu$ M	14 $\mu$ M	14 $\mu$ M

We next attempted to obtain resistance using a single drug selection step. Lineages 2, 3 and 4 were inoculated at a concentration of 14 $\mu$ M. These cultures took 3, 8 and 7 days to grow to saturation, respectively. Clones isolated from these cultures demonstrated 60x, 13x and 58x resistance respectively. DNA was isolated from one single clones from each of the four independently grown resistant cultures, as well as from one parental clone, and whole genome sequencing was performed to a depth of >50x coverage.

	Parent	Erg11 V154G	Erg11 T318N	Erg25 D234E	Erg25 H156N
IC50 ( $\mu\text{M}$ )	0.28	11.39	17.04	3.62	16.45
SEM	0.16	6.58	9.84	2.09	9.50



**Figure 4.6 IC50 Changes in MMV001239 Resistant Strains.** IC50s measured using a 16-point dilution after 18 hrs at 30 deg C shaking as described in the methods. *P*-values were determined using a paired single-tail t-test with (\*) indicating  $P < 0.05$  and (\*\*) indicating  $P < 0.01$ .

Analysis was performed to identify new single nucleotide variants (SNVs), insertions, and deletions. Each resistant clone had only one nonsynonymous single nucleotide variant change (SNV) (Table 2), in two genes of the *S. cerevisiae* ergosterol biosynthesis pathway, Erg11 and Erg25, with an additional three to five intergenic or silent mutations. No lineage contained insertions or deletions relative to the parental strain.

**Table 4.3 Summary of directed evolution sequencing results.** This table summarizes the whole genome sequencing results for the four MMV001239 resistant strains generated. Note, that each of the strains have only one or two missense mutations in the entire genome, making it very likely that these mutations are the cause of the resistance phenotype. It is also important to note that *FLO9* encodes a highly variable lectin-like protein expressed on the cells surface and likely involved in cellular defense (Chrispeels and Raikhel, 1991). These genes are highly variable and are commonly mutated in our experiments. However, they do not appear to confer any resistance phenotype.

	<b>Mutant 1</b>	<b>Mutant 2</b>	<b>Mutant 3</b>	<b>Mutant 4</b>
Total SNVs	5	6	4	4
Intergenic	2	2	1	2
Silent SNVs	2	3	1	1
<b>Missense</b>	<b>1</b>	<b>1</b>	<b>2</b>	<b>1</b>
<b>Mutated protein</b>	<b>Erg11</b>	<b>Erg11</b>	<b>Erg25, Flo9</b>	<b>Erg25</b>
Amino Acid Change	V154G	T318N	D234E (ERG25) I519M (FLO9)	H156N
ERG Protein Function	Lanosterol 14-alpha- demethylase	Lanosterol 14-alpha- demethylase	C-4 methyl sterol oxidase	C-4 methyl sterol oxidase
Insertions or Deletions	None	None	None	None



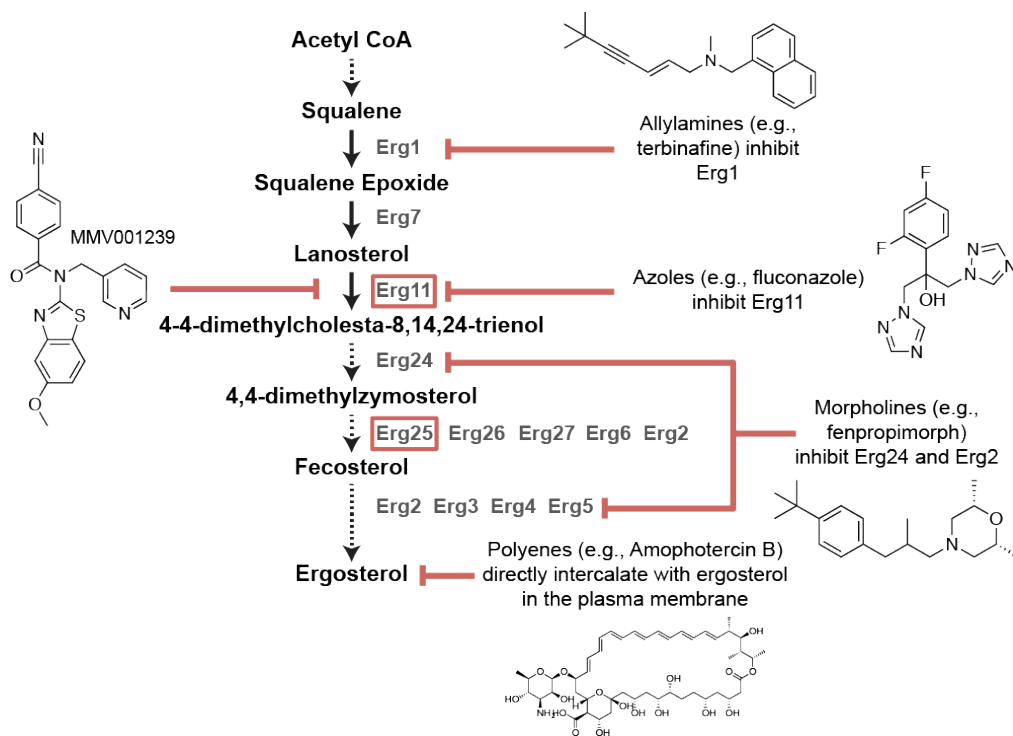
**Table 4.4 Sequencing statistics.** This table provides basic quality control statistics for the sequencing that was performed. \*Our pipeline did not include the analysis of insertions that aligned to other species, but other markers that are part of the ABC<sub>16</sub>-Monster strain from other species should also theoretically be present. The ABC<sub>16</sub>-Monster and all of its derived strains have a defective version of the HO gene, whereas the wild-type strain has a full deletion. This difference is not included in the table, for the sake of clarity. In addition, mutations in retrotransposons or flocculation genes have been excluded. \*\*The 16 deletions in all strains are the 16 ABC-transporters expected to be absent.

	Mutant 1	Mutant 2	Mutant 3	Mutant 4
Amino Acid Change	V154G	T318N	D234E (ERG25) I519M (FLO9)	H156N
Total reads	13,808,666	11,222,878	7,258,098	15,871,458
Aligned reads	13,639,301	1,1067,115	7,144,002	15,710,448
Percent Aligned Reads	98.7735	98.6121	98.4280	98.9855
Mean Coverage	109.84	89.22	57.68	126.04
Percent bases covered by 5 or more reads	99.4	99.3	99.4	99.4
Mean read length	100	100	100	100
Median insert size	249	238	234	232
Insertions relative to wild-type*	<i>URA3 (16x), CYC1 (3-4x)</i>			
Deletions relative to wild-type**	<i>adp1Δ snq2Δ ycf1Δ pdr15Δ yor1Δ vmr1Δ pdr11Δ nft1Δ bpt1Δ ybt1Δ ynr070wΔ yol075cΔ aus1Δ pdr5Δ pdr10Δ pdr12Δ</i>			

### **MMV001239 binds to *T. cruzi* CYP51**

Given that all resistance-conferring missense mutations occurred in *ERG11* or *ERG25*, two genes involved in ergosterol biosynthesis, it is very likely that MMV001239 targets this pathway. Ergosterol is a sterol component of the cellular membrane in fungi and some protozoa that plays a structural role similar to cholesterol in mammalian cells. The pathway in fungi has been a source of several important drug targets. Both the pathways and the classes of drugs that target it are summarized in (Fig 4.7).

*ERG11* is an essential gene that encodes lanosterol 14- $\alpha$ -demethylase, an cytochrome p450 enzyme that catalyzes the conversion of lanosterol by C-14 demethylation of lanosterol to form 4,4"-dimethyl cholesta-8,14,24-triene-3- $\beta$ -ol. This protein is the target of the azole drugs, a large class of antifungal small molecules that all share the same mechanism of action (Kalb et al., 1987). Each azole contains a nitrogenous heterocyclic ring that directly binds the central iron atom of the heme prosthetic group that sits in a central pocket of the enzyme. Triazoles, such as fluconazole, contain five-membered rings with 3 nitrogens and 2 carbons. Imidazole antifungals, such as ketoconazole, contain a five-membered ring with only 2 nitrogens. MMV001239 contains a 6 membered aromatic ring with a single nitrogen atom, or pyridine, which we hypothesized may bind to the central heme molecule, similar to the azole antifungals.



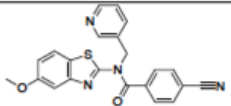
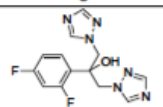
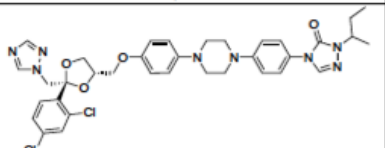
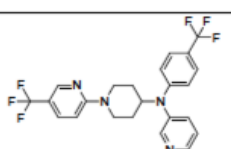
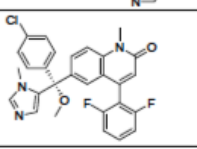
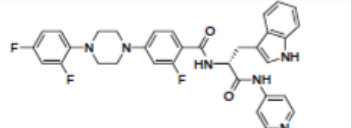
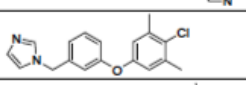
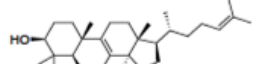
**Figure 4.7 The ergosterol biosynthetic pathway.** This diagram depicts the major steps of the ergosterol biosynthetic pathway in *S. cerevisiae*, as well as the drug classes that inhibit each step. The polyenes, which directly intercalate with ergosterol, and the azoles, which inhibit Erg11p are the most important in human medicine, and are used as staples of antifungal treatment. The allylamines such as terbinafine, accumulate in human nail tissues, so are primarily used to treat nail infections. Morpholines are used agriculturally. In the red squares the proteins encoded by genes that we found mutated in response to MMV001239 treatment. In this diagram, MMV001239 is depicted as inhibiting the same enzyme as the azoles. This diagram is adapted from (Onyewu et al., 2003).

Like fungi, ergosterol biosynthesis is also essential in *T. cruzi* (Buckner et al., 2003). In *T. cruzi*, the homologue of *ERG11* is named *CYP51*. Cyp51 is a validated drug target for *T. cruzi* and azole antifungals are known to bind and inhibit this enzyme. In addition, recombinant *TcCyp51* has been developed and used for high throughput target-based screens that have led the development of a number of next generation anti-trypanosomal small molecules, some of which

contain pyridine groups that have been shown through co-crystallography to bind to the central heme molecule of *TcCyp51* including the N-Indolyl-oxopyr-idinyl-4-aminopropanyl-based analogues and the fenarimol analogues Table 4.5 (Calvet et al., 2014). We hypothesized that our molecule most likely inhibits *TcCyp51* in a manner similar to these molecules.

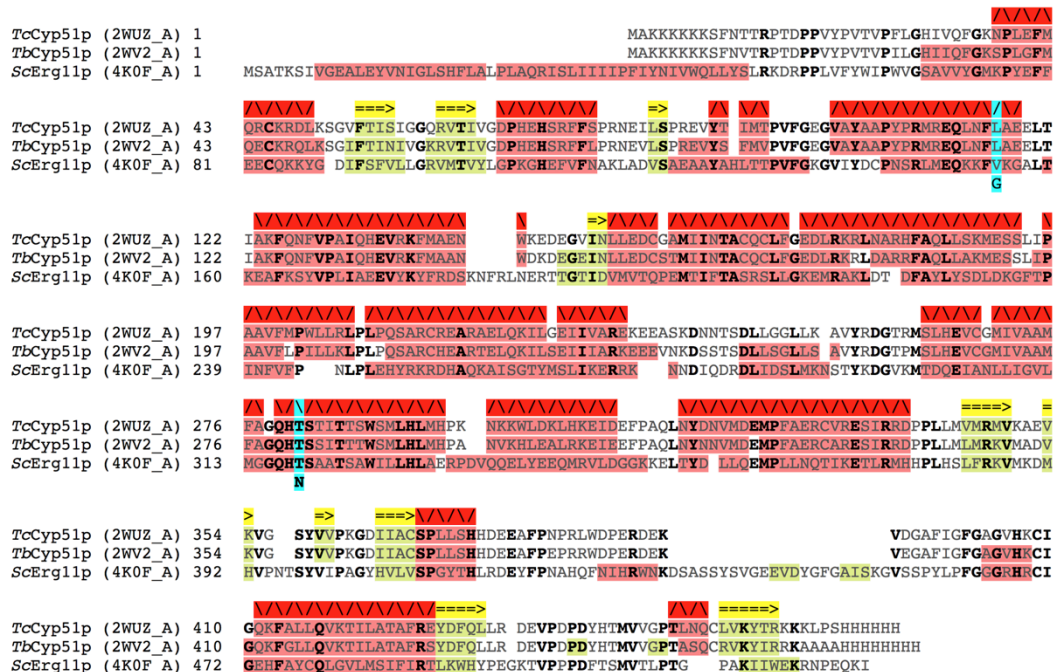
**Table 4.5 Chemical structures of varied Cyp51p/Erg11p inhibitors.**

Fluconazole and itraconazole are azole-class antifungals currently in clinical use. UDD, JKF, WVH and NEU321 are preclinical anti-trypanosomals under investigation (Vieira et al., 2014). WVH is part of the N-Indolyl-oxopyr-idinyl-4-aminopropanyl class under development at UCSF/UCSD/Scripps Florida. UDD is a fenarimol analogue under development by DNDI. NEU321 is from Northeastern University. JKF is Tipifarnib and is under development by University of Washington. Lanosterol is the native ligand of both Cyp51p and Erg11p

MMV001239	
Fluconazole	
Itraconazole	
UDD (PDBID: 3ZG3)	
JKF (PDBID: 3TIK)	
WVH (PDBID: 4C0C)	
NEU321 (PDBID: 4H6O)	
Lanosterol	

In addition, recently a library of over 100,000 compounds was screened for inhibition of *TcCyp51* (Gunatilleke et al., 2012). Of the 185 top hits, all contained aromatic nitrogenous heterocycles, including imidazoles, triazoles and pyridines, in spite of these compound these structures being present in less than 0.2% of the initial library. Furthermore, 92 of the 185 top hits (49.75%) contained pyridine rings similar to MMV001239.

A sequence alignment between *ScErg11p* and *TcCyp51p* is shown in Fig 4.8. These proteins show a high degree of amino acid conservation. In fact, *TcCYP51* complements the function of the yeast *Erg11* protein when expressed in an *Erg11*-deficient background (Buckner et al., 2003). Both identified MMV001239-resistance mutations in *ScERG11* map to highly conserved alpha helical domains. For the T318N substitution, the homologous residue in *T. cruzi* is also a threonine. The mutation from threonine to asparagine, replaces one polar uncharged amino acid for a slightly bulkier polar uncharged amino acid. This amino acid is on the side of the central enzymatic pocket and does not directly interact with the catalytic site, but is homologous to an amino acid that lies adjacent to the parts of the binding site of both the pyridine-containing experimental anti-trypanosomal small molecules and also imidazole anti-trypanosomals under development such as the University of Washington's tipifarnib analogs, Vanderbilt's VNI/VNF compounds and Northeastern University's NEU321 (Table 4.5) (Vieira et al., 2014).



**Figure 4.8 14-alpha demethylase sequence alignment.** This figure shows the amino acid sequence of Cyp51p from *T. cruzi* and *T. brucei* aligned to Erg11p from *S. cerevisiae*. The mutated amino acids found in our study are highlighted in blue. Amino acids that are conserved across all three organisms are in bold. Predicted alpha helices are in red and beta sheets are in yellow. *T. brucei* Cyp51p is included in this alignment to demonstrate its similarity to *T. cruzi* Cyp51p, since some crystal structures shown later in the paper are of *TbCyp51p*.

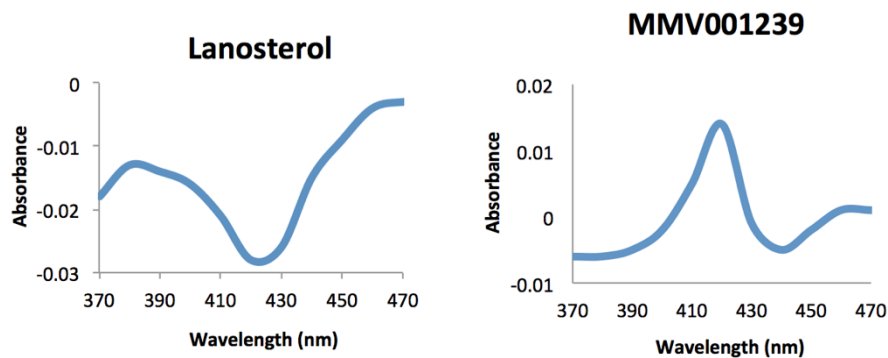
For the V154G mutation, the homologous amino acid in *T. cruzi* is a leucine. The conversion of a valine to a leucine in evolution is a conservative mutation since these are both small, aliphatic amino acids, as is the amino acid in the mutant, glycine. Examining the crystal structure, this amino acid is one of the closest to the center of the heme prosthetic group in the enzymatic core of the protein, so any changes at this site would be expected to be disruptive to the functioning of the enzyme and the binding of inhibitors to the heme molecule. Specifically, the mutation is in the middle of a highly conserved alpha helix known

as the GQHTS alpha-helical region (I-band) that runs adjacent to the site of catalysis.

*TcCYP51* is a cytochrome P450 enzyme and so has a characteristic spectrophotometric peak at 450 nm when reduced and bound to carbon monoxide. When bound to the endogenous ligand lanosterol, spectrophotometry yields a Type 1 spectrum, characterized by a peak and trough around 390 nm and 420 nm, respectively. In contrast, a flipped Type 2 spectrum is seen when an inhibitor such as an azole antifungal is bound, with a trough and peak around 390 nm and 430 nm, respectively, depending on the specific P450 enzyme being studied. Ligand binding can thus be evaluated using spectrophotometric assays (von Kries et al., 2010) (Podust et al., 2007) (Podust et al., 2007, von Kries et al., 2010).

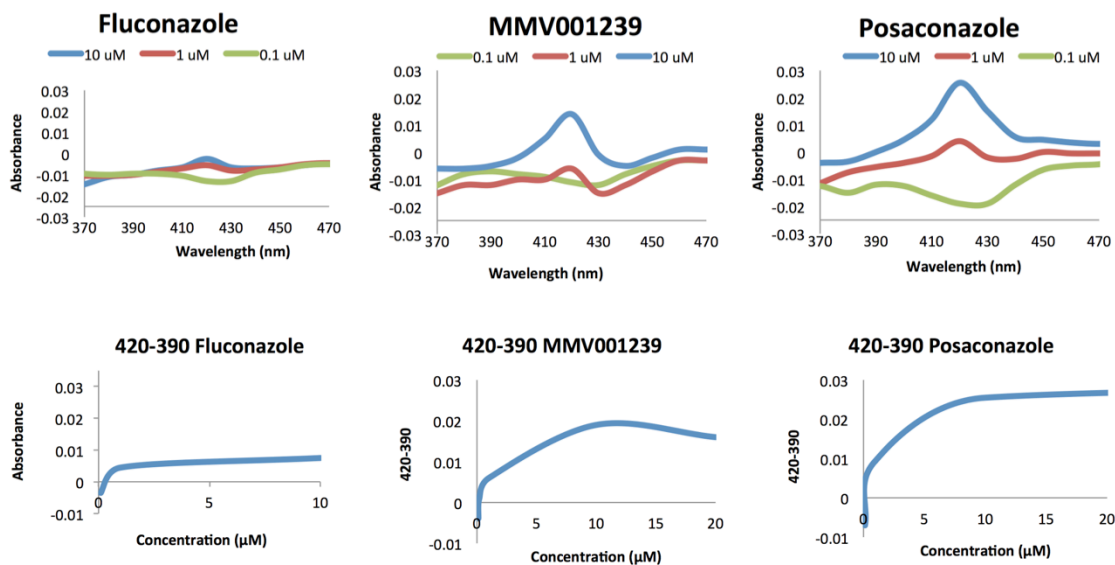
To evaluate MMV001239 binding to *TcCYP51*, we recorded the spectra of the enzyme bound to lanosterol and MMV001239 (Fig. 4.9), respectively, in 96-well format. As expected, lanosterol binding produced the characteristic Type 1 spectrum (Fig 4.9a). In contrast, MMV001239 exhibits the Type 2 spectrum characteristic of an inhibitor (Fig 4.9b). This confirmed our hypothesis that Cyp51 is the target of MMV001239 in *T. cruzi*.





**Figure 4.9 *TcCyp51p* absorbance difference spectra.** A) *TcCYP51* in complex with lanosterol shows Type 1 absorbance difference characteristic of the natural ligand of cytochrome P450 enzymes. B) MMV001239 shows Type 2 difference spectrum characteristic of P450 inhibitors. 0.5  $\mu$ l of compounds at 10  $\mu$ M in DMSO were assayed in 100 mM potassium phosphate buffer (pH 7.5) containing 10% glycerol in the absence or presence of 2.4  $\mu$ M *TcCyp51p* (Vieira et al., 2014).

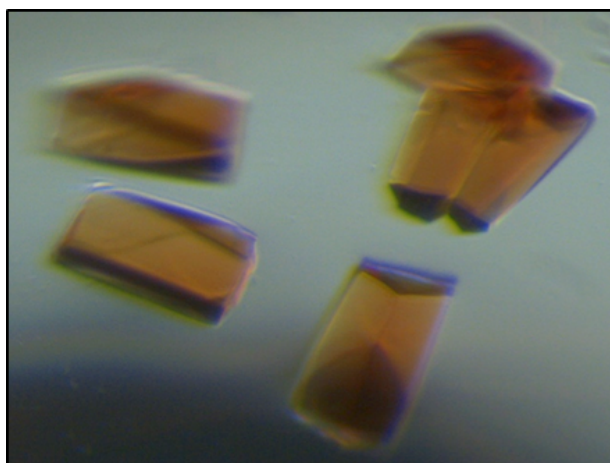
To evaluate the strength of MMV001239/*TcCYP51* binding, spectral analysis was performed across a range of concentrations for MMV001239, fluconazole (a weak inhibitor), and posaconazole (a potent inhibitor) (Fig. 4.10). By comparing the absorbance-peak amplitudes of the various spectra, we observed that the MMV001239 binding is more similar to that of posaconazole, a potent *TcCYP51* binder, rather than fluconazole, a weak *TcCYP51* binder. However, the binding constants cannot be accurately deduced from these experiments due to the sensitivity limit of the UV-visible spectroscopy performed in the medium-throughput microtiter plate format, when high protein-target-concentration limits the range of the measurement (Gunatilleke et al., 2012). Thus, we can only conclude that MMV001239 binds *TcCYP51* notably stronger than fluconazole and weaker than posaconazole.



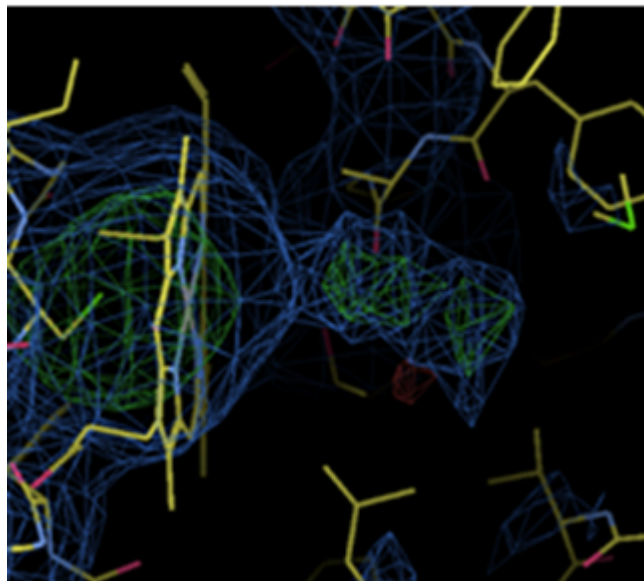
**Figure 4.10 Determining binding potency with absorbance difference spectra.** 0.5  $\mu\text{l}$  of compounds at concentrations (0.1 to 100  $\mu\text{M}$  in DMSO) were assayed in 100 mM potassium phosphate buffer (pH 7.5) containing 10% glycerol in the absence or presence of 2.4  $\mu\text{M}$  *TcCyp51p* (Vieira et al., 2014) A) To estimate the relative affinity of MMV001239 to *TcCyp51p*, binding spectra were recorded across a range of compound concentrations. The spectra were compared to fluconazole and posaconazole, which are canonical weak and strong *TcCyp51p* inhibitors, respectively. B) Absorbance values at 390 were subtracted from those at 420 and plotted against inhibitor concentration. Based on the maximum difference and the concentration at which it is reached, relative potency can be determined. MMV001239 exhibits a spectral profile more similar to a strong inhibitor than a weak inhibitor.

### Characterizing the binding site through X-ray crystallography and molecular docking

As *TcCYP51* is difficult to crystallize in the absence of a strong ligand (Choi et al., 2014), we used high-throughput methods to aggressively and exhaustively explore possible crystallization conditions. These efforts ultimately yielded *TcCYP51*-MMV001239 co-crystals (Fig. 4.11). However, despite having a nice appearance, these crystals only diffracted to a resolution of 3.8 Å, allowing us to visualize only the protein backbone and a fragment of the electron density orthogonal to the heme macrocycle (Fig 4.12). The latter suggested that MMV001239 binds via heme-iron coordination, consistent with both 1) the shift of the iron Soret band in the UV-vis binding spectra and 2) the known binding modes of other heme-iron-coordinating nitrogenous heterocycles.

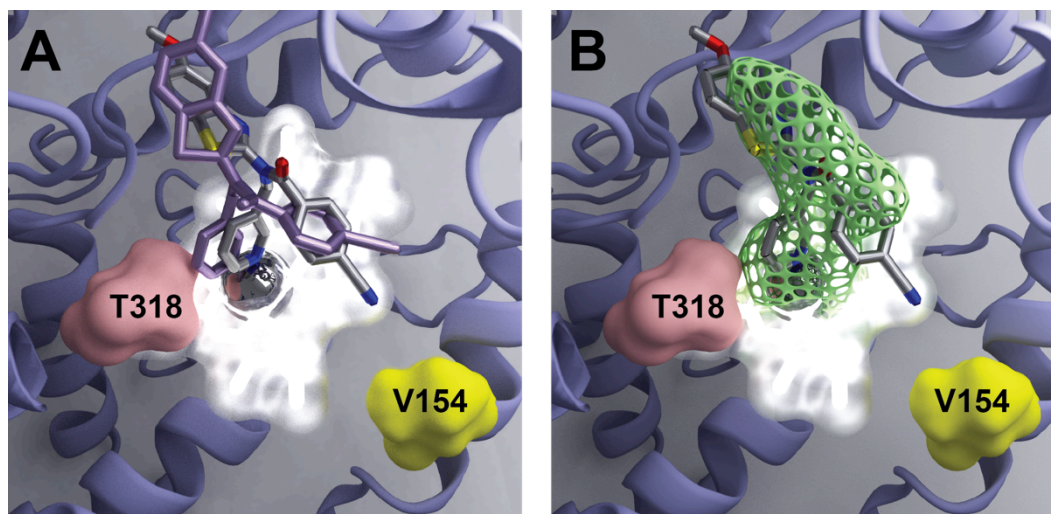


**Figure 4.11 Image of crystals used for X-RAY crystallography.** Light microscopic image of crystals from one of the wells used for imaging.



**Figure 4.12 Electron density diagram for MMV001239 in *TcCyp51*.** This image confirms the central role for the interaction between the heme prosthetic group and the nitrogen of the pyridine-methyl-benzamide group of MMV001239.

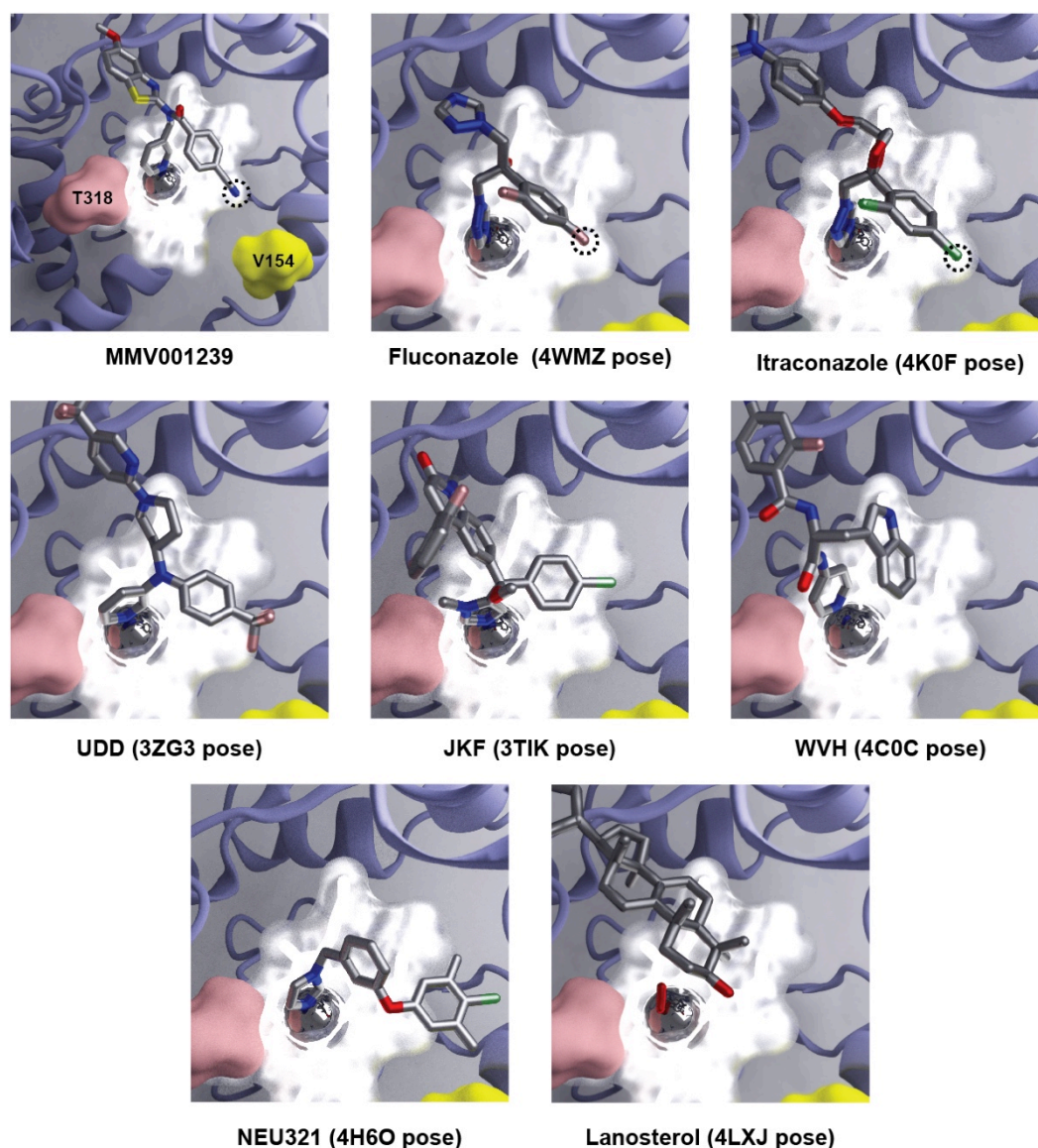
In an independent effort, we used physics-based molecular docking to predict the *ScErg11*/MMV001239 binding pose and compared that predicted pose to structures co-crystallized with lanosterol, fluconazole, and itraconazole (Fig. 4.13). The docked pose occupies the lanosterol catalytic pocket and is similar to the pose of the azole antifungals. The T318N mutation appears to sterically hinder inhibitor binding without interfering with the lanosterol-binding site. The V154G mutation may also disrupt the hydrophobic interaction between the cyanide nitrogen and the valine.



**Figure 4.13 Determining the binding pose of MMV001239 using computational docking.** A) An illustration of MMV001239 docked into CYP51 from *S. cerevisiae* and *T. brucei* (PDB IDs 4K0F and 3GW9), colored by atom name and in purple, respectively. The docked poses are similar for both homologs. For reference, the yeast protein is shown in blue, with residues that underwent directed evolution-induced changes shown as pink and yellow solid surfaces. Parts of the protein have been removed to facilitate visualization. Docking into *T. cruzi* CYP51 (PDB ID 4C27) was unsuccessful and so is omitted. B) The docked pose into the yeast homolog best matched the low-resolution crystallographic density. The electron density (in green mesh) was visualized at an isovalue of 0.04. Most of the density was removed to facilitate visualization.

To better understand how the evolved *ScERG11* mutations confer MMV001239 resistance, we mapped the evolved amino-acid changes onto *ScErg11* and *TcCyp51* crystal structures. The T318N mutation lies within a highly conserved GQHTS alpha-helical region (I-band) that is similar in both *S. cerevisiae* and *T. cruzi*, both in sequence and structure (Fig. 4.13). T318N maps to a location near the heme-iron-coordinated nitrogenous heterocycle characteristic of all known inhibitors and so may sterically hinder inhibitor binding without affecting the lanosterol binding site (Fig 4.13).

The V154G mutation lies adjacent to the inner pocket, but in a structurally conserved region that does not interact directly with the enzymatic site. However, the homologous amino acid in *T. cruzi* does lie adjacent to the crystallographic binding poses of many preclinical anti-trypanosomal CYP51 inhibitors (Fig. 4.14). We therefore hypothesize that the mechanism of MMV001239 resistance due to both mutations in Erg11 can be best explained if MMV001239 has a binding pose similar to the compounds listed in Fig. 4.14.



**Figure 4.14 Comparing the binding poses of a range of Cyp51p/Erg11p inhibitors.** In all cases, the protein shown is *S. cerevisiae* ERG11 (PDB ID 4K0F). The MMV001239 pose is docked, and the fluconazole, itraconazole, and lanosterol poses are crystallographic. Similarly, the remaining ligands were co-crystallized with the close *T. brucei* homologue CYP51. Atoms are circled to indicate that the MMV001239 nitrile nitrogen atom is a possible bioisostere for the equivalent fluconazole and itraconazole halogen atoms. Each of the inhibitors also contains a nitrogenous heterocycle that interacts with the iron atom at the center of the heme group.

## Discussion

Target identification is a major barrier to subsequent drug development that inevitably follows hit identification via phenotypic screens. Our method described here shows how directed evolution in drug-sensitive yeast can rapidly identify the molecular target of anti-trypanosomal compounds with unknown mechanisms of action. We hope this method will complement existing methods of target validation, including in vitro evolution in the parasite itself, proteomic approaches, haploinsufficiency profiling and others.

Although target identification through directed evolution can be performed in *T. cruzi*, it is very time consuming. A recent effort took eleven months, ultimately producing a single lineage with a mere 4-fold IC<sub>50</sub> shift (Khare et al., 2015). In that study, multiple independent lineages were not possible because maintaining *T. cruzi* cultures is particularly labor intensive. In comparison, directed evolution using drug-sensitive yeast to identify the MMV001239 target proved fast and effective. Selection in yeast required only 3-9 days, easily allowing for four selections. The only resistance-conferring mutations in these four lineages occurred in two genes that both encode proteins of the ergosterol biosynthetic pathway. Furthermore, half of the mutations were in the binding site of the drug's molecular target.

The high specificity of this approach is especially striking when compared with high-throughput approaches such as haploinsufficiency profiling (Winzeler et al., 1999, Giaever et al., 2002), chemical pull-downs, and other biochemical



assays, which often implicate many genes rather than just the single target (Zheng et al., 2004). We further showed that MMV001239 acts through direct binding to and inhibition of ScErg11p in *S. cerevisiae*, as well as its *T. cruzi* homologue TcCYP51. Finally, we identified a binding pose for the compound using a combination of x-ray crystallography and computational docking.

We found no mutations in genes associated with pleiotropic drug resistance, demonstrating the benefit of working with a yeast strain that lacks the major drug efflux pumps. Even if MMV001239 had been potent against an export-competent yeast strain, drug resistance may have emerged via mutations in genes associated with efflux pumps and their regulation, making target identification via directed evolution challenging because mutations in these secondary genes would not have revealed any information about the true target pathway. MMV001239 can now be added to the growing number of anti-trypanosomal CYP51 inhibitors. While its binding pose is similar to other investigational compounds, its unique structure may now be further optimized to improve potency and/or reduce toxicity, in preparation for preclinical studies.

In summary, we have shown how to use directed evolution in an engineered, drug-sensitive yeast strain to rapidly identify the target of anti-trypanosomal compounds discovered through high-throughput phenotypic screening. This approach will accelerate mechanistic studies in an era when high-throughput phenotypic screens are becoming standard in both academic and industry labs. Given the number of libraries that are being screened against this pathogen, a rapid method of drug target identification is critical to prevent a

bottleneck in the drug development process. In addition, we believe this method will be useful to other researchers studying compounds that are cytotoxic to eukaryotic cells, such as other single-celled parasites, pathogenic worms, and human cancer cells

### **Acknowledgments**

Gregory Mark Goldgof (G.M.G.) and Dr. Sabine Otilie (S.O.) designed the experiments and wrote the manuscript. Directed evolution experiments were performed by G.M.G. IC50 experiments were performed by G.M.G., S.O., Edger Vigil (E.V.), Jennifer Yang (J.Y.), Jack Schenken (J.S.), Rebecca Stanhope (R.S.), and Maximo Prescott (M.P.) Whole genome sequence analysis was performed by G.M.G., Felicia Gunawan (F.G.), and Micah J. Manary (M.J.M). Homology bioinformatics studies were designed and performed by G.M.G. Spectrophotometric assays were designed and performed by G.M.G., S.O. and Dr. Larisa Podust (L.P.). IC50s on live *Trypanosoma cruzi* were designed and performed by Laura-Isobel McCall (L.M.) and Jair Lage De Siqueira Neto (J.L.S.N.). X-ray crystallography experiments were designed and performed by L.P. and Eduardo Soares Constantino Lopes (E.S.C.L.). Computational docking studies were designed and performed by G.M.G, J.D.D. and Dr. Rommie E. Amaro (R.E.A.).

Chapter 4, in part, has been submitted for publication and may appear in the journal *ACS Chemical Biology*, 2016, Goldgof GM, Sabine Otilie, Claudia Calvet, Gareth K Jennings, Gregory M. LaMonte, Jake Schenken, Edgar Vigil,

Priyanka Kumar, Laura-Isobel McCall, Eduardo Soares Constantino Lopes, Felicia Gunawan, Jennifer Yang, Yo Suzuki, Jair Lage De Siqueira Neto, James H. McKerrow, Rommie E. Amaro, Larissa M. Podust, Jacob D. Durrant, and Elizabeth Winzeler. The dissertation author was the primary investigator and author of this paper.

## 5. The Phenyl-Amino-Methyl-Quinolinols (PAMQs) Inhibit Cyclic AMP Signaling in *P. falciparum*

### Abstract

Cyclic AMP (cAMP) is an essential secondary messenger molecule, and the cyclic AMP signaling pathway is evolutionarily conserved in most organisms, from prokaryotes to multicellular eukaryotes, including both humans, *Plasmodium*, and *Saccharomyces*. We use directed evolution and comparative chemogenomics to show that a class of highly potent antimalarial compounds, the phenyl-amino-methyl-quinolinols, or PAMQs, inhibit replication of *P. falciparum* through inhibition of the production of cAMP by adenylyl cyclase (AC). We show that this effect is selective for homodimeric ACs with the F1718 genotype and that PAMQs do not significantly affect cAMP levels in human cells. In addition, these compounds also show activity against other human pathogens that contain ACs with the F1718 genotype, including fungi, kinetoplastid parasites, such as *Trypanosoma cruzi* and *Leshmania infantum*, helminthic pathogens such as *Schistosoma mansoni*, and the mycobacterium *M. tuberculosis*. These findings suggest that PAMQs represent a new class of antimicrobial compounds that can be optimized to work against malaria and other eukaryotic and mycobacterial pathogens and that the cAMP pathway is a chemically validated target pathway.

## Introduction

Cyclic AMP, (3',5'-cyclic adenosine monophosphate) is a secondary messenger important in many biological processes that transduce signals across the plasma membrane, which result in the activation of downstream protein kinases. Cyclic AMP is synthesized from ATP by adenylyl cyclases, a family of enzymes that consists of six classes. Class III ACs include many prokaryotic and all known eukaryotic ACs and fall into two broad categories: those that are integral to the cell membrane and have at least one transmembrane helix (tmAC) and the soluble AC (sAC) that possess no transmembrane domains and are found in the cytosol, mitochondria, nucleus, and centriole (Linder, 2006) (Chen et al., 2000) (Steegborn, 2014).

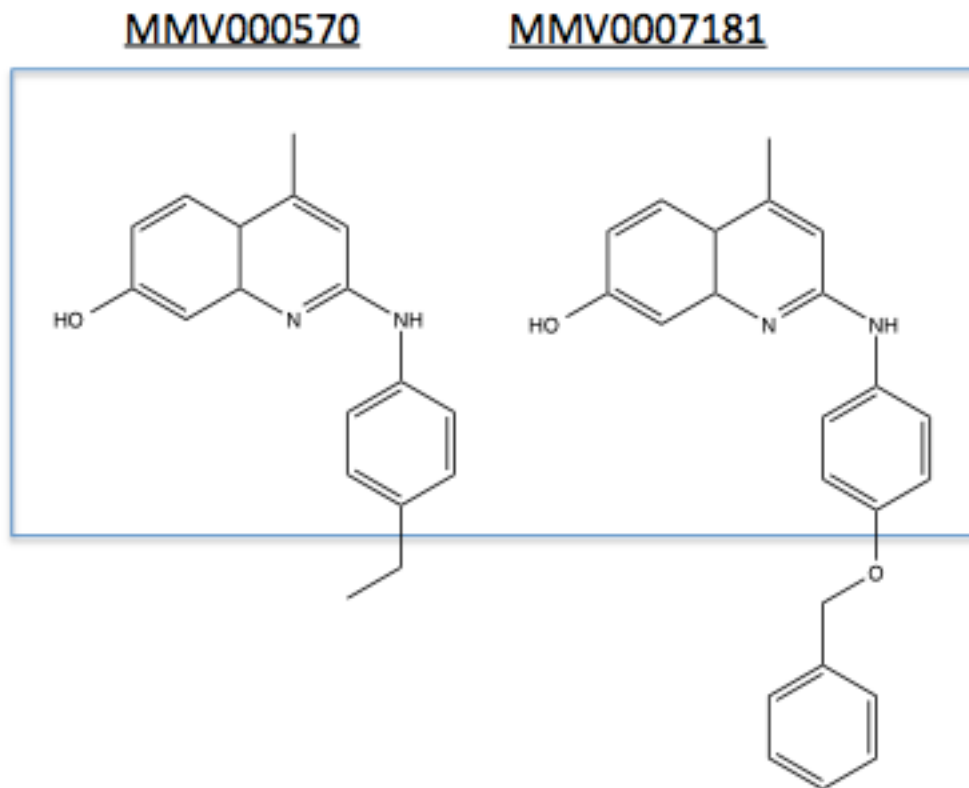
Fungal ACs are highly conserved between species and contain a single C-terminal catalytic domain, a serine/threonine protein phosphatase-like domain, a leucine zipper motifs and an N-terminal Ras-associating domain. Heterotrimeric G proteins activate some of these enzymes, even though they lack the characteristic transmembrane domains of mammalian ACs. *S. cerevisiae* contains a single adenylyl cyclase, *CYR1p*, which converts ATP into cAMP (Casperson et al., 1985). *CYR1* is an essential gene, and its deletion causes arrest in G1 (Matsumoto et al., 1982). Cyr1p activity is stimulated by binding to the small GTP-binding proteins Ras1p and Ras2p, and by the glucose-sensing G-protein-coupled receptor Gpr1p and the G-like protein Gpa2p. Increases in cAMP are stimulated in response to glucose and intracellular acidification (Linder, 2006).

Upon generation of cAMP by Cyr1p, it binds to the regulatory subunit of Protein Kinase A (PKA), Bcy1p (Cannon and Tatchell, 1987). Before binding of cAMP, PKA is an inactive tetrameric apoenzyme, composed of two regulatory subunits and two catalytic subunits, which can be encoded by one of three *TPK* genes. Binding of cAMP to Bcy1p causes the regulatory subunits to detach from the catalytic subunits. Once detached the catalytic subunit phosphorylates many different proteins crucial to signal transduction. cAMP levels are regulated by phosphodiesterases, encoded by *PDE1* and *PDE2*, which convert cAMP back into AMP.

Cyclic AMP (cAMP) signaling is also an essential process in *Plasmodium*, the parasite that causes malaria (Haste et al., 2012) and many other human pathogens (McDonough and Rodriguez, 2012). Perturbations in cyclic AMP levels in *Plasmodium* can cause cell death and interfere with pathogenic processes including chemotaxis, sexual commitment and erythrocyte invasion (Dawn et al., 2014) (Choi and Mikkelsen, 1990, Read and Mikkelsen, 1990). For this reason, phosphodiesterase inhibitors such as sildenafil, the erectile dysfunction medication better known by its trade name Viagra, have been proposed as a way to reduce the severity of malaria infection, however, their lack of selectivity for the parasitic enzymes may result in human toxicity that prevents such use (Martins et al., 2012).

The PAMQs that we chose to study are members of the open access MMV Malaria Box, a set of 400 compounds with potent blood stage anti-malarial activity that have been made publically available to promote open source drug

discovery (Spangenberg et al., 2013). As a result of the open-access nature of the MMV Malaria Box, these compounds have been tested against wide range of pathogenic organisms, which is how MMV007181 and MMV000570 (Fig. 5.1) were originally identified as having broad activity against a range of human pathogens (Table 5.1).



**Figure 5.1 Structure of PAMQs.** MMV000570 and MMV007181 share a common substructure that includes a phenyl-amino-methyl-quinolinol (PAMQ) structure. They differ in that MMV007181 has a methoxybenzyl group where MMV000570 has an ethyl group at the bottom of the above image. MMV007181 was discovered through our previous blood stage screen in collaboration with Novartis and has an IC<sub>50</sub> of 74nM against 3D7 and 12nM against W2 (PubChem CID:20867078) (Plouffe et al., 2008). MMV000570 was identified through a blood stage screen performed by a group at St. Jude's Children's Hospital and has an IC<sub>50</sub> of 684nM against 3D7 and 520nM against K1 (PubChem CID: 45489345) (Guiguemde et al., 2010). The discrepancy in potency may in part be to the method of screening. Side-by-side testing against the Dd2 parasite performed in our lab revealed an IC<sub>50</sub> of 41.2nM ± 10.8 for MMV000570 and 22.5nM ± 3.6 for MMV007181.



**Table 5.1 Summary of PAMQ activity against varied pathogens.** This table summarizes the whole cell screens that have been performed with MMV000570 and MMV007181. These compounds demonstrate broad activity against a range of human pathogens, with MMV007181 having broader activity than MMV000570, including toxicity against *Schistosoma* somules and adults, as well as *Mycobacterium*. MMV007181 is also more toxic to human liver cell lines.

\*Source: personal correspondence (data not shown).

Assay ID PubChem	Assay Source	Assay Organism	MMV000570 IC50 (uM)	MMV007181 IC50 (uM)	AC Type	AC F1718 Genotype ?
n/a	Goldgof	S cer ABC16-Monster	4.74	3.51	Homodimer	Yes
660866	MMV	P fal 3D7	0.194	0.684	Homodimer	Yes
660867	MMV	P fal K1	0.016	0.52	Homodimer	Yes
449704	Novartis	P fal W2	no data	0.012	Homodimer	Yes
449703	Novartis	P fal 3D7	no data	0.074	Homodimer	Yes
660868	MMV	T cruzi	3.72	4.39	Homodimer	Yes
660870	MMV	T brucei rhodesiense	11.89	3.54	Homodimer	Yes
660869	MMV	T brucei brucei	12.22	11.89	Homodimer	Yes
660871	MMV	L infantum	4.32	16.2	Homodimer	Yes
1159402	MMV	T gondi	Inactive	Inactive	Homodimer	No
1159401	MMV	E histolytica	Inactive	Inactive	?	No
1159414	MMV	Babesia	no data	3.45	?	No
1159413	MMV	Babesia	no data	0.17	?	No
1159416	MMV	Babesia	no data	4.91	?	No
1159415	MMV	Babesia	no data	6.31	?	No
660872	MMV	MRC-5 (Human Fibroblasts)	7.77	14.02	Heterodimer	No
685525	MMV	HepG2 (Human Hepatocytes)	no data	8.16	Heterodimer	No
449705	Novartis	Huh7 (Human Hepatocytes)	no data	17.2	Heterodimer	No
1079819-21	MMV	Onchocerca lienalis	Inactive	Inactive	?	No
685521-24	MMV	Mtb	inactive	active		Yes
n/a	Caffrey Lab*	Schistosoma mansoni somules	active	active	?	Yes

## Results

### Directed Evolution

MMV000570 has an IC<sub>50</sub> of  $7.0\mu\text{M}\pm 0.79$  against ABC<sub>16</sub>-Monster and a an IC<sub>50</sub> of  $47.3\mu\text{M}\pm 2.39$  for the wild type strain, suggesting that directed evolution was much better suited to the ABC<sub>16</sub>-Monster strain. Since the minimum inhibitory concentration (MIC) was around  $30\mu\text{M}$ , we performed three selections at that dosage and one at a higher dosage of  $50\mu\text{M}$ . All four selections grew after only one day, suggesting only weak selection. However, clones isolated from three out of four cultures demonstrated 4-6x IC<sub>50</sub> shifts relative to the parent strain. DNA was isolated from the three resistant strains, independently selected clones and prepared for sequence analysis.

MMV007181 has an IC<sub>50</sub> of  $2.48\mu\text{M}$  against ABC<sub>16</sub>-Monster. Selections were performed at 4, 7, 10 and  $25\mu\text{M}$ , which grew in 1, 4, 5, and 17 days, respectively. Clones isolated from the first two lineages each demonstrated approximately 10-fold resistance and so were selected for sequencing. Table 4.2 summarizes the selections that were performed.

**Table 5.2 A summary of PAMQ selections.** In total, four selections were performed for each PAMQ compound. Five of eight tubes selected for drug resistant clones after a single round of drug selection. Four out of these five grew in only a single day. Increased length of drug selection did not correlate with improved chances selecting for a resistant clone, as only one of three selections that took greater than one day to grow resulted in isolation of a drug resistant strain. The 17 days it took MMV007181-R2 to grow most likely are the result of the significant growth inhibition and cytotoxicity of the compound on non-resistant cells.

Lineage	Compound	Selection Concentration (uM)	Days To Grow	Drug Resistance Shift
R1	MMV000570	30	1	5.6x
R2	MMV000570	30	1	none
R3	MMV000570	30	1	5.8x
R4	MMV000570	50	1	4.8x
R1	MMV007181	10	5	None
R2	MMV007181	25	17	None
R3	MMV007181	4	1	9.8x
R4	MMV007181	7	4	9.7x

### Whole Genome Sequencing Results

All five resistant strains were sequenced to a read depth of 60-100x with 99.4% of bases covered by five or more reads. Table 5.3 summarizes the results of the sequencing. Each resistant clone had only one or two mutations that resulted in changes in amino acids. No insertions or deletions were detected. The most striking finding was that five out of five resistant strains had mutations in cAMP binding proteins and each mutation affected the part of the protein known to interact with cAMP. Table 4.4 provides summaries of the gene description for each gene that was mutated. Table 4.5 provides basic quality control statistics for the sequencing that was performed.

**Table 5.3 Summary of directed evolution sequencing results.** This table contains all detected mutations that confer changes to amino acids sequences. Mutations detected in retrotransposons and flocculation genes were not included, since they are often nonspecific and are either highly variable or appear to be so as a result of sequencing artifact. No insertions or deletions, relative to the parent strain were detected. \*Genes that interact with cAMP are in **bold**. \*\* Amino acid changes that change cAMP binding sites are in **bold**.

	Mutant 1	Mutant 2	Mutant 3	Mutant 4	Mutant 5
Compound	MMV000570	MMV000570	MMV000570	MMV007181	MMV007181
Lineage	R1	R3	R4	R3	R4
Mutated Genes*	<b>PDE2</b>	<b>CYR1</b>	<b>BCY1</b>	<i>YMR310C</i> , <b>PDE2</b>	<i>DAD2</i> , <b>PDE2</b>
Amino Acid Change**	<b>W278-Stop</b>	<b>F1718V</b>	<b>R275M</b>	R113T, <b>H335D</b>	T97A, <b>D400E</b>

**Table 5.4 Summary of genes with amino acid changes in drug resistant clones.** This table provides summaries of the gene description for each gene that was mutated (Source: *Saccharomyces Genome Database*). Genes in the cyclic AMP second messenger pathway are in **bold**.

Gene Name	Description
<b>BCY1</b>	<b>cAMP-dependent protein kinase regulatory subunit. Binding of cAMP to Bcy1p causes it to release from the catalytic subunit of PKA. Once unbound, the catalytic subunit becomes active and phosphorylates a range of proteins</b>
<b>CYR1</b>	<b>Adenylate cyclase. Converts ATP to cAMP.</b>
<i>DAD2</i>	Essential subunit of the Dam1 complex (aka DASH complex); complex couples kinetochores to the force produced by MT depolymerization thereby aiding in chromosome segregation; is transferred to the kinetochore prior to mitosis
<b>PDE2</b>	<b>3',5'-cyclic-nucleotide phosphodiesterase. One of two genes encoding PDEs that break down cAMP into AMP</b>
<i>YMR310C</i>	Putative methyltransferase

**Table 5.5 Sequencing statistics.** This table provides basic quality control statistics for the sequencing that was performed. \*, \*\* All of the observed insertions and deletions relative to wild-type are those expected in the ABC<sub>16</sub>-Monster strain. We include this data as validation that the bioinformatics pipeline is working.

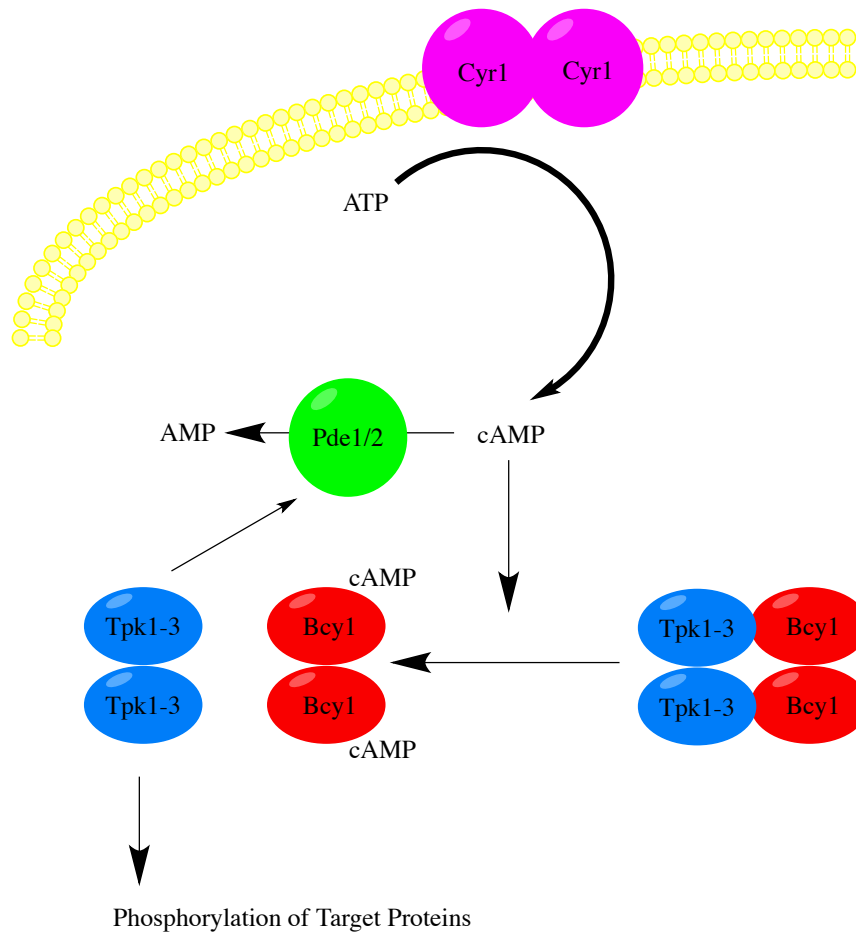
	Mutant 1	Mutant 2	Mutant 3	Mutant 4	Mutant 5
Total reads	9,961,476	10,346,606	8,315,002	9,424,474	11614328
Aligned reads	9,798,766	10,157,962	8,262,729	9,256,282	11280732
Percent Aligned Reads	98.3666	98.1768	99.3713	98.2154	97.1277
Mean Coverage	79.28	81.77	66.65	74.63	91.01
Percent bases covered by 5 or more reads	99.4	99.4	99.4	99.4	99.4
Mean read length	100	100	100	100	100
Insertions relative to wild-type*	<i>URA3</i> (16x), <i>CYC1</i> (3-4x)	<i>URA3</i> (16x), <i>CYC1</i> (3-4x)	<i>URA3</i> (16x), <i>CYC1</i> (3-4x)	<i>URA3</i> (16x), <i>CYC1</i> (3-4x)	<i>URA3</i> (16x), <i>CYC1</i> (3-4x)
Deletions relative to wild-type**	<i>adp1Δ</i> <i>snq2Δ</i> <i>ycf1Δ</i> <i>pdr15Δ</i> <i>yor1Δ</i> <i>vmr1Δ</i> <i>pdr11Δ</i> <i>nft1Δ</i> <i>bpt1Δ</i> <i>ybt1Δ</i> <i>ynr070wΔ</i> <i>yol075cΔ</i> <i>aus1Δ</i> <i>pdr5Δ</i> <i>pdr10Δ</i> <i>pdr12Δ</i>	<i>adp1Δ</i> <i>snq2Δ</i> <i>ycf1Δ</i> <i>pdr15Δ</i> <i>yor1Δ</i> <i>vmr1Δ</i> <i>pdr11Δ</i> <i>nft1Δ bpt1Δ</i> <i>ybt1Δ</i> <i>ynr070wΔ</i> <i>yol075cΔ</i> <i>aus1Δ</i> <i>pdr5Δ</i> <i>pdr10Δ</i> <i>pdr12Δ</i>	<i>adp1Δ</i> <i>snq2Δ</i> <i>ycf1Δ</i> <i>pdr15Δ</i> <i>yor1Δ</i> <i>vmr1Δ</i> <i>pdr11Δ</i> <i>nft1Δ</i> <i>bpt1Δ</i> <i>ybt1Δ</i> <i>ynr070wΔ</i> <i>yol075cΔ</i> <i>aus1Δ</i> <i>pdr5Δ</i> <i>pdr10Δ</i> <i>pdr12Δ</i>	<i>adp1Δ</i> <i>snq2Δ</i> <i>ycf1Δ</i> <i>pdr15Δ</i> <i>yor1Δ</i> <i>vmr1Δ</i> <i>pdr11Δ</i> <i>nft1Δ</i> <i>bpt1Δ</i> <i>ybt1Δ</i> <i>ynr070wΔ</i> <i>yol075cΔ</i> <i>aus1Δ</i> <i>pdr5Δ</i> <i>pdr10Δ</i> <i>pdr12Δ</i>	<i>adp1Δ</i> <i>snq2Δ</i> <i>ycf1Δ</i> <i>pdr15Δ</i> <i>yor1Δ</i> <i>vmr1Δ</i> <i>pdr11Δ</i> <i>nft1Δ</i> <i>bpt1Δ</i> <i>ybt1Δ</i> <i>ynr070wΔ</i> <i>yol075cΔ</i> <i>aus1Δ</i> <i>pdr5Δ</i> <i>pdr10Δ</i> <i>pdr12Δ</i>

Three of the five clones had mutations in phosphodiesterase 2 (*PDE2*), the enzyme responsible for cAMP degradation and negative feedback in the canonical signaling pathway. Both MMV007181 resistant mutants had mutations that cause amino acid changes in the cAMP-binding domain of Pde2p. In addition, one of the MMV000570 resistant clones (R1) had as its only coding mutation, an early stop codon that truncates Pde2p so that it does not contain the domain that binds cAMP. This mutation should reduce the cell's ability to breakdown cAMP, suggesting that the compound treated cells are experiencing a reduction in cAMP levels that needs to be compensated for.

The other two MMV000570 resistant clones also had mutations that affect cAMP signaling. R2's only coding mutation is in adenylyl cyclase (*BCY1*), also known as adenylate cyclase, the enzyme that produces cAMP from ATP. The mutation is found in the core of the homodimer adjacent to where ATP is converted to cAMP. Adenylyl cyclase is the only gene in the cAMP pathway that is considered essential for cell survival, presumably due to redundancy in the genes that encode the target of the second messenger, protein kinase A.

R3's only coding mutation is in the regulatory subunit of protein kinase A (*BCY1*), also known as *PKA-R*. When unbound to cAMP, PKA-R binds to the catalytic subunit of PKA, also known as PKA-C, which can be encoded by one of three redundant genes, named *TPK1-3*. When cAMP binds PKA-R, they are released from PKA-C, enabling it to phosphorylate its downstream partners. cAMP bound PKA-C has an activating effect on phosphodiesterase, enabling a negative feedback loop to regulate the amount of signal produced by this

messenger pathway (Conrad et al., 2014). The pathway is summarized in figure 5.2.



**Figure 5.2 The cyclic AMP second messenger system in *S. cerevisiae*.** The adenylyl cyclase, Cyr1, converts ATP to cAMP. Cyclic AMP then binds to the Protein Kinase A tetramer, which is composed of two inhibitory subunits (Bcy1) and two catalytic subunits (encoded by *TPK1*, *TPK2* and *TPK3*). Once a cAMP binds each of the two Bcy1 molecules, they detach from the catalytic subunits, allowing them to phosphorylate downstream messengers. One of these downstream targets includes phosphodiesterase (encoded by *PDE1* and *PDE2*). Pde1 and Pde2 degrade cAMP, creating a negative feedback loop that regulates the messaging system (Santangelo, 2006).

Growth, metabolism, and nutrient sensing (especially glucose), are regulated by this pathway. Adenylate cyclase, (Cyr1 in pink), activation is dependent on localization to the membrane and binding to RAS proteins. Activation allows the homodimer to convert ATP to cAMP. cAMP may then bind to the regulatory subunit (Bcy1 or PKA-R in red) of Protein Kinase A (PKA), also known as cAMP-dependent Protein Kinase. Once bound to both Bcy1 subunits of the tetrameric PKA have bound cAMP, they release the catalytic subunits (PKA-C in blue), encoded by genes *TPK1-3*, which then phosphorylate serine and threonine residues on a range of target proteins. One such set of target proteins are the phosphodiesterases (encoded by *PDE1 and PDE2* in green), which are activated by PKA-C. These enzymes break the cyclic phosphodiester bond of cAMP, converting it to AMP and providing negative feedback on the signaling pathway.

None of the PKA-C genes are essential to cell survival, most likely due to their redundancy. In addition, PKA-R is not essential, suggesting that under normal growth conditions, excessive cAMP signaling is not fatal in *S. cerevisiae* or that the cell might have other mechanisms for compensating. The fact that functional knockout of *PDE2* confers resistance to MMV000570 suggests that these compounds repress cAMP signaling, either through inhibition of its production or through inhibition of its downstream signal. If production by adenylyl cyclase is inhibited, we would expect a decrease in cellular cAMP levels in compound treated cells. However, if a downstream signal, such as PKA-C, is

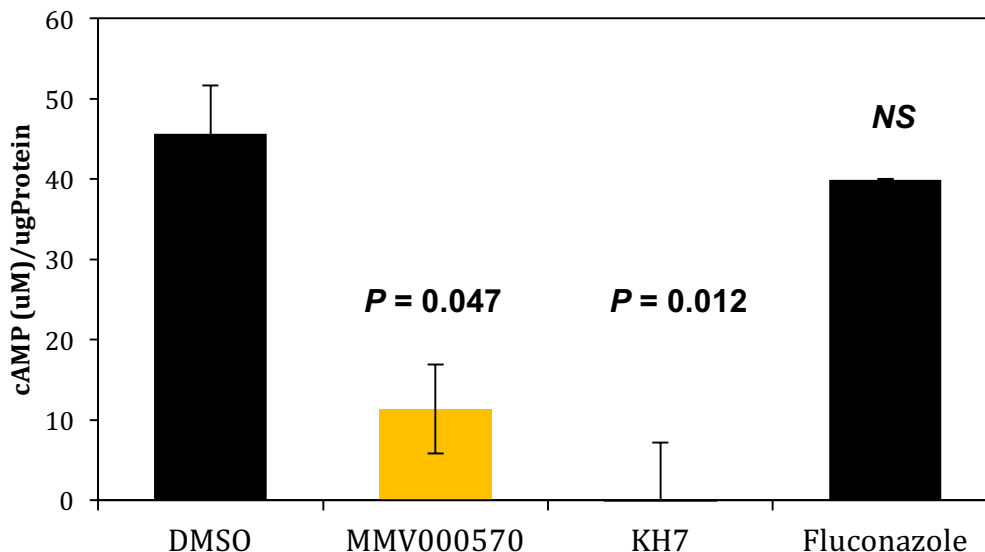


inhibited, we would expect a build up of cAMP in the cell as a compensation mechanism.

### **PAMQs Inhibit Intracellular cAMP Signaling**

To evaluate the hypothesis that the PAMQs inhibit cAMP signaling, we decided to measure intracellular cAMP levels after treatment with MMV000570 using the cAMP-glo assay (Kumar et al., 2007). Cells were treated with 100uM MMV000570 for four hours in rich media and then compared with DMSO controls. As a positive control, the adenylyl cyclase inhibitor KH7 was used. KH7 strongly inhibits (class III) adenylyl cyclases across a range of species, including humans, *Plasmodium* and *S. cerevisiae* (Salazar et al., 2012). As a negative control, we used DMSO. For the yeast assays, we used fluconazole as a control for cytotoxicity, an antifungal that inhibits ergosterol biosynthesis. This cytotoxic control was used to make sure that a decrease in cAMP signaling is not a general byproduct of poisoning yeast cells. Each of the drug conditions was tested at 100µM.

### *S. cerevisiae* cAMP Levels



**Figure 5.3** The effect of PAMQs on cAMP levels in *S. cerevisiae*. DMSO was used as a negative control, KH7 as a positive control, and fluconazole as a control cytotoxic compound with a mechanism of action unrelated to cAMP signaling. Both MMV000570 and KH7 showed significant inhibition of intracellular cAMP levels. All compounds were tested at 100µM.

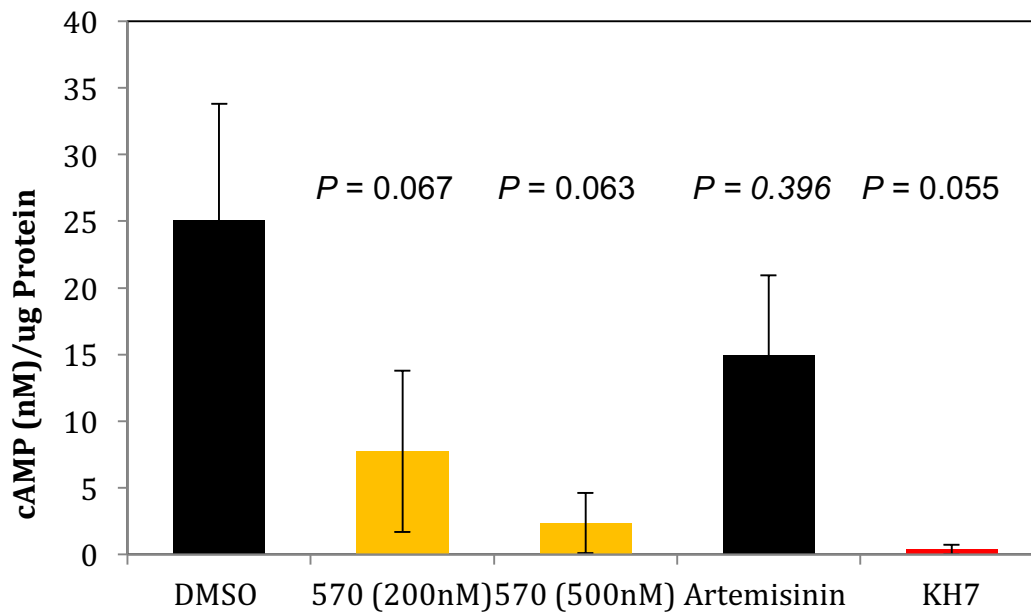
Treatment with MMV000570 induced a 75.1% reduction in intracellular cAMP levels ( $P = 0.047$ ) (Fig. 5.3). The positive control showed 100% reduction in cAMP levels ( $P = 0.012$ ). Treatment with fluconazole resulted in a 12.6% reduction in intracellular cAMP levels, relative to the DMSO control, which was not statistically significant (NS).

To verify that this inhibition in cAMP signaling also occurs in *P. falciparum*, we next repeated the assay in parasites isolated from infected human red blood cells. Human erythrocytes have much higher cAMP levels than *P. falciparum* parasites, so separating them is a necessary step for determining cAMP

inhibition for any compound that selectively inhibits parasite adenylyl cyclase, and not the host's ability to produce cAMP.

In this experiment, MMV000570 was tested at two different concentrations to determine whether or not the effect on intracellular cAMP is dose dependent. Since fluconazole is not cytotoxic to *Plasmodium*, artemisinin was used as a cytotoxic control. Once again, KH7 was used as a positive control. Treatment with 200nM MMV000570 resulted in a 69.2% reduction in the intracellular cAMP levels ( $P = 0.067$ ) (Fig. 5.4). Treatment with 500nM MMV000570 resulted in a 90.7% reduction in intracellular cAMP levels ( $P = 0.063$ ). KH7 treatment resulted in a 98.6% drop in cAMP levels ( $P = 0.055$ ). Artemisinin treatment induced a 40% drop in cAMP levels ( $P = 0.396$ ). This drop may suggest that general cytotoxicity causes some decrease in intracellular cAMP levels in *Plasmodium*. The relatively high  $P$ -values were due to high variability in the DMSO control. This experiment will have to be repeated as a result. However, the trend suggests dose-dependent inhibition of intracellular cAMP signaling in *Plasmodium* and a conserved mechanism of action for the PAMQs between fungi and apicomplexa.

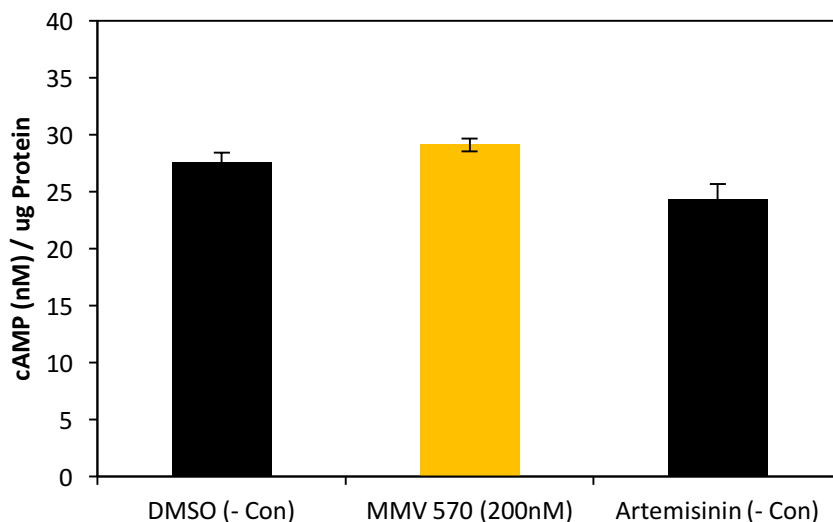
### *P. falciparum* cAMP Levels



**Figure 5.4** The effect of PAMQs on cAMP levels in *P. falciparum*. DMSO was used as a negative control, KH7 as a positive control, and artemisinin as a control cytotoxic compound with a mechanism of action unrelated to cAMP signaling. Both MMV000570 and KH7 showed inhibition of intracellular cAMP levels, although high *P-values* occurred due to internal variability.

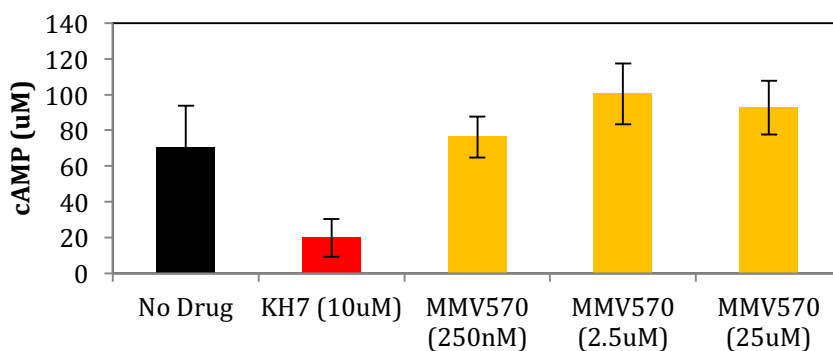
However, PAMQs will not be useful as a therapeutic lead compounds unless they have some degree of selectivity for parasite, versus human, cAMP inhibition. Consequently, we next tested the effects of MMV000570 against human red blood cells. We also wanted to discount the possibility that the effect we were observing was due to effects on host cAMP levels, since the parasites were isolated from drug treated red blood cells. Figure 5.4 demonstrates that MMV000570 has no significant effect on human red blood cells. Figure 5.5 further shows that MMV000570 has no effect on intracellular cAMP levels in human liver cells (Huh7 cell line) even at 125 times the dosages that we saw an effect in *P. falciparum*.

### Human Erythrocyte cAMP Levels



**Figure 5.5** The effect of PAMQs on cAMP levels in human red blood cells. MMV000570 does not inhibit cAMP signaling in human red blood cells.

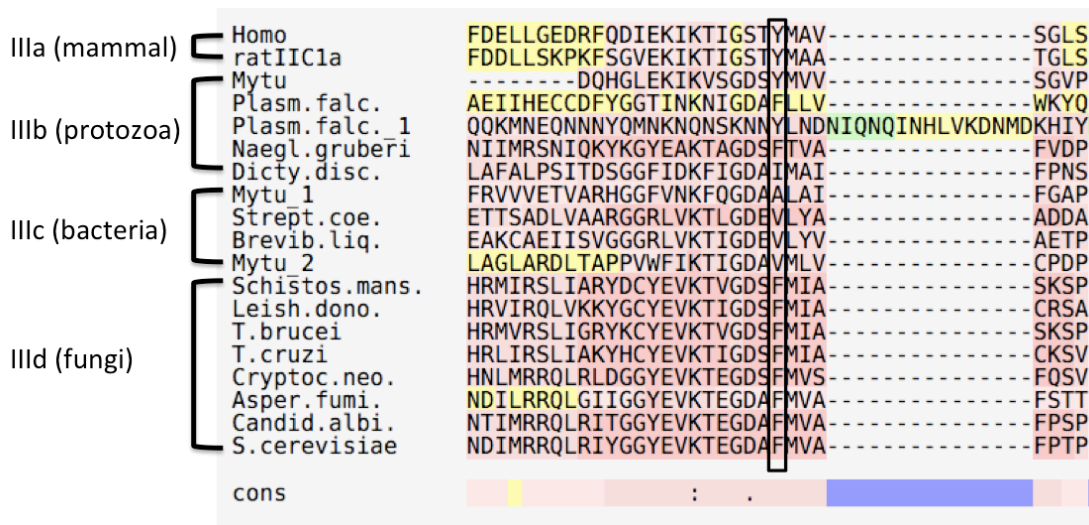
### Human Hepatocyte cAMP Levels



**Figure 5.6** The effect of PAMQs on cAMP levels in human liver cells. MMV000570 does not inhibit cAMP signaling in human liver cells from the Huh7 cell line. DMSO was used as a no drug control. KH7 was used as a positive control.

### **Evolutionary Sequence Analysis and Structural Biology**

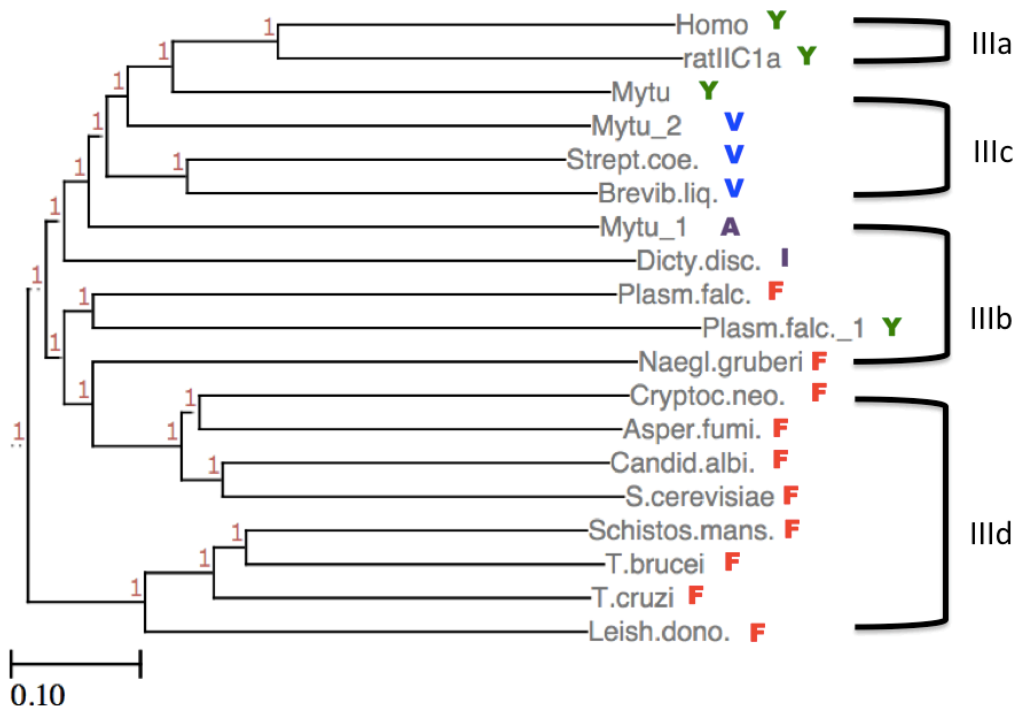
One explanation for how the compound reduces intracellular cAMP levels is through direct inhibition of adenylyl cyclase. By mapping the F1718 mutation in adenylyl cyclase onto a multiple sequence alignment we were able to identify a possible explanation for the selectivity of the compounds in different species (Fig 5.7). Many of the organisms the compound has activity against (Table 5.1) also have a phenylalanine at the homologous position.



**Figure 5.7 Sequence alignments of class III adenylyl cyclases.** M-Coffee was used to perform a multiple sequence alignment between class III ACs from a range of organisms (Wallace et al., 2006). In the black box are the amino acids that are homologues to the mutated amino acid that confers resistance to MMV000570. As follows are the information for each sequence formatted as “species:name in figure [gene] (NCBI Protein ID).” *H. sapiens*:Homo [ADCY2] (CAA52282), *R. norvegicus*:ratIIC1a [ADCY2] (P26769), *M. tuberculosis* Rv1625c:Mytu [CYA1 (O30820), *P. falciparum*:Plasm.falc. [gc-alpha] (AJ245435), *P. falciparum*:Plasm.falc. 1 [gc-beta] (AJ249165), *N. gruberi*:Naegl.gruberi [Adenylate Cyclase] XP\_002669895.1, *D. discoideum*:Dicty.disc. [Adenylyl Cyclase] AAD50121.1, *M. tuberculosis* Rv3645:Mytu 1 (CAB08835), *S. coelicolor*:Strept.coe. [CYA] (P40135), *B. liquifaciens* [CYA] Brevib.liq. (P27580), *M. tuberculosis* Rv1264:Mytu 2 (CAB00890), *S. mansoni*:Schistos.mans. [CAY18707] (CAY18707.1), *L. danovani*:Leish.dono. [RAC-A] (Q27675), *T. brucei*:T.brucei [S14201] (S14201), *T. cruzi* [Adenylate Cyclase] (AAC32776.1), *C. neoformans*:Cryptoc.neo. [Adenylate Cyclase] (XP\_012050872.1), *A. fumigatus*:Asper.fumi. [ACyA] KMK63181.1, *C. albicans*:Candid.albi. [Adenylate Cyclase] (CAB60230.2), *S. cerevisiae*:S.cerevisiae [Adenylate Cyclase] NP\_012529.3.

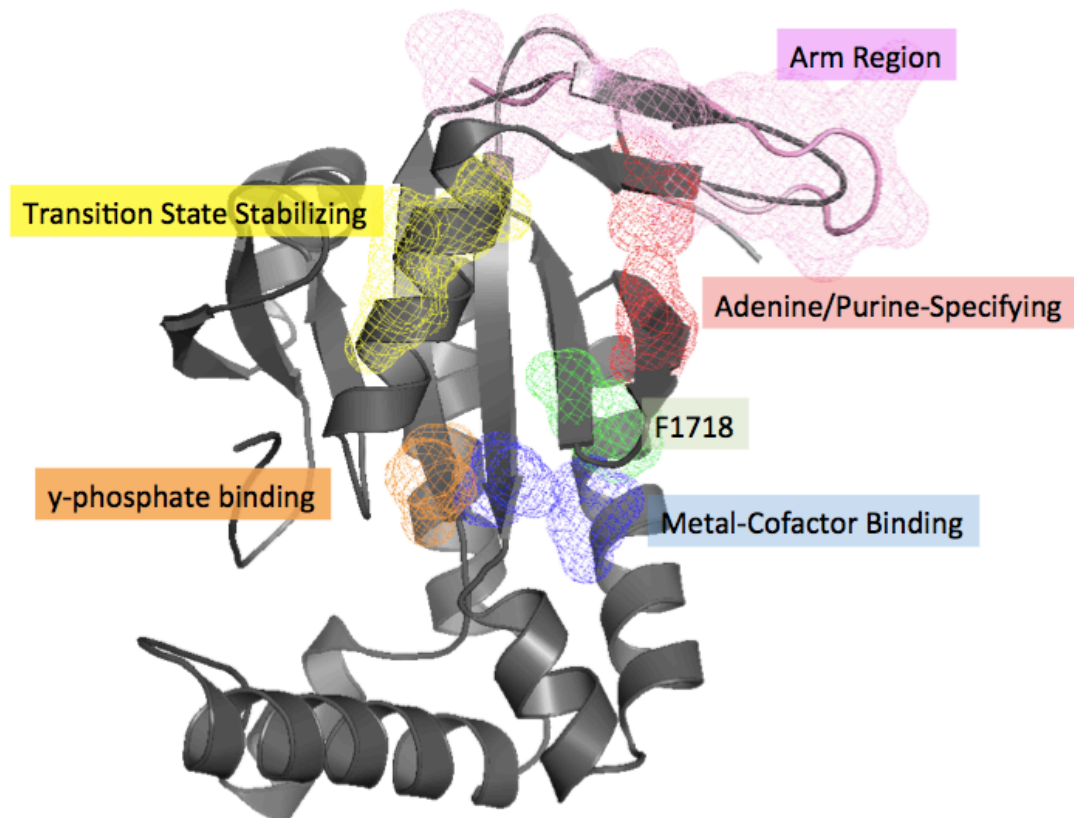


Class III adenylyl cyclases can be organized into 4 subclasses, or clades, a-d. Clade a includes mammalian ACs, clade b includes apicomplexa such as *Plasmodia*, clade c includes most prokaryotes and clade d include fungi and kinetoplastids. The mutation in AC we identified that confers resistance to our compounds is F1718V. This site is highly conserved within clades and roughly separates them. Clade a commonly contains a tyrosine (Y) at this site, clade b a leucine (L), clade c a valine (V) and clade d a phenylalanine (F).



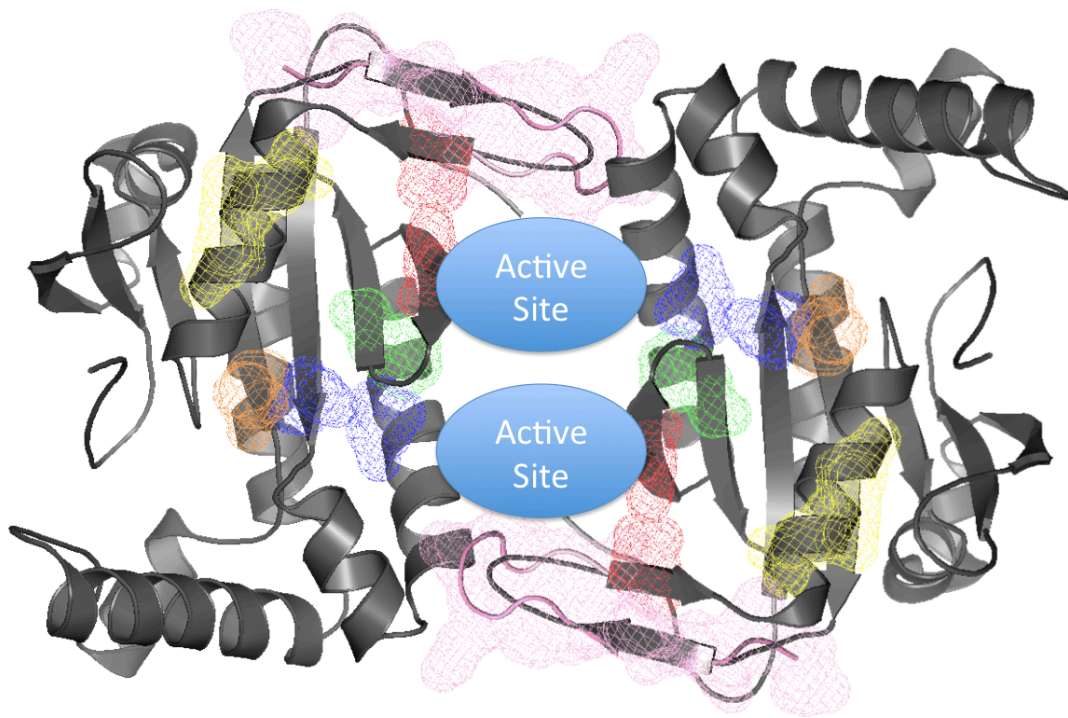
**Figure 5.8 Phylogenetic tree of class III adenylyl cyclases.** The colored letters to the right of each species name indicates the amino acid at the position homologous to the F1718V mutation. The naming species naming convention is the same as figure 5.7. The alignment from Fig. 5.7 was converted into a tree using TreeView (Huerta-Cepas et al., 2010).

In fact, many of the organisms that have susceptibility to this compound encode a phenylalanine at this site, including *Schistosoma mansoni* (Fig 5.8) It is interesting to note that the identified mutation converts the phenylalanine residue to a valine residue, which is typical of clade c, or prokaryotic ACs. Based on this, we would not expect PAMQs to be active against bacteria or organisms with a valine at the site homologous to F1718.



**Figure 5.9 Structural biology of adenylyl cyclase monomer.** The crystal structure is from a *T. brucei* adenylyl cyclase (RCSB: 1FX4). In green is the amino acid homologous to F1718. Other important domains are highlighted. This image was produced using PyMol (Bieger and Essen, 2001, Linder and Schultz, 2003, PyMOL, 2016 #359, Bairoch et al., 2005).

The amino acid at the F1718 homologous site also correlates with the structure of ACs. Mammalian ACs, which have a tyrosine at the homologous site are formed by a heterodimer encoded by multiple genes and a single enzymatic site. On the other hand, the ACs of fungi, apicomplexa and kinetoplastids are homodimers with two symmetrical active sites. The active sites of all class III ACs sits at interface of the two monomers of the dimer. This correlation is unsurprising, since the 1718 amino acid sits at the interface between the two homodimers, so that a mutation at this site would cause two changes to the interface site, one from each monomer.



**Figure 5.10 Structural biology of adenyl cyclase homodimer.** The above diagram depicts the structure of the dimerized AC. The two symmetrical active sites can be seen on either side of the mutated amino acids (in green). No crystal structure is available for the AC in the dimerized form. This is meant as a representational figure. Other important domains are highlighted using the same scheme as Fig 5.9.

The mutation also lies adjacent to both enzymatic sites. The simplest explanation for the compound's inhibition of cAMP production is that it binds to AC's active site and therefore prevents conversion of ATP to cAMP. We hypothesize that this binding is disrupted by the mutation, which alters the shape of the enzymatic site to make it more similar to a prokaryotic AC.

### **Discussion**

Here we use directed evolution and comparative chemogenomics to identify that a class of antimalarial compounds, the phenyl-amino-methyl-quinolinols, or PAMQs, inhibit replication of *P. falciparum* through inhibition of the production of cAMP by class III adenylyl cyclase (AC). We show that this affect is selective for homodimeric ACs with the F1718 genotype and that PAMQs do not affect cAMP levels in human cells. These findings suggest that PAMQs represent a new class of antimicrobial compounds that should be optimized for application against malaria, fungi and kinetoplastids.

### **Acknowledgments**

Gregory Mark Goldgof (G.M.G.) and Dr. Sabine Otilie (S.O.) designed the experiments and wrote the manuscript. Directed evolution experiments were performed by G.M.G. IC50 experiments were performed by G.M.G., S.O., Edger Vigil (E.V.), Jennifer Yang (J.Y.), Jack Schenken (J.S.), Rebecca Stanhope (R.S.), and Maximo Prescott (M.P.) Whole genome sequence analysis was

performed by G.M.G., Felicia Gunawan (F.G.), and Micah J. Manary (M.J.M). Sequence homology and phylogenetic analysis was performed by G.M.G and S.O. Cyclic AMP-glo assays were designed and performed by G.M.G., Gregory M. LaMonte (G.M.L.), E.V., J.Y. and Nelissa Figueroa (N.F). IC50s against *Plasmodium falciparum* were performed by G.M.L. Structural biology analysis was performed by G.M.G.

Chapter 5 is being prepared for submission for publication. Co-authors include Sabine Otilie, Edgar Vigil, Nelissa Figuero, Felicia Gunawan, Jennifer Yang, Jake Schenken, Gregory M. LaMonte, Yo Suzuki and Elizabeth A. Winzeler. The dissertation author was the primary investigator and author of this paper.

## 6. Methods

### Yeast strains

Wild-Type = Strain name: SY025. Genotype: *MATa hoΔ::[tetO<sub>2</sub>pr-GFP, URA3] can1Δ::GMToolkit-a lyp1Δ his3Δ1 leu2Δ0 ura3Δ0 met15Δ0*

ABC<sub>16</sub>-Monster = *MATa adp1Δ snq2Δ ycf1Δ pdr15Δ yor1Δ vmr1Δ pdr11Δ nft1Δ bpt1Δ ybt1Δ ynr070wΔ yol075cΔ aus1Δ pdr5Δ pdr10Δ pdr12can1Δ::GMToolkit-a lyp1Δ his3Δ1 leu2Δ0 ura3Δ0 met15Δ0* (deletions for the ABC transporter genes are marked with *[tetO<sub>2</sub>pr-GFP, URA3]*).

The strain containing cytoplasmically expressed ratiometric pHluorin (UCC9633) was created by amplifying a fragment of plasmid pADH1pr-RMP (Henderson et al., 2014) using primers URA3-tTA-intChr1 F and R to integrate ratiometric pHluorin (Miesenbock et al., 1998) into an empty region of chromosome 1 (17068-17161) under the control of the ADH1 promoter.

UCC4925:

*MATa/MATα his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 ura3Δ0/ura3Δ0*

*lys2Δ0/+ trp1Δ63/+ hoΔ::PSCW11-cre-EBD78-NatMX/hoΔ::PSCW11-cre-EBD78-NatMX loxP-CDC20-Intron-loxP-HphMX/loxP-CDC20-Intron-loxP-HphMX loxP-UBC9-loxp-LEU2/loxP-UBC9-loxp-LEU2*

UCC9633:

*UCC4925 chrI(17068-17161)PADH1-pHluorin-URA3/ chrI(17068-17161)PADH1-pHluorin-URA3*

### Chemical Information

The IUPAC name of MMV001239 compound is 4-cyano-N-(5-methoxy-1,3-benzothiazol-2-yl)-N-(pyridin-3-ylmethyl)benzamide. The chemical was purchased from Life Chemicals Inc. (catalog #F2515-1937) through MolPort.

The IUPAC name of MMV000570 compound is 2-(4-ethylanilino)-4-methyl-1H-quinolin-7-ol. The chemical was purchased from ChemDiv Inc. (catalog #C614-0191) through MolPort (catalog #MolPort-007-649-676).

The IUPAC name of MMV007181 compound is 4-methyl-2-(4-phenylmethoxyanilino)-1H-quinolin-7-ol. The chemical was purchased from ChemDiv Inc. (catalog #C614-0124) through MolPort (catalog #MolPort-007-649-655).

### **S. cerevisiae susceptibility assay**

ABC<sub>16</sub>-Monster yeast cells were inoculated from agar plates into 2 ml of liquid YPD media and grown to saturation (OD<sub>600</sub> > 1.0) overnight at 250 RPM in a shaking incubator at 30°C. Cultures were diluted to OD<sub>600</sub> 0.01 in 3.5 ml of YPD and grown to log phase. 30 µl were added to the wells of a 384-well plate. Compounds were added starting with a concentration of 150 µM, followed by 15 1:2 serial dilutions. An initial reading of OD<sub>600</sub> (t = 0 hrs.) was taken using a Synergy HT spectrophotometer, and cells were incubated for a period of 18 hours at 30°C. After incubation, plates were shaken for 30 seconds on the “high” setting and read at OD<sub>600</sub>. IC<sub>50</sub> values were determined by first subtracting the

OD600 values at  $t=0$  from those of the final reading, and utilizing Graphpad Prism to calculate nonlinear regression on  $\log(\text{inhibitor})$  vs. response with variable slope (four parameters).

### **Selection of MMV001239-resistant *S. cerevisiae***

Various concentrations of the inhibitor were added to 50 ml conical tubes containing 20  $\mu\text{l}$  of saturated ABC<sub>16</sub>-Monster cells in 20 ml of YPD. Each selection was cultured under vigorous shaking until the culture reached saturation. Saturated cultures were diluted into fresh media with the inhibitors, and multiple rounds of selection were performed at increasingly higher concentrations. After reaching a compound concentration that was substantially higher than that of the initial IC<sub>50</sub> concentration, polyclonal cultures were streaked onto agar plates containing the appropriate inhibitor concentration. Single colonies were isolated, and IC<sub>50</sub> assays were performed to determine the degree of evolved resistance vs. that of the parental strain.

### **Cross-Resistance Testing**

Each mutant lineage was grown to saturation in YPD media overnight in a shaker at 30°C and 250 rpm. The following day the culture was diluted to OD600 0.01 and allowed time to enter log phase for about 2 hrs. 30  $\mu\text{L}$  of each lineage were dispensed onto a 384-well plate and incubated in the presence of inhibitors, with a top concentration of 150  $\mu\text{M}$  and 16 serial dilutions for 18 hours. Viability



was normalized to a no-inhibitor control, based on OD600 measures at 0 and 18 hours.

### **Yeast dose response growth assay**

Yeast cells at a concentration of 0.01 optical densities were treated with the described drug concentrations. Cells grown in the absence of drugs were used as a negative control. After an 18 hour incubation at 30<sup>0</sup>C in a 96-well plate, cell density was determined by measuring absorbance at 600 nm using a spectrophotometer (Perkin-Elmer Wallac Envision). Percent growth was calculated using the formula ( $Ab_{600}^{treated}/Ab_{600}^{control} \times 100$ )

### **Statistical Tests Used for IC<sub>50</sub> Analysis**

IC<sub>50</sub> values were determined first by subtracting OD600 values at time 0 hours from time 18 hours. Nonlinear regression on log(inhibitor) vs. response with variable slope (four parameters) was then performed using Graphpad Prism. Minimum values were constrained to 0.0. Three experimental replicates of two technical duplicates were used to calculate each final IC<sub>50</sub> value. P values for IC<sub>50</sub> fold-changes were determined using a one-tailed ratio paired t-test comparing the ratio of the mutant-strain IC<sub>50</sub> value to that of the parental strain.

## Whole-Genome Sequencing and Analysis

DNA was extracted from yeast cells using the YeaStar Genomic DNA kit. For whole genome sequencing (WGS), genomic yeast DNA libraries were normalized to 0.2 ng/ $\mu$ L and prepared for sequencing according to the manufacturer's instructions using the Illumina Nextera XT kit whole-genome resequencing library (see the Illumina protocol of tagmentation followed by ligation, v. 2013, Illumina, Inc., San Diego). DNA libraries were clustered and run on an Illumina HiSeq as 2x100 paired end reads, according to the manufacturer's instructions. Base calls were made using the software CASAVA v1.8.2. Initial sequence alignments were done using the platypus software (Manary et al., 2014). Reads were aligned to the reference *S. cerevisiae* genome using BWA, and unmapped reads were filtered using SAMTools. SNPs were called using GATK and filtered using the platypus software. Copy number variants were analyzed using Control-FreeC (Boeva et al., 2012).

*Site-Directed Mutagenesis.* Point mutations were introduced in the ABC<sub>16</sub>-Monster strain using the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) system from *Streptococcus pyogenes* (DiCarlo et al., 2013). Plasmids p414-TEF1p-Cas9-CYC1t (p414) and p426-SNR52p-gRNA.CAN1.Y-SUP4t (p426) were obtained from Addgene (Cambridge, MA). Plasmid transformation markers were replaced with markers that are compatible with the yeast strains used in the current study. The *TRP1* marker of the p414 plasmid was replaced with the *MET15* marker. As the centromere-autonomous replication sequence (CEN-ARS) region of this plasmid

was difficult to amplify using Q5 DNA polymerase (New England Biolabs, Ipswich, MA), the CEN-ARS region of the original plasmid was also replaced with a CEN-ARS sequence from *Mycoplasma mycoides* JCVI-syn1.0 (Gibson et al., 2010). To create a p414 fragment containing the mycoplasma CEN-ARS sequence, three PCR products were generated: 1) an upstream 289-base-pair fragment amplified from p414 using the primers 414\_VecF and 414\_GB\_3R, 2) a 525-base-pair CEN-ARS fragment amplified from genomic DNA of the mycoplasma strain using the primers 414\_GB\_2F and 414\_GB\_2R, and 3) a downstream 644-base-pair fragment amplified from p414 using the primers 414\_GB\_1F and Pre-p426\_Halfway\_R (see Supplementary Table 2). These three fragments were combined using crossover PCR to generate a single, contiguous fragment (fragment 1). To prepare the remaining parts, a 2.1-kb fragment containing the *MET15* gene was amplified from a *MET15*-containing plasmid (Suzuki et al., 2015) using the primers 414\_MET15F and New\_414\_MET15\_R (fragment 2). The two additional fragments generated from the p414 plasmid were a 1.7-kb fragment amplified using the primers p426\_Halfway\_F and New\_414\_Vec\_Only\_HalfR (fragment 3), and a 5.6-kb fragment amplified using the primers 414\_Vec\_Half\_F and 414\_VecR (fragment 4). Fragments 1-4 were combined using homologous recombination in yeast to make the plasmid p414-TEF1p-Cas9-CYC1t-MET15.

The *URA3* marker of the p426 guide RNA plasmid was replaced with the *LEU2* marker. The p426 backbone was amplified as two pieces, a 2.2-kb piece amplified with the primers Attempt2\_426\_Vec\_F and p426\_Halfway\_R (fragment

5), and a 3.1-kb piece amplified with primers p426\_Halfway\_F and New\_426\_Vec\_R (fragment 6). A 2.1-kb *LEU2* fragment was amplified from a plasmid derived from pRS315 (ref. (Sikorski and Hieter, 1989); AM & YS, unpublished result) using the primers New\_426\_LEU2\_F and Attempt2\_426\_LEU2\_F (fragment 7). Fragments 5-7 were combined using Gibson assembly (Gibson et al., 2009) to make the plasmid p426-SNR52p-gRNA.CAN.Y-SUP4t-LEU2.

For the introduction of each mutation, a unique 20-base-pair guide RNA (gRNA) sequence was required. To generate the required variants of the gRNA plasmid, p426-SNR52p-gRNA.CAN.Y-SUP4t-LEU2 was PCR-amplified using a high-fidelity DNA polymerase PrimeSTAR Max (2× Master Mix, Takara Bio, Mountain View, CA) and primers containing the appropriate 20-base-pair gRNA sequence at the 5' end (see Supplementary Table 3). The two ends of the PCR product shared 25 base pairs of homology, replacing the original gRNA sequence.

After purification using NucleoSpin Gel and the PCR Clean-up kit (Macherey-Nagel, Bethlehem, PA), the linearized plasmid was introduced into High Efficiency NEB 5-alpha chemically competent cells (New England Biolabs, Ipswich, MA), per the manufacturer's protocol (Bugyi, 1972, Kostylev et al., 2015). Transformed colonies were selected on LB-ampicillin agar plates. Colonies were then cultured in liquid LB-ampicillin medium, and plasmids were isolated using a miniprep kit (Qiagen, Valencia, CA). The correct gRNA sequence was verified using Sanger sequencing. Repair fragments were prepared from complementary

90-base-pair oligonucleotides (IDT, Coralville, IA), which included the desired mutation (see Supplementary Table 4). To prepare double-stranded repair fragments, complementary oligonucleotides were mixed in equimolar amounts, denatured at 100°C for five minutes, and allowed to cool to 25°C with a ramp rate of 0.1°C per second.

The ABC<sub>16</sub>-Monster strain was initially transformed with ~200 ng of the p414-TEF1p-Cas9-CYC1t-MET15 plasmid using a standard lithium acetate transformation method (Schiestl and Gietz, 1989) and selected on synthetic complete (SC) minus Met agar plates. To introduce a desired mutation, Cas9-expressing cells were cultured overnight in SC-Met. Using lithium acetate transformation (Schiestl and Gietz, 1989), the cells were co-transformed with ~500 ng of the appropriate gRNA plasmid and 2.5 nmol of the corresponding double-stranded repair fragment. All of the transformed cells were selected on SC-Met-Leu agar plates.

To verify the correct introduction of the mutations, the genomic region around the mutation was PCR-amplified using a high-fidelity DNA polymerase Q5 (2× Master Mix, New England Biolabs) and sequenced using Sanger sequencing (see Supplementary Table 5). Before phenotype testing, strains were cultured overnight in a YPDA liquid medium, the resulting cells were plated on a YPDA agar medium to form isolated colonies, and the loss of both plasmids was confirmed by transferring the colonies to both rich and selective media.

**ScPMA1p ATPase Assay.**

ATP hydrolysis was assayed at 30°C in 0.5 mL of 50 mM MES/Tris, pH 6.25, 5 mM NaN<sub>3</sub>, 5 mM Na<sub>2</sub> ATP (Roche), 10 mM MgCl<sub>2</sub>, and an ATP regenerating system (5 mM phosphoenolpyruvate and 50 µg/mL pyruvate kinase). The reaction was terminated after 20 minutes by the addition of Fiske and Subbarow reagent, and the release of inorganic phosphate was measured at 660 nm after 45 minutes of color development.

**Cytosolic pH Measurement.**

Cells were cultured in YEPD (1% yeast extract, 2% peptone, 2% glucose) exponentially for 4 hours prior to 3 hours of treatment with 200 µM KAE609. For pH measurement, cells were transferred to low fluorescence medium (Orij et al., 2009) at a density of  $3 \times 10^7$  cells/mL after rinsing them in an equal volume of low fluorescence medium. Calibration curves were made as described (Henderson et al., 2014).

To quantify cytosolic pH and generate calibration curves, pHluorin fluorescence emission was measured at 512 nm using with a SpectraMax M5 microtitre plate spectrofluorometer (Molecular Devices, Sunnyvale, CA) providing excitation bands of 9 nm centered around 390 and 470 nm. Background fluorescence for wild-type cells not expressing pHluorin (UCC4925) was subtracted from the measurements. The ratio of emission intensity resulting from excitation at 390 and 470 nm was calculated. Ratios were fitted to a calibration

curve to derive cytosolic pH from three biological replicates for each pH measurement. pH values are represented as mean  $\pm$  sd.

### **ScPMA1p Homology Modeling.**

To better understand the KAE609 binding pose, a homology model of ScPMA1p was built with Schrödinger's Prime software(Jacobson et al., 2004) using the UniProt(Bairoch et al., 2005) sequence P05030 and a structure of the highly homologous *Sus scrofa* sodium-potassium pump (PDBID: 3N2F, chain C)(Laursen et al., 2015). The 3N2F and P05030 amino-acid sequences were first aligned using ClustalW(Larkin et al., 2007). The model was then constructed using Schrödinger's knowledge-based method. The structure was further processed with Schrödinger's Protein Preparation Wizard(Sastry et al., 2013). Hydrogen atoms were added at pH 7.0 using PROPKA(Olsson et al., 2011), water molecules were removed, and disulfide bonds were appropriately modeled. The system was then subjected to a restrained minimization using the OPLS\_2005 forcefield(Jorgensen et al., 1996, Kaminski et al., 2001), converging the heavy atoms to an RMSD of 0.30 Å.

### **KAE609 Docking**

A three-dimensional KAE609 model was prepared using Schrodinger's LigPrep module. The protonation state was determined for pH values in the 5.0 to 9.0 range using Epik(Shelley et al., 2007). The most likely tautomeric state and low-energy ring conformations were similarly determined. The molecular

geometry was optimized using the OPLS\_2005 forcefield(Jorgensen et al., 1996, Kaminski et al., 2001).

Glide XP(Friesner et al., 2006, Repasky et al., 2007) was used to dock KAE609 into the modeled E2-ScPMA1p pocket near the amino acids associated with evolved resistance. The ligand diameter midpoint box was 14 Å cubed. Glide was instructed to use flexible ligand sampling, to sample nitrogen inversions and ring conformations, to bias torsion sampling for amides only, and to add Epik state penalties to the docking scores. The van der Waals radii of atoms with assigned partial charges less than or equal to 0.15 e were scaled to 80%.

#### **CYP51 binding affinity assay by UV-visible spectroscopy**

0.5 µl of compounds at different concentrations (0.1 to 100 µM in DMSO) were assayed in 100 mM potassium phosphate buffer (pH 7.5) containing 10% glycerol in the absence or presence of 2.4 µM CYP51 (Vieira et al., 2014). Ligand binding was measured in 96-well plates, in duplicate. Absorption spectra were recorded from 350 nm to 500 nm in 10 nm increments using a scanning photometer. To correct for the organic-solvent effect, the same volume of DMSO was added to a reference well. Binding affinities were approximated from the titration curves by calculating the absorption shift ( $\Delta A$ ) between 420 and 390 nm for the indicated ligand concentrations, using the algorithm described in reference (Gunatilleke et al., 2012).



**Compound activity assessment against intracellular *T. cruzi* amastigotes**

The mouse myoblast cell line C2C12 (ATCC #CRL-1772) was maintained at 37°C with 5% CO<sub>2</sub> in Dulbecco's Modified Eagle's Medium (DMEM) containing 4.5 g/l glucose, supplemented with 5% fetal bovine serum (FBS), 100 U/ml penicillin, and 100 µg/ml streptomycin. *T. cruzi* CA-1/72 culture-derived trypomastigotes were obtained from infected C2C12 culture supernatants after six days of infection. Cells and parasites were seeded in 384-well black clear-bottom plates at 1.5 x 10<sup>5</sup> parasites/mL and 1.0 x 10<sup>4</sup> C2C12/mL density in 50 µl of DMEM media per well. Compounds were added immediately after infection, and plates were incubated for 72 hours at 37°C 5% CO<sub>2</sub>. Plates were fixed with 4% formaldehyde for at least one hour and stained with 0.5 µg/ml of 4',6-diamidino-2-phenylindole (DAPI) for four hours. Plates were then washed twice and imaged using ImageXpress Micro XL (Molecular Devices). Images were analyzed by dedicated algorithms developed using the MetaXpress software. Antiparasitic activity was normalized based on negative controls (vehicle wells) and positive controls (uninfected wells). Host-cell viability was determined based on the total number of C2C12 cells in each well relative to the average number of host cells in vehicle wells.

## Crystallography

To determine the MMV001239 binding mode, recombinant *TcCYP51*, modified by replacing the first 31 residues upstream of Pro32 with the fragment MAKKTSSKGKL (von Wachenfeldt et al., 1997) and by inserting a His<sub>6</sub>-tag at the C-terminus, was expressed and purified as described elsewhere (Chen et al., 2009). Concentrated, purified protein samples were stored at -80 °C and diluted to 0.1 mM prior to crystallization by mixing with 20 mM potassium phosphate (pH 7.5), 10% glycerol, 1 mM DTT, and 0.5 mM EDTA, supplemented with equimolar inhibitor.

Crystallization conditions were determined using commercial high-throughput screening kits available in deep-well format (Hampton Research, Aliso Viejo, CA), a nanoliter drop-setting Mosquito robot (TTP LabTech, Cambridge, MA) operating with 96-well plates, and a hanging drop crystallization protocol. Crystals were further optimized in 24-well plates for diffraction data collection and harvested from the 2- $\mu$ L drops containing 0.1 M HEPES, pH 7.5, 4% isopropanol, 6% PEG 3350, and 7.8 mM detergent (*n*-Nonyl-beta-D-maltoside). Prior to data collection, crystals were cryo-protected by first plunging them into a drop of reservoir solution supplemented with 20% ethylene glycol, and then flash freezing them in liquid nitrogen.

Diffraction data were collected at 100-110 K at beamline 8.3.1, Advanced Light Source, Lawrence Berkeley National Laboratory, USA. Data indexing, integration, and scaling were conducted using MOSFLM and the programs implemented in the ELVES software suite (Holton and Alber, 2004). The crystal

structures were determined by molecular replacement using diffraction data processed in the P3121 space groups, with the atomic coordinates of *T. cruzi* CYP51 (PDB ID code: 4C0C) serving as a search model. Model refinement was performed using the REFMAC5 software (Murshudov et al., 1997).

(Collaborative Computational Project, 1994, Murshudov et al., 1997) Data collection and refinement statistics are shown in table 6.1.

**Table 6.1. X-ray data collection and refinement statistics.** Parameters and statistics for the X-RAY crystallography study.

Protein	<b>TcCYP51</b>
Ligand	<b>MMV001239</b>
<b>Data collection</b>	
Space group	P3 <sub>1</sub> 21
Cell dimensions	
a, b, c (Å)	63.93, 63.93, 224.04
$\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 90.00, 120.00
Molecules in AU	1
Wavelength	1.11587
Resolution (Å)	3.78
R <sub>sym</sub> or R <sub>merge</sub> (%)	19
I / $\sigma$ I	7.3
Completeness (%)	99.8
Redundancy	9.5
<b>Refinement</b>	
No. reflections	5207
R <sub>work</sub> / R <sub>free</sub> (%)	28/35
B-factors	
Mean B value	153.9
Wilson plot B value	110.7
R.m.s deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.364

## Homology Modelling and Small-Molecule Docking

A crystal structure of *S. cerevisiae* lanosterol 14- $\alpha$  demethylase (PDB ID: 4K0F (Monk et al., 2014)) was obtained from the Protein Data Bank (Berman et al., 2000). This structure was processed with Schrödinger's Protein Preparation Wizard, which added hydrogen atoms at pH 7.0 using PROPKA (Olsson et al., 2011), converted selenomethionine residues to methionine residues, and optimized the protein geometry using the OPLS\_2005 force field (Jorgensen et al., 1996, Kaminski et al., 2001) until the heavy-atom RMSD converged to 0.30 Å. Models of two known ligands (itraconazole and MMV001239) were prepared using Schrödinger's LigPrep module. Ligand geometries were optimized using the OPLS\_2005 force field. All protonation states appropriate for pH 7.0 +/- 2.0 were similarly considered.

The known ligands were docked into the target structure using Schrödinger's Glide XP (default parameters) (Friesner et al., 2004, Halgren et al., 2004, Friesner et al., 2006, Repasky et al., 2007). A docking constraint was employed to encourage coordination with the heme metal atom where appropriate. Glide XP recaptured the crystallographic pose of itraconazole, included in the docking study as a positive control. The orientation of the MMV001239 nitrogenous heterocycle was manually adjusted only slightly to improve coordination with the heme iron atom.

***S. cerevisiae* cAMP Assay**

ABC<sub>16</sub>-Monster cells were grown overnight and then diluted to OD 0.1. They were then treated with the indicated concentrations of compound for one hour while shaking at 250RPM at 30C. They were next treated with zymolase and digestion buffer for 20 minutes at 37C. Cells were spun and the supernatant was decanted then frozen at -80C until the assay was run. cAMP levels were measured using the cAMPglo kit (Promega) and normalized via protein concentration (equal amounts of protein were loaded into each well before the assay). Samples were run in technical quadruplicate.

***P. falciparum* cAMP Assay**

Dd2 parasites were cultured via normal conditions until the day of the assay. 2.5ml of Dd2 parasites, at 4% parasitemia and 5% hemocrit, were treated with the indicated concentrations of compound for 4 hours at 37C. After 4 hours, parasites were washed with 1x PBS twice, then lysed with 0.15% saponin for 15min. Parasites were then washed twice with PBS, and then frozen at -80C until the assay was run. cAMP levels were measured using the cAMPglo kit (Promega) and normalized via protein concentration (equal amounts of protein were loaded into each well before the assay). Samples were run in technical quadruplicate.

### **Human Red Blood Cell cAMP Assay**

RBCs at 5% hemocrit were treated with the indicated concentrations of compound for 4 hours at 37C. Cyclic AMP levels were measured using the cAMPglo kit (Promega) and normalized via protein concentration (equal amounts of protein were loaded into each well before the assay). Samples were run in technical quadruplicate.

### **Human Liver Cell cAMP Assay**

Huh7 cells were cultured via normal conditions (DMEM, 10% serum) until the night before the assay. Cells were seeded at 25K cells per well, grown overnight in a 96 well plate, then treated with the indicated concentrations of MMV570, Art or KH7 (or other drugs...) for 4 hours at 37C. After 4 hours, cells were washed with 1x PBS twice. Cells were then lysed and cAMP levels measured using the cAMPglo kit (Promega). Samples were run in technical quadruplicate.

## 7. Conclusions and Future Studies

In this dissertation, we have used directed evolution in the yeast *S. cerevisiae* ABC<sub>16</sub>-Monster strain to identify the drug targets for a diverse set of compounds. As a proof-of-concept we correctly identified ScRpl28p, a ribosomal protein, as the target of cycloheximide and ScTop2p, DNA-Topoisomerase II, as the target of etoposide. We then proceeded to confirm that KAE609, an antimalarial in clinical trials, targets a p-type ATPase, which allowed us to study the binding site and SAR in the yeast system. In addition, we identified and characterized the target of an anti-trypanosomal compound MMV001239 as TcCyp51, an essential enzyme in the ergosterol biosynthetic pathway. We also discovered that the PAMQ antimalarial compounds MMV000570 and MMV007181 target the cAMP signaling pathway. This last identification demonstrates the strength of this approach since it can lead to the discovery of a pathway that can now be explored for further drug development.

More broadly, the comparative functional genomics approach presented here can be used as a tool to identify compound-target interactions in many other eukaryotic systems. It will be especially valuable to researchers who wish to identify targets in organisms that are not readily amenable to genetic experiments. The genetic and structural information obtained through directed evolution also provides insights into the binding pose, thus enabling subsequent chemical optimization. The method is not infallible, as resistance genes are sometimes found as well, but it will likely become more efficient over time, especially as more advanced drug-sensitive strains are designed and



engineered. As the whole-genome sequences of many evolved strains become available, we will be able to compile a general list of multiDrug resistant genes that can be readily eliminated as likely targets in future projects and deleted from improved Drug resistant strains.

One area in which this method may be especially useful in the near future is for target identification projects for antimalarial compounds that do not have activity against the asexual blood stages. Asexual blood stages are the only stage in which directed evolution can be performed in *Plasmodium*, because the other stages of the life cycle are difficult to culture in sufficient numbers to perform selection. Unlike blood stage asexuals, which can cycle through many generations in human erythrocytes, the other stages require a change in environment to reproduce. For example, the sexual blood stage parasites (gametocytes) are born in human erythrocytes, but require a mosquito's midgut for mating. The liver stage of the parasite, on the other hand, undergoes a limited number of divisions (shizogeny), before it bursts from the liver cell, at which point it can only infect erythrocytes.

The blood asexual stage of the parasite's life cycle contains the greatest number of individual parasites and therefore the greatest genetic diversity and most opportunity to select for drug resistance. As a result, some believe that compounds that specifically target other aspects of the *Plasmodium* life cycle will be less likely to lead to drug resistance, which has been the major challenge of the field over the past half century (Burrows et al., 2013). As a result, high throughput screens for compounds that specifically inhibit liver, sexual, and

mosquito stages of the parasite life cycle are currently underway. Target identification for these compounds may be an area in which drug sensitive yeast is especially useful.

In addition, there remain a number of important antimicrobials with unknown or controversial mechanisms of action, which are especially common in the field of neglected infectious disease. This list includes two of the most important antimalarials artemisinin and primaquine. Primaquine, and the 8-aminoquinolones, are interesting, because they are the only compounds that can kill the dormant hypnozoite stage of *P. vivax*, and more compounds with this ability are needed, due to the toxicity of the existing compounds. It also includes mainstays of antihelminthic treatment, like praziquantil. Comparative chemical genomics using the ABC<sub>16</sub>-Monster strain may be useful in determining the mechanism of action of these compounds as well, although with primaquine, the active metabolite may first need to be identified.

Beyond *Plasmodium*, there are many human parasitic diseases that are currently the subject of large high-throughput phenotypic screens that are not well suited for *in vitro* directed evolution. These include kinetoplastid parasites, such as *Trypanosoma cruzi*, *Trypanosoma brucei*, and *Leshmania*. They also include protozoans such as other apicomplexa as well as helminthic parasites. Finally, target identification for anti-cancer compounds discovered through phenotypic screens is another area that could benefit from this method. In the case of both antineoplastic and antimicrobial chemotherapies, this method may be valuable both for researchers with hits from high throughput screens, as well

as compounds isolated from natural products in academic laboratories following. This latter approach, which has been a major source of pharmaceutical innovation since the dawn of chemotherapy in the early 20<sup>th</sup> century, continues to be a source of innovation.

Finally, this work underscores the similarity in pathways between biological phyla, suggesting a path for transferring drugs or drug targets across different diseases. For example, ScPMA1p, the target discussed in Chapter 3, has been proposed as a potential pan-antifungal drug target, dating back to at least the 1990s (Perlin et al., 1997, Seto-Young et al., 1997, Monk et al., 2005). Now that we know that inhibitors of *Pf*Atp4p are potential ScPma1p inhibitors, the large amount of research into this target in the malaria community can now be applied to the development of a new class of antifungals, which are badly needed. Our work investigating binding site of the spiroindolones in the fungal protein should be extended to see whether spiroindolones, if approved as an antimalarial agent, may be useful as an adjunct to antifungal chemotherapy, or whether investigation of this binding site can be used to develop more potent inhibitor that are not as susceptible to export by fungal ABC-transporters.

The experience of open-source drug discovery typified by the MMV Malaria Box further underscores this possibility of transferring knowledge across disease domains. As part of MMV's strategy, they have tested the compounds in this antimalarial library against a wide range of eukaryotic pathogens. As a result, we already know that the PAMQ compounds discussed in Chapter 5 have activity against a range of other organisms, including the kinetoplastids *Trypanosoma*,

which causes Chagas disease and African sleeping sickness, and *Leshmania*, which causes leshmaniasis. It is also active against the helminth *Schistosoma mansoni*, which causes schistosomiasis. As more libraries are screened against arrays of human pathogens, the need for a model system for rapid drug target identification will only increase.

## 8. References

- Aboelsoud NH (2010) Herbal medicine in ancient Egypt. *J Med Plants Res* 4:82-86.
- Achan J, Talisuna AO, Erhart A, Yeka A, Tibenderana JK, Baliraine FN, Rosenthal PJ, D'Alessandro U (2011) Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. *Malaria J* 10.
- ACS (1999) American Chemical Society International Historic Chemical Landmarks. Discovery and Development of Penicillin.
- Altcheh J, Moscatelli G, Moroni S, Garcia-Bournissen F, Freilij H (2011) Adverse events after the use of benznidazole in infants and children with Chagas disease. *Pediatrics* 127:e212-218.
- Ambesi A, Allen KE, Slayman CW (1997) Isolation of transport-competent secretory vesicles from *Saccharomyces cerevisiae*. *Anal Biochem* 251:127-129.
- Anand N, Davis BD (1960) Damage by streptomycin to the cell membrane of *Escherichia coli*. *Nature* 185:22-23.
- Andersson B (2011) The *Trypanosoma cruzi* genome; conserved core genes and extremely variable surface molecule families. *Res Microbiol* 162:619-625.
- Armstrong-James D, Meintjes G, Brown GD (2014) A neglected epidemic: fungal infections in HIV/AIDS. *Trends Microbiol* 22:120-127.
- Ashley EA, Dhorda M, Fairhurst RM, Amaratunga C, Lim P, Suon S, Sreng S, Anderson JM, Mao S, Sam B, Sopha C, Chuor CM, Nguon C, Sovannaroeth S, Pukrittayakamee S, Jittamala P, Chotivanich K, Chutasmit K, Suchatsoonthorn C, Runchaoen R, Hien TT, Thuy-Nhien NT, Thanh NV, Phu NH, Htut Y, Han KT, Aye KH, Mokuolu OA, Olaosebikan RR, Folaranmi OO, Mayxay M, Khanthavong M, Hongvanthong B, Newton PN, Onyamboko MA, Fanello CI, Tshefu AK, Mishra N, Valecha N, Phyo AP, Nosten F, Yi P, Tripura R, Borrmann S, Bashraheil M, Peshu J, Faiz MA, Ghose A, Hossain MA, Samad R, Rahman MR, Hasan MM, Islam A, Miotto O, Amato R, MacInnis B, Stalker J, Kwiatkowski DP, Bozdech Z, Jeeyapant A, Cheah PY, Sakulthaew T, Chalk J, Intharabut B, Silamut K, Lee SJ, Vihokhern B, Kunasol C, Imwong M, Tarning J, Taylor WJ, Yeung S, Woodrow CJ, Flegg JA, Das D, Smith J, Venkatesan M, Plowe CV, Stepniewska K, Guerin PJ, Dondorp AM, Day NP, White NJ, Tracking Resistance to Artemisinin C (2014) Spread of artemisinin resistance in *Plasmodium falciparum* malaria. *N Engl J Med* 371:411-423.

- Aye KS, Nakajima C, Yamaguchi T, Win MM, Shwe MM, Win AA, Lwin T, Nyunt WW, Ti T, Suzuki Y (2016) Genotypic characterization of multi-drug-resistant *Mycobacterium tuberculosis* isolates in Myanmar. *J Infect Chemother* 22:174-179.
- Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh L-SL (2005) The Universal Protein Resource (UniProt). *Nucleic Acids Res* 33:D154-159.
- Balzi E, Goffeau A (1995) Yeast multidrug resistance: the PDR network. *J Bioenerg Biomembr* 27:71-76.
- Baragana B, Hallyburton I, Lee MC, Norcross NR, Grimaldi R, Otto TD, Proto WR, Blagborough AM, Meister S, Wirjanata G, Ruecker A, Upton LM, Abraham TS, Almeida MJ, Pradhan A, Porzelle A, Martinez MS, Bolscher JM, Woodland A, Norval S, Zuccotto F, Thomas J, Simeons F, Stojanovski L, Osuna-Cabello M, Brock PM, Churcher TS, Sala KA, Zakutansky SE, Jimenez-Diaz MB, Sanz LM, Riley J, Basak R, Campbell M, Avery VM, Sauerwein RW, Dechering KJ, Noviyanti R, Campo B, Frearson JA, Angulo-Barturen I, Ferrer-Bazaga S, Gamo FJ, Wyatt PG, Leroy D, Siegl P, Delves MJ, Kyle DE, Wittlin S, Marfurt J, Price RN, Sinden RE, Winzeler EA, Charman SA, Bebrevska L, Gray DW, Campbell S, Fairlamb AH, Willis PA, Rayner JC, Fidock DA, Read KD, Gilbert IH (2015) A novel multiple-stage antimalarial agent that inhibits protein synthesis. *Nature* 522:315-320.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235-242.
- Bern C (2015) Chagas' Disease. *N Engl J Med* 373:1882.
- Bieger B, Essen LO (2001) Structural analysis of adenylate cyclases from *Trypanosoma brucei* in their monomeric state. *EMBO J* 20:433-445.
- Bjorkman A, Phillips-Howard PA (1990) The epidemiology of drug-resistant malaria. *Trans R Soc Trop Med Hyg* 84:177-180.
- Boeva V, Popova T, Bleakley K, Chiche P, Cappo J, Schleiermacher G, Janoueix-Lerosey I, Delattre O, Barillot E (2012) Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* 28:423-425.
- Bousema T, Okell L, Shekalaghe S, Griffin JT, Omar S, Sawa P, Sutherland C, Sauerwein R, Ghani AC, Drakeley C (2010) Revisiting the circulation time

of *Plasmodium falciparum* gametocytes: molecular detection methods to estimate the duration of gametocyte carriage and the effect of gametocytocidal drugs. *Malar J* 9:136.

- Braunwald F, Kasper, Hauser, Longo, Jameson (2001) *Harrison's Principles of Internal Medicine 15th Edition*.
- Buckner FS, Joubert BM, Boyle SM, Eastman RT, Verlinde CL, Matsuda SP (2003) Cloning and analysis of *Trypanosoma cruzi* lanosterol 14 $\alpha$ -demethylase. *Mol Biochem Parasitol* 132:75-81.
- Bugyi B (1972) [Sir Joseph Barcroft and the Cambridge physiology school]. *Orv Hetil* 113:2663-2665.
- Burdine L, Kodadek T (2004) Target identification in chemical genetics: the (often) missing link. *Chem Biol* 11:593-597.
- Burrows JN, van Huijsduijnen RH, Mohrle JJ, Oeuvray C, Wells TN (2013) Designing the next generation of medicines for malaria control and eradication. *Malar J* 12:187.
- Calvet CM, Vieira DF, Choi JY, Kellar D, Cameron MD, Siqueira-Neto JL, Gut J, Johnston JB, Lin L, Khan S, McKerrow JH, Roush WR, Podust LM (2014) 4-Aminopyridyl-based CYP51 inhibitors as anti-*Trypanosoma cruzi* drug leads with improved pharmacokinetic profile and in vivo potency. *J Med Chem* 57:6989-7005.
- Cannon JF, Tatchell K (1987) Characterization of *Saccharomyces cerevisiae* genes encoding subunits of cyclic AMP-dependent protein kinase. *Mol Cell Biol* 7:2653-2663.
- Cappellini MD, Fiorelli G (2008) Glucose-6-phosphate dehydrogenase deficiency. *Lancet* 371:64-74.
- Carrara VI, Lwin KM, Phyo AP, Ashley E, Wiladphaingern J, Sriprawat K, Rijken M, Boel M, McGready R, Proux S, Chu C, Singhasivanon P, White N, Nosten F (2013) Malaria Burden and Artemisinin Resistance in the Mobile and Migrant Population on the Thai-Myanmar Border, 1999-2011: An Observational Study. *Plos Med* 10.
- Casperson GF, Walker N, Bourne HR (1985) Isolation of the gene encoding adenylate cyclase in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 82:5060-5063.
- Chen CK, Doyle PS, Yermalitskaya LV, Mackey ZB, Ang KK, McKerrow JH, Podust LM (2009) *Trypanosoma cruzi* CYP51 inhibitor derived from a *Mycobacterium tuberculosis* screen hit. *PLoS Negl Trop Dis* 3:e372.

- Chen KH, Miyazaki T, Tsai HF, Bennett JE (2007) The bZip transcription factor Cgap1p is involved in multidrug resistance and required for activation of multidrug transporter gene CgFLR1 in *Candida glabrata*. *Gene* 386:63-72.
- Chen Y, Cann MJ, Litvin TN, Iourgenko V, Sinclair ML, Levin LR, Buck J (2000) Soluble adenylyl cyclase as an evolutionarily conserved bicarbonate sensor. *Science* 289:625-628.
- Cherry JM, Ball C, Weng S, Juvik G, Schmidt R, Adler C, Dunn B, Dwight S, Riles L, Mortimer RK, Botstein D (1997a) Genetic and physical maps of *Saccharomyces cerevisiae*. *Nature* 387:67-73.
- Cherry JM, Ball C, Weng S, Juvik G, Schmidt R, Adler C, Dunn B, Dwight S, Riles L, Mortimer RK, Botstein D (1997b) Genetic and physical maps of *Saccharomyces cerevisiae*. *Nature* 387:67-73.
- Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR, Costanzo MC, Dwight SS, Engel SR, Fisk DG, Hirschman JE, Hitz BC, Karra K, Krieger CJ, Miyasato SR, Nash RS, Park J, Skrzypek MS, Simison M, Weng S, Wong ED (2012) *Saccharomyces* Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res* 40:D700-D705.
- Choi I, Mikkelsen RB (1990) *Plasmodium falciparum*: ATP/ADP transport across the parasitophorous vacuolar and plasma membranes. *Exp Parasitol* 71:452-462.
- Choi JY, Calvet CM, Vieira DF, Gunatilleke SS, Cameron MD, McKerrow JH, Podust LM, Roush WR (2014) R-Configuration of 4-Aminopyridyl-Based Inhibitors of CYP51 Confers Superior Efficacy Against *Trypanosoma cruzi*. *ACS Med Chem Lett* 5:434-439.
- Chrispeels MJ, Raikhel NV (1991) Lectins, lectin genes, and their role in plant defense. *Plant Cell* 3:1-9.
- ClinicalTrials.gov (2016a) Effectiveness of KAE609 in Reducing Asexual & Sexual Blood-stage *P. Falciparum* Infection & Infectivity to Mosquitos. ClinicalTrials.gov Identifier: NCT02543086.
- ClinicalTrials.gov (2016b) Efficacy, Safety, Tolerability and Pharmacokinetics of KAF156 in Adult Patients With Acute, Uncomplicated *Plasmodium Falciparum* or *Vivax* Malaria Mono-infection. ClinicalTrials.gov Identifier:NCT01753323.



- ClinicalTrials.gov (2016c) A Study to Assess Efficacy, Safety of KAE609 in Adult Patients With Acute Malaria Mono-infection. ClinicalTrialsgov Identifier:NCT01860989.
- Collaborative Computational Project N (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50:760-763.
- Comroe JH, Jr. (1978) Pay dirt: the story of streptomycin. Part I. From Waksman to Waksman. *Am Rev Respir Dis* 117:773-781.
- Conrad M, Schothorst J, Kankipati HN, Van Zeebroeck G, Rubio-Teixeira M, Thevelein JM (2014) Nutrient sensing and signaling in the yeast *Saccharomyces cerevisiae*. *FEMS Microbiol Rev* 38:254-299.
- Cox E, White JR, Flaks JG (1964) STREPTOMYCIN ACTION AND THE RIBOSOME. *Proc Natl Acad Sci U S A* 51.
- DaRocha WD, Otsu K, Teixeira SM, Donelson JE (2004) Tests of cytoplasmic RNA interference (RNAi) and construction of a tetracycline-inducible T7 promoter system in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 133:175-186.
- Davies J, Gilbert W, Gorino L (1964) STREPTOMYCIN, SUPPRESSION, AND THE CODE. *Proc Natl Acad Sci U S A* 51.
- Dawn A, Singh S, More KR, Siddiqui FA, Pachikara N, Ramdani G, Langsley G, Chitnis CE (2014) The central role of cAMP in regulating *Plasmodium falciparum* merozoite invasion of human erythrocytes. *PLoS Pathog* 10:e1004520.
- de Loubresse NG, Prokhorova I, Holtkamp W, Rodnina MV, Yusupova G, Yusupov M (2014) Structural basis for the inhibition of the eukaryotic ribosome. *Nature* 513:517-+.
- De Rycker M, O'Neill S, Joshi D, Campbell L, Gray DW, Fairlamb AH (2012) A static-cidal assay for *Trypanosoma brucei* to aid hit prioritisation for progression into drug discovery programmes. *PLoS Negl Trop Dis* 6:e1932.
- DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM (2013) Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res* 41:4336-4343.
- Dunitz JD, Taylor R (1997) Organic fluorine hardly ever accepts hydrogen bonds. *Chem - Eur J* 3:89-98.

- Durrant JD, Amaro RE, Xie L, Urbaniak MD, Ferguson MA, Haapalainen A, Chen Z, Di Guilmi AM, Wunder F, Bourne PE, McCammon JA (2010) A multidimensional strategy to detect polypharmacological targets in the absence of structural and sequence homology. *PLoS Comput Biol* 6:e1000648.
- Dutcher JD (1968) The discovery and development of amphotericin B. *Dis Chest* 54:Suppl 1:296-298.
- El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, Ghedin E, Worthey EA, Delcher AL, Blandin G, Westenberger SJ, Caler E, Cerqueira GC, Branche C, Haas B, Anupama A, Arner E, Aslund L, Attipoe P, Bontempi E, Bringaud F, Burton P, Cadag E, Campbell DA, Carrington M, Crabtree J, Darban H, da Silveira JF, de Jong P, Edwards K, Englund PT, Fazelina G, Feldblyum T, Ferella M, Frasch AC, Gull K, Horn D, Hou L, Huang Y, Kindlund E, Klingbeil M, Kluge S, Koo H, Lacerda D, Levin MJ, Lorenzi H, Louie T, Machado CR, McCulloch R, McKenna A, Mizuno Y, Mottram JC, Nelson S, Ochaya S, Osoegawa K, Pai G, Parsons M, Pentony M, Pettersson U, Pop M, Ramirez JL, Rinta J, Robertson L, Salzberg SL, Sanchez DO, Seyler A, Sharma R, Shetty J, Simpson AJ, Sisk E, Tammi MT, Tarleton R, Teixeira S, Van Aken S, Vogt C, Ward PN, Wickstead B, Wortman J, White O, Fraser CM, Stuart KD, Andersson B (2005) The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309:409-415.
- Elias MC, Faria M (2009) Are there epigenetic controls in *Trypanosoma cruzi*? *Ann N Y Acad Sci* 1178:285-290.
- Engel LW, Straus SE (2002) Development of therapeutics: opportunities within complementary and alternative medicine. *Nat Rev Drug Discov* 1:229-237.
- Ferrero L, Cameron B, Manse B, Lagneaux D, Crouzet J, Famechon A, Blanche F (1994) Cloning and primary structure of *Staphylococcus aureus* DNA topoisomerase IV: a primary target of fluoroquinolones. *Mol Microbiol* 13:641-653.
- Flannery EL, Chatterjee AK, Winzeler EA (2013) Antimalarial drug discovery - approaches and progress towards new medicines. *Nat Rev Microbiol* 11:849-862.
- Flannery EL, McNamara CW, Kim SW, Kato TS, Li F, Teng CH, Gagaring K, Manary MJ, Barboa R, Meister S, Kuhlen K, Vinetz JM, Chatterjee AK, Winzeler EA (2015) Mutations in the P-type cation-transporter ATPase 4, PfATP4, mediate resistance to both aminopyrazole and spiroindolone antimalarials. *ACS Chem Biol* 10:413-420.

- Flemming a (1929) On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzae. Br J Exp Pathol 10:226–236.
- Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47:1739-1749.
- Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. J Med Chem 49:6177-6196.
- Gamo FJ, Sanz LM, Vidal J, de Cozar C, Alvarez E, Lavandera JL, Vanderwall DE, Green DV, Kumar V, Hasan S, Brown JR, Peishoff CE, Cardon LR, Garcia-Bustos JF (2010) Thousands of chemical starting points for antimalarial lead identification. Nature 465:305-310.
- Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Perteza M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DMA, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B (2002) Genome sequence of the human malaria parasite Plasmodium falciparum. Nature 419:498-511.
- Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K, Deutschbauer A, Entian KD, Flaherty P, Foury F, Garfinkel DJ, Gerstein M, Gotte D, Guldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kotter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang CY, Ward TR, Wilhelmy J, Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M, Philippsen P, Davis RW, Johnston M (2002) Functional profiling of the Saccharomyces cerevisiae genome. Nature 418:387-391.
- Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM, Merryman C, Vashee S, Krishnakumar R, Assad-Garcia N, Andrews-Pfannkoch C, Denisova EA,

- Young L, Qi ZQ, Segall-Shapiro TH, Calvey CH, Parmar PP, Hutchison CA, 3rd, Smith HO, Venter JC (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* 329:52-56.
- Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6:343-345.
- Guiguemde WA, Shelat AA, Bouck D, Duffy S, Crowther GJ, Davis PH, Smithson DC, Connelly M, Clark J, Zhu F, Jimenez-Diaz MB, Martinez MS, Wilson EB, Tripathi AK, Gut J, Sharlow ER, Bathurst I, El Mazouni F, Fowble JW, Forquer I, McGinley PL, Castro S, Angulo-Barturen I, Ferrer S, Rosenthal PJ, Derisi JL, Sullivan DJ, Lazo JS, Roos DS, Riscoe MK, Phillips MA, Rathod PK, Van Voorhis WC, Avery VM, Guy RK (2010) Chemical genetics of *Plasmodium falciparum*. *Nature* 465:311-315.
- Gunatilleke SS, Calvet CM, Johnston JB, Chen CK, Erenburg G, Gut J, Engel JC, Ang KK, Mulvaney J, Chen S, Arkin MR, McKerrow JH, Podust LM (2012) Diverse inhibitor chemotypes targeting *Trypanosoma cruzi* CYP51. *PLoS Negl Trop Dis* 6:e1736.
- Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem* 47:1750-1759.
- Hane MW, Wood TH (1969) *Escherichia coli* K-12 mutants resistant to nalidixic acid: genetic mapping and dominance studies. *J Bacteriol* 99:238-241.
- Hao GF, Wang F, Li H, Zhu XL, Yang WC, Huang LS, Wu JW, Berry EA, Yang GF (2012) Computational Discovery of Picomolar Q(o) Site Inhibitors of Cytochrome bc(1) Complex. *J Am Chem Soc* 134:11168-11176.
- Haste NM, Talabani H, Doo A, Merckx A, Langsley G, Taylor SS (2012) Exploring the *Plasmodium falciparum* cyclic-adenosine monophosphate (cAMP)-dependent protein kinase (PfPKA) as a therapeutic target. *Microbes Infect* 14:838-850.
- Henderson KA, Hughes AL, Gottschling DE (2014) Mother-daughter asymmetry of pH underlies aging and rejuvenation in yeast. *Elife* 3:e03504.
- Hoepfner D, McNamara CW, Lim CS, Studer C, Riedl R, Aust T, McCormack SL, Plouffe DM, Meister S, Schuierer S, Plikat U, Hartmann N, Staedtler F, Cotesta S, Schmitt EK, Petersen F, Supek F, Glynne RJ, Tallarico JA, Porter JA, Fishman MC, Bodenreider C, Diagana TT, Movva NR, Winzeler EA (2012a) Selective and Specific Inhibition of the *Plasmodium falciparum*

Lysyl-tRNA Synthetase by the Fungal Secondary Metabolite Cladosporin. *Cell Host Microbe* 11:654-663.

Hoepfner D, McNamara CW, Lim CS, Studer C, Riedl R, Aust T, McCormack SL, Plouffe DM, Meister S, Schuierer S, Plikat U, Hartmann N, Staedtler F, Cotesta S, Schmitt EK, Petersen F, Supek F, Glynne RJ, Tallarico JA, Porter JA, Fishman MC, Bodenreider C, Diagana TT, Movva NR, Winzeler EA (2012b) Selective and specific inhibition of the plasmodium falciparum lysyl-tRNA synthetase by the fungal secondary metabolite cladosporin. *Cell Host Microbe* 11:654-663.

Holton J, Alber T (2004) Automated protein crystal structure determination using ELVES. *Proc Natl Acad Sci U S A* 101:1537-1542.

Huang Z, Chen K, Zhang J, Li Y, Wang H, Cui D, Tang J, Liu Y, Shi X, Li W, Liu D, Chen R, Sugang RS, Pan X (2013) A functional variomics tool for discovering drug-resistance genes and drug targets. *Cell Rep* 3:577-585.

Huerta-Cepas J, Dopazo J, Gabaldon T (2010) ETE: a python Environment for Tree Exploration. *Bmc Bioinformatics* 11:24.

Iwata K, Yamaguchi H, Hiratani T (1973) Mode of action of clotrimazole. *Sabouraudia* 11:158-166.

Jacobson MP, Pincus DL, Rapp CS, Day TJF, Honig B, Shaw DE, Friesner RA (2004) A Hierarchical Approach to All-Atom Protein Loop Prediction. *Proteins: Struct, Funct, Bioinf* 55:351-367.

Jimenez-Diaz MB, Ebert D, Salinas Y, Pradhan A, Lehane AM, Myrand-Lapierre ME, O'Loughlin KG, Shackelford DM, de Almeida MJ, Carrillo AK, Clark JA, Dennis ASM, Diep J, Deng XY, Duffy S, Endsley AN, Fedewa G, Guiguemde WA, Gomez MG, Holbrook G, Horst J, Kim CC, Liu J, Lee MCS, Matheny A, Martinez MS, Miller G, Rodriguez-Alejandre A, Sanz L, Sigal M, Spillman NJ, Stein PD, Wang Z, Zhu FY, Waterson D, Knapp S, Shelat A, Avery VM, Fidock DA, Gamo FJ, Charman SA, Mirsalis JC, Ma HS, Ferrer S, Kirk K, Angulo-Barturen I, Kyle DE, DeRisi JL, Floyd DM, Guy RK (2014a) (+)-SJ733, a clinical candidate for malaria that acts through ATP4 to induce rapid host-mediated clearance of Plasmodium. *Proc Natl Acad Sci U S A* 111:E5455-E5462.

Jimenez-Diaz MB, Ebert D, Salinas Y, Pradhan A, Lehane AM, Myrand-Lapierre ME, O'Loughlin KG, Shackelford DM, Justino de Almeida M, Carrillo AK, Clark JA, Dennis AS, Diep J, Deng X, Duffy S, Endsley AN, Fedewa G, Guiguemde WA, Gomez MG, Holbrook G, Horst J, Kim CC, Liu J, Lee MC, Matheny A, Martinez MS, Miller G, Rodriguez-Alejandre A, Sanz L, Sigal M, Spillman NJ, Stein PD, Wang Z, Zhu F, Waterson D, Knapp S,

- Shelat A, Avery VM, Fidock DA, Gamo FJ, Charman SA, Mirsalis JC, Ma H, Ferrer S, Kirk K, Angulo-Barturen I, Kyle DE, DeRisi JL, Floyd DM, Guy RK (2014b) (+)-SJ733, a clinical candidate for malaria that acts through ATP4 to induce rapid host-mediated clearance of Plasmodium. *Proc Natl Acad Sci U S A* 111:E5455-5462.
- Jo S, Lim JB, Klauda JB, Im W (2009) CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. *Biophys J* 97:50-58.
- Jorgensen WL, Maxwell DS, TiradoRives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225-11236.
- Kaiser M, Maes L, Tadoori LP, Spangenberg T, Ioset JR (2015a) Repurposing of the Open Access Malaria Box for Kinetoplastid Diseases Identifies Novel Active Scaffolds against Trypanosomatids. *J Biomol Screen* 20:634-645.
- Kaiser M, Maser P, Tadoori LP, Ioset JR, Brun R (2015b) Antiprotozoal Activity Profiling of Approved Drugs: A Starting Point toward Drug Repositioning. *PLoS One* 10:e0135556.
- Kalb VF, Woods CW, Turi TG, Dey CR, Sutter TR, Loper JC (1987) Primary structure of the P450 lanosterol demethylase gene from *Saccharomyces cerevisiae*. *DNA* 6:529-537.
- Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J Phys Chem B* 105:6474-6487.
- Kato J, Nishimura Y, Imamura R, Niki H, Hiraga S, Suzuki H (1990) New topoisomerase essential for chromosome segregation in *E. coli*. *Cell* 63:393-404.
- Kaufer NF, Fried HM, Schwindinger WF, Jasin M, Warner JR (1983) Cycloheximide resistance in yeast: the gene and its protein. *Nucleic Acids Res* 11:3123-3135.
- Khare S, Roach SL, Barnes SW, Hoepfner D, Walker JR, Chatterjee AK, Neitz RJ, Arkin MR, McNamara CW, Ballard J, Lai Y, Fu Y, Molteni V, Yeh V, McKerrow JH, Glynn RJ, Supek F (2015) Utilizing Chemical Genomics to Identify Cytochrome b as a Novel Drug Target for Chagas Disease. *PLoS Pathog* 11:e1005058.

- Kostylev M, Otwell AE, Richardson RE, Suzuki Y (2015) Cloning Should Be Simple: *Escherichia coli* DH5alpha-Mediated Assembly of Multiple DNA Fragments with Short End Homologies. *PLoS One* 10:e0137466.
- Kuhen KL, Chatterjee AK, Rottmann M, Gagaring K, Borboa R, Buenviaje J, Chen Z, Francek C, Wu T, Nagle A, Barnes SW, Plouffe D, Lee MC, Fidock DA, Graumans W, van de Vegte-Bolmer M, van Gemert GJ, Wirjanata G, Sebayang B, Marfurt J, Russell B, Suwanarusk R, Price RN, Nosten F, Tungtaeng A, Gettayacamin M, Sattabongkot J, Taylor J, Walker JR, Tully D, Patra KP, Flannery EL, Vinetz JM, Renia L, Sauerwein RW, Winzeler EA, Glynn RJ, Diagana TT (2014) KAF156 is an antimalarial clinical candidate with potential for use in prophylaxis, treatment, and prevention of disease transmission. *Antimicrob Agents Chemother* 58:5060-5067.
- Kuhlbrandt W, Zeelen J, Dietrich J (2002) Structure, mechanism, and regulation of the *Neurospora* plasma membrane H<sup>+</sup>-ATPase. *Science* 297:1692-1696.
- Kumar M, Hsiao K, Vidugiriene J, Goueli SA (2007) A bioluminescent-based, HTS-compatible assay to monitor G-protein-coupled receptor modulation of cellular cyclic AMP. *Assay Drug Dev Technol* 5:237-245.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947-2948.
- Laursen M, Gregersen JL, Yatime L, Nissen P, Fedosova NU (2015) Structures and characterization of digoxin- and bufalin-bound Na<sup>+</sup>, K<sup>+</sup>-ATPase compared with the ouabain-bound complex. *Proc Natl Acad Sci U S A* 112:1755-1760.
- Leang R, Barrette A, Bouth DM, Menard D, Abdur R, Duong S, Ringwald P (2013a) Efficacy of Dihydroartemisinin-Piperaquine for Treatment of Uncomplicated *Plasmodium falciparum* and *Plasmodium vivax* in Cambodia, 2008 to 2010. *Antimicrob Agents Ch* 57:818-826.
- Leang R, Ros S, Duong S, Navaratnam V, Lim P, Arie F, Kiechel JR, Menard D, Taylor WRJ (2013b) Therapeutic efficacy of fixed dose artesunate-mefloquine for the treatment of acute, uncomplicated *Plasmodium falciparum* malaria in Kampong Speu, Cambodia. *Malaria J* 12.
- Lee BY, Bacon KM, Bottazzi ME, Hotez PJ (2013) Global economic burden of Chagas disease: a computational simulation model. *Lancet Infect Dis* 13:342-348.

- Lehane AM, Ridgway MC, Baker E, Kirk K (2014) Diverse chemotypes disrupt ion homeostasis in the Malaria parasite. *Mol Microbiol* 94:327-339.
- Leong FJ, Li RB, Jain JP, Lefevre G, Magnusson B, Diagana TT, Pertel P (2014) A First-in-Human Randomized, Double-Blind, Placebo-Controlled, Single- and Multiple-Ascending Oral Dose Study of Novel Antimalarial Spiroindolone KAE609 (Cipargamin) To Assess Its Safety, Tolerability, and Pharmacokinetics in Healthy Adult Volunteers. *Antimicrob Agents Ch* 58:6209-6214.
- Linder JU (2006) Class III adenylyl cyclases: molecular mechanisms of catalysis and regulation. *Cell Mol Life Sci* 63:1736-1751.
- Linder JU, Schultz JE (2003) The class III adenylyl cyclases: multi-purpose signalling modules. *Cell Signal* 15:1081-1089.
- Llurba Montesino N, Kaiser M, Brun R, Schmidt TJ (2015) Search for Antiprotozoal Activity in Herbal Medicinal Preparations; New Natural Leads against Neglected Tropical Diseases. *Molecules* 20:14118-14138.
- Lomize MA, Lomize AL, Pogozheva ID, Mosberg HI (2006) OPM: Orientations of proteins in membranes database. *Bioinformatics* 22:623-625.
- Lucantoni L, Silvestrini F, Signore M, Siciliano G, Eldering M, Dechering KJ, Avery VM, Alano P (2015) A simple and predictive phenotypic High Content Imaging assay for *Plasmodium falciparum* mature gametocytes to identify malaria transmission blocking compounds. *Sci Rep* 5:16414.
- Luzzato L, Apirion D, Schlessinger D (1968) STREPTOMYCIN IN E. COLI:INTERRUPTION OF THE RIBOSOME CYCLE AT THE INITIATION OF PROTEIN SYNTHESIS. *Proc Natl Acad Sci U S A*.
- Mager WH, Planta RJ, Ballesta JG, Lee JC, Mizuta K, Suzuki K, Warner JR, Woolford J (1997) A new nomenclature for the cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae*. *Nucleic Acids Res* 25:4872-4875.
- Manary MJ, Singhakul SS, Flannery EL, Bopp SER, Corey VC, Bright AT, McNamara CW, Walker JR, Winzeler EA (2014) Identification of pathogen genomic variants through an integrated pipeline. *Bmc Bioinformatics* 15.
- Marin-Neto JA, Rassi A, Jr., Morillo CA, Avezum A, Connolly SJ, Sosa-Estani S, Rosas F, Yusuf S, Investigators B (2008) Rationale and design of a randomized placebo-controlled trial assessing the effects of etiologic treatment in Chagas' cardiomyopathy: the BENznidazole Evaluation For Interrupting Trypanosomiasis (BENEFIT). *Am Heart J* 156:37-43.



- Martins YC, Zanini GM, Frangos JA, Carvalho LJ (2012) Efficacy of different nitric oxide-based strategies in preventing experimental cerebral malaria by *Plasmodium berghei* ANKA. *PLoS One* 7:e32048.
- Matsumoto K, Uno I, Oshima Y, Ishikawa T (1982) Isolation and characterization of yeast mutants deficient in adenylate cyclase and cAMP-dependent protein kinase. *Proc Natl Acad Sci U S A* 79:2355-2359.
- McClure NS, Day T (2014) A theoretical examination of the relative importance of evolution management and drug development for managing resistance. *Proc Biol Sci* 281.
- McDonough KA, Rodriguez A (2012) The myriad roles of cyclic AMP in microbial pathogens: from signal to sword. *Nat Rev Microbiol* 10:27-38.
- McNamara CW, Lee MC, Lim CS, Lim SH, Roland J, Nagle A, Simon O, Yeung BK, Chatterjee AK, McCormack SL, Manary MJ, Zeeman AM, Dechering KJ, Kumar TR, Henrich PP, Gagaring K, Ibanez M, Kato N, Kuhen KL, Fischli C, Rottmann M, Plouffe DM, Bursulaya B, Meister S, Rameh L, Trappe J, Haasen D, Timmerman M, Sauerwein RW, Suwanarusk R, Russell B, Renia L, Nosten F, Tully DC, Kocken CH, Glynn RJ, Bodenreider C, Fidock DA, Diagana TT, Winzeler EA (2013) Targeting *Plasmodium* PI(4)K to eliminate malaria. *Nature* 504:248-253.
- Meister S, Plouffe DM, Kuhen KL, Bonamy GM, Wu T, Barnes SW, Bopp SE, Borboa R, Bright AT, Che J, Cohen S, Dharia NV, Gagaring K, Gettayacamin M, Gordon P, Groessl T, Kato N, Lee MC, McNamara CW, Fidock DA, Nagle A, Nam TG, Richmond W, Roland J, Rottmann M, Zhou B, Froissard P, Glynn RJ, Mazier D, Sattabongkot J, Schultz PG, Tuntland T, Walker JR, Zhou Y, Chatterjee A, Diagana TT, Winzeler EA (2011) Imaging of *Plasmodium* liver stages to drive next-generation antimalarial drug discovery. *Science* 334:1372-1377.
- Merritt C, Silva LE, Tanner AL, Stuart K, Pollastri MP (2014) Kinases as druggable targets in trypanosomatid protozoan parasites. *Chem Rev* 114:11280-11304.
- Miesenbock G, De Angelis DA, Rothman JE (1998) Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature* 394:192-195.
- Monk BC, Niimi K, Lin S, Knight A, Kardos TB, Cannon RD, Parshot R, King A, Lun D, Harding DR (2005) Surface-active fungicidal D-peptide inhibitors of the plasma membrane proton pump that block azole resistance. *Antimicrob Agents Chemother* 49:57-70.

- Monk BC, Tomasiak TM, Keniya MV, Huschmann FU, Tyndall JDA, O'Connell JD, Cannon RD, McDonald JG, Rodriguez A, Finer-Moore JS, Stroud RM (2014) Architecture of a single membrane spanning cytochrome P450 suggests constraints that orient the catalytic domain relative to a bilayer. *Proc Natl Acad Sci U S A* 111:3865-3870.
- Montecucco A, Zanetta F, Biamonti G (2015) Molecular Mechanisms of Etoposide. *EXCLI J* 14:95-108.
- Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53:240-255.
- Nakamura S, Nakamura M, Kojima T, Yoshida H (1989) gyrA and gyrB mutations in quinolone-resistant strains of *Escherichia coli*. *Antimicrob Agents Chemother* 33:254-255.
- Nes WR, Sekula BC, Nes WD, Adler JH (1978) The functional importance of structural features of ergosterol in yeast. *J Biol Chem* 253:6218-6225.
- Ng AWK, Wasan KM, Lopez-Berestein G (2003) Development of liposomal polyene antibiotics: an historical perspective. *J Pharm Pharm Sci* 6:67-83.
- NIAID (2016) *The Life Cycle of the Malaria Parasite*.
- Olsson MHM, Sondergaard CR, Rostkowski M, Jensen JH (2011) PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK(a) Predictions. *J Chem Theory Comput* 7:525-537.
- Onda M, Ota K, Chiba T, Sakaki Y, Ito T (2004) Analysis of gene network regulating yeast multidrug resistance by artificial activation of transcription factors: involvement of Pdr3 in salt tolerance. *Gene* 332:51-59.
- Onyewu C, Blankenship JR, Del Poeta M, Heitman J (2003) Ergosterol biosynthesis inhibitors become fungicidal when combined with calcineurin inhibitors against *Candida albicans*, *Candida glabrata*, and *Candida krusei*. *Antimicrob Agents Chemother* 47:956-964.
- Ostrosky-Zeichner L, Casadevall A, Galgiani JN, Odds FC, Rex JH (2010) An insight into the antifungal pipeline: selected new molecules and beyond. *Nat Rev Drug Discov* 9:719-727.
- Otten H (1986) Domagk and the development of the sulphonamides. *J Antimicrob Chemother* 17:689-696.
- Palumbi SR (2001) Humans as the world's greatest evolutionary force. *Science* 293:1786-1790.

- Park JT, Strominger JL (1957) Mode of action of penicillin. *Science* 125:99-101.
- Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nature Reviews Drug Discovery* 6:29-40.
- Payne DJ, Gwynn MN, Holmes DJ, Rosenberg M (2004) Genomic approaches to antibacterial discovery. *Methods Mol Biol* 266:231-259.
- Pena I, Manzano MP, Cantizani J, Kessler A, Alonso-Padilla J, Bardera AI, Alvarez E, Colmenarejo G, Cotillo I, Roquero I, de Dios-Anton F, Barroso V, Rodriguez A, Gray DW, Navarro M, Kumar V, Sherstnev A, Drewry DH, Brown JR, Fiandor JM, Martin JJ (2015) New Compound Sets Identified from High Throughput Phenotypic Screening Against Three Kinetoplastid Parasites: An Open Resource. *Sci Rep-Uk* 5.
- Perlin DS, Seto-Young D, Monk BC (1997) The plasma membrane H(+)-ATPase of fungi. A candidate drug target? *Ann N Y Acad Sci* 834:609-617.
- Pierce SE, Davis RW, Nislow C, Giaever G (2007) Genome-wide analysis of barcoded *Saccharomyces cerevisiae* gene-deletion mutants in pooled cultures. *Nat Protoc* 2:2958-2974.
- Pierce SE, Fung EL, Jaramillo DF, Chu AM, Davis RW, Nislow C, Giaever G (2006) A unique and universal molecular barcode array. *Nat Methods* 3:601-603.
- Plouffe D, Brinker A, McNamara C, Henson K, Kato N, Kuhen K, Nagle A, Adrian F, Matzen JT, Anderson P, Nam TG, Gray NS, Chatterjee A, Janes J, Yan SF, Trager R, Caldwell JS, Schultz PG, Zhou Y, Winzeler EA (2008) In silico activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen. *Proc Natl Acad Sci U S A* 105:9059-9064.
- Plouffe DM, Wree M, Du AY, Meister S, Li F, Patra K, Lubar A, Okitsu SL, Flannery EL, Kato N, Tanaseichuk O, Comer E, Zhou B, Kuhen K, Zhou Y, Leroy D, Schreiber SL, Scherer CA, Vinetz J, Winzeler EA (2016) High-Throughput Assay and Discovery of Small Molecules that Interrupt Malaria Transmission. *Cell Host Microbe* 19:114-126.
- Podust LM, von Kries JP, Eddine AN, Kim Y, Yermalitskaya LV, Kuehne R, Ouellet H, Warriar T, Altekoster M, Lee JS, Rademann J, Oschkinat H, Kaufmann SH, Waterman MR (2007) Small-molecule scaffolds for CYP51 inhibitors identified by high-throughput screening and defined by X-ray crystallography. *Antimicrob Agents Chemother* 51:3915-3923.

- Rassi A, Jr., Rassi A, Marin-Neto JA (2010) Chagas disease. *Lancet* 375:1388-1402.
- Read LK, Mikkelsen RB (1990) Cyclic AMP- and Ca<sup>2+</sup>(+)-dependent protein kinases in *Plasmodium falciparum*. *Exp Parasitol* 71:39-48.
- Repasky MP, Shelley M, Friesner RA (2007) Flexible ligand docking with Glide. *Curr Protoc Bioinformatics Chapter 8:Unit 8 12*.
- Rottmann M, McNamara C, Yeung BK, Lee MC, Zou B, Russell B, Seitz P, Plouffe DM, Dharia NV, Tan J, Cohen SB, Spencer KR, Gonzalez-Paez GE, Lakshminarayana SB, Goh A, Suwanarusk R, Jegla T, Schmitt EK, Beck HP, Brun R, Nosten F, Renia L, Dartois V, Keller TH, Fidock DA, Winzeler EA, Diagana TT (2010a) Spiroindolones, a potent compound class for the treatment of malaria. *Science* 329:1175-1180.
- Rottmann M, McNamara C, Yeung BKS, Lee MCS, Zou B, Russell B, Seitz P, Plouffe DM, Dharia NV, Tan J, Cohen SB, Spencer KR, Gonzalez-Paez GE, Lakshminarayana SB, Goh A, Suwanarusk R, Jegla T, Schmitt EK, Beck HP, Brun R, Nosten F, Renia L, Dartois V, Keller TH, Fidock DA, Winzeler EA, Diagana TT (2010b) Spiroindolones, a Potent Compound Class for the Treatment of Malaria. *Science* 329:1175-1180.
- Rubin RP (2007) A brief history of great discoveries in pharmacology: In celebration of the centennial anniversary of the founding of the American society of pharmacology and experimental therapeutics. *Pharmacol Rev* 59:289-359.
- Salazar E, Bank EM, Ramsey N, Hess KC, Deitsch KW, Levin LR, Buck J (2012) Characterization of *Plasmodium falciparum* adenylyl cyclase-beta and its role in erythrocytic stage parasites. *PLoS One* 7:e39769.
- Sandhu P, Akhter Y (2016) The drug binding sites and transport mechanism of the RND pumps from *Mycobacterium tuberculosis*: Insights from molecular dynamics simulations. *Arch Biochem Biophys* 592:38-49.
- Santangelo GM (2006) Glucose signaling in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* 70:253-282.
- Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aid Mol Des* 27:221-234.
- Schiestl RH, Gietz RD (1989) High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet* 16:339-346.

- Schmidt BH, Osheroff N, Berger JM (2012) Structure of a topoisomerase II-DNA-nucleotide complex reveals a new control mechanism for ATPase activity. *Nat Struct Mol Biol* 19:1147-+.
- Serrano R, Kiellandbrandt MC, Fink GR (1986) Yeast Plasma-Membrane Atpase Is Essential for Growth and Has Homology with (Na<sup>+</sup>+K<sup>+</sup>), K<sup>+</sup>- and Ca<sup>2+</sup>-Atpases. *Nature* 319:689-693.
- Seto-Young D, Monk B, Mason AB, Perlin DS (1997) Exploring an antifungal target in the plasma membrane H<sup>(+)</sup>-ATPase of fungi. *Biochim Biophys Acta* 1326:249-256.
- Seydel KB, Kampondeni SD, Valim C, Potchen MJ, Milner DA, Muwalo FW, Birbeck GL, Bradley WG, Fox LL, Glover SJ, Hammond CA, Heyderman RS, Chilingulo CA, Molyneux ME, Taylor TE (2015) Brain swelling and death in children with cerebral malaria. *N Engl J Med* 372:1126-1137.
- Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M (2007) Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. *J Comput Aid Mol Des* 21:681-691.
- Sikorski RS, Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122:19-27.
- Spangenberg T, Burrows JN, Kowalczyk P, McDonald S, Wells TN, Willis P (2013) The open access malaria box: a drug discovery catalyst for neglected diseases. *PLoS One* 8:e62906.
- Spillman NJ, Allen RJW, McNamara CW, Yeung BKS, Winzeler EA, Diagana TT, Kirk K (2013) Na<sup>+</sup> Regulation in the Malaria Parasite *Plasmodium falciparum* Involves the Cation ATPase PfATP4 and Is a Target of the Spiroindolone Antimalarials. *Cell Host Microbe* 13:227-237.
- Spillman NJ, Kirk K (2015) The malaria parasite cation ATPase PfATP4 and its role in the mechanism of action of a new arsenal of antimalarial drugs. *Int J Parasitol Drugs Drug Resist* 5:149-162.
- Steegborn C (2014) Structure, mechanism, and regulation of soluble adenylyl cyclases - similarities and differences to transmembrane adenylyl cyclases. *Biochim Biophys Acta* 1842:2535-2547.
- Strominger JL (1957) Microbial uridine-5'-pyrophosphate N-acetylamino sugar compounds. II. Incorporation of uracil-2-C<sup>14</sup> into nucleotide and nucleic acid. *J Biol Chem* 224:525-532.

- Sud IJ, Feingold DS (1981) Mechanisms of action of the antimycotic imidazoles. *J Invest Dermatol* 76:438-441.
- Suzuki Y, Assad-Garcia N, Kostylev M, Noskov VN, Wise KS, Karas BJ, Stam J, Montague MG, Hanly TJ, Enriquez NJ, Ramon A, Goldgof GM, Richter RA, Vashee S, Chuang RY, Winzeler EA, Hutchison CA, 3rd, Gibson DG, Smith HO, Glass JI, Venter JC (2015) Bacterial genome reduction using the progressive clustering of deletions via yeast sexual cycling. *Genome Res* 25:435-444.
- Suzuki Y, St Onge RP, Mani R, King OD, Heilbut A, Labunskyy VM, Chen W, Pham L, Zhang LV, Tong AH, Nislow C, Giaever G, Gladyshev VN, Vidal M, Schow P, Lehar J, Roth FP (2011) Knocking out multigene redundancies via cycles of sexual assortment and fluorescence selection. *Nat Methods* 8:159-164.
- Swinney DC, Anthony J (2011) How were new medicines discovered? *Nat Rev Drug Discov* 10:507-519.
- Sykes ML, Baell JB, Kaiser M, Chatelain E, Moawad SR, Ganame D, Ioset JR, Avery VM (2012) Identification of compounds with anti-proliferative activity against *Trypanosoma brucei brucei* strain 427 by a whole cell viability based HTS campaign. *PLoS Negl Trop Dis* 6:e1896.
- Tanaka K, Matsumoto K, Toh EA (1989) IRA1, an inhibitory regulator of the RAS-cyclic AMP pathway in *Saccharomyces cerevisiae*. *Mol Cell Biol* 9:757-768.
- Tanner M, de Savigny D (2008) Malaria eradication back on the table. *Bull World Health Organ* 86:82.
- Tipper DJ, Strominger JL (1965) Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc Natl Acad Sci U S A* 54:1133-1141.
- Tu Y (2011) The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. *Nat Med* 17:1217-1220.
- Vaidya AB, Morrisey JM, Zhang Z, Das S, Daly TM, Otto TD, Spillman NJ, Wyvratt M, Siegl P, Marfurt J, Wirjanata G, Sebayang BF, Price RN, Chatterjee A, Nagle A, Stasiak M, Charman SA, Angulo-Barturen I, Ferrer S, Belen Jimenez-Diaz M, Martinez MS, Gamo FJ, Avery VM, Ruecker A, Delves M, Kirk K, Berriman M, Kortagere S, Burrows J, Fan E, Bergman LW (2014) Pyrazoleamide compounds are potent antimalarials that target Na<sup>+</sup> homeostasis in intraerythrocytic *Plasmodium falciparum*. *Nat Commun* 5:5521.

- van den Bossche H, Willemsens G, Cools W, Lauwers WF, Le Jeune L (1978) Biochemical effects of miconazole on fungi. II. Inhibition of ergosterol biosynthesis in *Candida albicans*. *Chem Biol Interact* 21:59-78.
- van Pelt-Koops JC, Pett HE, Graumans W, van der Vegte-Bolmer M, van Gemert GJ, Rottmann M, Yeung BKS, Diagana TT, Sauerwein RW (2012) The Spiroindolone Drug Candidate NITD609 Potently Inhibits Gametocytogenesis and Blocks *Plasmodium falciparum* Transmission to Anopheles Mosquito Vector. *Antimicrob Agents Ch* 56:3544-3548.
- van Staveren WC, Solis DY, Hebrant A, Detours V, Dumont JE, Maenhaut C (2009) Human cancer cell lines: Experimental models for cancer cells in situ? For cancer stem cells? *Biochim Biophys Acta* 1795:92-103.
- Vieira DF, Choi JY, Calvet CM, Siqueira-Neto JL, Johnston JB, Kellar D, Gut J, Cameron MD, McKerrow JH, Roush WR, Podust LM (2014) Binding mode and potency of N-indolyloxopyridinyl-4-aminopropanyl-based inhibitors targeting *Trypanosoma cruzi* CYP51. *J Med Chem* 57:10162-10175.
- von Kries JP, Warrier T, Podust LM (2010) Identification of small-molecule scaffolds for p450 inhibitors. *Curr Protoc Microbiol* Chapter 17:Unit17 14.
- von Wachenfeldt C, Richardson TH, Cosme J, Johnson EF (1997) Microsomal P450 2C3 is expressed as a soluble dimer in *Escherichia coli* following modification of its N-terminus. *Arch Biochem Biophys* 339:107-114.
- Wagner BKS, S.L. (2016) The Power of Sophisticated Phenotypic Screening and Modern Mechanism-of-Action Methods. *Cell Chemical Biology*.
- Wallace IM, O'Sullivan O, Higgins DG, Notredame C (2006) M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res* 34:1692-1699.
- Weibull C (1953) The isolation of protoplasts from *Bacillus megaterium* by controlled treatment with lysozyme. *J Bacteriol* 66:688-695.
- Wernsdorfer WH, Payne D (1991) The dynamics of drug resistance in *Plasmodium falciparum*. *Pharmacol Ther* 50:95-121.
- White NJ, Pukrittayakamee S, Phyo AP, Rueangweeraayut R, Nosten F, Jittamala P, Jeeyapant A, Jain JP, Lefevre G, Li RB, Magnusson B, Diagana TT, Leong FJ (2014) Spiroindolone KAE609 for *Falciparum* and *Vivax* Malaria. *New Engl J Med* 371:403-410.
- WHO (2014) Status Report on Artemisinin Resistance.
- WHO (2015) World Malaria Report 2015.

- Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta JL, Riles L, Roberts CJ, Ross-MacDonald P, Scherens B, Snyder M, Sookhai-Mahadeo S, Storms RK, Veronneau S, Voet M, Volckaert G, Ward TR, Wysocki R, Yen GS, Yu K, Zimmermann K, Philippsen P, Johnston M, Davis RW (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285:901-906.
- Wongsrichanalai C, Sirichaisinthop J, Karwacki JJ, Congpuong K, Miller RS, Pang L, Thimasarn K (2001) Drug resistant malaria on the Thai-Myanmar and Thai-Cambodian borders. *Southeast Asian J Trop Med Public Health* 32:41-49.
- Wu CC, Li TK, Farh L, Lin LY, Lin TS, Yu YJ, Yen TJ, Chiang CW, Chan NL (2011) Structural Basis of Type II Topoisomerase Inhibition by the Anticancer Drug Etoposide. *Science* 333:459-462.
- Yeung BK, Zou B, Rottmann M, Lakshminarayana SB, Ang SH, Leong SY, Tan J, Wong J, Keller-Maerki S, Fischli C, Goh A, Schmitt EK, Krastel P, Francotte E, Kuhen K, Plouffe D, Henson K, Wagner T, Winzeler EA, Petersen F, Brun R, Dartois V, Diagana TT, Keller TH (2010a) Spirotetrahydro beta-carbolines (spiroindolones): a new class of potent and orally efficacious compounds for the treatment of malaria. *J Med Chem* 53:5155-5164.
- Yeung BKS, Zou B, Rottmann M, Lakshminarayana SB, Ang SH, Leong SY, Tan J, Wong J, Keller-Maerki S, Fischli C, Goh A, Schmitt EK, Krastel P, Francotte E, Kuhen K, Plouffe D, Henson K, Wagner T, Winzeler EA, Petersen F, Brun R, Dartois V, Diagana TT, Keller TH (2010b) Spirotetrahydro beta-Carbolines (Spiroindolones): A New Class of Potent and Orally Efficacious Compounds for the Treatment of Malaria. *J Med Chem* 53:5155-5164.
- Zheng XS, Chan TF, Zhou HH (2004) Genetic and genomic approaches to identify and study the targets of bioactive small molecules. *Chem Biol* 11:609-618.