TOWARD THE SYNTHESIS AND EVALUATION OF NOVEL N-ACYLATED FR-900098 ANALOGS

BY

RYAN ELLIS COBB

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Chemical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Adviser:

Professor Huimin Zhao
Abstract

Malaria represents one of the most significant threats to global health today, with half of the world’s population currently at risk. Although there exist several antimalarial drugs to treat the disease, many of them have been used and over-used for decades, leading to widespread resistance in the causative *Plasmodium* parasite. As a result, new antimalarials are needed not only to replace the old, but also to supplement each other in novel combination therapies. Two such candidates are fosmidomycin and its *N*-acetyl analog FR-900098, *Streptomyces* natural products first identified in the 1980s. These compounds were later found to be potent inhibitors of 1-deoxy-D-xylulose 5-phosphate reductoisomerase, a necessary enzyme in *Plasmodium* for isoprenoid biosynthesis via the non-mevalonate pathway, which does not exist in humans. Although FR-900098 was determined to be about twice as potent as fosmidomycin, it is important to note that, as a *Streptomyces* secondary metabolite, it is by no means optimized as a human antimalarial drug.

In the work presented here, we have studied and characterized the *N*-acetyltransferase FrbF from the FR-900098 biosynthetic pathway, which was recently completely reconstituted in *E. coli*. This enzyme is responsible for transfer of the acetyl group from acetyl-CoA to the phosphonate substrate, representing the penultimate step in FR-900098 synthesis. By exploiting the substrate specificity of FrbF, we can synthesize novel derivatives of FR-900098 with alternate *N*-acyl functionalities both *in vivo* and *in vitro*, such as the *N*-propionyl analog FR-900098P, which was found to be even more potent than its parent compound. We have also heterologously expressed and characterized the target enzyme from *P. falciparum*, which serves as a tool to evaluate the effectiveness of novel antimalarial compounds. Finally, we have developed a screening method for the directed evolution of FrbF to identify mutants with altered substrate specificity.
This is dedicated to my parents
Acknowledgments

I first wish to thank my advisor, Huimin Zhao, for giving me the opportunity to work on these projects, and for the constant guidance and support he has provided throughout my graduate career. I also thank the Zhao research group for their invaluable advice and useful discussions, and in particular Matt DeSieno for all that he has taught me from my very first day in the lab. Finally, I am indebted to my family for the love and understanding they have always provided me.
# Table of Contents

Chapter 1: Introduction to the antimalarial phosphonic acid FR-900098 ........................................ 1  
1.1 Malaria: an ongoing global health issue .................................................................................. 1  
  1.1.1 The malaria threat ............................................................................................................. 1  
  1.1.2 Past eradication efforts .................................................................................................... 1  
  1.1.3 Current control efforts ...................................................................................................... 2  
1.2 Phosphonic acids: new solutions from natural products ...................................................... 4  
  1.2.1 Pharmaceuticals from natural products ............................................................................ 4  
  1.2.2 Fosmidomycin and FR-900098 .......................................................................................... 6  
    1.2.2.1 Initial discovery and studies ...................................................................................... 6  
    1.2.2.2 Renewed interest and clinical trials .......................................................................... 7  
    1.2.2.3 Heterologous expression and characterization ......................................................... 9  
    1.2.2.4 Total synthesis ........................................................................................................... 12  
    1.2.2.5 Synthesis of analogs and derivatives ......................................................................... 13  
1.3 Conclusions .......................................................................................................................... 17  
1.4 References .......................................................................................................................... 17  
1.5 Figures .................................................................................................................................... 23  

Chapter 2: Characterization of the N-Acetyltransferase FrbF from the FR-900098 Biosynthetic Pathway ................................................................................................................................. 29  
2.1 Introduction ........................................................................................................................... 29  
2.2 Results and Discussion ......................................................................................................... 32  
  2.2.1 Kinetics characterization with phosphonate substrates .................................................. 32  
    2.2.1.1 CMP-5’-3APn synthesis and kinetics ................................................................... 32  
    2.2.1.2 CMP-5’-H3APn, the native phosphonate substrate for FR-900098 production .......................................................................................................... 33  
  2.2.2 Kinetics characterization with coenzyme A substrates .................................................. 35  
    2.2.2.1 Acetyl-CoA, the native acetyl donor for FR-900098 production ......................... 35  
    2.2.2.2 Evaluation of alternate acyl-CoAs ........................................................................ 36  
  2.2.3 Elucidation of the structure and catalytic mechanism .................................................... 38  
    2.2.3.1 Analysis of the crystal structure ............................................................................ 38  
    2.2.3.2 Mutational analysis of the active site ..................................................................... 39  
2.3 Conclusions and Outlook ...................................................................................................... 41  
2.4 Materials and Methods ......................................................................................................... 42  
  2.4.1 Materials ......................................................................................................................... 42  
  2.4.2 Protein expression and purification ................................................................................. 43  
  2.4.3 Generation of FrbF mutants ............................................................................................. 44  
  2.4.4 Preparation of CMP-5’-3APn ......................................................................................... 45  
  2.4.5 CMP-5’-H3APn oxidation assays ..................................................................................... 45  
  2.4.6 Kinetic assays .................................................................................................................. 46  
  2.4.7 Relative activity assays .................................................................................................... 47  
2.5 References ............................................................................................................................ 47  
2.6 Tables ..................................................................................................................................... 49  
2.7 Figures .................................................................................................................................. 50  

Chapter 3: Directed Evolution of FrbF for Synthesis of FR-900098P and Characterization with its Target Enzyme from *Plasmodium falciparum* ........................................................................... 57  
3.1 Introduction ............................................................................................................................ 57
3.2 Results and Discussion ................................................................. 60
   3.2.1 Pfdxr: the target enzyme from Plasmodium falciparum .................. 60
      3.2.1.1 Recombinant expression in E. coli .................................. 60
      3.2.1.2 Kinetic characterization .............................................. 62
      3.2.1.3 Structure elucidation with fosmidomycin and FR-900098 .......... 64
   3.2.2 FR-900098P ......................................................................... 65
      3.2.2.1 Synthesis in vitro .......................................................... 65
      3.2.2.2 Synthesis in vivo .......................................................... 67
      3.2.2.3 Pfdxr inhibition studies .................................................. 69
   3.2.3 Directed evolution of FrbF ....................................................... 71
      3.2.3.1 Screening method development ........................................ 71
      3.2.3.2 Library creation and preliminary results ............................. 73
3.3 Conclusions and Outlook ............................................................ 74
3.4 Materials and Methods .............................................................. 75
   3.4.1 Materials .............................................................................. 75
   3.4.2 Cloning, expression, and purification ....................................... 76
   3.4.3 Generation of Pfdxr mutants .................................................. 77
   3.4.4 Kinetic assays ....................................................................... 77
   3.4.5 Inhibition assays ................................................................. 78
   3.4.6 In vitro FR-900098P synthesis ............................................... 78
   3.4.7 In vivo FR-900098P synthesis ................................................. 79
   3.4.8 FrbF library creation ............................................................ 80
   3.4.9 FrbF screening method ......................................................... 81
3.5 References .................................................................................. 82
3.6 Tables ......................................................................................... 84
3.7 Figures ....................................................................................... 86
Chapter 1: Introduction to the antimalarial phosphonic acid FR-900098

1.1 Malaria: an ongoing global health issue

1.1.1 The malaria threat

With 225 million cases and 781,000 deaths estimated in 2009 alone, malaria remains one of the most significant threats to global health today [1]. Despite past and current control efforts, about half of the world’s population still remains at risk. The disease is caused by any of four different species of Plasmodia – Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae, and Plasmodium ovale – all of which are transmitted to human hosts through mosquitoes of the genus Anopheles. Of these four parasite species, Plasmodium falciparum and Plasmodium vivax are the most common. Common symptoms of malaria include headache, fever, chills, and vomiting, and can become significantly more severe without treatment, leading to anemia, respiratory distress, encephalopathy (cerebral malaria), and death [2]. Children under five years of age are particularly susceptible to the disease, accounting for 85% of the reported deaths.

1.1.2 Past eradication efforts

Efforts to stop malaria began around the turn of the 20th century, when the distribution of the disease included significant portions of North America, Europe, and northern Asia (Figure 1.1). During that time, the focus of these efforts was environmental control of the mosquito vector and its breeding sites [3]. The discovery of the insecticidal properties of dichlorodiphenyltrichloroethane (DDT) in the 1940s greatly aided this cause. In 1955, the World Health Organization ambitiously launched its Global Malaria Eradication campaign, which relied
on a combination of widespread insecticide usage and mass drug administration. In the developed world, this approach proved to be quite successful (Figure 1.1). However, it eventually became evident that the goal of complete eradication was not feasible. Notably, this effort ignored the majority of sub-Saharan Africa, where the infrastructure did not exist to follow through on the plan [4].

Despite the great general effectiveness of insecticides, insecticide resistance in Anopheles mosquitoes was first reported in the mid-1950s [5]. Further, mass drug administration was carried out primarily using only one of two antimalarial compounds: chloroquine or pyrimethamine (Figure 1.2). As a result, early effectiveness was hampered by the development of resistance as high selection pressure was applied, and reports of chloroquine-resistant and pyrimethamine-resistant malaria followed [6-8]. In 1969, the WHO concluded that global malaria eradication was no longer a feasible goal, and so financial support for the Campaign was rapidly reduced. Subsequently, the malaria problem continued to grow throughout the 1970s, 1980s, and early 1990s, up to the point where 300 – 500 million cases and 1.5 – 2.7 million deaths were estimated in 1994 [9].

1.1.3 Current control efforts

The current era of malaria control began in 1993 with the WHO’s adoption of the new Global Malaria Control Strategy. This plan called for more disease-focused action, advocating rational, locally-tailored control efforts [4]. By 1998, the Roll Back Malaria Partnership was formed and established clear goals for malaria control in 2010, 2015, and beyond. These goals call for the parallel implementation of locally appropriate vector control methods (including insecticide-
treated bed nets and indoor residual spraying) and effective diagnosis and treatment of infected individuals [10]. Chief among the malaria treatments are so-called ACTs, or artemisinin combination therapies.

Artemisinin, a sesquiterpene lactone derived from the plant *Artemisia annua* (Figure 1.2), has a long history of human therapeutic use. Known as qing hao in China, the plant was first referenced in 168 BC as a treatment for hemorrhoids, and in 340 AD was noted as a treatment for fever, first suggesting its antimalarial capabilities [11]. In 1972, artemisinin was first isolated from *Artemisia annua*. Subsequent assays against *Plasmodium falciparum* demonstrated significant potency, comparable to that of chloroquine and mefloquine. Trials in mice also showed good efficacy in treating *P. berghei*, and a 1979 human trial cured 2099 patients, although moderate recrudescence was observed [12].

Although the cultivated native producer continues to be the primary source of artemisinin, demand for the compound still exceeds the practical output of this effort [13]. To solve this problem, total synthesis of the compound was attempted by several research groups in the 1980s and 1990s. Although multiple total syntheses were achieved, none were suitable for commercialization [14-17]. As vast improvements were made in recombinant DNA technology around the turn of the century, biosynthesis became a much more attractive alternative. Efforts toward the biosynthesis of artemisinin were spearheaded by the research group of Jay Keasling. In 2003, they demonstrated synthesis of the artemisinin precursor amorphadiene in *Escherichia coli* via an engineered mevalonate pathway from *Saccharomyces cerevisiae* and a synthetic amorpha-4,11-diene synthase gene [18]. In 2006, artemisinic acid titers of up to 100 mg•L$^{-1}$
were achieved in an engineered *S. cerevisae*, paving the way to commercially relevant biosynthesis [19]. By 2009, further optimization of both the heterologous mevalonate pathway and the overall fermentation process significantly increased amorphadiene titers in *E. coli* to an average of 27.4 g•L⁻¹ [20].

Although artemisinin (and its derivatives, such as artesunate or artemether) remain the key component in the current malaria treatment campaign, they do not necessarily represent a complete solution to the ongoing malaria problem. Evidence of artemisinin resistance has already been observed in field isolates, calling for urgent containment measures to stop its spread [21-23]. This underscores the need for new antimalarial drugs, not only to replace the old but to supplement each other in novel combination therapies as the fight against malaria continues.

1.2 Phosphonic acids: new solutions from natural products

1.2.1 Pharmaceuticals from natural products

For centuries, mankind has turned to nature for ways to improve quality of life. Microbial fermentation, for example, has been extensively utilized for the production of beer, bread, and wine, with the oldest known fermentations occurring at least 9000 years ago [24]. In the field of human health, use of natural products began in earnest in 1929 with Alexander Fleming’s discovery of the antibacterial penicillin [25]. In the ensuing decades, over one million natural products have been discovered from plants, animals, and microorganisms, of which 20 – 25 % exhibit biological activity [26]. At present, natural products continue to see extensive pharmaceutical use. In the period from 1981 – 2006, approximately half of all small molecule NCEs (new chemical entities) were natural products, natural product derivatives, natural product
mimics, or synthetic analogues of natural product pharmacophores [27]. However, it is noteworthy that (i) the vast majority (>99%) of microorganisms remain uncultivable in the laboratory setting [28] and (ii) of the sequenced microbial genomes, most contain several gene clusters for which the corresponding natural product is either unknown or undetected [29]. Taking these two facts together, it can be concluded that there still exists a great wealth of as yet undiscovered natural products for the treatment of any number of human diseases.

One class of natural products underexploited for use as human therapeutics is the phosphonic acids. These compounds are characterized by the presence of a highly stable carbon-phosphorus (C-P) bond, which can withstand harsh treatment in strong acid or base. These compounds are capable of mimicking phosphate esters and carboxylates, two of the most ubiquitous chemical moieties in all of biology. As a result, they have the potential to serve as analogues of these compounds, and thus as potent inhibitors of a great number of enzymes [30].

Multiple phosphonic acid (and phosphinic acid, containing a C-P-C bond) natural products with interesting bioactivities have been discovered and characterized (Figure 1.3). Bialaphos, for example, is an herbicide, the gene cluster of which was identified and characterized in 1986 [31]. Fosfomycin, originally called phosphonomycin, was first identified from multiple *Streptomyces* species and later found to be produced by *Pseudomonas syringae* [32, 33]. Heterologous expression in *Streptomyces lividans* enabled identification of the minimal of the gene cluster from *Streptomyces fradiae* [34]. Currently, fosfomycin is a clinically-approved antibiotic for urinary tract infections and has also been found to be effective against resistant strains of *Staphylococcus aureus*. Another phosphonate antibiotic is dehydrophos, first identified in 1984
This compound demonstrates broad-spectrum activity, and its biosynthetic cluster has been heterologously expressed and characterized [36]. These three examples represent only a fraction of the known naturally occurring small molecule phosphonates.

**1.2.2 Fosmidomycin and FR-900098**

1.2.2.1 Initial discovery and studies

Two phosphonic acid compounds of particular interest in the antimalarial field are fosmidomycin (originally called FR-31564) and FR-900098. These two compounds were first isolated in the late 1970s by the Fujisawa Pharmaceutical Co., Ltd. from *Streptomyces lavendulae* and *S. rubellomurinus* sp. nov., respectively [37, 38]. They were discovered in a screening program of soil isolates against the nocardicin C-supersensitive mutant of *Pseudomonas aeruginosa* NCTC 10490, designed to identify novel inhibitors of cell wall biosynthesis. The chemical structures of both compounds were solved at the time of isolation. They were found to differ only in their N-acyl groups, with fosmidomycin containing a formyl group and FR-900098 containing an acetyl moiety. Antibiotic spectra for these compounds were determined using a variety of microorganisms in a serial dilution agar plate assay, and both were found to be much more effective against Gram-negative bacteria than Gram-positive. FR-900098, in particular, was found to be most effective against *Salmonella typhi* O-901 and a *Pseudomonas aeruginosa* mutant (MICs of 2 and 1.6 µg/mL, respectively) while fosmidomycin was most effective against *Sarcina lutea* PCI-1001 and the same *P. aeruginosa* mutant (MICs of 0.1 and 0.05 µg/mL, respectively). However, fosmidomycin was found to be at least as effective as FR-900098 against all strains tested, often exhibiting a significantly lower MIC. As a result, further development efforts and clinical trials proceeded with fosmidomycin in the 1980s [39-43]. It
was found to have good pharmacokinetic properties and was initially targeted as a treatment of urinary tract infections. Though initial results showed promise, the drug was found to be poorly effective against recurrent infections, and so further development was halted [44].

1.2.2.2 Renewed interest and clinical trials

Following its initial discovery, little subsequent research interest was expressed in FR-900098, as it was thought to be the inferior of the two compounds from an antibiotic standpoint. However, interest in both compounds was rekindled in 1993 with the elucidation of the non-mevalonate pathway for the biosynthesis of isoprenoid precursors [45]. As isoprenoid synthesis was previously thought to be generated solely through the mevalonate pathway, the discovery of this alternate pathway shed light on novel targets for herbicides and antibiotics. In 1989, it had already been demonstrated that fosmidomycin derived its antibiotic activity through inhibition of isoprenoid synthesis [46]. Kuzuyama and co-workers demonstrated this inhibition in vitro with purified deoxyxylulose-5-phosphate reductoisomerase (Dxr) from the non-mevalonate pathway in *E. coli*, observing mixed-mode inhibition with a $K_i$ value of 38 nM for fosmidomycin [47]. The full resurgence of fosmidomycin and FR-900098 was not underway until 1999, however, when the Jomaa research group first identified the two compounds as potential antimalarials [48]. In this study, bioinformatic analysis was employed to identify a possible Dxr enzyme in the recently sequenced genome of *Plasmodium falciparum*. Fosmidomycin and FR-900098 were both assayed against three strains of *P. falciparum*. While both compounds exhibited low IC$_{50}$ values ($\leq$370 nM), FR-900098 that proved to be roughly twice as effective as fosmidomycin. This also proved true in *in vitro* inhibition assays against the purified recombinant target enzyme. Finally, animal studies were conducted in mice infected with *P. vinckei*. Again, FR-900098
proved to be the most effective antimalarial drug, exhibiting <1% parasitemia with a 2 mg/kg dosage.

Likely owing to the difficulty in preparing large quantities of FR-900098, antimalarial clinical trials in humans instead proceeded with fosmidomycin. The first of these trials were carried out in Gabon by Missinou and co-workers [49]. In this study, 27 adults with uncomplicated \textit{P. falciparum} malaria were administered fosmidomycin (1.2 g every 8 h) for 3, 4, or 5 days. In the 4- and 5-day treatment groups, almost 90\% of patients had negative blood smears at day 14, indicating effective treatment of the disease. In the 3-day group, however, the observed cure rate was only 60\%. Nevertheless, a rapid parasite clearance time of 48 h was observed, indicating the effectiveness of this drug. A second study carried out in both Gabon and Thailand utilized a 7-day treatment regimen [50]. After 7 days, all 20 subjects were found to be cured of the disease. Twenty-eight days after the onset of treatment, however, recrudescence was observed in two of nine patients in Gabon (one patient was unavailable for follow-up), and seven of nine patients in Thailand (one patient had developed \textit{P. vivax} parasitemia). These results suggested the limits of fosmidomycin monotherapy, inspiring subsequent trials to consider combination therapies. Fosmidomycin-clindamycin was then evaluated in children of various age groups [51-53]. Initial studies focused on children aged 7 – 14. Patients were treated for either 1, 2, 3, 4, or 5 days and re-evaluated at day 14. In each of the multi-day treatment groups, 100\% cure rates were observed by day 14, while a cure rate of 50\% was observed with only 1 day of treatment. Comparison with 5-day regimens of either fosmidomycin alone or clindamycin alone clearly demonstrated the advantage of the combination therapy. Extension of the subject age range showed similar cure rates in children as young as 3, but a significant decrease in children aged 1
– 2. An alternate combination therapy of fosmidomycin-arteunate was also evaluated in children aged 6 – 15, revealing 100 % 14-day cure rates with dosing regimens of ≥2 days. Follow-up at 28 days revealed ≥90 % cure rates with dosing regimens of ≥3 days [54].

1.2.2.3 Heterologous expression and characterization

With renewed interest in FR-900098 came renewed interest in its biosynthesis, first to elucidate the potentially unique biochemical transformations involved, and second to facilitate industrially relevant over-production through protein and metabolic engineering efforts. Although few phosphonic acid gene clusters had previously been characterized, a common initial step in their biosyntheses was noted to be conversion of phosphoenolpyruvate (PEP) to phosphonopyruvate through the activity of a PEP mutase (PEPM) gene [55]. As a result, identification of the FR-900098 biosynthetic cluster from the native producer S. rubellomurinus began with PCR-amplification of the PEPM gene from a genomic DNA fosmid library using degenerate primers. One hit from this screening process was the fosmid 4G7, which was subsequently integrated into the chromosome of the heterologous host Streptomyces lividans. Evaluation of the culture supernatant from this host against the purified E. coli Dxr and a phosphonate-sensitive E. coli strain confirmed that the cluster produced an active phosphonate Dxr inhibitor, and 31P-NMR and LC-MS analysis confirmed this compound to be FR-900098. Sequencing and deletion analysis led to the identification of eight open reading frames (ORFs), designated frbA – frbH, which were found to be necessary for FR-900098 production; two additional ORFs, frbI and frbJ, immediately downstream of the minimal cluster but not necessary for FR-900098 production; and one final ORF, dxrB, which appeared to be encode a Dxr homologue. Of the eight requisite ORFs, frbD was found to encode the PEPM originally identified in the fosmid
library screening, and thus likely catalyzes the first step in FR-900098 biosynthesis. However, no phosphonopyruvate decarboxylase gene was identified, which catalyzes the second step in other phosphonate biosyntheses. Instead, it was proposed that the early steps of FR-900098 biosynthesis parallel the TCA cycle, catalyzed by FrbA, FrbB, FrbC, and FrbE. *In vitro* assays with purified FrbD and FrbC clearly demonstrated the formation of 2-phosphonomethylmalate, confirming FrbC to be phosphonomethylmalate synthase. FrbA and either FrbB or FrbE (or both) were predicted to catalyze the subsequent TCA-analogous steps, yielding 2-oxo-4-phosphonobutyrate. To complete FR-900098 synthesis, FrbH was predicted to catalyze either the decarboxylation or transamination (or both) necessary to form 3-aminopropylphosphonate. Finally, acetylation of the amine was proposed to be catalyzed by FrbF, while hydroxylation of the amine was unassigned. As N-acetyl-3-aminopropylphosphonate was detected in the supernatant of the heterologous producer, while N-hydroxy-3-aminopropylphosphonate was not, acetylation was presumed to occur first.

As the final steps in FR-900098 biosynthesis were assigned purely based on *in silico* analysis, further research was necessary to confirm these predictions with *in vitro* and *in vivo* data. To accomplish this goal, the entire FR-900098 pathway was transferred to a more tractable host: *E. coli*. Additionally, this host would be more desirable for subsequent engineering and over-production efforts. The eight requisite ORFs and *dxrB*, thought to be involved in self-resistance, were placed on three compatible plasmids, each under the control of an IPTG-inducible T7 promoter. The result was a strain capable of reaching FR-900098 titers of 6.3 mg/L in shake-flask cultures in LB media [56]. Additional strains containing only part of the downstream pathway were constructed for use in whole-cell feeding studies. Curiously, when these strains
were fed 2-amino-4-phosphonobutyrate (2APn), the intermediates (detected by LC-MS) included CMP-5’-3-aminopropylphosphonate (CMP-5’-3APn), CMP-5’-N-acetyl-3-aminopropylphosphonate (CMP-5’-Ac3APn), and CMP-5’-FR-900098, suggesting the inclusion of a nucleotide transferase in the pathway. These feeding studies also suggested the necessity of \textit{frbG}, the function of which had been unassigned after the initial heterologous expression in \textit{S. lividans}.

Further \textit{in silico} analysis of FrbH revealed the presence of both a PLP-dependent aminotransferase/decarboxylase domain and a nucleotide transferase domain. \textit{In vitro} assays with the purified enzyme demonstrated the functionality of both of these domains in the conversion of 2APn to CMP-5’-3APn. Subsequent \textit{in vitro} assays were carried out with FrbF and FrbG, the flavin-dependent monooxygenase now thought to catalyze amine hydroxylation. Interestingly, both substrates were found to be active toward the FrbH product CMP-5’-3APn; FrbF, however, was also found to be active toward the hydroxylamine derivative, while FrbG was not active toward the \textit{N}-acetyl derivative. Thus, the path to CMP-5’-FR-900098 was concluded to proceed first via hydroxylation and then acetylation, with the \textit{N}-acetyl derivative appearing as a dead-end side product (Figure 1.4). As demonstrated in the heterologous hosts, formation of the final product FR-900098 could be achieved simply through a ubiquitous nucleotide hydrolase within the cell. Nevertheless, the \textit{frbI} gene in the extended FR-900098 cluster was found to encode a nucleotide hydrolase as well, and was demonstrated to convert CMP-5’-FR-900098 to FR-900098. Finally, a possible function for FrbJ was demonstrated \textit{in vitro} in the conversion of FR-900098 to FR-33289, a derivative first reported by Okuhara and
co-workers at the same time that FR-900098 was originally identified [38]. However, *in vivo* production of this compound was not detected.

1.2.2.4 Total synthesis

The first efforts toward the total synthesis of fosmidomycin and FR-900098 were undertaken by researchers at the Fujisawa Chemical Co., Ltd. at the time of their original isolation [37]. In this case, the goal of the syntheses was to confirm the predicted chemical structures of the two novel compounds, and so economics and scale were not priorities. Nevertheless, both compounds were successfully synthesized in three steps. First, 3-(N-tosyl-N-benzyloxyamino)propyl bromide was coupled to sodium diethylphosphonate through a Michaelis-Becker reaction. Second, the product was hydrolyzed in concentrated HCl-AcOH to give the hydroxylamine. Finally, the complete FR-900098 or fosmidomycin were obtained through acylation with Ac₂O/H₂O or acetic-formic anhydride, respectively.

Although syntheses of certain fosmidomycin and FR-900098 derivatives were pursued in subsequent years (see next section), total synthesis of the parent compounds themselves was not revisited until 2007, when two new reaction schemes were proposed [57, 58]. The method of Perruchon and co-workers featured as its initial step a one-pot conversion of acrolein to *N*- (benzyloxy)-*N*-(3-iodopropyl)acetamide through the sequential addition of several components. The C-P bond is then installed through reaction with dibenzyl phosphonate, yielding tribenzyl-protected FR-900098. Finally, hydrogenolytic removal of the protecting groups gives the final product. In contrast, the method of Fokin and co-workers begins with a nitroso-ene reaction between commercially available diethyl allylphosphonate and *in situ*-prepared nitrosocarboxyl
methane to yield the unsaturated diethyl ester derivative of FR-900098. Hydrogenation then
gives the diethyl ester derivative, which can be converted to FR-900098 through subsequent ester
hydrolysis and partial neutralization. Alternatively, dibenzyl allylphosphonate can be utilized as
the starting reagent. In this case, the product of the nitroso-ene reaction can be converted
directly to the final product through hydrogenation and neutralization, although in lower overall
yield.

1.2.2.5 Synthesis of analogs and derivatives

While fosmidomycin and FR-900098 have been demonstrated to be potent antimalarial
compounds, the fact remains that they are not optimized for use as human therapeutics. As a
result, several efforts have been directed toward the synthesis of a variety of analogs and
derivatives, both in a search for greater efficacy and to facilitate structure-activity relationship
(SAR) studies. The earliest of these was again carried out by researchers at the Fujisawa
Chemical Co., Ltd. [59]. Through their synthetic efforts, they were able to generate four
different types of compounds: fosmidomycin and FR-900098 with an additional hydroxyl group
on the β-carbon; 1-trans-propenyl analogs of both compounds; ethylene analogs of both
compounds, containing only a two-carbon spacer; and methylphosphinic acids of both
compounds. As the antimalarial capabilities of this class of compounds had not been established,
they were instead assayed against a broad spectrum of microbes for antibiotic behavior.
Although some of the derivatives showed minor improvements against particular species,
fosmidomycin was again found to be the most potent overall antibiotic. As a result, interest in
synthesis of derivatives and analogs remained minimal for the next several years. One
exception, however, is the work of Öhler and Kanzler, who in 1995 published an alternate
synthetic route to the 1-trans-propenyl analogs [60]. Through subsequent hydrogenation, the corresponding propylphosphonates could be generated for further derivitization.

With the discovery of fosmidomycin and FR-900098 as potent antimalarials and subsequent clinical trials, interest in derivative synthesis boomed. Among the goals of these efforts was improvement of the 20 – 40% bioavailability observed for fosmidomycin [61]. To this end, multiple syntheses of ester prodrugs were reported (Figure 1.6a). These derivatives were designed to increase uptake by masking the charged phosphonate moiety until the compounds enter the bloodstream, at which point ubiquitous esterases release the active drug. In 2001, Reichenberg and co-workers synthesized three diaryl ester prodrugs of FR-900098: the diphenyl ester, the bis-(2-methyl-phenyl) ester, and the bis-(4-methoxy-phenyl) ester [62]. Of these, oral administration of the diphenyl ester was found to be more effective than oral FR-900098 in the P. vinckei mouse model, while oral administration of the bis-(4-methoxy-phenyl) ester was most effective, performing comparably to intraperitoneally dosed FR-900098. However, concern over the toxicity of the phenol product of ester bond cleavage led to the exploration of alternate alkyl masking groups. Synthesis of a variety of acyloxyalkyl prodrugs and evaluation in the P. vinckei mouse model revealed the pivaloyloxymethyl, acetylxyethyl, and propionyloxyethyl esters to be more effective than FR-900098. Additionally, evaluation of plasma concentrations of FR-900098 after dosage confirmed the improved bioavailability of the acetylxyethyl ester [63]. In a follow-up study, alkoxy carbonyloxy esters were also synthesized and evaluated [64]. Both the methoxycarbonyloxyethyl and ethoxycarbonyloxyethyl esters were found to be more effective than FR-900098. Single ester prodrugs were also synthesized and evaluated, as these crystalline compounds would lend themselves more easily to capsule formulation than the oily diester
compounds. However, no improvement in efficacy over FR-900098 was observed with these compounds.

The next area of derivitization to be explored in depth was substitution on the α-carbon (Figure 1.6b), with several examples published within the last five years [65-69]. The van Calenbergh group, for example, explored a variety of substituted phenyl groups in the α-position and assayed them both against purified E. coli Dxr and two P. falciparum strains. Curiously, none of the derivatives proved to be better inhibitors of E. coli Dxr than fosmidomycin or FR-900098, but three of the compounds (one fosmidomycin derivative and two FR-900098 derivatives, each featuring chlorine substituents on the phenyl moiety) did show significant improvement in IC₅₀ against one or both of the P. falciparum strains. The Kurz group similarly tested a series of arylmethyl fosmidomycin and FR-900098 derivatives against P. falciparum, also including pivaloyloxymethyl ester protecting groups on the phosphonate. Although they were able to identify one compound (containing a 3,4-dichlorobenzyl substituent) with improved activity when compared to the fosmidomycin prodrug, none of their analogs could outperform the FR-900098 prodrug. Finally, a different study by Kurz and co-workers looked at various alkyl substitutions at the α-position of the pivaloyloxymethyl ester prodrugs. It was found that ethyl, propyl, isopropyl, and dimethyl substitutions significantly decreased antimalarial efficacy, while methyl substitution produced a compound of comparable antimalarial activity to the FR-900098 prodrug. Some substituted phenyl derivatives were also included in this study, with results similar to those described above. More recently, the van Calenbergh group examined halogenation at the α-position. Substitution of FR-900098 with fluorine or chlorine was analyzed by growth inhibition of P. falciparum and in the P. berghei mouse model. Both
compounds performed comparably to FR-900098 in *P. falciparum* inhibition, and the fluorinated derivative performed better in the *in vivo* trial.

Beyond just α-substituted carbon spacers, cyclized and unsaturated derivatives have also been synthesized and analyzed for antimalarial efficacy. Another study from the van Calenbergh group looked at cyclopropyl fosmidoycin and FR-900098 derivatives, restricting rotational freedom of the spacer by bridging the α- and β-carbons [70]. Unfortunately, none of the derivatives synthesized could outperform FR-900098 in inhibition of *E. coli* Dxr or *P. falciparum* growth. Similarly, incorporation of the three-carbon spacer into a cyclopentane ring failed to match the inhibition levels of fosmidomycin and FR-900098 toward *E. coli* Dxr [71]. Finally, synthesis of α,β-unsaturated derivatives with a variety of α-aryl substitutions was carried out, but the resulting compounds were all significantly worse inhibitors of *E. coli* Dxr [72].

Among the most recent attempts at the synthesis of novel derivatives are the β- and γ-oxa isosteres and the so-called “reverse analogs,” replacing the N-terminal moiety with a reversed hydroxamic acid group (Figure 1.6c). Of the β- and γ-oxa isosteres synthesized, γ-oxa substitution was found to significantly reduce activity, while β-oxa substitution of FR-900098 or the N-methyl hydroxamate reverse analog yielded slightly better inhibitors of *P. falciparum* growth [73]. Alternatively, the Kurz group synthesized the α-phenyl-substituted hydroxamate and N-methyl hydroxamate FR-900098 reverse analogs, revealing the latter to be a particularly potent inhibitor of *P. falciparum* and its recombinant Dxr enzyme [74]. The α-fluorinated N-methyl hydroxamate reverse analog was also found to have high antimalarial activity when
assayed in the *P. berghei* mouse model, further illustrating the promise of reverse analogs as more potent antimalarials [69].

**1.3 Conclusions**

Although worldwide efforts to stop malaria have proceeded for several decades, the disease remains to this day a significant global health issue that threatens half of the world’s population. In recent years, the spread of the disease has been held in check, primarily through the distribution of insecticide-treated bed nets, residual insecticide spraying, and the development of artemisinin-based combination therapies. Nevertheless, new drugs are still needed to combat the ever-present threat of microbial drug resistance. Phosphonic acids, an under-utilized class of natural products, may prove to be a valuable ally in this fight.

Fosmidomycin and its N-acetyl analog FR-900098, though first discovered in 1980, have only in the past decade shown their potential as highly effective antimalarial drugs. As *Streptomycete* natural products, however, these compounds are by no means optimized as human therapeutics. Thus, through the synthesis of a variety of derivatives, it may be possible to identify even more potent antimalarials. With the recent elucidation and heterologous expression of the FR-900098 biosynthetic pathway, the possibility to generate new derivatives through protein engineering, metabolic engineering, and combinatorial biosynthesis is opened, allowing access to novel derivatives difficult to access by synthetic means.

**1.4 References**


1.5 Figures

Figure 1.1: Global scope of malaria from 1900 to the present, from ref. 3.
Figure 1.2: Chemical structures of some of the most widely-used antimalarial drugs.
Figure 1.3: A selection of naturally occurring phosphonate compounds with interesting bioactivities.
Figure 1.4: The biosynthetic pathway of FR-900098, adapted from ref. 5.
Figure 1.5: Overview of synthetic strategies for FR-900098 production from (a) Kamiya et al. (ref. 36), (b) Perruchon et al. (ref. 57), and (c) Fokin et al. (ref. 56). For complete reaction details, refer to the respective manuscripts.
Figure 1.6: Overview of FR-900098 derivatives explored as potential antimalarial drugs. (a) Ester protecting groups found to increase the antimalarial efficacy of FR-900098. (b) Substitutions at the α position with significantly improved antimalarial activity compared to FR-900098. (c) Potent reverse analogs of FR-900098, including the β-oxa reverse analog with pivaloyloxyethyl ester protecting groups and two α-substituted compounds.
Chapter 2: Characterization of the \( N \)-Acetyltransferase FrbF from the FR-900098 Biosynthetic Pathway

2.1 Introduction

The FR-900098 biosynthetic pathway from \textit{Streptomyces rubellomurinus} has been well studied in the heterologous hosts \textit{Streptomyces lividans} and \textit{Escherichia coli} through the collaborative efforts of the Kelleher, Metcalf, and Zhao groups at the University of Illinois at Urbana-Champaign [1, 2]. As a result, the specific biochemical transformations catalyzed by each enzyme in the pathway have been established, laying the groundwork for both the overproduction of FR-900098 in a heterologous host and for the combinatorial biosynthesis of novel FR-900098 derivatives. The latter aim can be achieved through two distinct means: introduction of novel enzyme(s) from different pathways into the native biosynthetic framework, or retooling of the native biosynthetic enzymes to introduce novel functionalities. Both of these approaches have been successfully demonstrated in the literature. For example, McDaniel and co-workers applied the first approach to synthesize several derivatives of the polyketide natural product erythromycin through the substitution of polyketide synthase (PKS) domains from the rapamycin biosynthetic gene cluster [3]. Bernhardt and co-workers, on the other hand, applied the second approach, using a novel colorimetric screening method to broaden the substrate specificity of strictosidine synthase (STS) from \textit{Catharanthus roseus} [4]. As a result, they were able to synthesize new indole alkaloid derivatives not accessible via the native pathway.
The focus of the work described here is to apply the latter approach to the biosynthetic pathway of FR-900098 for the synthesis of novel derivatives. To this end, the N-acetyltransferase FrbF was selected as the target for kinetic and structural characterization as well as directed evolution for multiple reasons. First of all, FrbF catalyzes the penultimate step of FR-900098 biosynthesis, followed only by nucleotide cleavage to give the final product. As a result, none of the upstream pathway intermediates will be altered in the context of a mutated FrbF, indicating that all of the upstream enzymes will still be able to function normally. Mutating an enzyme further upstream in the pathway, in contrast, would yield a series of altered pathway intermediates. Given the high substrate specificity observed for some of the pathway enzymes, such as the bifunctional nucleotide transferase/decarboxylase FrbH, this could significantly slow or completely halt the biosynthesis of any novel derivatives. The only enzyme downstream of FrbF, the nucleotide hydrolase FrbI, exhibits very broad substrate specificity, and thus would likely be unaffected by the introduction of any FrbF mutant. Further, even in the absence of FrbI, ubiquitous nucleotide hydrolases within the cell can also serve to catalyze the final step. Second, although numerous derivatives of FR-900098 have been synthesized as described in the previous chapter, variation in the N-acyl substituent remains largely unexplored, with only the naturally produced N-formyl (fosmidomycin) and N-acetyl (FR-900098) moieties typically considered. Moreover, variation at this end of the molecule is tolerated and can lead to more potent Dxr inhibitors, as exemplified by the N-methyl hydroxamate “reverse analogs” described in Chapter 1. Third, although FrbF demonstrates activity analogous to enzymes of the well-studied GCN5-like N-acetyltransferase (GNAT) superfamily, FrbF shares no sequence similarity with any members thereof. Rather, FrbF shares sequence similarity with a number of uncharacterized, putative aminoglycoside N(3’)-acetyltransferases. As a result, characterization of FrbF could grant useful biochemical
insights into a new class of N-acetyltransferases, and could facilitate further studies of new uncharacterized secondary metabolite pathways.

The previous research efforts of Ty Johannes and Matt DeSieno in the Zhao group have demonstrated that FrbF exhibits activity only toward cytidine monophosphate-conjugated substrates [2]. Further, they concluded that FrbG-catalyzed hydroxylation precedes FrbF-catalyzed acetylation, as in vitro analysis of purified FrbG shows no activity toward the N-acetylated substrate, CMP-5’-N-acetyl-3-aminopropylphosphonate (CMP-5’-Ac3APn). FrbF, however, was found to catalyze acetylation of either the primary amine substrate, CMP-5’-3-aminopropylphosphonate (CMP-5’-3APn), or the hydroxylamine substrate, CMP-5’-N-hydroxy-3-aminopropylphosphonate (CMP-5’-H3APn). Interestingly, FrbF was observed to more readily catalyze acetylation of the primary amine than the hydroxylamine. This is a surprising finding in the context of the biosynthetic pathway, given that CMP-5’-H3APn is the true substrate for FR-900098 synthesis, and also biochemically, as the hydroxylamine moiety would be expected to be more reactive toward alkylation.

Here, it is our goal to confirm the qualitative studies previously performed with FrbF by alternate methods, and further to quantify the kinetic behavior of this enzyme with its native substrates. Additionally, thorough analysis of the active site will be applied to elucidate the catalytic mechanism of this enzyme. Finally, the substrate specificity of the wild type enzyme will be explored with respect to its acyl-donor to inform future directed evolution efforts targeting the substrate specificity of FrbF for biosynthesis of novel FR-900098 derivatives.
2.2 Results and Discussion

2.2.1 Kinetics characterization with phosphonate substrates

2.2.1.1 CMP-5’-3APn synthesis and kinetics

Although the FR-900098 biosynthetic pathway has been definitively demonstrated to proceed via N-hydroxylation followed by N-acetylation [2], FrbF is nevertheless able to acetylate both the primary amine substrate CMP-5’-3APn and the hydroxylamine substrate CMP-5’-H3APn (shown in Figure 2.1a). As a result, CMP-5’-3APn was enzymatically synthesized using FrbH and purified for kinetic assays. Kinetic parameters with CMP-5’-3APn have previously been reported using a continuous assay with 5,5’-dithiobis-(2-nitrobenzoic acid) (DTNB), but numerous efforts to replicate these results under identical conditions proved unsuccessful. This is most likely due to the limited sensitivity of DTNB-based quantitation, especially given the low activity level of FrbF. Further, DTNB can react with the protein itself, increasing the noise level. As a result, alternate methods were pursued to definitively ascertain the correct kinetic parameters.

The two methods considered for measuring kinetics were an enzyme-coupled assay with α-ketoglutarate dehydrogenase, in which the free CoA product of FrbF catalysis is rapidly converted to succinyl-CoA by α-ketoglutarate dehydrogenase with concomitant reduction of NAD⁺, and an HPLC-based assay, in which the CMP-conjugated product can be directly assayed at a UV absorbance wavelength of 254 nm. Evaluation of both methods found HPLC-based quantitation to have a much better signal-to-noise ratio, and so this method was chosen for subsequent assays. Surprisingly, linear dependence on substrate concentration was observed far beyond the $K_m$ value of 35 µM previously reported, ultimately yielding a $K_m$ value of 391 µM.
with a $k_{cat}$ value of 2.042 min$^{-1}$ (Table 2.1). Given that CMP-$5’$-$3$AP$n$ is not the native substrate for FrbF, however, it is not surprising to see this relatively low level of activity. Further evaluation of the enzyme with the His$_6$ tag at the C-terminal position or with the N-terminal His$_6$ tag removed via thrombin cleavage did not show significantly altered activity (data not shown).

2.2.1.2 CMP-$5’$-$H3$AP$n$, the native phosphonate substrate for FR-900098 production

The native substrate for FR-90098 synthesis, CMP-$5’$-$H3$AP$n$, can be synthesized in vitro using the purified FrbH and FrbG enzymes. However, numerous efforts to observe FrbF catalysis with the purified reaction product revealed no reproducible activity. Subsequent LC-MS analysis of the purified substrate revealed a dominant signal corresponding to a parent ion m/z of 457 in negative mode, with only a minimal signal at the expected m/z of 459. It was then proposed that the m/z = 457 signal comes from the nitroso derivative CMP-$5’$-$NO3$AP$n$, which could potentially be formed via a second hydroxylation of the hydroxylamine to the dihydroxy intermediate, followed by a spontaneous dehydration.

Two schemes were proposed for the oxidation of the hydroxylamine to the nitroso. First, the hydroxylamine could become a surrogate substrate for FrbG and pass through a second catalytic cycle, yielding the nitroso after spontaneous dehydration. Second, the hydroxylamine could simply undergo a non-enzymatic oxidation, either through reaction with oxygen or through a disproportionation reaction with another hydroxylamine molecule. To study the first possibility, the FrbG reaction was monitored over time by LC-MS, and the signals for the substrate and both products were followed with time. In this case, we would expect to see an initial formation of the hydroxylamine, followed by consumption of this compound and a lagging increase in
formation of the nitroso. However, concomitant formation of both compounds was observed, even at the earliest time points assayed, indicating that the second possibility may be at play.

To further study the second possibility, two methods were pursued. First, the FrbG reaction was run for a specified amount of time, after which the enzyme was removed by diafiltration. Subsequently, the filtered reaction mixture was incubated at a constant temperature and periodically analyzed by LC-MS. At 30 °C, rapid loss of the hydroxylamine compound was observed, with half of the sample depleted in ~90 min (Figure 2.2a). Even at 5 °C, the m/z = 459 signal was completely lost from the LC-MS trace following overnight incubation. To corroborate this observation, attempts were made to purify the hydroxylamine from FrbG reaction mixtures, but this could not be achieved. Instead, the nitroso compound was purified and subsequently reduced back to the hydroxylamine under anaerobic conditions using samarium diiodide. Following anaerobic incubation for several hours, little reformation of the nitroso was observed by LC-MS. Once exposed to air, however, rapid reformation occurred, with the hydroxylamine signal almost completely lost after 30 min (Figure 2.2b). Thus, it can be concluded that the conversion of the hydroxylamine to the nitroso is a rapid non-enzymatic process, as has been observed previously for other hydroxylamine compounds [5].

To assess the kinetic behavior of FrbF with this substrate, in situ generation via concomitant FrbG catalysis was attempted. Interestingly, incubation of equimolar amounts of FrbF and FrbG showed no significant difference in FrbF activity, both via HPLC analysis and DTNB endpoint assay. This is likely due to the fact that even though FrbG can supply the true hydroxylamine substrate for FrbF, both enzymes can also compete for the primary amine substrate, and some of
the hydroxylamine can be lost to the non-reactive nitroso. Increasing the FrbG concentration, however, led to an increase in observed FrbF activity, up to a 10:1 molar ratio of FrbG to FrbF. Beyond this ratio, FrbF activity did not increase further. Given the same initial concentration of FrbF, the observed activity level of FrbF increased 6.7-fold when supplemented with a 10-fold molar excess of FrbG. This indicates that FrbF does in fact show greater activity with CMP-5’-H3APn than CMP-5’-3APn, which correlates with its understood role in FR-900098 biosynthesis.

2.2.2 Kinetics characterization with coenzyme A substrates

2.2.2.1 Acetyl-CoA, the native acetyl donor for FR-900098 production

For the biosynthesis of FR-900098, FrbF installs an acetyl moiety on the hydroxylamine substrate by transferring the respective functional group from acetyl-CoA. To determine the kinetic parameters with respect to this substrate, we employed the HPLC-based assay described above. In this case, the CMP-5’-3APn surrogate substrate was used to ensure that the concentration of this substrate remained constant from trial to trial. Kinetic characterization with acetyl-CoA reveals a 20-fold lower $K_m$ value than that observed for the CMP-5’-3APn substrate. This is likely due to the significant number of favorable contacts that the enzyme can form with this rather large substrate molecule in its binding cleft, as will be described in the Section 2.2.3.1. The $k_{cat}$ observed with acetyl-CoA is approximately 55 % of the $k_{cat}$ observed for CMP-5’-3APn. This is a reasonable observation given that the acetyl-CoA kinetic parameters were measured with the CMP-5’-3APn concentration near its $K_m$ value, at which point half of the maximum rate would be expected.
2.2.2.2 Evaluation of alternate acyl-CoAs

As described above in the Introduction section, it is our goal not only to characterize the behavior of FrbF toward its native substrates, but also to determine the scope of acyl-donor substrates that it will accept for the synthesis of novel N-acylated FR-900098 derivatives. In the literature, substrate specificity varies significantly between different acyltransferase enzymes. For example, Walker and co-workers studied the N-benzyoltransferase from Taxus cuspidata, which is involved in Taxol biosynthesis [6]. They assayed this enzyme with its native substrate, benzoyl-CoA, as well as the similar analog phenylacetyl-CoA and dissimilar analog acetyl-CoA, but only saw product formation with the native substrate. In contrast, Morell and co-workers studied the relative activity of N-acyltransferases involved in ceramide biosynthesis isolated from mouse brain microsomes [7]. Although the highest activity was observed toward stearoyl-CoA, lower activity was also observed toward lignoceroyl-CoA, palmitoyl-CoA, and oleoyl-CoA, at a ratio of 60:12:3:1, respectively. More recently, Magalhaes and Blanchard studied the AAC(3)-IV aminoglycoside acetyltransferase from E. coli, which is involved in antibiotic resistance [8]. They looked at the shorter acyl-donors acetyl-, propionyl-, malonyl-, and butyryl-CoA, and determined kinetic constants for all four. Interestingly, acetyl- and propionyl-CoA showed similar $k_{cat}$ values, but a difference in $K_m$ of ~6-fold in favor of acetyl-CoA. Extending the chain further, propionyl- and malonyl-CoA showed indistinguishable $K_m$ values, but a 3.7-fold difference in $k_{cat}$ in favor of propionyl-CoA. Butyryl-CoA was by far the least preferred substrate, with both the highest $K_m$ and lowest $k_{cat}$. Finally, as an example of substrate promiscuity, Killenberg and Jordan showed approximately equivalent $K_m$ and $V_{max}$ values for the bile acid-CoA:amino acid N-acyltransferase from rat liver toward four different bile acid-CoA substrates [9].
For our studies, we wanted to determine the breadth of substrate promiscuity exhibited by FrbF toward its acyl-donor. As a result, seven alternate acyl-CoAs were selected for relative activity assays, including propionyl-, isobutyryl-, β-hydroxybutyryl-, acetoacetyl-, malonyl-, succinyl-, and glutaryl-CoA (Figure 2.1b). These substrates were selected to gradually expand the size of the acyl substituent, such that only small changes to the FR-900098 scaffold will be made. Additionally, these acyl-CoA substrates are biologically relevant, and their overproduction could potentially be engineered in future host strains for the overproduction of specific FR-900098 derivatives.

Activity assays with alternate acyl-CoA substrates were first carried out using the DTNB-based assay method. However, low activity levels and competing non-enzymatic hydrolysis of the acyl-CoA thioester bonds hindered accurate determination of activity, and so HPLC-MS analysis was instead employed. Initial analyses showed no detectable activity by UV absorbance or MS for glutaryl-CoA, which is not surprising given that this was the largest acyl substituent evaluated. Surprising, however, was the absence of detectable activity for β-hydroxybutyryl-CoA, given that activity could be detected for the similar acetoacetyl- and malonyl-CoA substrates. This suggests that the second carbonyl moiety not present in β-hydroxybutyryl-CoA may be necessary for favorable interactions with the active site to support catalysis. Of the other substrates evaluated, both isobutyryl- and succinyl-CoA showed activity by MS analysis, as indicated by the presence of the expected m/z parent ion with the corresponding m/z = 322 daughter ion due to loss of CMP. However, no detectable UV peak was observed for either of these two products, preventing accurate quantitation of relative activity toward these novel substrates. Propionyl-, malonyl-, and acetoacetyl-CoA all showed activity both by UV
absorbance and MS analysis (Figure 2.3). As a result, activity levels were measured for each of these substrates relative to acetyl-CoA. Additionally, kinetic parameters were determined for propionyl-CoA, which exhibited the highest relative activity. This is not surprising, given that the propionyl moiety is the most similar to the acetyl moiety.

Interestingly, the $K_m$ value obtained for propionyl-CoA is approximately the same as that obtained for acetyl-CoA. This likely indicates that binding of the acyl-CoA substrate relies predominantly on interactions with the CoA backbone itself, and not with the acyl functionality. The ~12-fold drop in $k_{cat}$ for this substrate relative to acetyl-CoA, then, is most likely caused by the increased steric hindrance introduced by the bulkier acyl chain, which could potentially block the amine or hydroxylamine of the substrate from attacking the thioester carbonyl. This is further evinced by the decreasing activity observed with the bulkier malonyl-CoA (2.7 % relative activity) and acetoacetyl-CoA (0.4 % relative activity), and the lack of quantifiable activity observed when the isobutyryl-CoA or succinyl-CoA substrates are utilized.

### 2.2.3 Elucidation of the structure and catalytic mechanism

#### 2.2.3.1 Analysis of the crystal structure

The x-ray crystal structure of FrbF in complex with acetyl-CoA was solved in collaboration with Brian Bae of Satish Nair’s group at the University of Illinois, Urbana-Champaign. As anticipated based on the lack of sequence similarity, FrbF shares no structural similarity to members of the GNAT superfamily of enzymes. Instead, the most similar structural homologs are the proteins YokD from *Bacillus subtilis* and BA2930 from *Bacillus anthracis*, both of which are biochemically uncharacterized. Querying the GenBank database reveals several other
uncharacterized gene sequences, mostly from *Bacillus*. As the functions of these genes are as yet undetermined, it is possible that they too could be involved in the synthetic pathways of heretofore undiscovered secondary metabolites.

The overall structure of FrbF is comprised of a large domain and a small domain connected by multiple linkers without secondary structure (Figure 2.4a). This forms a significant cleft into which the acetyl-CoA molecule is aligned. At the adenosine end of the molecule, a hydrophobic pocket composed of Leu-48, Trp-53, and Val-54 interacts favorably with the adenine ring. Additionally, hydrogen bonds are formed with residue Arg-19 and the backbone carbonyls of Gly-52 and Val-54. Further favorable interactions are formed between the phosphopantetheine arm and residues Ser-188 and His-45. Finally, multiple polar residues are located in the vicinity of the acetyl-group thioester, the catalytic relevance of which will be explored in the next section. Adjacent to the acetyl-CoA binding cleft is another site where the CMP-conjugated phosphonate substrate is likely to bind (Figure 2.4b). Although the co-crystallized structure with this substrate was not obtained, it is likely that basic residues within this region help to favorably interact with the diphosphate motif. This type of interaction would also help to explain the absence of FrbF activity observed toward the non-CMP-conjugated 3APn substrate in previous studies.

2.2.3.2 Mutational analysis of the active site

In order to elucidate the catalytic mechanism of FrbF, the residues located within 6 Å of the acyl moiety of the bound acetyl-CoA were analyzed (Figure 2.5). In principle, one could expect the reaction mechanism to proceed either via direct attack of the amine or hydroxylamine substrate
on the carbonyl group of acetyl-CoA, as has been observed for the majority of GNAT superfamily members, or via an acyl-enzyme intermediate, as demonstrated in the yeast histone acetyltransferase Esa1 and in multiple mycobacteria [10, 11]. Of the residues present in the active site, Glu-187, Ser-188, Thr-190, and Glu-231 were identified as possible general acids, while only His-193 was identified as a potential general base. Alanine mutants at each of these five positions were created, and their relative activities were measured in comparison to the wild type enzyme (Figure 2.6). The results indicate that the Glu-187→Ala and Glu231→Ala mutants demonstrate activities similar to the wild-type enzyme, confirming that these two residues do not play a role in catalysis and further ruling out mechanisms that require formation of a covalent acyl intermediate with an active site carboxylate. The Ser-188→Ala mutant exhibited a modest decrease in activity, as anticipated based on its interaction with the phosphopantetheine arm of the acetyl-CoA substrate. In contrast, the activity of the His-193→Ala mutants was less than 1% of that of the wild-type, and no activity could be observed for the Thr-190→Ala mutant. These results strongly suggest the involvement of Thr-190 and His-193 in acid/base chemistry during acetyl-group transfer. These results are consistent with the mechanistic route proposed in Figure 2.7. Following the binding of the substrate conjugate CMP-5'-3APn and acetyl donor acetyl-CoA at the active site, His-193 functions as a general base to facilitate the attack of the amine or hydroxylamine on the electrophilic carbonyl of the acetyl group. Protonation of the resultant tetrahedral intermediate results in its collapse to yield the hydroxamate and CoA. Alternatively, it is worth noting that His-193 could instead serve to stabilize the tetrahedral intermediate, as the N-O distance varies from 2.9 Å to 4.2 Å between the two pairs of FrbF dimers.
Further analysis was carried out at the Thr-190 position through mutations to both Ser and Cys, as well as mutation of His-45 to Ala. His-45 is the nearest residue to Thr-190, and so we speculated that this residue might serve to activate Thr-190 for a novel reaction mechanism involving attack of Thr-190 on the carbonyl of the acetyl group to form an acyl-enzyme intermediate. The His-45→Ala mutation reduced the relative activity of FrbF to 1.5% of the wild type, supporting the catalytic relevance of this residue. However, high-resolution mass spectrometry provided no evidence of an acyl-enzyme intermediate (data not shown). Further analysis of the crystal structure suggests that His-45 can participate in hydrogen bonding interactions both with a carbonyl group of the acetyl-CoA phosphopantetheine arm and with the backbone amide nitrogen of either Thr-190 or Ser-191, suggesting that this residue is critical for the positioning of acetyl-CoA and the catalytically requisite Thr-190 residue. Further, the Thr-190→Cys mutation, which would more closely reflect the active site of NAT enzymes that utilize an acyl-enzyme intermediate, lowered the activity dramatically, discrediting the prospect of a novel, Thr-mediated ping pong mechanism. Finally, the Thr-190→Ser mutation had no detectable effect on the relative activity.

2.3 Conclusions and Outlook

In this study, we structurally and biochemically characterized the N-acetyltransferase FrbF from the FR-900098 biosynthetic cluster. Kinetic parameters were determined with acetyl-CoA, the native acyl-donor, and CMP-5’-3APn, a surrogate phosphonate substrate. Additionally, we identified a rapid, non-enzymatic oxidation of the native hydroxylamine substrate, CMP-5’-H3APn, to its nitroso derivative. Although this prevented the measurement of kinetic parameters with this substrate, we nevertheless demonstrated via in situ generation of CMP-5’-H3APn that
this compound is in fact the preferred substrate for FrbF, both through HPLC analysis and a chemically-coupled assay. Further, we demonstrated that FrbF exhibits reasonably high substrate specificity for acetyl-CoA as its acyl-donor, as relative activity levels fell significantly with increasing substrate size. For propionyl-CoA, this was due predominantly to a decrease in $k_{cat}$, indicating that it is likely hindrance of attack by the nucleophilic substrate, and not of binding, that lowers activity with novel acyl-donors. Finally, we employed a site-directed mutagenesis approach to elucidate the roles of key active site residues in catalysis.

The knowledge gained here will be useful in the development of new FR-900098 derivatives. Understanding of the kinetic behavior of FrbF will aid in the selection of appropriate screening conditions for novel FrbF substrates. Moreover, by studying the acyl-donor scope of wild type FrbF, we can better select target substrates for future directed evolution studies. With knowledge of the crystal structure, too, we can more rationally target our directed evolution approach to specific residues that may be involved in limiting the acyl-donor size, while leaving intact the residues necessary for catalysis. Finally, characterization of this representative of a new class of $N$-acetyltransferases may facilitate future studies of previously uncharacterized members of this family, which could lead to the discovery of new and interesting secondary metabolite gene clusters.

### 2.4 Materials and Methods

#### 2.4.1 Materials

The *Escherichia coli* cloning strain DH5α, expression strain BL21(DE3) and the expression vector pET28a were obtained from EMD Biosciences (San Diego, CA). Kanamycin and
isopropyl-β-D-1-thiogalactopyranoside (IPTG) were purchased from Gold Biotechnology (St. Louis, MO). All chemicals, including reaction buffers and the reagents samarium diiodide, cytidine triphosphate (CTP), 2-amino-4-phosphonobutyrate (2APn), and all acyl-CoAs, were purchased from Sigma-Aldrich (St. Louis, MO) except for the co-factor NADPH, which was purchased from Applichem (Darmstadt, Germany), and organic solvents, which were purchased from Thermo-Fisher Scientific (Pittsburgh, PA). All PCR reagents and restriction enzymes were purchased from New England Biolabs (Ipswich, MA), and DNA Miniprep and Gel Purification Kits were purchased from Qiagen (Valencia, CA). All primers were synthesized by Integrated DNA Technologies (Coralville, IA). Talon Cobalt immobilized metal affinity chromatography (IMAC) resin was purchased from Clontech Laboratories (Mountain View, CA). All other materials were purchased from Thermo-Fisher Scientific.

2.4.2 Protein expression and purification

The *E. coli* BL21(DE3) strains containing pET28a-FrbF, pET28a-FrbG, and pET28a-FrbH were previously prepared via ligation of the respective ORFs between the NdeI and HindIII restriction sites. Cells were grown in Terrific Broth (TB) media supplemented with kanamycin (50 μg/mL) at 37 °C to an OD$_{600}$ of ~0.8, after which induction was carried out by addition of 0.3 mM IPTG at 25 °C. In the case of FrbH, IPTG induction was accompanied by addition of 50 μg/mL pyridoxal-5’-phosphate (PLP). After 18 hr, the cells were harvested by centrifugation at 7500 rpm for 15 min and resuspended in 20 mM Tris-HCl (pH 7.65), 0.5 M NaCl, and 15 % glycerol supplemented with 1 mg/mL lysozyme. After a freeze-thaw cycle at -80 °C, the cell suspension was sonicated to ensure sufficient lysis. The lysate was clarified multiple times by centrifugation at 15000 rpm for 15 min, after which the His$_6$-tagged proteins were purified by affinity
chromatography on TALON Superflow Co\textsuperscript{2+} resin coupled to fast-performance liquid chromatography. The eluted proteins were washed three times in 50 mM HEPES (pH 7.25), concentrated, and stored in 15 % glycerol at -80 °C.

2.4.3 Generation of FrbF mutants

FrbF amino acid candidates for mutation were visualized using Molecular Operating Environment software (Chemical Computing Group, Montreal, Canada). Mutants of the FrbF enzyme were generated by the megaprimer PCR method. Briefly, primers were designed at the site of interest containing the desired mutation such that a larger primer (a “megaprimer”) containing the mutation could be generated by PCR, using the regular forward or reverse primer as the second primer (whichever yielded the shortest PCR product). Next, PCR amplification of the full-length mutant gene was carried out using the megaprimer at one end of the gene and the regular primer at the opposite end. Finally, the full length gene containing the mutation was digested with NdeI and HindIII and ligated into the correspondingly digested pET28a vector backbone. The ligation product was then used to transform the E. coli cloning strain DH5\textalpha. Transformants were picked and grown in 5 mL LB + 50 µg/mL kanamycin. Their plasmids were isolated and sequenced by ACGT, Inc. (Wheeling, IL), and Sequencher software (Gene Codes Corporation, Ann Arbor, MI) was used to confirm the presence of only the desired mutation. The correct plasmids were then used to transform the expression strain E. coli BL21(DE3), and frozen stocks of the resultant strains were stored in 15 % glycerol at -80 °C.
2.4.4 Preparation of CMP-5’-3APn

CMP-5’-3APn for kinetic and relative activity assays was synthesized enzymatically using purified N-His₆-FrbH. Synthesis was carried out in 50 mM HEPES buffer (pH 7.25) with 1.5 mM CTP, 3 mM 2APn, 10 mM MgCl₂, and ~10 µM FrbH at 30 °C for 4 hr. Following removal of FrbH with a 10 kDa cutoff Amicon centrifugal filter unit (Millipore, Billerica, MA), the product was purified by fractionation on an Agilent 1100 Series HPLC (Agilent, Palo Alto, CA) using an Alltech Prevail C18 reverse phase column (Grace Davison Discovery Sciences, Deerfield, IL) with isocratic flow of 15 mM ammonium formate buffer. Fractions containing the product were pooled, lyophilized, and resuspended in double-distilled water. The concentration of the purified CMP-5’-3APn product was determined by UV absorbance at 254 nm.

2.4.5 CMP-5’-H3APn oxidation assays

Enzymatic synthesis of CMP-5’-H3APn was performed using purified N-His₆-FrbG. Reactions were carried out in 50 mM HEPES buffer (pH 7.25) with 500 µM CMP-5’-3APn, 500 µM NADPH, and ~30 µM FrbG at 30 °C for 15 min. Following removal of FrbG with a 10 kDa cutoff Amicon centrifugal filter unit, the product mixture was incubated at 30 °C. Samples were collected at specified time points and kept on dry ice until analysis using an Agilent XCT ion-trap MSD mass spectrometer at the Roy J. Carver Metabolomics Center (University of Illinois, Urbana, IL). Chemical synthesis of CMP-5’-H3APn was performed using the oxidation product CMP-5’-NO₃APn, which was purified from FrbG reaction mixtures via HPLC fractionation. Reduction of the nitroso to the hydroxylamine was performed anaerobically following the method of Kende and Mendoza [12]. In an anaerobic chamber, the lyophilized nitroso compound was resuspended in a mixture of tetrahydrofuran (THF) and methanol (95%/5% v/v).
To this solution was added a 4-fold molar excess of samarium diiodide in THF. Following agitation for 3 min at room temperature, the reaction was acidified with 0.1 % trifluoroacetic acid and extracted with ethyl acetate. The product solution was then analyzed by LC-MS both before and after aerobic exposure.

2.4.6 Kinetic assays

All activity assays were performed with N-His$_6$-tagged FrbF in 50 mM HEPES buffer (pH 7.25) at 30 °C. For determination of kinetic parameters, the concentration of one substrate was maintained at 400 µM, while the other substrate concentration was varied from 0 – 1000 µM. Samples were collected over a 7 minute time course for product quantitation and determination of initial rate. Reactions were initiated with the addition of 1.3 µM FrbF and quenched with 1 % trifluoroacetic acid. All assays were performed in triplicate. Kinetic parameters were fit to the data using nonlinear least-squares regression in OriginPro 8 (OriginLab Corporation, Northampton, MA). Relative activity assays with alternate CoA substrates were performed with 400 µM CMP-5’-3APn and 200 µM of the CoA compound. Reactions were initiated with 15 µM FrbF, and samples were collected over a 60 minute time course. For substrates for which no activity was detected in this time, the assay time was extended up to 12 hr. For all assays, products were detected by UV absorbance at 254 nm on an Agilent 1100 Series HPLC (Agilent, Palo Alto, CA), with retention times confirmed using an Agilent XCT ion-trap MSD mass spectrometer at the Roy J. Carver Metabolomics Center (University of Illinois, Urbana, IL). For all LC-MS analyses with alternate CoA substrates, product identities were confirmed based on detection of the expected m/z parent ion with a corresponding m/z $= 322^-$ (CMP) daughter ion in the MS/MS profile.
2.4.7 Relative activity assays

Relative activity assays with purified N-His$_6$-tagged FrbF mutants were performed with 400 µM CMP-5'3APn and 200 µM acetyl-CoA over a 7 minute time course. For those mutants that exhibited significantly reduced activity, the assay time was extended up to 60 minutes. Samples were collected and analyzed as described above.

2.5 References


### 2.6 Tables

**Table 2.1:** FrbF kinetic parameters (mean ± standard deviation of three independent trials)

<table>
<thead>
<tr>
<th>Substrate</th>
<th>$K_m$ (µM)</th>
<th>$k_{cat}$ (min$^{-1}$)</th>
<th>Relative activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMP-5'-3APn</td>
<td>391 ± 61</td>
<td>2.04 ± 0.14</td>
<td>1.0</td>
</tr>
<tr>
<td>CMP-5'-H3APn</td>
<td>ND</td>
<td>ND</td>
<td>6.7$^a$</td>
</tr>
<tr>
<td>Acetyl-CoA</td>
<td>19.5 ± 0.4</td>
<td>1.12 ± 0.03$^b$</td>
<td>1.0</td>
</tr>
<tr>
<td>Propionyl-CoA</td>
<td>23.6 ± 4.2</td>
<td>0.089 ± 0.024$^b$</td>
<td>0.080 ± 0.022</td>
</tr>
<tr>
<td>Malonyl-CoA</td>
<td>ND</td>
<td>ND</td>
<td>0.027 ± 0.006</td>
</tr>
<tr>
<td>Acetoacetyl-CoA</td>
<td>ND</td>
<td>ND</td>
<td>0.0044 ± 0.0013</td>
</tr>
<tr>
<td>Isobutyryl-CoA</td>
<td>ND</td>
<td>ND</td>
<td>MS</td>
</tr>
<tr>
<td>Succinyl-CoA</td>
<td>ND</td>
<td>ND</td>
<td>MS</td>
</tr>
</tbody>
</table>

$^a$: substrate generated *in situ* with 10-fold excess FrbG

$^b$: measured at 400µM CMP-5'-3APn

ND: not determined

MS: below UV detection limit; product observed by MS/MS only
2.7 Figures

Figure 2.1: Substrates investigated in this study. (a) Acyl-group acceptors. (b) Coenzyme A acyl-donors.
Figure 2.2: Analysis of the non-enzymatic oxidation of the hydroxylamine product to its nitroso derivative. (a) FrbG reaction mixtures were monitored by LC-MS following removal of the enzyme by diafiltration for loss of the hydroxylamine product ($m/z = 459$). (b) The nitroso derivative was reduced back to the hydroxylamine under anaerobic conditions using SmI$_2$, but rapidly reformed upon exposure to air. The left trace shows extracted ion chromatograms for the hydroxylamine (red curve) and the nitroso (blue curve), along with a significant CMP side product (green curve), following anaerobic reduction. The right trace shows reformation of the nitroso (blue) following 30 minutes of aerobic exposure.
Figure 2.3: HPLC UV traces from reactions of FrbF with alternate acyl-CoA substrates. The inset MS/MS trace on each plot shows fragmentation (at the position indicated by the arrow) of the expected parent ion for each product to the characteristic m/z = 322 (CMP) daughter ion.
Figure 2.4: (a) The backbone of FrbF in complex with acetyl-CoA (Ac-CoA), showing the two-domain architecture. (b) Surface rendering of FrbF showing the acetyl-CoA binding cleft and the likely binding site for the phosphonate substrate. Images generated by Satish Nair.
Figure 2.5: Active site residues within 6 Å of the acetyl-CoA thioester (center) selected for site-directed mutagenesis.
Figure 2.6: Activity levels of each of the FrbF point mutants toward acetyl-CoA and CMP-5’-3APn relative to the wild type enzyme.
Figure 2.7: The proposed mechanism for the $N$-acetylation of CMP-5'-H3APn catalyzed by FrbF.
Chapter 3: Directed Evolution of FrbF for Synthesis of FR-900098P and Characterization with its Target Enzyme from *Plasmodium falciparum*

3.1 Introduction

In the field of pharmaceutical development, the most desirable drug candidate is one that comes closest to being a so-called magic bullet; in other words, a drug that targets the root of the diseased state with no off-target effects. It follows, then, that the best drugs for a parasitic disease such as malaria are those that are simultaneously toxic to the parasite itself but innocuous to the human host. In order to achieve this sort of specificity for antimalarials, one can exploit the innate differences between the physiology of *Plasmodium* and humans. One such difference is the pathway by which the isoprenoid precursors isopentynyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) are synthesized.

Isoprenoids are a very large and diverse family of biological compounds necessary to sustain life. Included among them are sterols, which can help regulate membrane fluidity and serve as signaling molecules, and ubiquinones involved in electron transport for aerobic respiration. In humans, the isoprenoid precursors IPP and DMAPP are synthesized via the mevalonate pathway, which was discovered in the 1950s (Figure 3.1a). This pathway begins with acetyl-CoA, derived from central metabolism. Condensation of two acetyl-CoA molecules yields acetoacetyl-CoA, which condenses with a third acetyl-CoA molecule to generate 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). A subsequent reduction liberates mevalonic acid, which is then twice
phosphorylated. Finally, decarboxylation yields IPP, which can be isomerized into DMAPP. In *Plasmodium falciparum*, the enzymes necessary for the mevalonate pathway are absent; instead, orthologs of the non-mevalonate pathway enzymes are found [1]. Although discovered several decades after the mevalonate pathway, the non-mevalonate pathway, alternatively known as the 2C-methyl-D-erythritol 4-phosphate (MEP) pathway or the 1-deoxy-D-xylulose 5-phosphate (DXP) pathway for its key intermediates, has been studied and characterized in great detail [2]. This pathway begins with the condensation of pyruvate and glyceraldehyde 3-phosphate from central metabolism to produce DXP (Figure 3.1b). Following a concomitant reduction and isomerization to MEP, the compound is conjugated to a cytidine monophosphate (CMP) moiety. Next, phosphorylation occurs at the 2-position, and the compound is subsequently converted to a cyclic diphosphate via loss of CMP. Finally, a reductive ring opening occurs, and IPP and DMAPP are formed.

The two pathways described above feature distinct biochemical transformations catalyzed by unique enzymes; as a result, each enzyme in the non-mevalonate pathway represents an ideal target for therapeutic intervention with minimal risk to the human host. Recent efforts toward the discovery and development of non-mevalonate pathway inhibitors have focused on many of these enzymes as targets. For example, the first enzyme in the pathway, DXP synthase (Dxs), can be inhibited by 5-keto clomazone, which is consequently used as an herbicide to block isoprenoid biosynthesis in the chloroplast [3]. Additionally, 3-(4-chloro-phenyl)-5-benzyl-4H-pyrazolo[1,5-a]pyrimidin-7-one, a known inhibitor of mammalian transketolase superfamily members, has been explored as a scaffold for the design of novel Dxs inhibitors for tuberculosis treatment [4]. Inhibitors of other pathways enzymes, including 4-diphosphocytidyl-2C-methyl-
D-erythritol synthase (Cms), 4-diphosphocytidyl-2C-methyl-D-erythritol kinase (Cmk), and 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (Mecds), have also been reported, and are reviewed in the recent publication by Obiol-Pardo and co-workers [5].

The discovery and design of inhibitors of the second non-mevalonate pathway enzyme, DXP reductoisomerase (Dxr), have predominantly focused on the fosmidomycin scaffold; these efforts were reviewed in the previous chapter. To evaluate the effectiveness of fosmidomycin and FR-900098 as Dxr inhibitors, inhibition parameters have been determined with Dxr enzymes from multiple species (Table 3.1). Kuzuyama and co-workers first determined the inhibition constant of fosmidomycin against the Dxr from E. coli (EcDxr) to be 38 nM, and characterized the compound as a competitive inhibitor [6]. Later on, evidence was found for slow, tight-binding inhibition of EcDxr by fosmidomycin, with a subsequent $K_i^*$ value of 21 nM [7]. Fosmidomycin inhibition constants have also been measured for Zymomonas mobilis Dxr (ZmDxr), where $K_i$ was found to be significantly greater [8], and Synechocystis sp. PCC6803, which was found to have a comparable $K_i$ value to EcDxr but a lower $K_i^*$ [9]. This study also measured the $K_i^*$ for FR-900098 with Synechocystis sp. PCC6803 Dxr, and found it to be ~5-fold lower than that of fosmidomycin. Finally, $IC_{50}$ values have been measured for both fosmidomycin and FR-900098 against Dxr from E. coli, Pseudomonas aeruginosa (PaDxr), and P. falciparum (PfDxr) [10, 11]. Of these three targets, both compounds were most effective against PfDxr. Nevertheless, $K_i$ values have not yet been determined with PfDxr.

Also useful in the study of new Dxr inhibitors is X-ray crystallographic data of the enzyme-inhibitor complex, such that the molecular basis for inhibition can be established. Multiple Dxr-
fosmidomycin co-crystal structures have been solved, including those of *E. coli* [12-14], *Mycobacterium tuberculosis* [15], and *Thermotoga maritima* [16]. However, these enzymes share only modest sequence identity with that of *P. falciparum*, and so their utility in the design of novel antimalarials is limited.

Here, it is our goal to develop a screening method suitable for the identification of mutants of the N-acetyltransferase FrbF capable of utilizing new acyl-donor substrates for the synthesis of novel FR-900098 derivatives. We will utilize the PfDxr enzyme as a tool to assess the efficacy of novel FR-900098 derivatives relative to the parent compound. The enzyme will be characterized kinetically and inhibition constants will be measured with fosmidomycin, FR-900098, and novel derivatives thereof. Further, insights gained from solution of the crystal structure in complex with its inhibitors will aid in the design of new PfDxr inhibitors for future drug development studies.

### 3.2 Results and Discussion

#### 3.2.1 PfDxr: the target enzyme from *Plasmodium falciparum*

**3.2.1.1 Recombinant expression in E. coli**

As described in the previous chapter, it is our goal to harness the natural biosynthetic pathway of FR-900098 production to generate novel *N*-acylated derivatives. The ultimate goal of this project, however, is not just to synthesize new compounds, but to discover among these derivatives more potent antimalarial drugs. It is important to remember that fosmidomycin and FR-900098 are naturally produced as *Streptomycese* secondary metabolites, and not as human antimalarials. Thus, these compounds are not optimized as PfDxr inhibitors, and so there is
likely room for improvement on this front. Studies by Jomaa and co-workers first demonstrated via in vivo studies in three different P. falciparum strains that although they differ only in their N-acylations, FR-900098 is a more potent inhibitor of parasite growth than fosmidomycin, exhibiting IC$_{50}$ values 2- to 3-fold lower than those of fosmidomycin [17]. Further, although they did not quantify any inhibition values, they demonstrated inhibition of the purified recombinant PfDxr enzyme by both compounds, with FR-900098 again showing greater effectiveness.

To develop a means to evaluate the antimalarial potential of novel FR-900098 derivatives, we elected to heterologously express and purify the PfDxr target enzyme in E. coli. Deviating slightly from the literature precedent, we first obtained a complete dxr gene codon-optimized for expression in E. coli, including the coding sequence for the 72 amino acid leader peptide. Following transformation of the BL21(DE3) expression strain with this construct, numerous efforts were made to express the heterologous protein. However, SDS-PAGE analysis revealed neither soluble nor insoluble expression, and subsequent efforts to purify even trace amounts of the protein were unsuccessful (data not shown).

As a result, the core of the gene without the first 216 nucleotides, which appear to code for an endoplasmic reticulum signal peptide and an apicoplast targeting sequence [17], was cloned into a separate vector and used to transform BL21(DE3). SDS-PAGE analysis of expression cultures revealed apparent overexpression of the truncated protein, but predominantly in insoluble form. Following optimization of expression conditions, including expression temperature, IPTG concentration, and total culture volume, the pure PfDxr protein was successfully purified, as
confirmed by SDS-PAGE. To demonstrate activity of the pure protein, an assay with the native substrates NADPH and DXP was carried out, demonstrating reduction in absorbance at 340 nm indicative of NADPH oxidation (Figure 3.2). To prove that this activity was due to the PfDxr enzyme itself and not a contaminating protein, the assay was repeated in the presence of fosmidomycin, demonstrating nearly complete inhibition.

3.2.1.2 Kinetic characterization

Although the PfDxr enzyme has previously been studied [17], it has yet to be kinetically and structurally characterized. Nevertheless, limited kinetic parameters have been determined for other DXP reductoisomerase enzymes from different species. For example, Kuzuyama and co-workers determined the DXP \( K_m \) and \( k_{cat} \) for the \( E. coli \) Dxr to be 250 \( \mu \)M and 59 min\(^{-1} \), respectively, in the presence of \( \text{Mn}^{2+} \) as the divalent cation [18]. Interestingly, the \( K_m \) value decreased in the presence of \( \text{Mg}^{2+} \) or \( \text{Co}^{2+} \), although \( k_{cat} \) decreased as well. Grolle and co-workers reported the Dxr from \( Z. mobilis \) to have a comparable \( K_m \) of 300 \( \mu \)M with a \( V_{\max} \) of 19.5 U mg protein\(^{-1} \), which is equivalent to a \( k_{cat} \) of \( \sim 13 \text{ s}^{-1} \) assuming a molecular weight of 40 kDa [8]. For \( P. aeruginosa \), a lower \( V_{\max} \) of 11.9 U mg protein\(^{-1} \) (\( k_{cat} \) of \( \sim 8 \text{ s}^{-1} \)) was observed [10]. Argyrou and Blanchard studied divalent cation specificity of the Dxr from \( M. tuberculosis \), finding little difference between \( \text{Co}^{2+} \) and \( \text{Mn}^{2+} \), which both outperformed \( \text{Mg}^{2+} \) [19]. Nevertheless, they did not rule out the biological relevance of \( \text{Mg}^{2+} \), as intracellular levels of this cation are typically much greater than those of the other two. Depending on the metal ion, they determined the \( K_m \) for DXP to be between 4 and 100 \( \mu \)M, with \( k_{cat} \) ranging from 1.6 to 5.0 s\(^{-1} \).
Based on the conditions presented in the Dxr kinetic studies and inhibition studies described above and in the subsequent section, we elected to measure the kinetics of PfDxr in the presence of Mn\(^{2+}\), as the greatest or near-greatest activities have typically been observed with this divalent cation [8, 19, 20]. Only NADPH was evaluated as the co-factor, as no Dxr protein has been identified with preference toward NADH. To conserve materials and purified enzyme, assays were first attempted in a microtiter plate format in a total volume of 100 \(\mu\)L per well. However, results obtained in this format were largely inconsistent (data not shown), and so a more robust platform was sought. Activity measured in a cuvette in a total volume of 300 \(\mu\)L proved to be much more reproducible. However, it was observed that subsequent kinetic trials carried out using the same aliquot of purified enzyme yielded \(k_{\text{cat}}\) values that decreased with time. In contrast, the calculated \(K_m\) values did not change significantly with time, indicating that the enzyme is not highly stable in solution, even when kept on ice. As a result, a fresh aliquot of PfDxr (stored at -80 °C in 15 % glycerol) was used for all subsequently performed assays to try to minimize loss of activity.

The resulting kinetic parameters are shown in Table 3.2. Interestingly, the \(K_m\) value of 52 \(\mu\)M obtained for PfDxr is significantly lower than the value reported for EcDxr in the presence of the same divalent cation. This suggests that the PfDxr enzyme is particularly well suited to bind DXP, which may have implications for the binding of inhibitors as well. The observed \(k_{\text{cat}}\) value of 120 min\(^{-1}\) is approximately twice that of EcDxr, but not as great as those obtained for the other species described above. Overall, these kinetic parameters are most similar to those obtained for MtDxr with Mg\(^{2+}\) (\(K_m = 42 \mu\)M, \(k_{\text{cat}} = 2.1 \text{s}^{-1}\)) [19].
Along with biochemical characterization of PfDxr activity, it is our goal to better understand the molecular basis for PfDxr inhibition by fosmidomycin and FR-900098. As a result, the x-ray crystal structure of PfDxr in complex with each of these inhibitors, along with the co-factor NADPH and the divalent cation Mn$^{2+}$, was solved in collaboration with Brian Bae and Zhi Li of Satish Nair’s group at the University of Illinois, Urbana-Champaign. As mentioned in the Introduction, there are multiple structures of Dxr enzymes available, some of which have been co-crystallized with the fosmidomycin inhibitor. Nevertheless, utilization of these structures for drug design has yet to generate any new leads, and so it is likely that there are key differences between PfDxr and the currently available structures.

The co-crystal structures with fosmidomycin and FR-900098 were solved to resolutions of 1.9 Å and 2.2 Å, respectively. In both structures, the hydroxamate moiety of the inhibitor coordinates to the Mn$^{2+}$ ion via its two oxygen atoms, as shown in Figure 3.3. Additionally, the N-acyl moiety inserts into a hydrophobic binding pocket formed by Trp-296, Met-298, Met-360, and the nicotinamide ring of the cofactor. The presence of this pocket justifies the increased potency of FR-900098 relative to fosmidomycin as a PfDxr inhibitor, and suggests that further hydrophobic extension of the N-acyl side chain could provide even better binding to the active site. At the other end of the molecule, the phosphonate moiety interacts favorably with residues Ser-269 and His-293 via hydrogen bonding interactions.

Comparison of our PfDxr structure to the EcDxr co-crystal structure with the co-factor and fosmidomycin reveals key differences that suggest increased potency of the inhibitor toward
PfDxr. For example, PfDxr possesses a narrower substrate-binding cavity, with movement of Met-298 toward the N-acyl moiety to form the aforementioned hydrophobic pocket. Concomitantly, H-293 moves into a position in which it can hydrogen bond with the phosphonate end of the inhibitor. The equivalent residues in EcDxr (His-208 and Met-213) are not positioned well for either of these interactions. An additional difference is seen at the position of Ser-269 in the PfDxr structure, which can also hydrogen bond with the phosphonate. In EcDxr, Gly-184 occupies this space and thus cannot hydrogen bond.

To evaluate these interactions biochemically, the Ser-269→Ala and Met-298→Ala mutants were constructed. Following purification, both mutants were assayed for activity. While the S269A mutant was capable of turning over NADPH and was inhibited by fosmidomycin, the M298A mutant was inactive. A repeated attempt at purification of the M298A mutant again yielded the inactive enzyme. This likely indicates that this active site residue is catalytically requisite, and might be involved in maintaining structural integrity of the active site or positioning the substrate for catalysis. Inhibition assays were carried out with the active S269A mutant, and an increase in the $K_i$ value for fosmidomycin from 7.5 nM to 9.8 nM was observed. This demonstrates that this residue’s interaction with the inhibitor is partly responsible for the superior inhibition of PfDxr by fosmidomycin relative to EcDxr.

### 3.2.2 FR-900098P

#### 3.2.2.1 Synthesis in vitro

As described in the previous chapter, we evaluated several alternate acyl-CoA substrates with FrbF to determine the ability of the wild type enzyme to synthesize novel FR-900098
derivatives. Unsurprisingly, the greatest relative activity level (~8 % compared to acetyl-CoA) was observed with the structurally most similar propionyl-CoA substrate. Based on our observations of a hydrophobic binding pocket to accommodate the N-acyl moiety in the PfDxr co-crystal structure, we selected the N-propionyl FR-900098 derivative, which is here called FR-900098P, as a target for synthesis and subsequent characterization.

Synthesis of FR-900098P was first attempted through sequential reactions with the purified FrbH, FrbG, and FrbF enzymes. After each step, the enzyme was removed via diafiltration and products were isolated via HPLC fractionation. However, this procedure resulted in almost undetectable amounts of the penultimate product, CMP-5’-FR-900098P. This was likely due to the non-enzymatic oxidation of the hydroxylamine intermediate, CMP-5’-H3APn, to its nitroso derivative, as described in the previous chapter. In an attempt to prevent this undesired side reaction, the FrbG and FrbF reactions were instead performed simultaneously, such that the hydroxylamine product of FrbG could immediately be acylated by FrbF. This was found to increase the amount of CMP-FR-900098P generated, such that it became easily detectable by HPLC (Figure 3.4a). Following purification of this product, CMP cleavage was performed by addition of FrbI. Previously, FrbI had been revealed to be highly promiscuous in its nucleotide hydrolase activity, and so cleavage of CMP from the final product was always carried out in a separate reaction [21]. The identity of the newly synthesized product was confirmed by tandem MS/MS (Figure 3.4b). In positive mode, FR-900098 undergoes a characteristic 198→138 fragmentation, indicating loss of both the N-acetyl and N-hydroxyl moieties. As expected, we observed the same characteristic fragmentation (212→138) for FR-900098P.
3.2.2.2 Synthesis in vivo

While in vitro synthesis with purified enzymes is suitable for the production of small amounts of FR-900098P, it is not suitable for large-scale production given the cost of protein purification and of the substrate materials, particularly propionyl-CoA and NADPH. Instead, a commercially viable process would likely follow either a completely synthetic approach or a biocatalytic approach using whole cells as “microbial chemical factories.” Between these two alternatives, there are many advantages to utilizing a biosynthetic framework for the synthesis of value added products [22]. For example, chemical syntheses often require expensive, hazardous, or non-renewable raw materials, and can also require harsh operating conditions. Microbial biosyntheses, in contrast, can be performed using renewable biomass as a primary feedstock, and require only very mild operating conditions. Additionally, biosynthetic machinery can access complicated reactions and specific stereochemistry that would otherwise be very difficult or even impossible to achieve by synthetic means.

As a result, we sought to demonstrate the total synthesis of FR-900098P in vivo using an E. coli host. Previous work in the Zhao lab has already demonstrated synthesis of FR-900098 in E. coli [21], which represents the first heterologous synthesis of a phosphonate in this host. In this approach, the biosynthetic genes frbA through frbH were cloned into three compatible plasmids with three different antibiotic resistance markers, and each gene was placed under the control of an IPTG-inducible T7 promoter. Following transformation into strain BL21(DE3), production of FR-900098 up to 6 mg/L was observed in selective LB media after 40 hours of growth at 30 °C. In the synthesis of FR-900098P, we first attempted to identify the product from the FR-900098 producing strain described above, given that E. coli can naturally produce propionyl-CoA and
that the wild type FrbF enzyme can accept propionyl-CoA as a substrate. However, no FR-900098P was detected, likely indicating that intracellular propionyl-CoA concentrations are too low to generate detectable amounts of the product. To overcome this problem, we investigated a mutasynthetic approach by which sodium propionate was fed to the culture at the time of induction. Once taken up by the cell, propionate can be converted to propionyl-CoA via the propionyl-CoA synthetase gene prpE, which can then be utilized by FrbF. As shown in Figure 3.5a, this approach proved to be successful, and FR-900098P was detected in the cell culture broth by MS/MS. FR-900098, however, was still the dominant phosphonate product, and FR-900098P represented only ~5 – 8 % of the total phosphonate yield.

In an attempt to further increase FR-900098P production in vivo using the native FR-900098 biosynthetic machinery, a metabolically engineered E. coli host was investigated. In general, to increase flux through a secondary metabolite pathway, a driving force needs to be created. One method by which to accomplish this goal is to engineer the host’s metabolism to over-produce a key substrate for a rate-limiting step. In the case of FR-900098P, it is likely that the FrbF-catalyzed acylation is slow, given the relatively low activity observed with the non-native substrate. To overcome this, then, propionyl-CoA concentrations in the cell can be increased via two modifications. First, over-expression of prpE can increase the rate at which propionyl-CoA is synthesized. Second, deletion of the prpRBCD operon prevents propionyl-CoA from being consumed in primary metabolism via conversion to succinate and pyruvate. Both of these modifications have been reported in the literature and successfully utilized in the heterologous production of secondary metabolites [23]. As a result, we obtained this engineered strain, called BAP1, from the Blaine Pfeifer lab at Tufts University. To engineer the BAP1 strain for FR-
900098P production, it was simultaneously transformed with the three compatible plasmids containing the FR-900098 pathway via electroporation. Following selection on LB media with three antibiotics, a colony was isolated containing all three plasmids. Diagnostic PCR amplification of a pathway gene from each plasmid confirmed the presence of all three in the strain. Production of FR-900098P in this host was analyzed under identical growth conditions as the BL21 host. Relative to cell culture density, the FR-900098P titer increased 3.5-fold in the BAP1 strain as compared to the BL21 strain (Figure 3.5b). However, it was observed that the BAP1 strain did not reach as high a density as the BL21 strain, resulting in an overall increase in endpoint FR-900098P concentration of only ~50%. This is likely due to the increased metabolic burden placed on the BAP1 strain by over-expressing prpE under a T7 promoter while simultaneously limiting its ability to utilize the exogenous propionate for central metabolism. As observed in the BL21 strain, the BAP1 strain also produced FR-900098 as its primary phosphonate product. However, the relative amount of FR-900098P was significantly greater in BAP1, comprising ~22 – 25% of the total phosphonate yield.

3.2.2.3 PfDxr inhibition studies

Having demonstrated synthesis of FR-900098P both in vitro and in vivo, it was our goal to determine whether or not this compound is actually a promising antimalarial drug candidate. Thus, we carried out inhibition studies with the purified PfDxr target enzyme. For comparison, inhibition constants for both fosmidomycin and FR-900098 were first determined (Table 3.2). To calculate these values, rate measurements were made for multiple substrate concentrations at multiple concentrations of each inhibitor. Double-reciprocal plots for each inhibitor showed intersection at the vertical axis, indicating competitive inhibition made manifest by an increase in
$K_m$ only. From the measured $K_m$ values, the inhibition constants for fosmidomycin and FR-900098 were determined to be 7.5 nM and 3.7 nM, respectively. These values correlate very well with the in vivo IC$_{50}$ values measured against both the HB3 strain (350 nM for fosmidomycin, 170 nM for FR-900098) and A2 strain (370 nM for fosmidomycin, 170 nM for FR-900098) of *P. falciparum* [17]. Both values are significantly lower than the 38 nM inhibition constant reported for EcDxr, and are more comparable to the values reported for *Synechocystis* sp. PCC6803 (4 nM for fosmidomycin, 0.9 nM for FR-900098) [6, 9]. After we had carried out these measurements, Behrendt and co-workers published very similar PfDxr inhibition constants of 8.4 nM for fosmidomycin and 2.6 nM for FR-900098, which further justify the validity of our inhibition assay and the results derived therefrom [24].

Preparation of pure FR-900098P for inhibition studies was performed from the in vitro reaction mixtures, as the in vivo culture broth contains numerous contaminating compounds that are difficult to separate completely from the desired product. The amount of product generated was quantified by weighing the lyophilized product, and was further checked via measurement of the LC-MS extracted ion peak area and by comparison to the concentration of the CMP-conjugated intermediate prior to treatment with the nucleotide hydrolase FrbI, which could be quantified via measurement of the UV absorbance peak area at 254 nm. Kinetic measurements at multiple inhibitor concentration yielded a sub-nanomolar inhibition constant of 0.92 nM, indicating a significantly more potent inhibitor than fosmidomycin and FR-900098. This result correlates very well with the observation of a hydrophobic binding pocket in the PfDxr crystal structures.
3.2.3 Directed evolution of FrbF

3.2.3.1 Screening method development

As described above and in the previous chapter, the N-acetyltransferase FrbF in the FR-900098 biosynthetic pathway can accept, to a limited extent, alternate acyl-CoA substrates to synthesize alternate FR-900098 derivatives. Nevertheless, activity is significantly reduced when alternate substrates are used, limiting the ability to synthesize novel derivatives both in vitro, where yield is poor, and in vivo, where competition with acetyl-CoA leads to a predominance of FR-900098 in the total phosphonate product pool. As a result, it would be desirable to identify mutants of FrbF that demonstrate increased activity toward non-native acyl-CoA substrates.

To achieve this goal, we developed a screening method to detect FrbF activity in cell lysates in a high-throughput manner. This method relies on 5,5'-dithiobis-(2-nitrobenzoic acid), known as DTNB or Ellman’s reagent, to detect the free coenzyme A product of the acyltransfer reaction [25]. Following a FrbF catalytic cycle, coenzyme A is generated, which possesses a free sulfhydryl terminus. DTNB can participate in a rapid, non-enzymatic disulfide exchange with this product, liberating 2-nitro-5-thiobenzoic acid (TNB). TNB has a large extinction coefficient at a wavelength of 412 nm, and as a result its formation can readily be measured. The overall flowchart of the FrbF screening procedure is outlined in Figure 3.6.

Initially, a short, kinetic assay was envisioned, with TNB formation measured continuously following addition of DTNB and the substrates to the assay plate. However, numerous efforts to optimize assay conditions, including substrate concentrations, DTNB concentration, and assay time, repeatedly yielded inconsistent results with poor signal-to-noise ratios. This is likely due to
side reactions between DTNB and other thiol-containing compounds present in cell lysate. Partial purification of FrbF via in-well capture on Ni-NTA resin was attempted to lower background, but yielded similarly inconsistent results. As a result, the kinetic assay was abandoned in favor of an endpoint assay. To determine an ideal assay time, reaction mixtures containing the substrates CMP-5′-3APn and acetyl-CoA as well as DTNB were incubated with cell lysates harboring either the empty pET28a vector or a pET28a-FrbF construct containing the wild type enzyme. As shown in Figure 3.7a, the background begins to level out at ~30 min, while the FrbF signal continues to increase until ~120 min. Thus, it was deemed that a two-point assay, with measurements taken at 30 min and 120 min and subtracted, would give the best signal-to-noise ratio. Initial tests of this assay yielded significant variation about the mean (as high as 42 %), but with refinement of technique the variation was reduced significantly to ~14 % of the mean.

To further improve the assay, it would be best to screen with the native hydroxylamine phosphonate substrate, CMP-5′-H3APn, rather than CMP-5′-3APn, the primary amine surrogate. As described in the previous chapter, however, this substrate is oxygen-sensitive, and as a result is impractical to prepare. Thus, we instead evaluated in situ generation of this substrate via addition of FrbG to the assay mixture. Curiously, FrbG activity was significantly reduced in the presence of DTNB, likely due to interactions with sulphydryl groups in the protein itself. As a result, the two-point assay protocol was amended such that the cell lysate was split between two plates, and DTNB was only added immediately prior to measurement. In situ CMP-5′-H3APn generation was found to significantly increase the output signal, such that the total assay time could be shortened to 60 min. In a half-plate assay with the wild type enzyme and either acetyl-
CoA or propionyl-CoA, variation in output signal was found to be very low (~6 % relative to the mean), with a signal-to-noise level of ~4 – 7 depending on the substrate (Figure 3.7b).

3.2.3.2 Library creation and preliminary results

To limit the size of the library of FrbF mutants that would need to be screened, we have chosen a site-directed approach to target only those residues that might be actively involved in blocking the incorporation of a larger acyl-CoA substrate. Having determined the crystal structure of FrbF in complex with acetyl-CoA, we selected for saturation mutagenesis those residues located within 6 Å of the substrate acetyl group, with the exception of Thr-190 and His-193, which were found to be catalytically requisite. The resulting sites targeted for saturation mutagenesis are Phe-80, Glu-187, Ser-188, Asn-189, Glu-231, and Phe-234. Degenerate primers were designed at each of these sites such that a saturation mutagenesis library could be constructed via megaprimer PCR, as described below in the Materials and Methods section. Following library transformation, multiple colonies from the Phe-80, Glu-231, and Phe-234 libraries were chosen and sequenced to confirm mutation only at the desired site. The results of the sequencing indicated a different codon in each construct successfully sequenced with no off-target mutations.

Screening is currently underway at each of the six positions identified above. Propionyl-CoA has been chosen as the first acyl-CoA substrate due to the promising in vitro inhibition results obtained with FR-900098P. At present, the Phe-80 and Phe-234 saturation mutagenesis libraries have been screened; however, only false positives have been obtained (data not shown). Screening with propionyl-CoA will continue at the remaining four positions. If no positive hits
are obtained, a different acyl-CoA substrate can be explored. Alternatively, a broader library creation method can be employed to incorporate mutations at other positions.

### 3.3 Conclusions and Outlook

In this study, we successfully expressed the target enzyme for FR-900098 and its derivatives, PfDxr, in *E. coli* and characterized its kinetic behavior with the sugar substrate DXP in the presence of the co-factor NADPH and the divalent cation Mn$^{2+}$. We also determined inhibition constants for both fosmidomycin and FR-900098, revealing the latter to be twice as potent an inhibitor. Additionally, we structurally characterized the enzyme in complex with either fosmidomycin or FR-900098 in collaboration with Satish Nair’s lab. Analysis of these structures revealed key differences between the Dxr enzymes from *P. falciparum* and *E. coli*, explaining the greater potency of the phosphonate inhibitors toward PfDxr. Further, we discovered a hydrophobic binding pocket at the hydroxamate end of the inhibitor, suggesting that a hydrophobic extension at this position could provide even stronger inhibition. Based on this finding, we synthesized and evaluated the novel *N*-propionyl derivative FR-900098P, which was found to have a significantly lower inhibition constant than the parent compound. Synthesis of this compound was demonstrated both *in vitro* and *in vivo*, and a metabolically engineered *E. coli* host was employed to further increase the product titer. Finally, a colorimetric screening method was developed to identify FrbF mutants with greater activity toward novel acyl-CoA substrates.

Future efforts in this direction will focus both on identifying the most promising antimalarial drug candidates and improving their production in the heterologous host. The screening assay
developed here can be utilized to identify FrbF activity toward any novel acyl-CoA substrate, as it relies only on detection of the free coenzyme A product. As a result, it can be employed to pursue other FR-900098 derivatives beyond FR-900098P. We are currently investigating other such derivatives to identify targets beyond propionyl-CoA for FrbF directed evolution of substrate specificity. The best FrbF mutants will then be integrated into the FR-900098 gene cluster to improve heterologous production. Additionally, further metabolic engineering efforts are actively being pursued in the Zhao lab. Currently, the heterologous FR-900098 pathway in *E. coli* utilizes a strong T7 promoter for every pathway gene, which introduces a heavy metabolic burden on the cell. To alleviate this burden, the expression levels of each gene need to be balanced. Further metabolic engineering is also being investigated to increase overall flux through the FR-900098 synthetic pathway. The results of both of these approaches will be readily applicable to the production of novel FR-900098 derivatives.

### 3.4 Materials and Methods

#### 3.4.1 Materials

The *Escherichia coli* cloning strain DH5α, expression strain BL21(DE3) and the expression vector pET28a were obtained from EMD Biosciences (San Diego, CA). Antibiotics and isopropyl-β-D-1-thiogalactopyranoside (IPTG) were purchased from Gold Biotechnology (St. Louis, MO). All chemicals, including all buffers, the reagents cytidine triphosphate (CTP) and 2-amino-4-phosphonobutyrate (2APn), and all acyl-CoAs were purchased from Sigma-Aldrich (St. Louis, MO) except for the co-factor NADPH, which was purchased from Applichem (Darmstadt, Germany), and 1-deoxy-D-xylulose 5-phosphate (DXP), which was purchased from Echelon Biosciences (Salt Lake City, UT). All PCR reagents and restriction enzymes were
purchased from New England Biolabs (Ipswich, MA), and DNA Miniprep and Gel Purification Kits were purchased from Qiagen (Valencia, CA). All primers were synthesized by Integrated DNA Technologies (Coralville, IA). Talon Cobalt immobilized metal affinity chromatography (IMAC) resin was purchased from Clontech Laboratories (Mountain View, CA). All other materials were purchased from Thermo-Fisher Scientific (Pittsburgh, PA).

3.4.2 Cloning, expression, and purification

The complete dxr gene from *Plasmodium falciparum* was codon-optimized for expression in *E. coli* by DNA 2.0 (Menlo Park, CA). The core enzyme sequence (without nucleotides 1-216, corresponding to the 72 amino acid leader peptide) was PCR amplified with added NdeI and HindIII restriction sites and cloned into the respective sites of vector pET28a, introducing an N-terminal His$_6$ tag to the protein. *E. coli* BL21(DE3) was transformed with the resulting construct by electroporation. Cells were grown in LB media supplemented with kanamycin (50 μg/mL) at 37 °C to an OD$_{600}$ of ~0.8, after which induction was carried out by addition of 0.5 mM IPTG at 25 °C. After 18 hr, the cells were harvested by centrifugation at 7500 rpm for 15 min and resuspended in 20 mM Tris-HCl (pH 7.65), 0.5 M NaCl, and 15 % glycerol supplemented with 1 mg/mL lysozyme. After a freeze-thaw cycle at -80 °C, the cell suspension was sonicated to ensure sufficient lysis. The lysate was clarified by centrifugation, after which the His$_6$-tagged protein was purified by affinity chromatography on TALON Superflow Co$^{2+}$ resin (Clontech, Mountain View, CA) coupled to fast-performance liquid chromatography. The eluted protein was washed three times in 50 mM HEPES (pH 7.25), concentrated, and stored in 15% glycerol at -80°C.
3.4.3 Generation of PfDxr mutants

Mutants of the PfDxr enzyme were generated by the megaprimer PCR method. Briefly, primers were designed at the site of interest containing the desired mutation such that a larger primer (a “megaprimer”) containing the mutation could be generated by PCR, using the regular forward or reverse primer as the second primer (whichever yielded the shortest PCR product). Next, PCR amplification of the full-length mutant gene was carried out using the megaprimer at one end of the gene and the regular primer at the opposite end. Finally, the full length gene containing the mutation was digested with NdeI and HindIII and ligated into the correspondingly digested pET28a vector backbone. The ligation product was then used to transform the *E. coli* cloning strain DH5α. Transformants were picked and grown in 5 mL LB + 50 µg/mL kanamycin. Their plasmids were isolated and sequenced by the UIUC Core Sequencing Facility (Urbana, IL), and Sequencher software (Gene Codes Corporation, Ann Arbor, MI) was used to confirm the presence of only the desired mutation. The correct plasmids were then used to transform the expression strain *E. coli* BL21(DE3), and frozen stocks of the resultant strains were stored in 15% glycerol at -80 °C.

3.4.4 Kinetic assays

All kinetic assays were carried out with N-His$_6$-tagged PfDxr in 100 mM Tris-HCl buffer (pH 7.5). Determination of kinetic parameters was carried out in a total volume of 300 µL with 1 mM MnCl$_2$, 0.3 mM NADPH, and 50 nM enzyme while the DXP concentration was varied from 0 – 250 µM. Oxidation of NADPH was continuously monitored via absorbance at 340 nm in a Nanodrop 2000c spectrophotometer (Thermo Scientific, Wilmington, DE) at 37 °C for a maximum of 5 min. All assays were performed in triplicate. Michaelis-Menten kinetic
parameters were fit to the data using OriginPro 8 software (OriginLab Corporation, Northampton, MA). Due to the instability of the PfDxr enzyme, a fresh aliquot was used for each independent assay.

3.4.5 Inhibition assays

All inhibition assays were carried out with N-His$_6$-tagged PfDxr in 100 mM Tris-HCl buffer (pH 7.5). To determine inhibition constants for fosmidomycin, FR-900098, and FR-900098P, kinetic assays were performed at multiple concentrations of each inhibitor ranging from 0 – 10 nM. At each inhibitor concentration, $K_m$ values were determined, with reactions carried out at 37 °C in a total volume of 300 µL with 1 mM MnCl$_2$, 0.3 mM NADPH, and 50 nM enzyme while the DXP concentration was varied from 0 – 250 µM. From the measured $K_m$ values, $K_i$ values were determined using the equation for competitive inhibition. This procedure was repeated in triplicate for each inhibitor.

3.4.6 *In vitro* FR-900098P synthesis

Synthesis of FR-900098P *in vitro* was carried out with the purified N-His$_6$-tagged FrbF, FrbG, FrbH, and FrbI enzymes and commercially available substrates. First, cytidine monophosphate-5’-3-aminopropylphosphonate (CMP-5’-3APn) was synthesized in 50 mM HEPES buffer, pH 7.25, containing 1.5 mM cytidine triphosphate (CTP), 3 mM 2-amino-4-phosphonobutyrate (2APn), 10 mM MgCl$_2$, and ~10 µM FrbH with incubation at 30 °C for 4 hours, after which nearly complete conversion was observed. Following removal of FrbH with a 10 kDa cutoff filter, CMP-5’-FR-900098P was synthesized by addition of 1.5 mM NADPH, 1.5 mM propionyl-CoA, ~10 µM FrbF, and ~15 µM FrbG with incubation at 30 °C for 4 hours. At this point, side
products such as CMP-5’-2-amino-4-phosphonobutyrate (CMP-5’-2APn), CMP-5’-N-propionyl-3-aminopropylphosphonate (CMP-5’-P3APn), CMP, CDP, and NADP+ were removed via fractionation on an Agilent 1100 Series HPLC using an Alltech Prevail C18 column with isocratic 15 mM ammonium formate elution. The resulting CMP-5’-FR-900098P was lyophilized and resuspended in double-distilled H2O. Finally, CMP cleavage was carried out in 10 mM Tris-HCl buffer, pH 7.5, by addition of ~10 µM FrbI with incubation at 30 °C for 4 hours. The final product titer was estimated from the concentration of CMP-5’-FR-900098P (measured by UV absorbance at 254 nm) and confirmed by dry weight of the final lyophilized product. Correct identity and purity of the final product was determined using an Agilent XCT ion-trap MSD mass spectrometer, where the anticipated m/z = 212+ parent ion and m/z = 138+ daughter ion were observed.

3.4.7 In vivo FR-900098P synthesis

Synthesis of FR-900098P in vivo was carried out in two different E. coli expression strains. The first was a BL21(DE3) strain transformed with three compatible plasmids containing the entire FR-900098 biosynthetic pathway except for the nucleotide hydrolase FrbI. This strain was previously prepared and described by Johannes and DeSieno, et al. [21]. The second FR-900098P expression strain was prepared from the metabolically engineered BAP1 strain provided by Dr. Blaine Pfeifer, which has been genetically modified to block conversion of propionyl-CoA to succinate and pyruvate for central metabolism [23]. This strain was transformed with the same three-plasmid FR-900098 production system. Successful transformation was confirmed by diagnostic PCR amplification of a pathway gene from each plasmid. Both strains were cultured in 50 mL LB medium supplemented with kanamycin (50
µg/mL), ampicillin (100 µg/mL), and chloramphenicol (25 µg/mL) at 37 °C to an OD$_{600}$ of ~0.8.

At this point, expression of the FR-900098 pathway enzymes was induced at 30 °C by addition of 0.5 mM IPTG, as well as 20 mM sodium propionate to increase propionyl-CoA synthesis. Samples of culture supernatant were collected at multiple time points up to 94 hr and clarified via centrifugation at 13,200 rpm for 10 min. Analysis of clarified samples was carried out on an Agilent XCT ion-trap MSD mass spectrometer with a 100 x 4.6 mm Synergi 4µ Fusion-RP 80A column (Phenomenex, Torrance, CA) using 0.1 % formic acid in H$_2$O (mobile phase A) and acetonitrile (mobile phase B) with the following elution program: 0 % B to 40 % B over 10 min; 40 % B to 100 % B over 3 min; and 100 % B for an additional 5 min. Products were identified in positive mode by their expected m/z values (198 for FR-900098, 212 for FR-900098P) with the corresponding m/z = 138 daughter ion.

3.4.8 FrbF library creation

The FrbF crystal structure was visualized using Molecular Operating Environment software (Chemical Computing Group, Montreal, Canada). Residues within 6 Å of the acetyl-CoA acyl group were selected as saturation mutagenesis candidates, with the exception of those identified as catalytically requisite. Saturation mutagenesis was carried out using primers containing a randomized codon at the site of interest, and the mutant FrbF constructs were generated using the megaprimmer PCR method described above. Libraries of FrbF mutants were digested with the restriction enzymes NdeI and HindIII and ligated into the respective sites of the correspondingly digested pET28a vector. Following transformation of BL21(DE3) and selection on LB plates containing kanamycin (50 µg/mL), plasmids were isolated from multiple colonies in each library and sequenced to confirm diversity at the targeted amino acid position.
3.4.9 FrbF screening method

Screening of FrbF mutant libraries was performed in a high-throughput microtiter plate assay. Using sterile toothpicks, individual colonies were picked from library transformation plates to the wells of a 96-well plate containing 100 µL of LB medium supplemented with kanamycin (50 µg/mL) per well. Following overnight growth to saturation at 37 °C, 5 µL from each well was used to inoculated 100 µL of fresh LB + kanamycin in a separate plate. This plate was then incubated at 37 °C until an average OD$_{600}$ of ~1.2 was reached, at which point 100 µL of LB + kanamycin supplemented with 0.6 mM IPTG was added to each well, yielding a final concentration of 0.3 mM IPTG. Induced plates were incubated at 30 °C overnight to allow FrbF overexpression. Afterward, the cells were pelleted by centrifugation at 4000 rpm for 10 min and resuspended in 100 µL of 50 mM HEPES buffer (pH 7.25) supplemented with 1 mg/mL lysozyme. Following incubation at room temperature for 30 min, the lysis plate was subjected to a freeze-thaw cycle at -80 °C. Immediately before the assay was performed, the plate was thawed and the cell lysate clarified via centrifugation at 4000 rpm for 10 min at 4 °C. To perform the assay, 30 µL of lysate from each well was added to two separate 96-well plates. To the first plate was added 50 µL of 10 mM DTNB and 70 µL of 50 mM HEPES buffer (pH 7.25), and the absorbance at 412 nm was measured immediately. To the second plate was added 70 µL of the assay mixture, which contained (final concentrations): 400 µM CMP-5’-3APn, 400 µM NADPH, 300 µM of the alternate CoA substrate, and 4.8 µM FrbG in 50 mM HEPES buffer. The second plate was incubated at 30 °C for 60 min, after which 50 µL of 10 mM DTNB was added to each well and the absorbance at 412 nm was measured. The difference between the two measured absorbance values was taken as an estimate of the total amount of product generated.
3.5 References


### 3.6 Tables

**Table 3.1:** Inhibition values reported with purified Dxr enzymes (all units are nM)

<table>
<thead>
<tr>
<th>Species</th>
<th>fosmidomycin</th>
<th>FR-900098</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\text{IC}_{50}$</td>
<td>$K_i$</td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td>50</td>
<td>38</td>
</tr>
<tr>
<td><em>P. falciparum</em></td>
<td>32</td>
<td>--</td>
</tr>
<tr>
<td><em>P. aeruginosa</em></td>
<td>150</td>
<td>--</td>
</tr>
<tr>
<td><em>Synechocystis</em> sp. PCC6803</td>
<td>--</td>
<td>57</td>
</tr>
<tr>
<td><em>Z. mobilis</em></td>
<td>--</td>
<td>600</td>
</tr>
</tbody>
</table>
Table 3.2: Kinetic parameters measured for PfDxr (mean ± standard deviation)

<table>
<thead>
<tr>
<th>$K_{m}$, DXP</th>
<th>$k_{cat}$</th>
<th>$K_{i}$ fosmidomycin</th>
<th>$K_{i}$ FR-900098</th>
<th>$K_{i}$ FR-900098P</th>
</tr>
</thead>
<tbody>
<tr>
<td>52 ± 13 μM</td>
<td>120 ± 30 min$^{-1}$</td>
<td>7.5 ± 2.0 nM</td>
<td>3.7 ± 0.9 nM</td>
<td>0.92 ± 0.19 nM</td>
</tr>
</tbody>
</table>
3.7 Figures

Figure 3.1: Synthesis of isoprenoid precursors via (a) the mevalonate pathway and (b) the non-mevalonate pathway.
Figure 3.2: (a) Purity of the recombinant PfDxr enzyme was confirmed via SDS-PAGE (expected MW = 47 kDa), while (b) activity was confirmed by tracking NADPH oxidation in the presence or absence of fosmidomycin.
Figure 3.3: Co-crystal structure of PfDxr with (a) fosmidomycin and (b) FR-900098 reveals the hydrophobic binding pocket. (c) Comparison of PfDxr (yellow) with EcDxr (blue) reveals more favorable interactions with the inhibitor in PfDxr. Images generated by Satish Nair.
Figure 3.4: Enzymatic synthesis of FR-900098P. (a) Co-incubation of FrbG and FrbF following the FrbH reaction yields a mix of UV active compounds, including: a, CDP; b, CMP; c, CMP-5’-NO3APn; d, CMP-5’-3APn; e, CMP-5’-P3APn; f, NADP⁺; and g, CMP-5’-FR-900098P. (b) MS trace showing synthesized FR-900098P (m/z = 212⁺) with the expected m/z = 138⁺ daughter ion.
Figure 3.5: In vivo biosynthesis of FR-900098P. (a) MS trace showing phosphonates produced (FR-900098 = 198\(^+\), FR-900098P = 212\(^+\)) in the presence or absence of exogenous sodium propionate. (b) Comparison of relative FR-900098P concentrations produced in BL21 and BAP1.
Figure 3.6: Flowchart of screening method for detection of FrbF activity toward novel substrates. Briefly, a saturation mutagenesis library of FrbF constructs is generated by megaprimer PCR and used to transform *E. coli* BL21(DE3). Individual colonies are grown, induced, and lysed in a microtiter plate. Following addition of lysate to the FrbF assay mixture, TNB formation is monitored at $\lambda = 412$ nm to detect FrbF activity.
**Figure 3.7**: Screening method development. (a) The time course of a lysate assay shows activity of FrbF over a significant time period. (b) Screening of several wells expressing wild type FrbF with either acetyl- or propionyl-CoA shows low variation about the mean and significant signal above background.