Interaction of α-thymidine Inhibitors and Natural Substrates with Thymidylate Kinase from *Plasmodium falciparum*

Mengshen Chen

2019

Carnegie Mellon University

Interaction of α-thymidine Inhibitors and Natural Substrates with Thymidylate Kinase from *Plasmodium falciparum*

by

Mengshen Chen

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

at

Carnegie Mellon University

Department of Biological Sciences

Pittsburgh, Pennsylvania

Advised by Dr. Danith Ly and Dr. Gordon S. Rule

September 20, 2019

I’d like to dedicate this thesis to my loving wife…

Acknowledgement

I’d like to express my sincere gratitude towards my advisers, Dr. Danith Ly and Dr. Gordon S. Rule, along with my research committee members, Dr. Marcel P. Bruchez and Dr. David D. Hackney for their continuous support, guidance and patience. This dissertation would not be in shape and I would never have been able to be at where I am now without their dedication and contribution.

I’d like to thank my wife and my parents for their contributions during my Ph. D. career. My wife, Wenjing Huang, made a tough decision to transfer to Pittsburgh when I decided to continue my education at CMU, and I can’t thank her enough for her company and partnership. My parents have always been supportive to my decisions, and I really appreciate their understanding and encouragement.

I was fortunate to have the opportunity to work between two labs during my Ph. D. career, where I met amazing people who would patiently help and train me as I grew as a researcher and a Ph. D. candidate. Dr. Shivaji Ashokrao Thadke and Dr. Wei-Che Hsieh were always there when I ran into organic chemistry problems; and I could always rely on Dr. Kaustubh Sinha and Ian Fucci when I needed help in biochemistry related subjects.

My cohorts are a group of enthusiastic, energizing and inspiring people. I learned a great deal from them and had a lot of fun. I am lucky to meet people like Rolando Cuevas, with whom I shared the same room during my interview at CMU and we became friends ever since; and Ian Fucci, who is the best co-worker one could possibly hope for.

I thank everyone whom I worked with during my years as a Ph. D. student and I thank the Bruchez lab, Roberto R. Gil and Virgil Simlaceanu for their help on equipment.

My experience as a Ph. D. candidate would not be complete without any of these people, and I feel extremely fortunate to have the privilege to be among them during this journey.

Abstract

Malaria is one of the most devastating parasite-borne infectious diseases and *Plasmodium falciparum* is associated with the deadliest form of malaria. Due to the emerging drug resistance against current treatments in these parasites, there is continuous need in developing novel anti-malarial drugs.

Thymidylate kinase from *Plasmodium falciparum* (*Pf*TMK) is an attractive druggable target for the development of anti-malarial agents due to its vital role in the *Plasmodium* life cycle and potential in drug specificity. Several groups have developed and reported *Pf*TMK inhibitors, however, none have entered clinical trials. The field has been limited by the lack of structure-activity relationship (SAR) studies to aid further structural modifications of existing inhibitors; furthermore, there has not been a study that comprehensively characterize *Pf*TMK-substrate interaction in order to elaborate on the unique substrate specificity of *Pf*TMK of being capable of phosphorylating both dTMP and dGMP.

In our study, we aim to address these concerns in the current field. We designed and synthesized small molecules derived from a reported *Pf*TMK inhibitor, and thoroughly characterized the synthesized compounds in terms of enzyme inhibition and binding property; we also probed for binding site of the synthesized derivatives to provide an

understanding on the SAR. We acquired methyl side-chain (ILV) assignments for *Pf*TMK. The *Pf*TMK-substrate interaction was comprehensively characterized in terms of steady-state kinetics, binding property, thermodynamics and dynamics. These studies provide significant insights on the unique substrate specificity of *Pf*TMK; and more importantly, revealed a unique half-site binding mechanism with dGMP, in which we showed that the binding of the first dGMP molecule results in intermonomer communication and possible conformation change that prohibits the binding of second dGMP in the other subunit of the *Pf*TMK dimer. These observations have significant implication on the future *Pf*TMK inhibitor design; guanosine-based small molecules could be more optimal than their thymidine-based counterparts in terms of *Pf*TMK inhibition and binding specificity.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Malaria is one of the most devastating parasite-borne infectious diseases with over 200 million infections and over 600,000 deaths each year. The majority of malaria patients are children in Africa under age of 5.[1] *Plasmodium falciparum*, the deadliest form of malaria parasite, is associated with, and responsible for the mortality and morbidity of malaria infection. Current treatments against *P. falciparum* malaria include Artemisinin Combination Therapies (ACTs). Although ACTs work quite well against malaria, there is serious concern that widespread resistance to ACTs is emerging in these parasites, and this is due to counterfeit or substandard treatments, unregulated administration, and the lack of complementary combination treatment (such as lumefantrine).[2] Consequently, there is continuous need in developing novel anti-malarial drugs.

Thymidylate kinase (TMK) is one of the key enzymes in all organisms; it is the only nucleotide kinase that facilitates the phosphorylation of deoxythymidine 5'-

monophosphate (dTMP) to deoxythymidine 5'-diphosphate (dTDP); the product is the precursor to deoxythymidine 5'-triphosphate (dTTP), which is an essential building block in the process of DNA replication and repair (Figure 1.1). Consequently, TMKs have been extensively pursued in the pharmaceutical industry as drug target in many organisms, such as human[3,4,5], and pathogens such as Gram-positive bacteria[6,7], *P. falciparum*[8,9,10,11], *Mycobacterium tuberculosis*[12,13], and *Pseudomonas aerugonosa*[14]. The development of malaria demands large amount of dTTP for DNA replication, this makes the TMK from *Plasmodium falciparum* (*Pf*TMK) a suitable drug target against malaria.

Several *Pf*TMK inhibitors have been developed up to date. However, none have entered clinical trials. Most developed inhibitors are thymidine-based competitive inhibitors; and very limited effort has been dedicated to developing guanosine-based competitive inhibitors or allosteric inhibitors. The relatively slow progress in the field could be the result of lack of systematic structure-activity relationship (SAR) studies and incomprehensive studies on the protein biochemistry, thermodynamics, and dynamics of *Pf*TMK.

In Chapter 1 I summarize the pathophysiology of malaria and current strategies for the treatment of malaria. In Chapter 2, I discuss the assignment of methyl-side chain residues of *Pf*TMK using nuclear magnetic resonance (NMR) spectrometry. In Chapter 3 I focused on improving one of the existing *Pf*TMK inhibitor in attempt to enhance its metabolic stability and efficacy, and to fully characterize the inhibition, binding properties, and binding site of the synthesized derivative. In Chapter 4 I investigated the effects of different substrates (dTMP and dGMP) on enzymatic activity, binding, and

dynamics of the protein-substrate complex. Overall, I provided a comprehensive understanding of *Pf*TMK-substrate interaction, in which the enzyme kinetics, binding properties, thermodynamic parameters and dynamic characters were studied to elaborate on the unique substrate specificity of *Pf*TMK, and to examine the SAR in hope to shed some light on future inhibitor design, as discussed in Chapter 5.

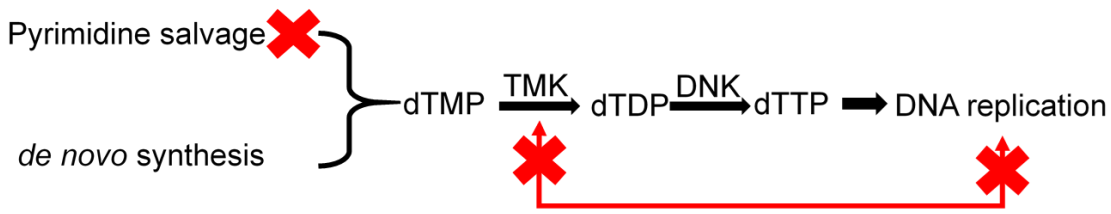


Figure 1.1. dTTP cycle of *Plasmodium falciparum*. Biosynthesis of dTTP involves pyrimidine salvage or de novo synthesis of dTMP, dTMP is then phosphorylated by thymidylate kinase to produce dTDP, and is subsequently phosphorylated by nucleoside diphosphate kinase to form dTTP, which is an essential building block for DNA replication and repair. *P. falciparum* is incapable of pyrimidine salvage, and inhibition of thymidylate kinase will lead to termination of DNA replication, which is lethal to the organism.

1.1 Malaria and Current Treatment

1.1.1 Malaria Pathophysiology

The spread of malaria is a cycle that involves humans and female *Anopheles* mosquitoes. In which the malaria parasites grow and multiply in the liver cells of human body and subsequently are transferred into red blood cells and destroy them. The burst of red blood cells releases merozoites, which can be harvested by female *Anopheles* mosquitoes during blood feeding to spread the disease in other hosts. After mating in the gut of mosquitoes and forming sporozoite, the sporozoite will then be transferred to another human during blood feeding, thus beginning another cycle and wide-spread of this infectious disease[15].

After the sporozoites are introduced into the human body by malaria-infected mosquitoes, malaria development enters exoerythrocytic stage, the sporozoites first infect the liver cells and develop into schizonts, rupturing liver cells and release merozoites. This marks the beginning of erythrocytic stage and invasion of red blood cells. The parasites will multiply sexually and asexually, mature into schizonts, and eventually lead to bursting of red blood cells and release of matured malaria-parasite, which will subsequently invade other healthy red blood cells[15](Figure 1.2).

Symptom of malaria includes fever, chill, sweats, headaches, nausea (uncomplicated malaria), and in severe cases, anemia, hemoglobinuria, acute respiratory distress syndrome, acute kidney injury, seizures, comas, and even death (severe malaria). There are over 200 million reported case of malaria infections each year, with over 0.6 million deaths.

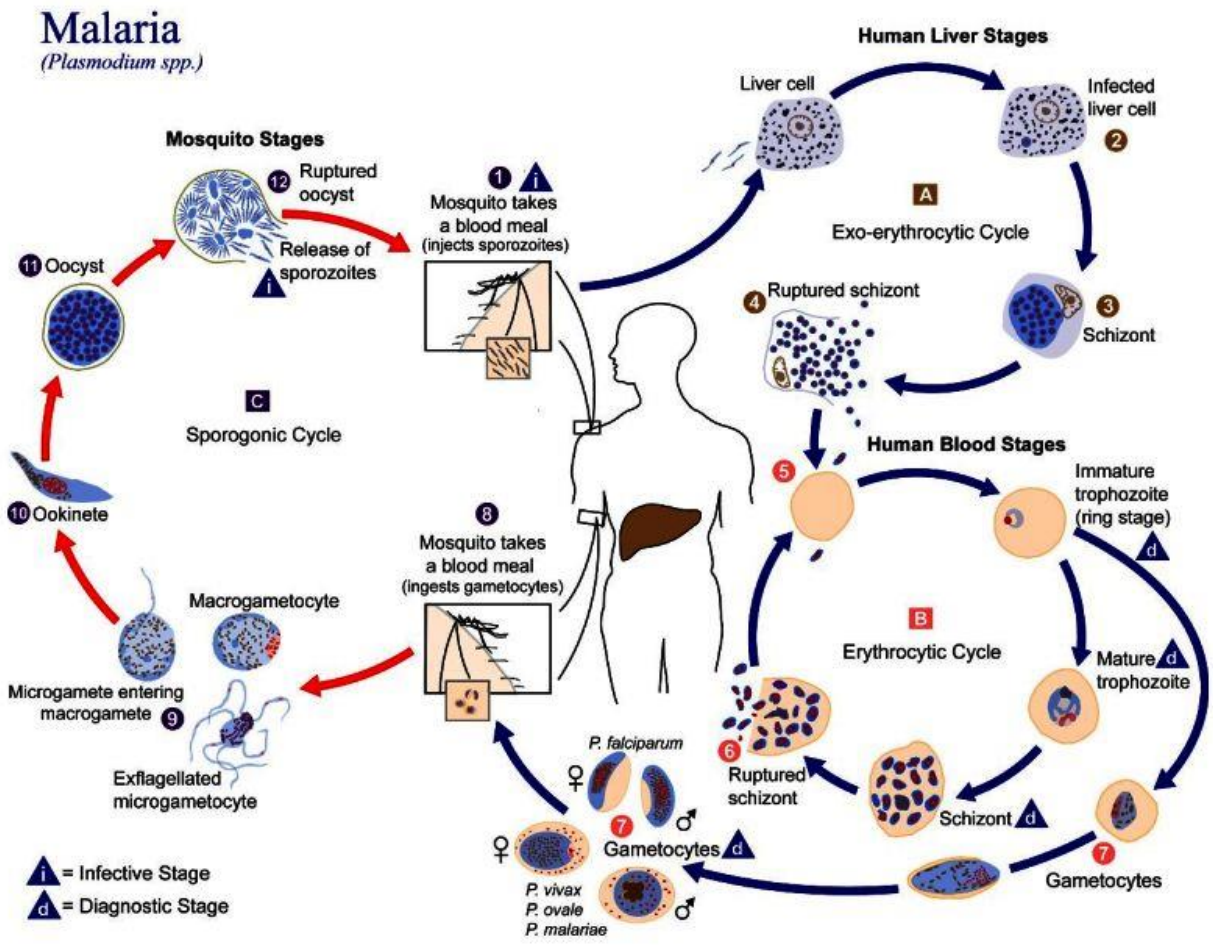


Figure 1.2. Malaria life cycle. Image obtained from: Beck T. Will malaria soon be a thing of the past?the potential of recombinant protein vaccines to control one of the world's most deadly diseases. 2006 Jun 26. In: Dean L, McEntyre J, editors. Coffee Break: Tutorials for NCBI Tools [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 1999-. Figure 1, [The malaria parasite life cycle...]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK5951/figure/malaria_LifeCycle/

1.1.2 Current Malaria Treatment and Mechanism

The current go-to treatment for malaria is Artemisinin-combine therapies (ACTs). Artemisinin is a chemical isolated from the plant *Artemisia annua*, that was first discovered by Youyou Tu in 1972. Artemisinin is a sesquiterpene lactone that contains a unique peroxide bridge, which plays a critical role in the mechanism of this compound[16] (Figure 1.3).

The exact mechanism of action of artemisinins is still unknown[17]. Artemisinin is a pre-drug that does not directly attack malarial parasites or cells. It is activated by the cleavage of the endoperoxide ring[18]. During the malaria cycle, infected red blood cells rupture, and expose the heme in hemoglobin which contains Fe^{2+} oxide. The endoperoxide ring is then cleaved by the Fe^{2+} oxide, resulting in production of free radicals that bind to and eliminate malaria-parasites[19,20] (Figure 1.3).

Although ACTs work quite well against malaria, drug resistance is slowly emerging among malarial parasites. The resistance has arisen because of counterfeit or substandard treatments, unregulated administration of drugs, and the lack of complementary combination treatment, such as lumefantrine[2], which is also an antimalarial agent that targets the erythrocytic stage that improves efficacy of artemisinin. Because of the emerging drug resistance, there is a continuing need to develop additional antimalarial compounds.

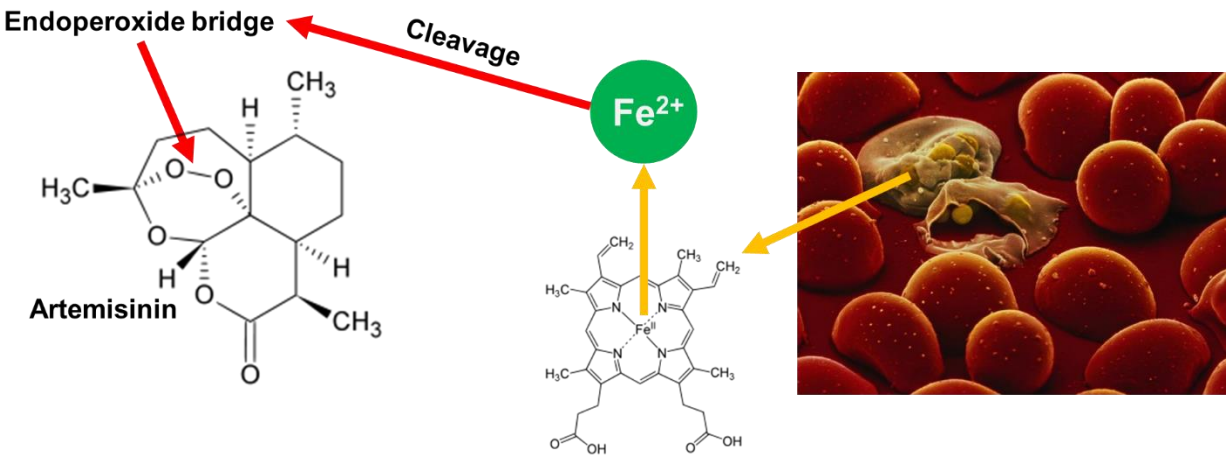


Figure 1.3. Artemisinin mechanism of action. Malaria causes burst of red blood cells and exposure of heme in hemoglobin which contains Fe^{2+} oxide; Fe^{2+} oxide leads to cleavage of the endoperoxide bridge in artemisinin and production of free radicals that target and kill parasite.

1.2 Thymidylate Kinase from *Plasmodium falciparum* (*Pf*TMK)

1.2.1 Structure and Catalytic Mechanism

X-ray crystal structures of TMKs have been studied in many organisms, including human[21], *Saccharomyces cerevisiae*[22], *Mycobacterium tuberculosis*[23], *Vaccinia virus*[24], *Plasmodium falciparum*[25], and *Candida albicans*[26]. The crystal structure studies revealed that all TMKs are homodimers, even in the absence of bound ligand (see Figure 1.4C). Each subunit consists of five stranded parallel β-sheets and 7−11 α-helices (Figure 1.4A, 1.4B). Two conserved motifs are present in the nucleotide binding domain: a P-loop, consisting of a few conserved sequences that are responsible for substrate recognition and catalysis, and a LID domain which partially covers bound ATP, and plays a role in catalysis. TMKs are classified into type I and type II based on the residue composition of P-loop: type I TMKs have additional basic residue in the P-loop, such as human, yeast, and Plasmodium TMKs, and type II TMKs lack this basic residue[25].

The general enzyme mechanism of TMKs is similar to that of adenylate kinases; after the substrates are bound to the protein, the phosphate transfer undergoes through an

associated mechanism, during which the migrating phosphate is bound to both substrates in the transition state. Magnesium ions and arginine residues around the active site stabilizes the transition state. Taking type I TMKs as an example, during the transition state, movements of LID and P-loop occur; an arginine residue in the P-loop interacts with magnesium ion and an aspartic acid residue in the LID domain, while the lysine residue interacts with the backbone of the P-loop, leading to a partially closed state of the protein.

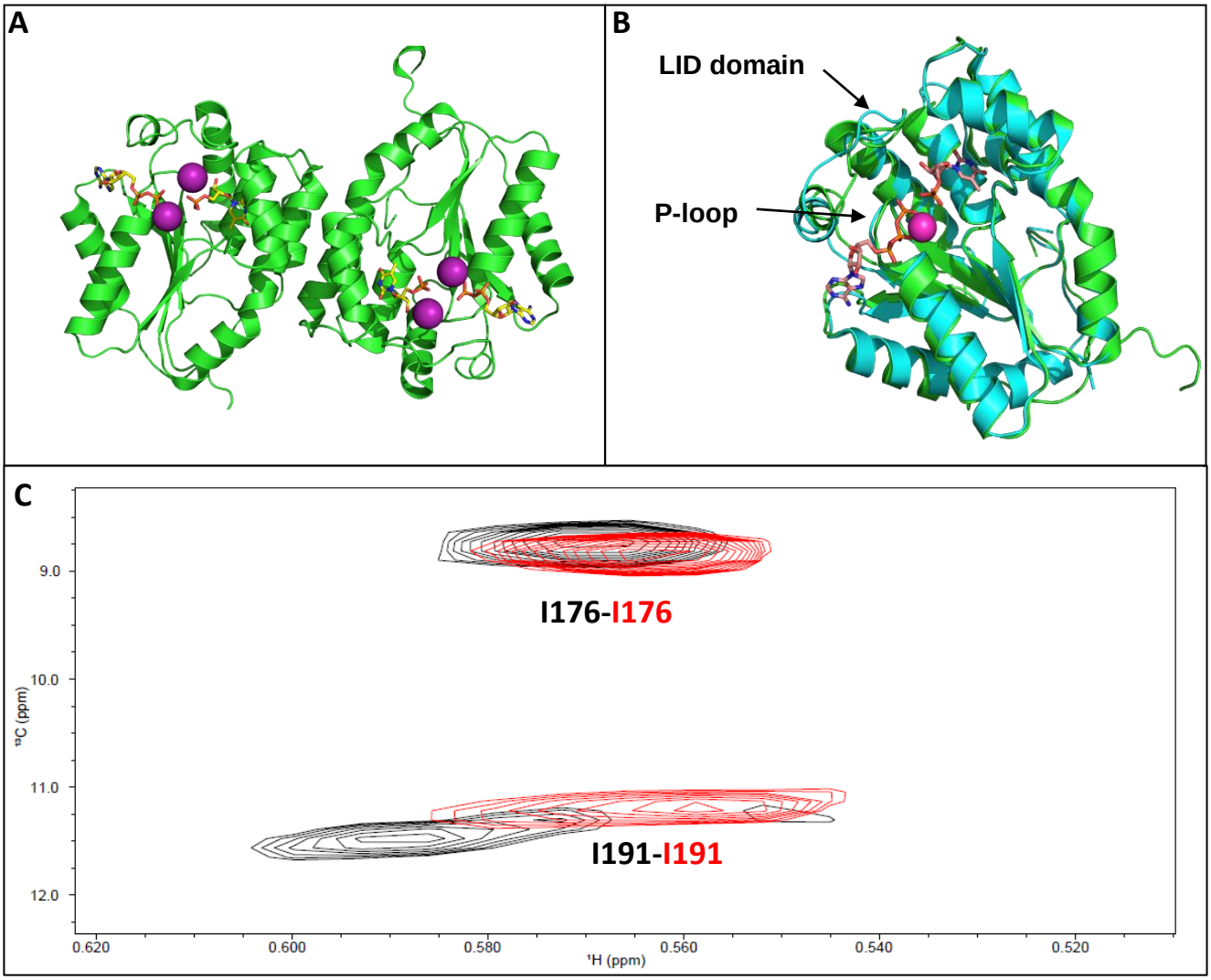


Figure 1.4. X-ray crystal structure of *Pf*TMK. Panel A shows the *Pf*TMK dimer complexed with ADP, dTMP, and Na+ (Protein Data Bank (pdb) entry 2WWF). Panel B shows an overlay of huTMK (green) and *Pf*TMK (cyan). X-ray crystal structure of huTMK (pdb entry 1NN0) is superimposable on that of *Pf*TMK (pdb entry 2WWF), except for the LID domain. Panel C shows the methyl resonance of I176 and I191 under no dTMP (black) and under saturating concentration of dTMP (red); this illustrates that *Pf*TMK exists as homodimer with or without substrate bound, the peak intensity and linewidth are consistent between the two conditions, the only difference is in the chemical shift. A monomer would be expected to show a significantly narrower linewidth than the dimer.

1.2.2 Unique Substrate Specificity and Potential in Drug Development

*Pf*TMK is one of the unique examples of TMKs that can also utilize dGMP as a substrate (Figure 1.5), such characteristic is not observed for the human version of the enzyme. However, it is not clear how dGMP is capable of binding to *Pf*TMK, and what residues are directly responsible for granting *Pf*TMK the capability of using dGMP as substrate.

The superposition of the enzyme−substrate complex for human TMK (huTMK) and *Pf*TMK suggests that the human and Plasmodium enzymes are highly similar in topology (Figure 1.4B). However, sequence alignment of the two proteins only exhibits a sequence similarity of 41% (Figure 1.6), it is worthwhile mentioning is that the LID domains are poorly conserved and cannot be superimposed.

In summary, *Pf*TMK has unique substrate specificity of being able to phosphorylate dGMP, and huTMK and *Pf*TMK are quite different in terms of amino acid sequences. These two features highly suggest that *Pf*TMK would be an appropriate druggable target to design specific inhibitors as effective antimalarial drugs.

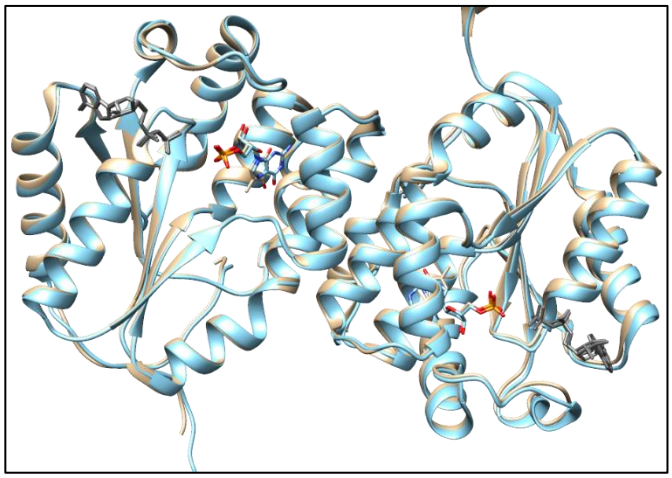


Figure 1.5 . Overlay of X-ray crystal structure of dTMP-*Pf*TMK (khaki) and dGMP-*Pf*TMK (cyan) complexes. Overlay of the crystal structure of dTMP-PfTMK (Protein Data Bank (pdb) entry 2WWF) and dGMP-PfTMK (pdb entry 2WWG) complexes.

```
PfTMK   6    KKGKFIVFEGLDRSGKSTQSKLLVEYLKNNNVEVKHLYFPNRETGIGQIISKYLKMENSM  65
             ++G  IV EG+DR+GKSTQS+ LVE L       + L FP R T IG+++S YL+ ++ +
huTMK   4    RRGALIVLEGVDRAGKSTQSRKLVEALCAAGHRAELLRFPERSTEIGKLLSSYLQKKSDV  63

PfTMK   66   SNETIHLLFSANRWEHMNEIKSLLLKGIWVVCDRYAYSGVAYSSGALNLNKTWCMNPDQG  125
              + ++HLLFSANRWE +  IK  L +G+ +V DRYA+SGVA++    N +  WC  PD G
huTMK   64   EDHSVHLLFSANRWEQVPLIKEKLSQGVTLVVDRYAFSGVAFTGAKENFSLDWCKQPDVG  123

PfTMK   126  LIKPDVVFYLNVPPNYAQNRSDYGEEIYEKVETQKKIYETYKHFAHEDY--WINIDATRK  183
             L KPD+V +L +    A  R  +G E YE    Q++    +     +    W  +DA++
huTMK   124  LPKPDLVLFLQLQLADAAKRGAFGHERYENGAFQERALRCFHQLMKDTTLNWKMVDASKS  183

PfTMK   184  IEDIHNDI  191
             IE +H DI
huTMK   184  IEAVHEDI  191
```

Figure 1.6. Sequence alignment of *Pf*TMK and huTMK. Sequence alignment showed 41% identity between the two amino acid sequences. The well-conserved P-loop is highlighted in red, and poorly-conserved LID domain is highlighted in cyan. This explains the low root-mean-squared-deviation (RMSD) for superposition of the P-loops and high RMSD for the LID domains.

1.3 Development of *Pf*TMK-inhibitor as Anti-malarial Agents

Several groups have developed *Pf*TMK-inhibitors (Figure 1.7): Kato et al.[9] explored racemic carbocyclic derivatives of thymidine and obtained K_i values for *Pf*TMK that were as low as 20 μM. Noguchi et al.[10] subsequently prepared pure enantiomers of the derivative reported by Kato et al, producing compounds with a slightly higher affinity. Cui et al. reported the development of a large number of inhibitors of *Pf*TMK. These compounds were based on an α-thymidine thiourea inhibitor of TMK from *M. tuberculosis*. They established that α-thymidine inhibitors are more potent than their β-thymidine counterparts and that replacement of the thiourea with urea improved inhibition. Cui et al. reported that N- (5′-deoxy-α-thymidin-5′-yl)-N′-[4-(2-chlorobenzyloxy)-phenyl]urea was the most effective inhibitor of *P. falciparum* growth, with an EC_{50} value of 28 nM and a CC_{50} of 29 μM[11].

A potential limitation of this particular compound is that it is only characterized *in-vivo*, and the ability of this compound to inhibit *Pf*TMK was not characterized; consequently, it's difficult to determine if the parasite growth inhibition is due to *Pf*TMK inhibition, and whether the compound binds to *Pf*TMK and functions as a competitive inhibitor or allosteric inhibitor. This compound is also metabolically unstable, likely because of the presence of the chlorobenzene substituent. In the liver, chlorobenzene quickly goes

through oxidation to form an epoxide, and the epoxide is either hydrolyzed or conjugated to glutathione to form pre-mercapturic acid[27], which can be easily eliminated by human body.

Another aspect that is lacking in this field is that although several groups have characterized the steady-state kinetics and thermodynamic parameters of *Pf*TMK, there have not been a combined study that comprehensively characterize the kinetics, binding property, thermodynamics and dynamic of *Pf*TMK to provide a comprehensive understanding of SAR for substrates and potential inhibitors.

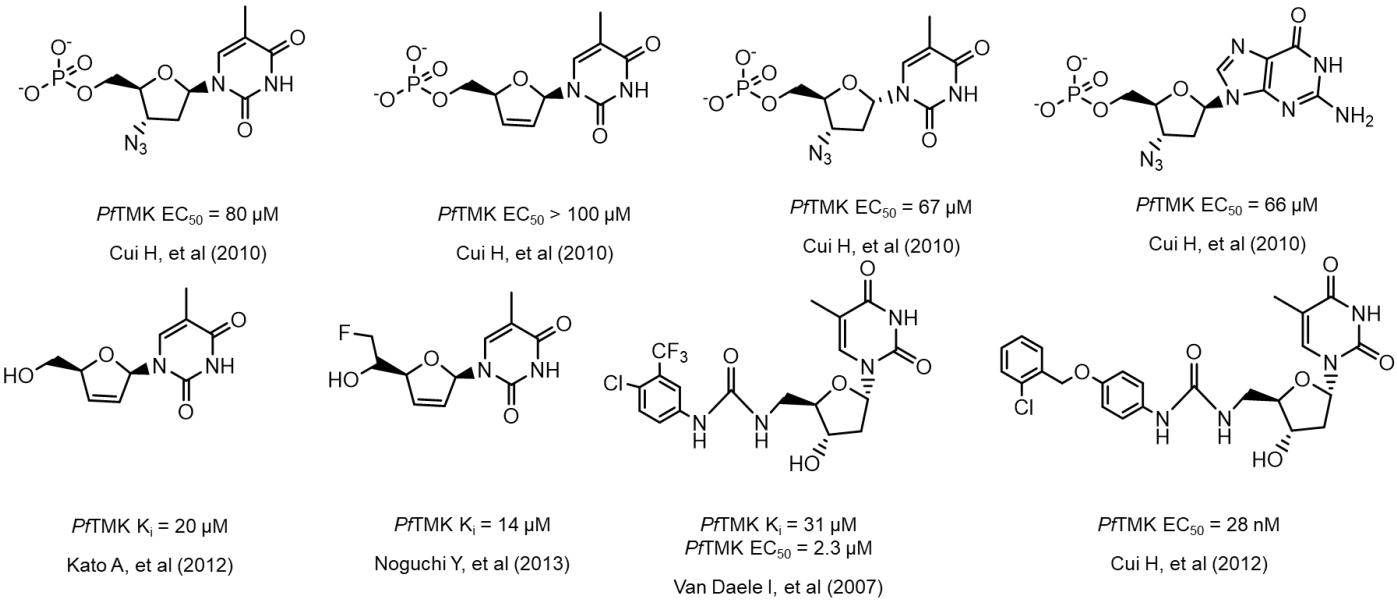


Figure 1.7 . Reported *Pf*TMK inhibitors. Most developed inhibitors are thymidine-based small molecules-and it is established that the α-thymidine variations are more effective than the β counterparts.

Chapter 2

Methyl Assignment of PfTMK

(NMR) spectroscopy is a power tool in the field of organic chemistry and biochemistry which exploits the magnetic properties of nuclei[28,29,30] to obtain information of interest. In the field of organic chemistry, NMR is the definitive method for the identification of monomolecular organic compounds. While in the field of biochemistry, NMR has more applications (usually for proteins smaller than 35 KDa), such as structural determination, protein dynamic studies and protein-substrate binding studies. NMR can also be used for drug discovery by exploiting the sensitivity of chemical shifts to binding of small ligands to the target; a previous study used NMR for high throughput screening of fragment library targeting thymidylate kinase from *M. tuberculosis*[31]. These features make NMR the perfect tool for our studies, as our interest lies in comprehensively studying the protein-substrate interaction of *Pf*TMK (~25 KDa),

developing small molecules that specifically target *Pf*TMK and potentially screening fragment libraries to identify promising leads against *Pf*TMK.

Methyl residue assignment of *Pf*TMK is a key bottle neck to all of our NMR experiments in the projects. Methyl side chains are positioned in the core of the protein and are sensitive to substrate binding and dynamic changes during binding events. These features make methyl residues ideal for detection of dynamic changes and probing residues that participate in binding[32,33].

Methyl residue assignment was achieved by a combination of site-directed mutagenesis and 4D NOESY experiment. In which the isoleucine residues were assigned by site-directed mutagenesis and 2D HMQC experiments; and the remaining leucine and valine residues were assigned by referencing between the isoleucine assignments, 4D NOE data and reported crystal structure. We were able to assign all visible isoleucine resonances and most of the leucine and valine resonances, to their corresponding amino acid in the protein.

2.1 Site-directed Mutagenesis

First, 2D HMQC experiment was carried out to obtain a complete wild-type *Pf*TMK (^{1}H-^{13}C) isoleucine methyl spectrum. Site-directed mutagenesis experiments were then carried out to substitute each of the isoleucine residues to valine residues (I11V, I51V, I54V, I55V, I70V, I85V, I93V, I127V, I162V, I172V, I176V, I178V, I184V, I187V, I191V, I198V). Substitution of isoleucine to valine is done to minimize the structural changes due to the point mutations[34]. 2D (^{1}H-^{13}C) methyl spectrum was obtained for each of the mutant *Pf*TMK, and each of the mutant spectrum will have one peak that does not show up in the isoleucine region (with the exception of I54 and I55, the two residues overlap i.e. have degenerative chemical shifts) in the wild-type spectrum, resulting in each mutant spectrum looking the same as the wild-type spectrum in terms of peak chemical shift, the only difference is the peak intensity between WT and mutant spectrum for this peak. Assignment of I184 is shown as an example in Figure 2.1, from the overlay of WT-*Pf*TMK and I184V spectrum, it can be observed that a peak is missing in the mutant spectrum compared to the WT spectrum, suggesting that this peak represents I184. This process was repeated for all isoleucine residues, and all isoleucine residues were successfully assigned (Figure 2.2). Some peaks in the 2D spectrum appear to have lower intensity compared to other resonances, this is due to intermediate exchange which causes linewidth broadening and results in lowered intensity.

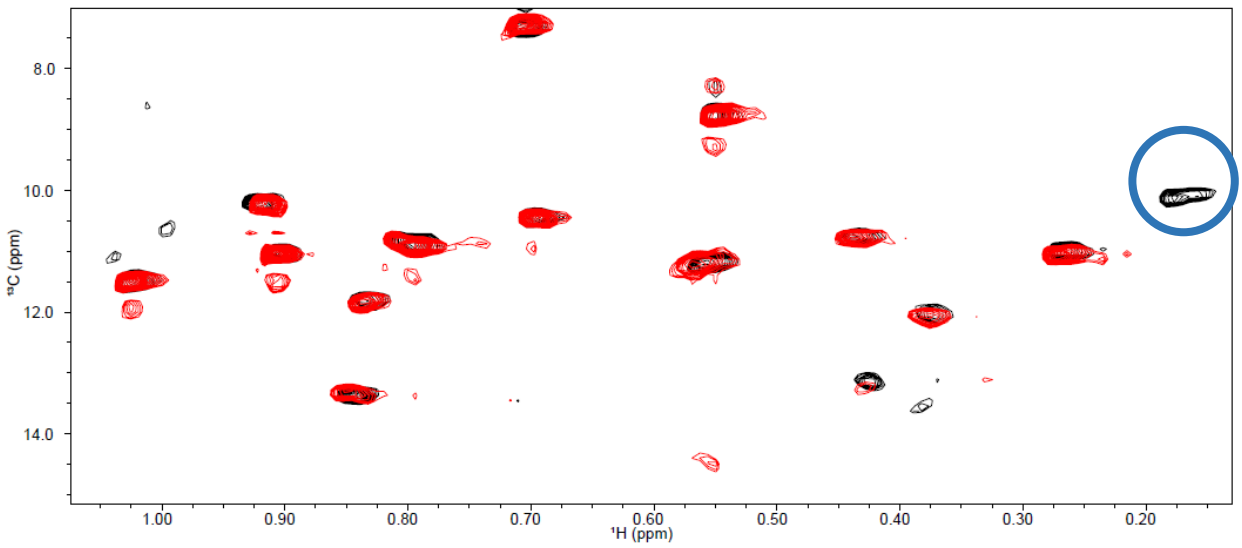


Figure 2.1. Isoleucine region of overlay of WT-*Pf*TMK (black) and I184V spectra (red). From the overly spectra, the WT spectrum exhibits one more peak than the I184V spectrum, this allowed us to assign this peak to I184. Both samples were saturated with dTMP and ADP.

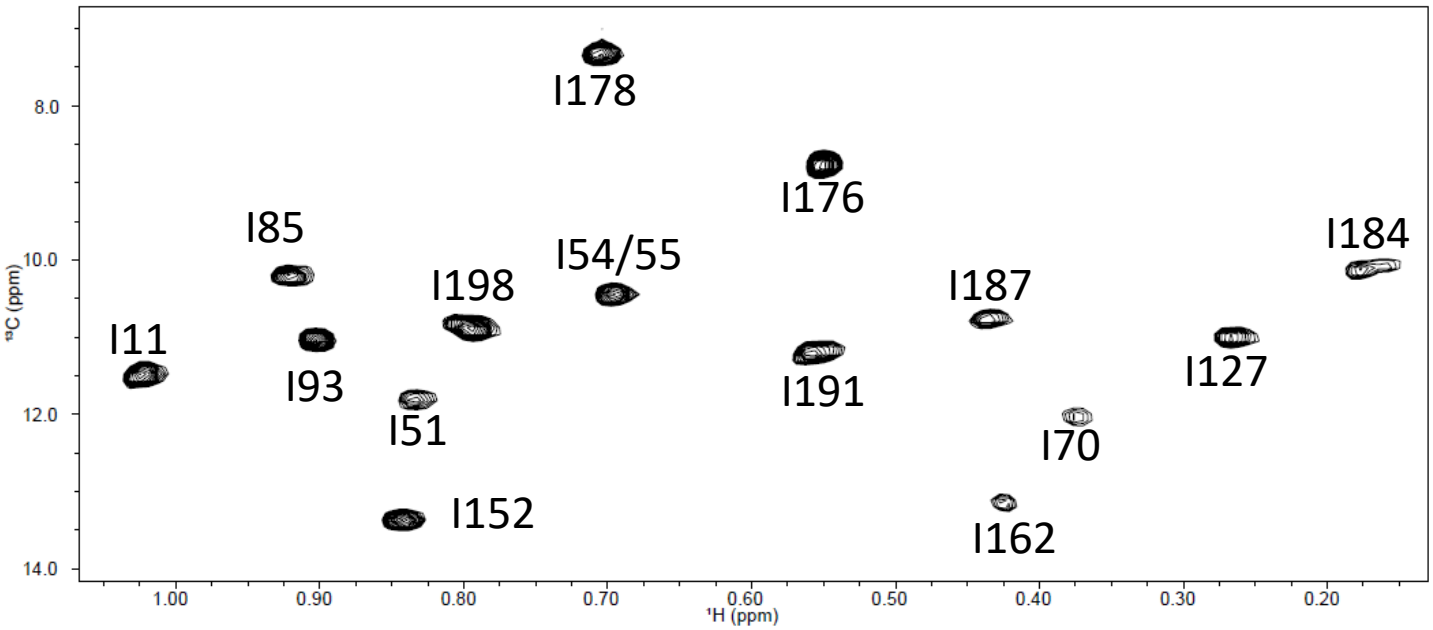


Figure 2.2. Assignments for all isoleucine residues in *Pf*TMK in saturated concentrations of dTMP and ADP.

2.2 4D Methyl-methyl NOESY

The 4D NOESY experiment examines the three-dimensional spatial correlation among the methyl groups in isoleucine, leucine and valine residues. In the 4D NOESY spectrum, each methyl will exhibit an intense self-peak along with other (if any) weaker cross-peaks due to correlation with any nearby methyl within the 5-7Å radius. The intensity of cross-peaks depends on the $1/r^6$.

The assignment of leucine and valine methyl groups were accomplished by cross referencing 4D NOESY spectrum, isoleucine assignments and reported crystal structure. More specifically, we start with looking at each of the assigned Isoleucine peaks in the 4D NOESY spectrum and identify resonance from methyls that are relatively close to that isoleucine residue that are represented by weaker cross-peaks. Then we would look at the crystal structure and examine which leucine/valine residues are near that isoleucine residue, and assign them based on measured distance and peak intensity from 4D NOESY spectrum (Figure 2.3).

This process is repeated for all isoleucine residues to assign as many leucine/valine residues as possible. The remainder leucine/valine residues that are not physically in close proximation to any isoleucine residues were assigned by using the peaks of assigned leucine/valine residues from the 4D NOESY spectrum.

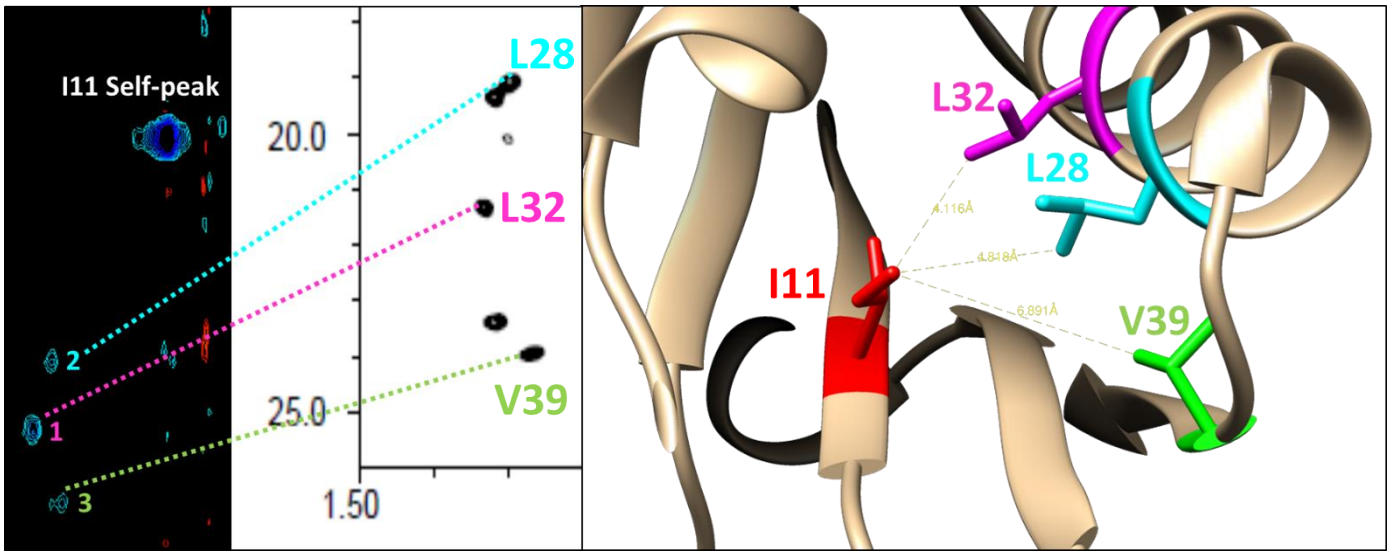


Figure 2.3. Illustration of leucine/valine residue assignments. Based on isoleucine assignment, we looked at its 4D NOE peak, where a strong self-peak and weaker cross-peaks (if any) can be observed. Note that the intensity difference in cross-peaks is a representation of their distance to the isoleucine of interest. Using the crystal structure of dTMP-PfTMK complex, we were able to measure the distance between I11 and nearby leucine/valine residues. Using these pieces of information, the L28, L32, V39 assignments were achieved.

2.3 Conclusion

All isoleucine residues were successfully assigned, and most visible leucine and valine residues were successfully assigned. Some leucine/valine residue could not be assigned due to the absence of cross-peaks in the NOESY spectrum; this make it difficult to determine which residues correspond to these peaks. Nevertheless, we obtained sufficient assignment information to establish the foundation of our subsequent NMR experiments: such as 2D HMQC titration experiments, CPMG relaxation dispersion experiments, and fragment library screening via chemical shift perturbation detection using NMR.

Chapter 3

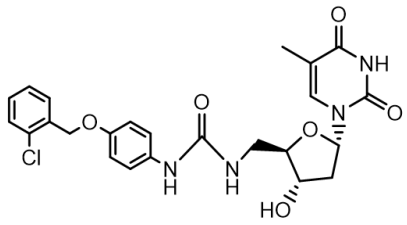
Synthesis and Characterization of Lead Compounds Against *Pf*TMK

Cui et al. reported their synthesized lead compound N-(5′-deoxy-α-thymidin-5′-yl)-N′-[4 (2-chlorobenzyloxy)- phenyl]urea (13b, Figure 3.1) to have an impressive EC_{50} value of 28 nM[11]. Although this compound is highly effective in inhibiting parasite growth, it remains unclear whether the growth inhibition is the result of *Pf*TMK inhibition or inhibition of another target.

More importantly, this compound is highly metabolically unstable due to the presence of the chlorobenzene substituent. In the liver, this substituent quickly goes through oxidation to form an epoxide, which subsequently goes through hydrolysis or

conjugation to glutathione to form pre-mercapturic acid. In other words, administering this compound would directly result in excretion and no pharmacological effect.

The rationale behind the design of our derivatives is to replace the metabolically unstable chlorobenzene substituent with more stable substituents to examine if chlorobenzene is required for binding and inhibition, and to attempt to enhance the inhibition effect and binding affinity. The characterization process included the Cl-derivative to examine whether its parasite growth inhibition is the result of *Pf*TMK inhibition or off-targeting effects.



*Pf*TMK EC_{50} = 28 nM

Cui H, et al (2012)

Figure 3.1. Structure of compound 13b, Cl-derivative.

3.1 Synthetic Strategy

The general synthetic strategy is presented in Figure 3.2. The process is broken down into three components: synthesis of α-thymidine side, isocyanate side and final coupling. The most critical step in the whole process is the racemization of β-thymidine to α-thymidine: this inverted isomer is reported to be more than 10-fold more effective than its β-thymidine counterparts. The racemization was done by protection of 3’ and 5’-OH, followed by isomerization reaction and deprotection reactions. Also note that the isocyanates are highly reactive and unsuitable for storage. Consequently, the final coupling step was carried out immediately after the purification of compounds 4a-e.

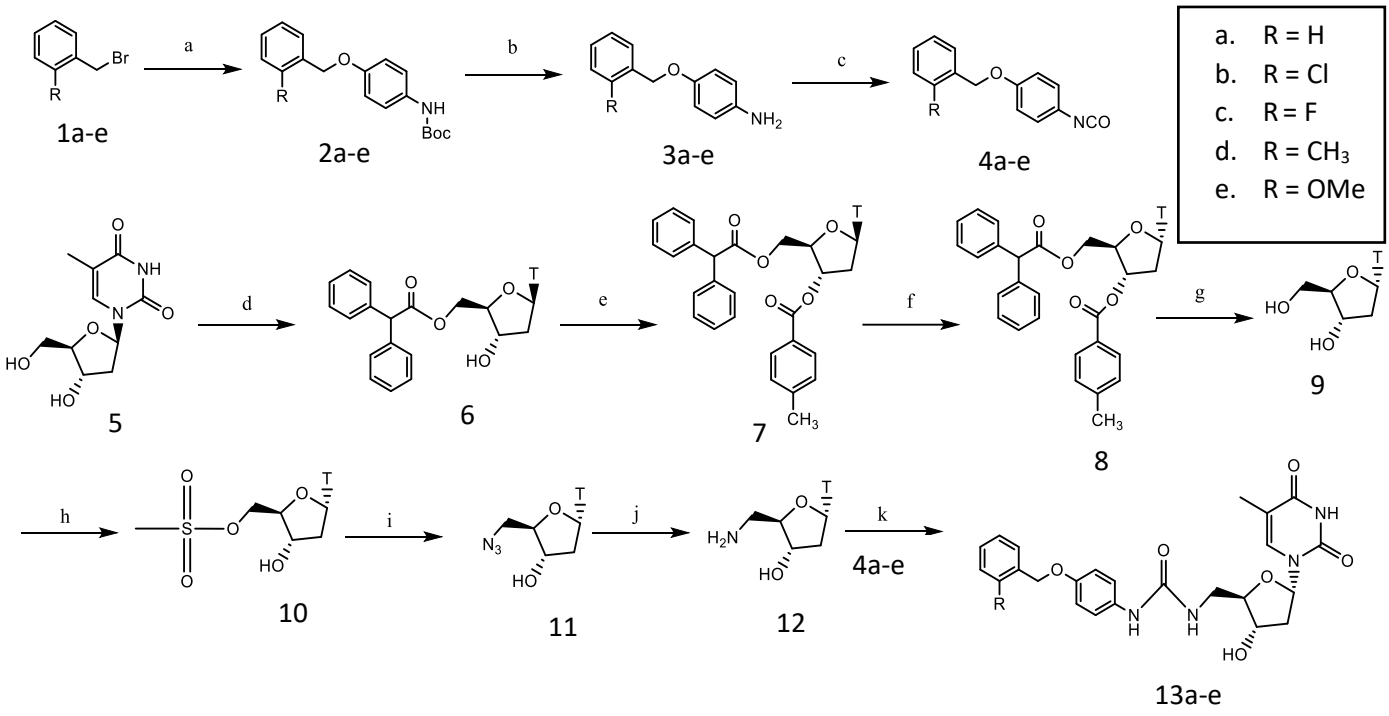


Figure 3.2. Library synthesis. Reagents and conditions: (a) NaH, N-Boc protected 4-aminophenol, DMF, 0°C, 1.5 hrs; (b) TMS-Cl, MeOH, RT, 3 hrs; (c) Et3N, triphosgene, EtOAc, 77°C, 3 hrs; (d) Diphenylacetyl chloride, pyridine, DCM, 0°C, overnight; (e) p-toluoyl chloride, pyridine, DCM, RT, overnight; (f) Acetic anhydride, H_2SO4, ACN, RT, 3 hrs; (g) NaOMe, MeOH, RT, 3 hrs; (h) MsCl, pyridine, -38°C, 1 hr, 0°C, 3hrs; (i) NaN_3, DMF, RT, 3 hrs; (j) 10% Pd/C, MeOH, RT, 3 hrs; (k) DMF, RT, 3 hrs

Synthesis of Isocyanate Side:

2-H-, 2-Cl-, 2-F-, and 2-CH3-benzyl bromide was reacted with N-Boc-4-aminophenol in DMF with addition of sodium hydride (0 °C for 2 hours) to yield 2a (96% yield), 2b (95% yield), 2c (90% yield), and 2d (95% yield).

The synthesis of compound 1e followed a slightly different procedure: 3-methoxybenzyl alcohol was reacted with phosphotribromide in DCM for 3 hours at 0 ºC to generate compound 2e.

The removal of Boc-protecting group was carried out by dissolving compounds 5a – e in methanol, then trimethylsaline chloride was slowly added to the reaction mixture, the reaction was run for 2 hours at room temperature, yielding compound 3a (96% yield), 3b (92% yield), 3c (92% yield), 3d (86% yield), and 3e (96% yield).

Compounds 3a – e were dissolved in ethyl acetate with addition of triethyl amine (1 equivalence) and triphosgene (2 equivalence), and the reaction was performed at 77 °C for 3 hours under reflux. The reaction mixture was flashed by a thin layer of silica gel and concentrated to yield compounds 4a - e.

Synthesis of α-thymidine side:

β-thymidine was reacted with diphenylacetyl chloride (1 equivalence) to yield 5'-O-diphenylacetyl-β-thymidine (6) in order to protect the 5’ hydroxyl group. The reaction was carried out in pyridine and methylene chloride (DCM) overnight at room temperature to yield 5'-O-diphenylacetyl-β-thymidine (6), and the residue was purified via silica gel chromatography (52% yield).

This is followed by the protection of 3’ hydroxyl with p-toluoyl chloride (2 equivalence): compound 10 with dissolved in pyridine/DCM, and the reaction was run overnight at room temperature. The organic phase was extracted with DCM (with water addition), the residual pyridine was quenched by HCl, which was subsequently saturated with sodium bicarbonate. The final organic phase was purified by column chromatography to yield 3'-O-p-toluoyl-5'-O-diphenylacetyl-β-thymidine (7) (95% yield).

The racemization reaction was then performed to convert β-thymidine into α-thymidine. Compound 7 was dissolved in acetonitrile followed by addition of sulfuric acid (1.5 equivalence) and acetic anhydride. After 3-hour runtime, the reaction was quenched with saturated sodium bicarbonate, and the residue was extracted with ethyl acetate. 3'-O-p-toluoyl-5'-O-diphenylacetyl-α-thymidine (8) was then crystallized in ethanol (47% yield).

The protecting groups on compound 8 were removed by reacting with sodium methoxide (2 equivalence) in methanol for 3 hours, and the reaction mixture was neutralized with Dowex ion exchange resin until pH reached 6, the resin was then filtered out after extensive washes with DCM, methanol and water. DCM and methanol

was first evaporated, and the water layer was collected and then concentrated to yield α-thymidine (9) (90% yield).

This is followed by mesylation of compound 9 at 5’-OH: compound 9 was dissolved in pyridine/DCM at -38 °C, and methanesulfonyl chloride (1 equivalence) was added. The reaction was kept at -38 °C for one hour, and two additional hours at 0 °C. The reaction was quenched with water, and compound was purified by column chromatography to yield 5'-O-methanesulfonyl-α-thymidine (10) (65% yield).

Azidation of compound 10 was then performed as follows: compound 10 was dissolved in dry DMF with addition of sodium azide (2 equivalence). The reaction was heated to 100 °C overnight. Then the solvents were evaporated, and the dried solid was re-dissolved in ethanol. The precipitate was filtered out, and the ethanol filtrate was concentrated and purified by column chromatography to yield 5'- azido-5’-deoxy-α-thymidine (11) (50% yield).

For the hydrogenation step, compound 11 was dissolved in methanol (with addition of Pd/C), and hydrogenated with hydrogen under atmospheric pressure for 3 hours. Pd/C was filtered out, the filtrate was concentrated to yield 5'-amino-5'-deoxy-α-thymidine (12) (91% yield).

Final Coupling:

The isocyanate intermediates (compounds 3a – e) are highly reactive. Consequently, the final coupling reaction of thymidine side to the isocyanate side was performed immediately after the production of the isocyanates. Compound 3a – e were dissolved in DMF and compound 12 was added, and the reaction was run for 3 hours at room temperature. The crude compounds were extracted with ethyl acetate and then purified by column chromatography to yield final compounds 13a (71% yield), 13b (86% yield), 13c (80% yield), 13d (68% yield), 13e (78% yield).

The end products, compounds 13a−e, were confirmed by both proton NMR and mass spectrometry (Figures 3.3−3.7). The Cl derivative (13b) was further evaluated through a series of two-dimensional (2D) NMR experiments, including HMBC, HSQC, and COSY.

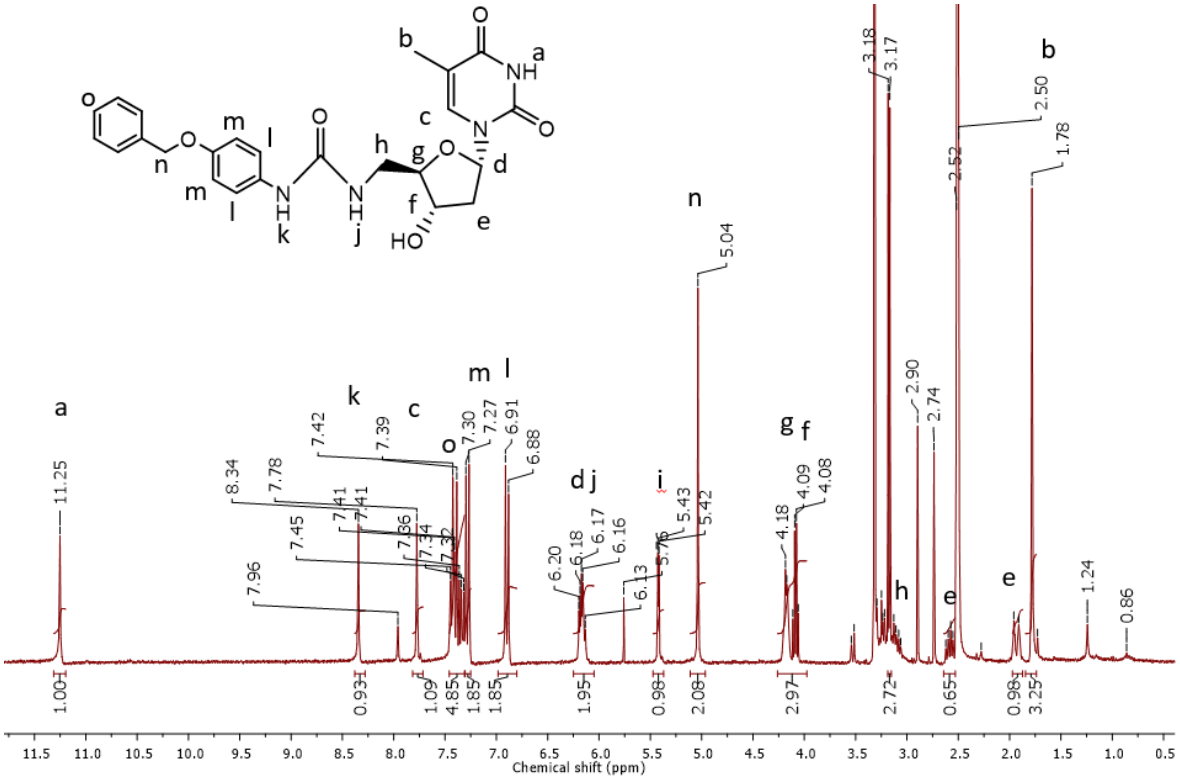


Figure 3.3. Proton NMR spectrum for compound 13a (H-derivative). All peaks are assigned to corresponding protons labeled in the chemical structure. Minor solvent peaks are observed at 5.76 ppm (DCM), 3.17 ppm (MeOH), and 2.90/2.74 ppm (DMF).

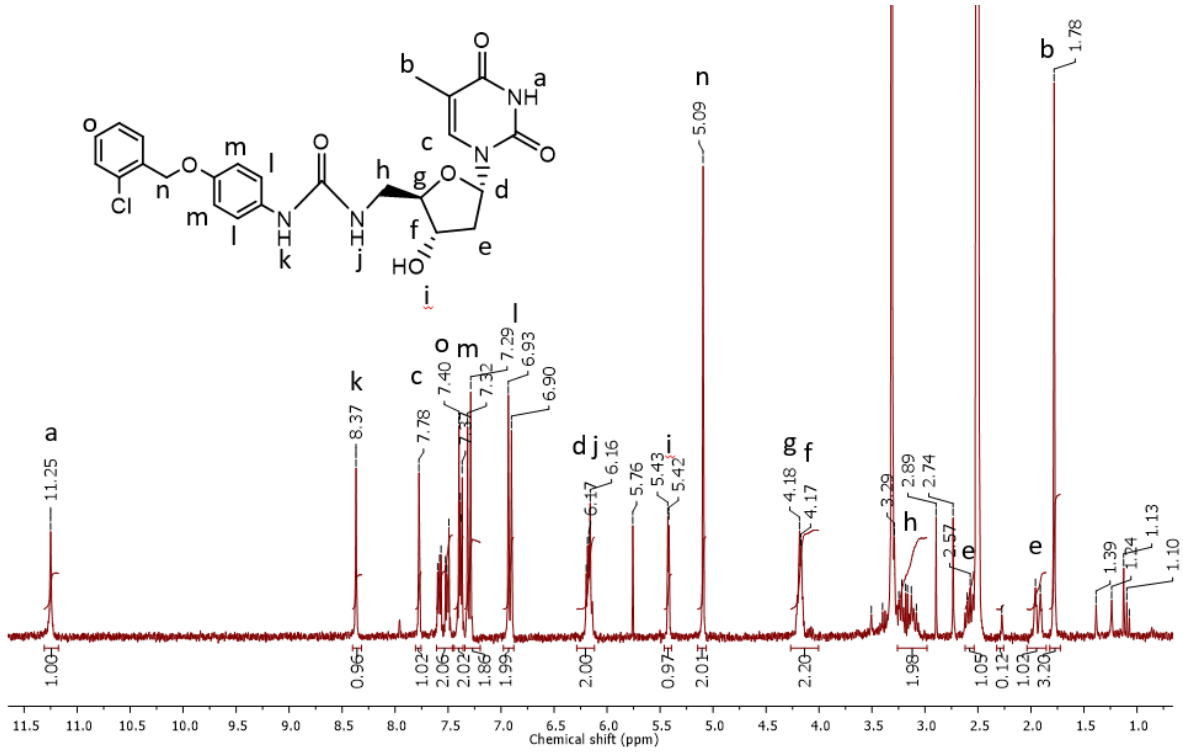


Figure 3.4. Proton NMR spectrum for compound 13b (Cl-derivative). All peaks are assigned to corresponding protons labeled in the chemical structure. Minor solvent peaks are observed at 5.76 ppm (DCM), and 2.90/2.74 ppm (DMF).

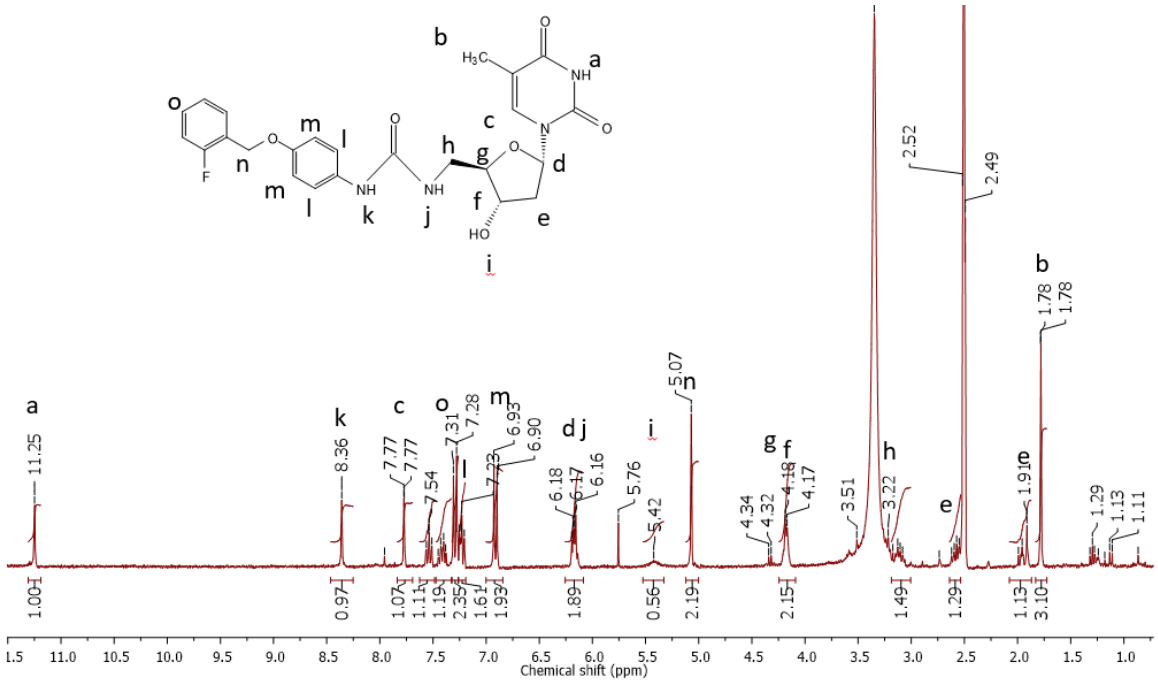


Figure 3.5. Proton NMR spectrum for compound 13c (F-derivative). All peaks are assigned to corresponding protons labeled in the chemical structure. Minor solvent peak is observed at 5.76 ppm (DCM).

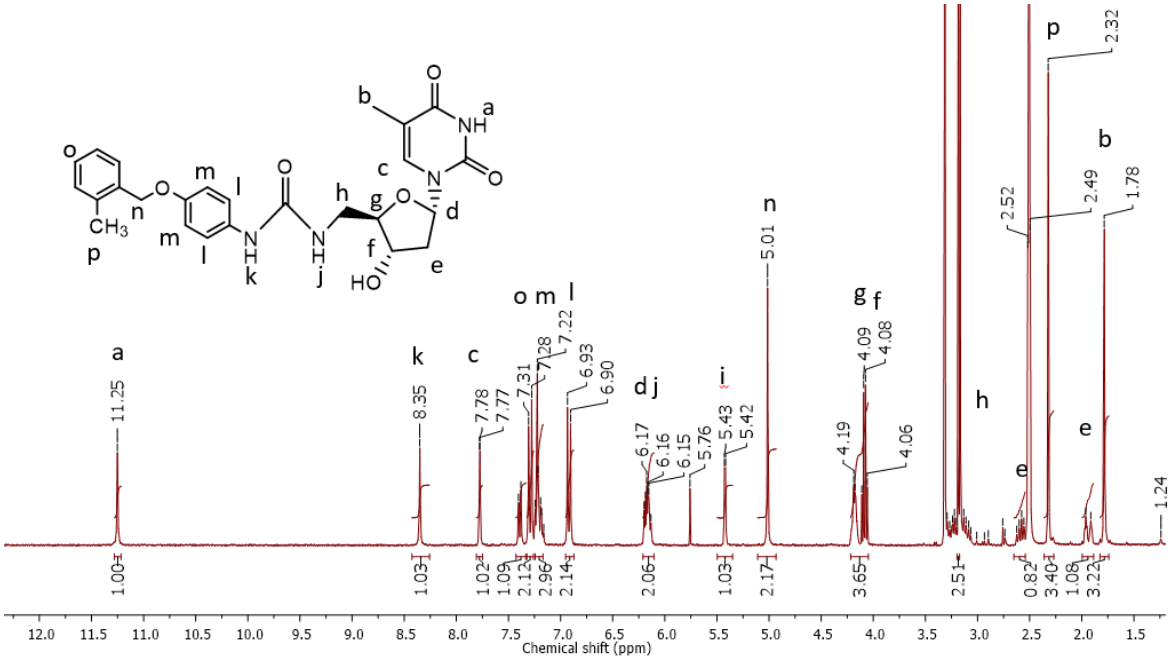


Figure 3.6. Proton NMR spectrum for compound 13d (CH3-derivative). All peaks are assigned to corresponding protons labeled in the chemical structure. Minor solvent peaks are observed at 5.76 ppm (DCM), 3.17 ppm (MeOH), and 2.90/2.74 ppm (DMF).

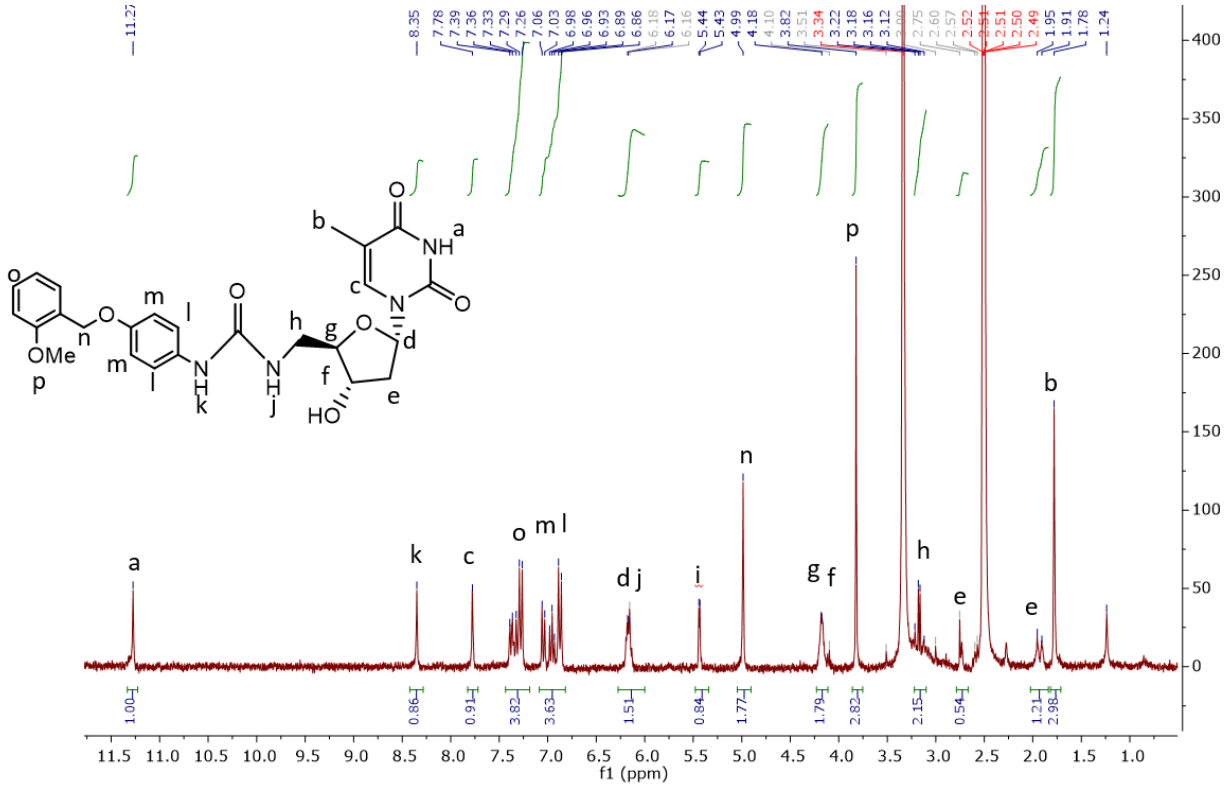


Figure 3.7. Proton NMR spectrum for 13e (OMe-derivative). All peaks are assigned to corresponding protons labeled in the chemical structure. No minor solvent peaks are observed in the spectrum.

3.2 Characterization of Synthesized Compounds

One of the major problems with compound 13b is that the compound was only characterized *in-vivo*. Consequently, it is difficult to evaluate whether the parasite growth inhibition is caused by *Pf*TMK inhibition or an off-target effect. In this set of experiments, we aimed to fully characterize compound 13b and our synthesized derivatives: compounds 13c and 13e. Compounds 13c and 13e were characterized because 13b contains a pharmacologically unstable chlorobenzene substituent, in compounds 13c and 13e, the unstable substituent was substituted with an electron-withdrawing or an electron-donating group (both are more metabolically stable), i.e., fluorobenzene, and methoxybenzene, respectively, to evaluate the role this terminal moiety plays in the binding and inhibition effect of these lead compounds.

More importantly, through the characterization process, we hope to obtain SAR information to shed some light on what structural modification could be made to further improve the binding affinity and inhibition effect of the lead compounds.

3.2.1 K_i Determination

The NADH-coupled assay was performed to experimentally measure the K_M and k_{cat} of *Pf*TMK for TMP and to examine the inhibition effect of compounds 13b, 13c, and 13e. The k_{cat} of wild-type (WT) *Pf*TMK was measured to be 3.1 ± 0.1 s^{-1}, and the K_M for TMP for *Pf*TMK was 25.5 ± 0.5 μM; our experimental values for both k_{cat} and K_M are consistent with previously published data[35]. Compounds 13b, 13c and 13e all behaved as competitive inhibitors: the introduction of these synthesized compounds only led to significant increase in apparent K_M, while the k_{cat} remained relatively unchanged.

The K_i values were obtained by using a Dixon plot: NADH-coupled experiments were carried out at 2, 3.5, 5, 10 and 20 μM of dTMP under the presence of 100, 200 and 300 μM of 13b to obtain the initial velocities. The Dixon plot was obtained by plotting the inverse of the initial rates against the concentration of the inhibitor, and the x-axis intercept will give an estimation of the K_i value. This procedure was repeated for compounds 13c and 13e, and the measured K_i values for these compounds are listed in Table 3.1. Dixon plots are shown in Figure 3.8B for compound 13c, Figure 3.9B for compound 13c, and Figure 3.10B for compound 13e.

13b showed a K_i of 220 ± 7μM, 13c showed a K_i of 217 ± 18 μM, and 13e exhibited a K_i of 230 ± 14 μM. The observed K_i values are significantly inconsistent with the reported EC_{50} value at 28 nM, this suggests that 13b may inhibit parasite growth through some

other off-target effect. Nevertheless, substitution of the chlorobenzene substituent exhibited very minor differences in terms of K_i; this indicates that the chlorobenzene moiety is not required for *Pf*TMK inhibition.

WT *Pf*TMK	K_M (μM)	K_D (μM)	K_i (μM)
TMP	25.5 ± 0.5	--	--
Cl-derivative (17b)	--	175.1 ± 1.6	220 ± 7
F-derivative (17c)	--	177.5 ± 15.0	217 ± 18
OMe-derivative (17e)	--	181.3 ± 1.0	230 ± 14

Table 3.1. Wild-type *Pf*TMK binding and kinetics parameters. K_D values were obtained from tryptophan fluorescence, K_i values from steady-state kinetics.

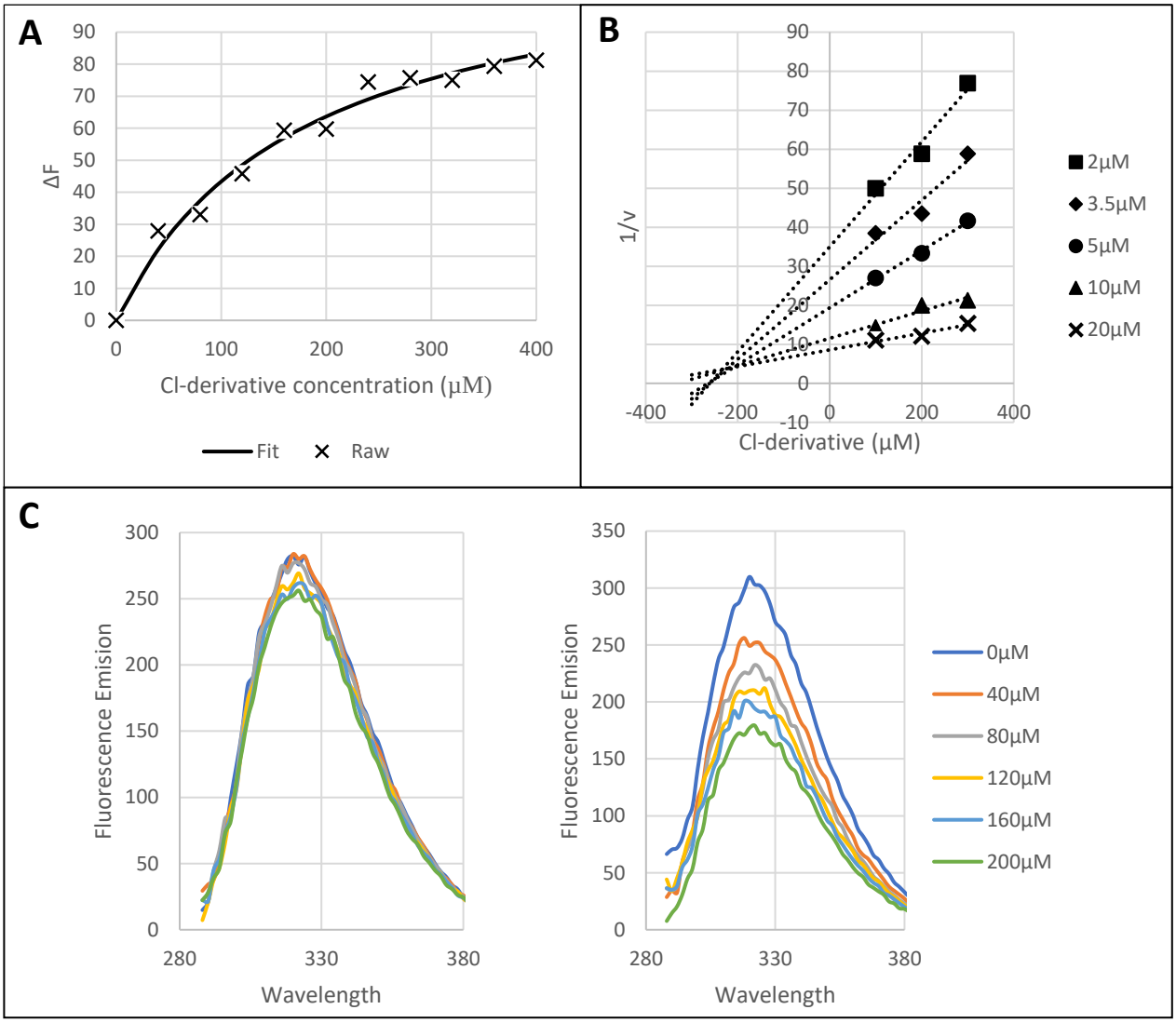


Figure 3.8. Characterization of Cl-derivative (13b). Panel A: K_D estimation: ΔF is plotted against the concentration of the synthesized compound and fitted to a hyperbolic curve for the estimation of the K_D value. The K_D of compound 13b is estimated to be 175.08±1.56μM. Panel B: Dixon plot for the estimation of K_i: 1/v is plotted against the concentration of the synthesized compound to estimate the K_i, in which the x-coordinate of the intercept of the trend lines = $-K_i$. Each line represents a different concentration of dTMP. The K_i for compound 13b is estimated to be ~225μM. Panel C: Left: Tryptophan fluorescence titration curve for F74A mutant. Right: Tryptophan fluorescence titration curve for Wild-type PfTMK. Significant difference is observed between the two set of titration curves. Compound 13b is unable to bind to the F74A mutant.

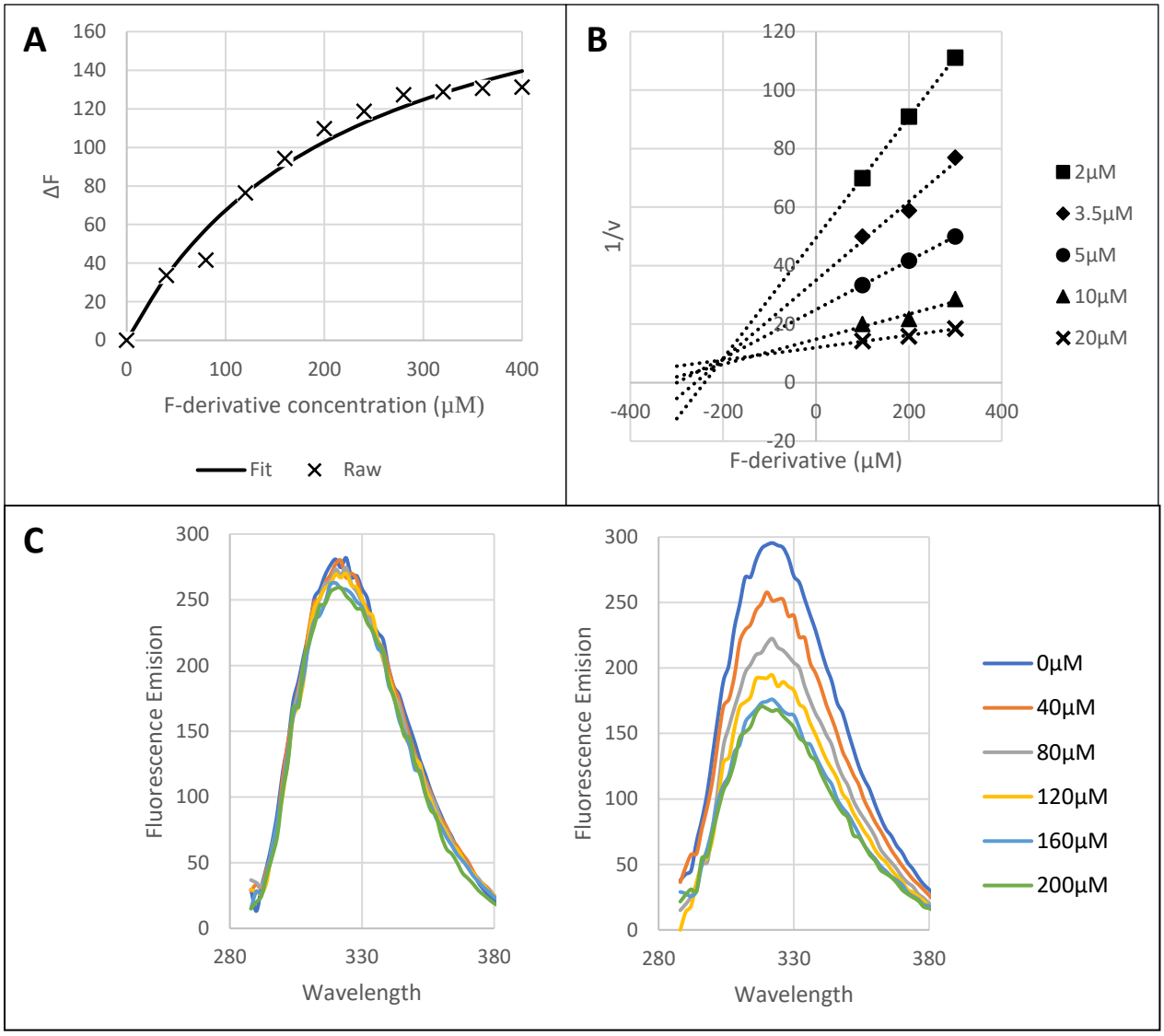


Figure 3.9. Characterization of F-derivative (13c). Panel A: K_D estimation: ΔF is plotted against the concentration of the synthesized compound and fitted to a hyperbolic curve for the estimation of the K_D value. The K_D of compound 13c is estimated to be 217.53±15.04μM. Panel B: Dixon plot for the estimation of K_i:1/v is plotted against the concentration of the synthesized compound to estimate the K_i, in which the x-coordinate of the intercept of the trend lines = - K_i. Each line represents a different concentration of dTMP. The K_i for compound 13c is estimated to be ~205μM. Panel C: Left: Tryptophan fluorescence titration curve for F74A mutant. Right: Tryptophan fluorescence titration curve for Wild-type PfTMK. Significant difference is observed between the two set of titration curves. Compound 13c is unable to bind to the F74A mutant.

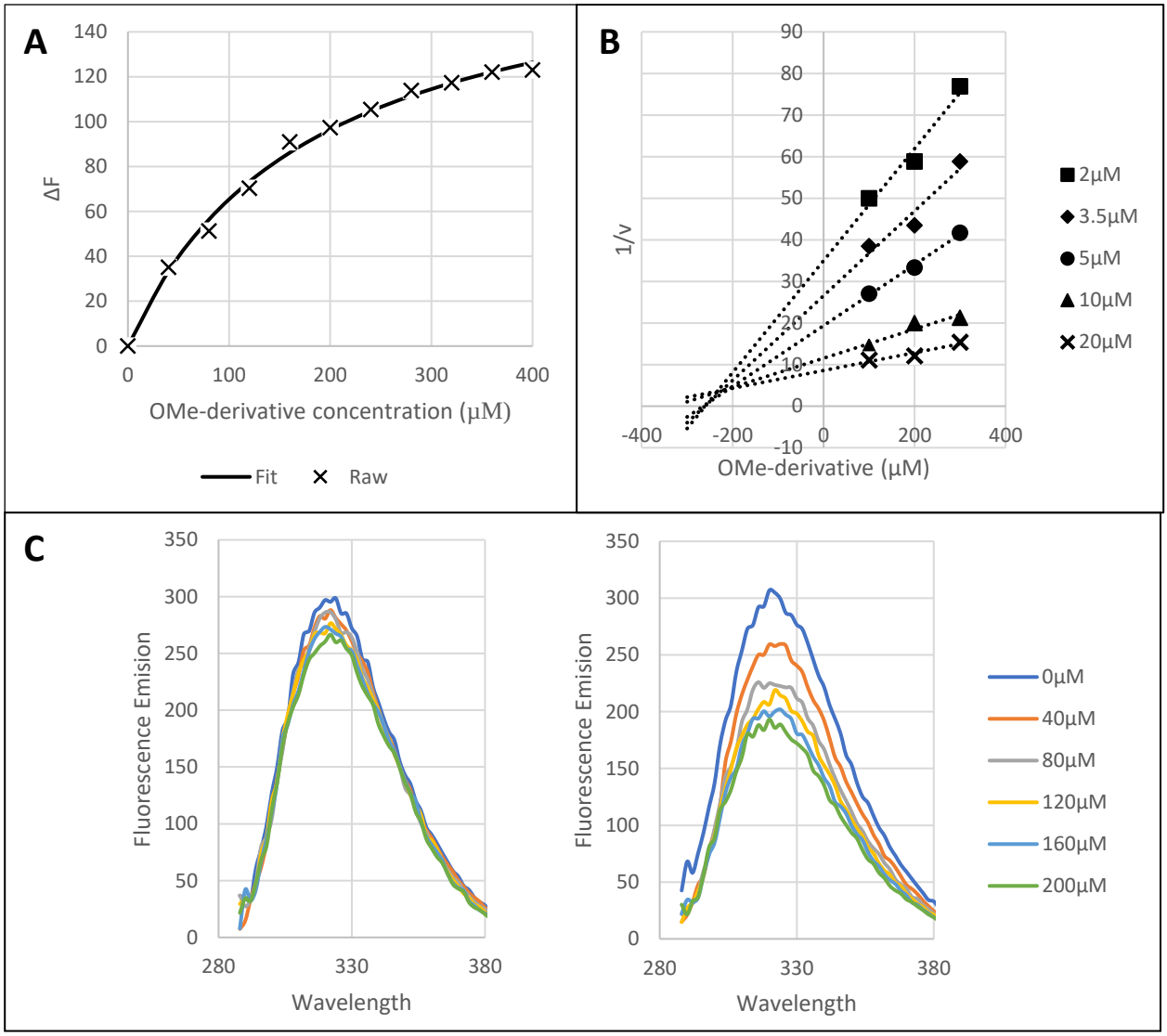


Figure 3.10. Characterization of OMe-derivative (13e). Panel A: K_D estimation: ΔF is plotted against the concentration of the synthesized compound and fitted to a hyperbolic curve for the estimation of the K_D value. The K_D of compound 13e is estimated to be 181.29±0.98. Panel B: Dixon plot for the estimation of K_i: 1/v is plotted against the concentration of the synthesized compound to estimate the K_i, in which the x-coordinate of the intercept of the trend lines = - K_i. Each line represents a different concentration of dTMP. The K_i for compound 13e is estimated to be ~220μM. Panel C: Left: Tryptophan fluorescence titration curve for F74A mutant. Right: Tryptophan fluorescence titration curve for Wild-type PfTMK. Significant difference is observed between the two set of titration curves. Compound 13e is unable to bind to the F74A mutant.

3.2.2 K_D Determination

To validate the K_i values from the NADH-coupled assay, K_D values were obtained from tryptophan fluorescence as a function of inhibitor concentration for each of the synthesized compounds (13b, 13c, and 13e). The fluorescence changes due to the binding of 13b are shown in Figure 3.8A, of 13c are shown in Figure 3.9A, and of 13e is shown in Figure 3.10A. The K_D values for each of the compounds were measured in duplicate and are listed in Table 3.1. The K_D values were measured at 175.1 ± 1.6, 177.5 ± 15.0 and 181.3 ± 1.0 µM for compound 13b, 13c and 13e, respectively. The measured K_D values are comparable to the measured K_i values. However, these values are not consistent with the reported EC50 value, this observation further suggests that the parasite growth inhibition by 13b is likely due to an off-target effect.

3.2.3 Modeling of Inhibitor Binding Site

Docking simulations were performed to confirm the binding site of compounds 13b, 13c and 13e. The docking simulation experiments were performed using Autodock Vina[36]. In our docking simulations, we used the X-ray crystal structure of *Pf*TMK in complex with TMP (Protein Data Bank entry 2WWF)[25]. The three-dimensional structures of the ligands were prepared with Chemsketch version 14.01[37], and the PDBQT formats of both the receptor and ligands were prepared with Autodock Tools[38,39].

All water and solvent atoms were removed, and the polar hydrogens were added prior to the docking simulation. Considering that our preliminary data suggested that 13b, 13c and 13e are behaving like competitive inhibitors, the grid box was generated by centering it on the TMP binding site of *Pf*TMK, and the coordinates of the grid box were then fed into Autodock Vina for the docking simulation calculations. Molecular structure images were prepared using Pymol version 1.8[40].

We observed that the synthesized compounds share the same binding pocket with dTMP and exhibit a similar binding mode (Figure 3.11, 3.12 and 3.13), and that the terminal benzyl substituents are outside of the active site. Compound 13b, 13c and 13e exhibited favorable and similar binding energies [Cl derivative (13b), −9.3 kcal/mol; F

derivative (13c), −9.3 kcal/mol; OMe derivative (13e), −9.1 kcal/mol], which suggests that substitution of the functional groups on the terminal benzyl group has a relatively small effect on the binding, this observation is consistent with our observations from both K_i and K_D measurements.

The three compounds fit into the dTMP binding pocket and retain the base stacking between the pyrimidine base and that of the aromatic ring of Phe74, which is consistent with their activity as competitive inhibitors. The docking simulations also show that the α configuration facilitates interactions with His71 and Arg78 residues, which improves the interaction of the 3′-OH with Tyr107 and Arg99. In addition to these residues, the docking simulation revealed that Asp17, Arg18, Lys21, Arg47, and Glu151 potentially interact with the synthesized compounds (Figure 3.11 and Table 3.2).

The critical role of Phe74 in binding thymidine is shown by the fact that the F74A mutant has no enzymatic activity[35]. Consequently, the replacement of Phe74 with Ala (F74A) should significantly weaken the binding of the inhibitors to *Pf*TMK. The structural integrity of the F74A mutant was previously investigated using circular dichroism. Those experiments indicated that there were no significant changes in the overall secondary structure of the F74A protein. To further validate the structural integrity of the F74A mutant, we compared 2D proton−carbon HMQC spectra of ILV-labeled F74A protein to the ILV spectrum of WT protein (Figure 3.14). A small number of methyl peak shifts were observed in the isoleucine region: I51, I54/55, I85 and I162; however, a majority of the peaks were in similar positions. Tryptophan fluorescence titration experiments were performed to determine the binding affinity of the inhibitors to the F74A protein for 13b (Figure 3.8C), as well as 13c (Figure 3.9C) and 13e (Figure 3.10C). We observed that

the binding affinity of the synthesized compounds is significantly reduced for the mutants, similarly to how F74A impacts dTMP binding. This further illustrate that compounds 13b, 13c and 13e share similar binding mode with dTMP, and function as competitive inhibitors.

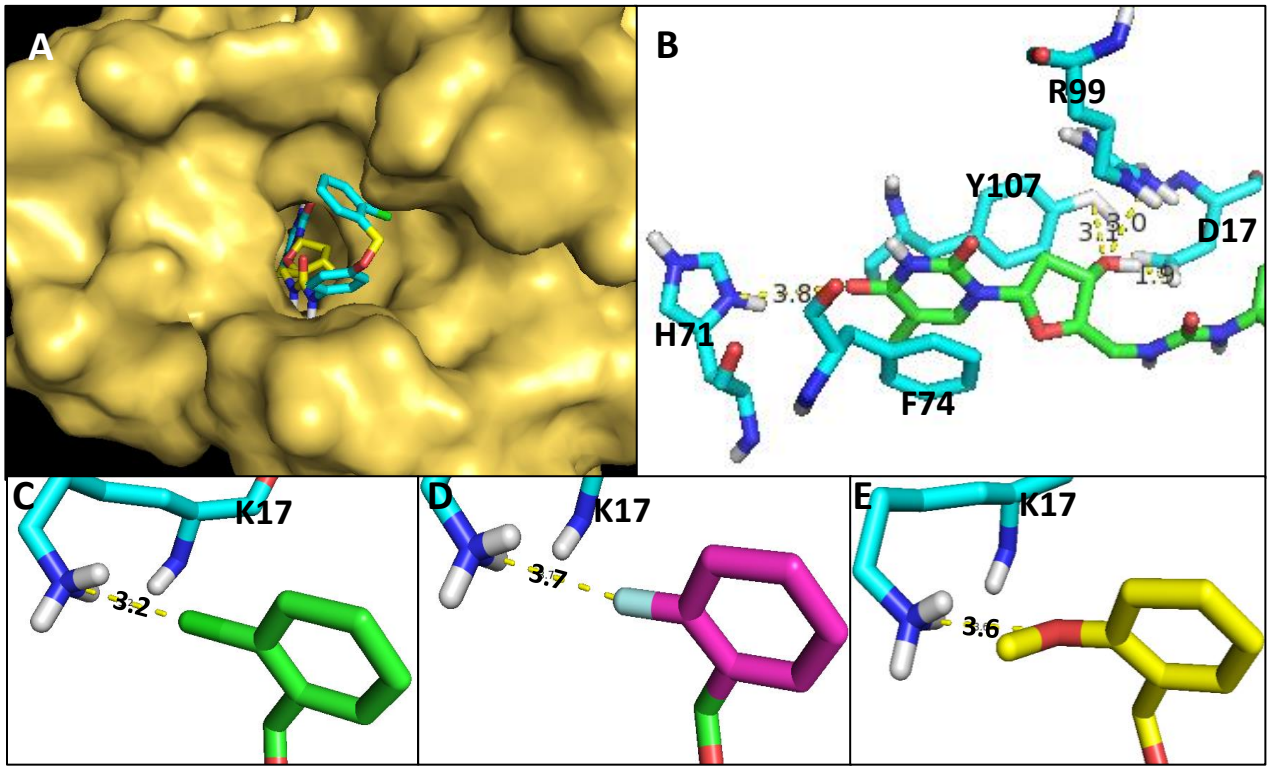


Figure 3.11. Docking simulation. Panel A: Compound 13b fits tightly into the TMP binding pocket with a binding energy of -9.3 Kcal/mol. Panel B: Atomic level detail of the α-thymidine part interacting with PfTMK. Panel C: Chloro-benzene substituent (compound 13b) interacting with Lys17. Panel D: Fluoro-benzene substituent (compound 13c) interacting with Lys17. Panel C: Methoxy-benzene substituent (compound 13e) interacting with Lys17.

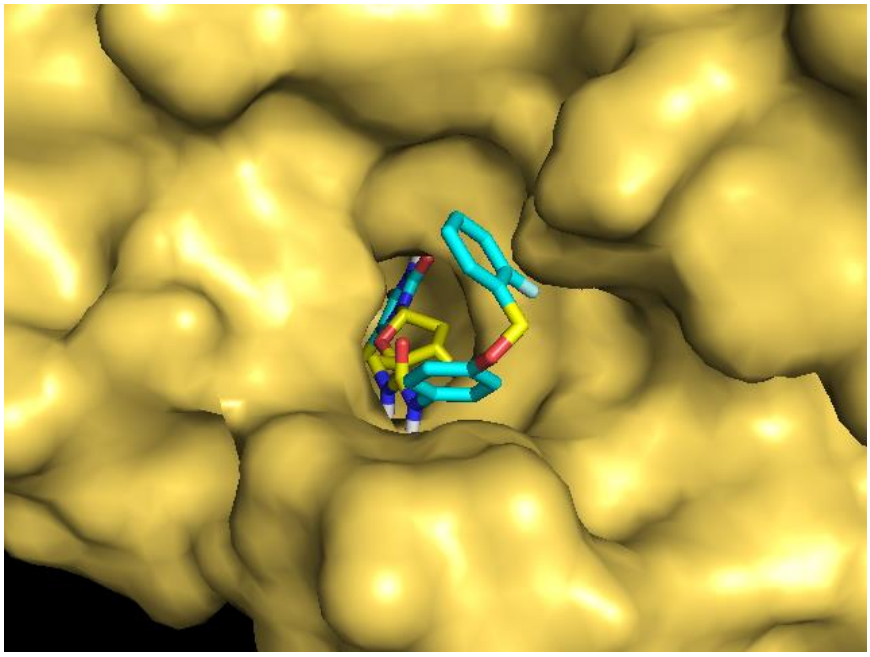


Figure 3.12. F-derivative (13c) docking simulation. Compound 21c fits tightly into the TMP binding pocket with a binding energy of -9.3 Kcal/mol.

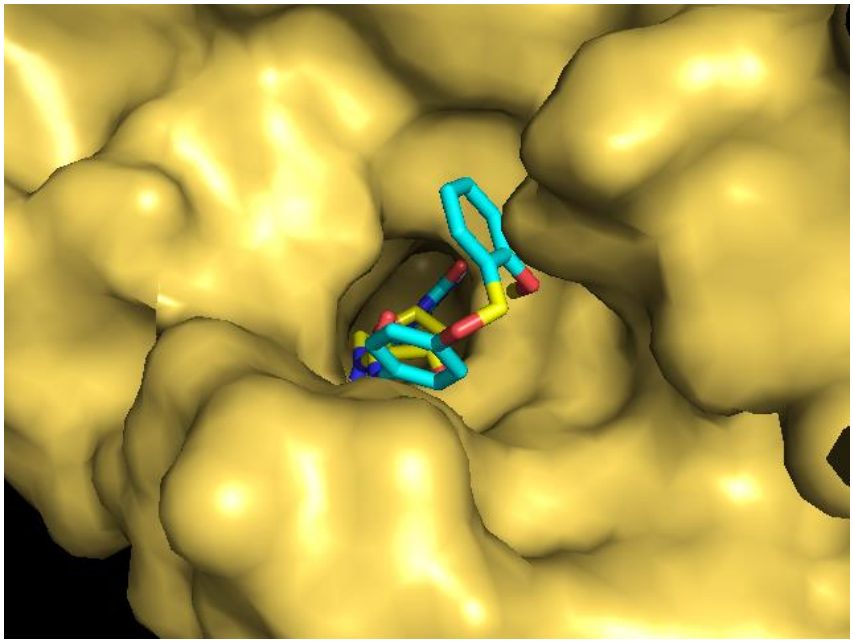


Figure 3.13. OMe-derivative (13e) docking simulation. Compound 21e fits tightly into the TMP binding pocket with a binding energy of -9.1 Kcal/mol.

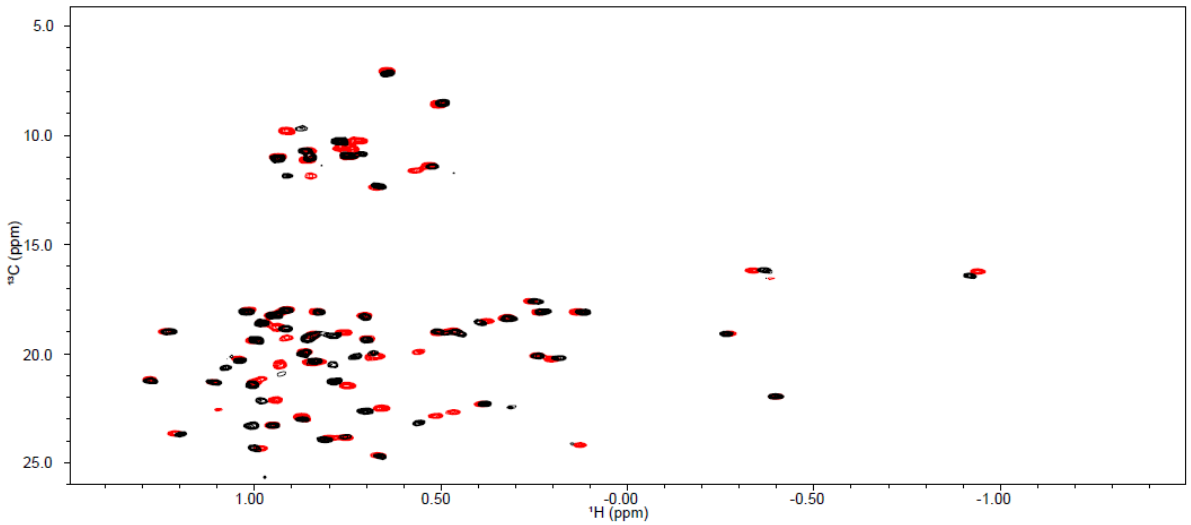


Figure 3.14. Overlay of PfTMK HMQC spectra. Wild-type PfTMK spectrum (black) is overlaid with the F74A mutant spectrum (red). The F74A is structurally intact. Both samples were saturated with dTMP.

Amino acid residue	Ligand
Tyr107	3’OH
Arg99	3’OH
Asp17	3’OH, urea linkage
F74	Pyrimidine base
His71	Pyrimidine base
Arg78	Pyrimidine base
Glu151	Urea linkage
Arg47	Urea linkage
Arg18	Phenol group
Lys17	Terminal benzyl functional group

Table 3.2. **Compound interaction with *Pf*TMK**

3.3 Conclusion

The enzyme inhibition, binding, docking simulation, and mutational studies showed that the Cl, F, and OMe derivatives (13b, 13c, and 13e, respectively) are all competitive inhibitors of *Pf*TMK with comparable K_i and K_D values. Consequently, the chlorobenzene can be substituted with electron-withdrawing or -donating groups (i.e., fluorine and methoxy) without significant changes in binding affinity.

Although the Cl derivative (13b) is an efficient inhibitor of parasite growth, with an impressive EC50 value of 28 nM, our experimental data show that the binding to *Pf*TMK is relatively weak, with K_i and K_D values around 200 μM. This suggests that the physiological target of 13b is unlikely to be *Pf*TMK, but some other enzyme in the cell. An alternative explanation is that the metabolite of the 13b consists of an epoxide, which could be lethal to the organism. One possible approach to examine epoxide toxicity is to investigate if compound 13c, which contains a stable fluorobenzene substituent and does not form epoxide as metabolite, exhibits similar EC50 with 13b. If significantly different EC50 values were observed between compound 13b and 13c, that would suggest the parasite growth inhibition of 13b is a result of epoxide accumulation.

Although the results from the docking simulation suggest that modifications could be made to the terminal benzyl group or the urea linkage that would improve binding of these inhibitors, the relatively high K_i values for these compounds suggest that

significant structural modifications will be required to increase their binding affinity. It is unclear if these changes would provide sufficient improvement to the binding energy such that they would become effective at inhibiting *Pf*TMK and consequently growth of the parasite. Moreover, considering that the inhibition of growth by 13b is likely due to off-target effects, it may not serve as an optimal scaffold for investigation of structure−activity relationships (SARs) for this group of compounds with respect to *Pf*TMK since there are quite a few other reported lead compounds that exhibited much better binding affinity and inhibition effect against *Pf*TMK. Rather, it may be worthwhile to identify the physiological target of 13b to aid the development of more potent *Pf*TMK inhibitors.

Chapter 4

Characterization of *Pf*TMK-substrate Interaction

The steady-state kinetics, thermodynamics, and crystal structure of *Pf*TMK have been examined in several studies. However, there has not been a study that comprehensively characterizes the protein biochemistry of *Pf*TMK and provide a comprehensive understanding on the binding mode and unique substrate specificity of *Pf*TMK. In this project, we aimed to characterize the steady-state kinetics, thermodynamics, binding properties, and dynamics difference between dTMP-*Pf*TMK and dGMP-*Pf*TMK complexes; and elaborate on the unique substrate specificity of *Pf*TMK (being able to utilize dGMP as substrate) as well as examining the binding mode of dTMP-*Pf*TMK and dGMP-*Pf*TMK complexes. Through these studies, we identified a residue that is

responsible for the unique substrate specificity of *Pf*TMK and revealed a unique half-site binding mode with dGMP. These findings provide insights on future development of *Pf*TMK-specific inhibitors using purine-based analogs as scaffolds.

4.1 Steady-state Enzyme Kinetics

NADH-coupled assays were carried out to examine the kinetics of *Pf*TMK with ATP and dTMP or dGMP as substrate. The experimental values are listed in table 4.1. We observed the K_M for dTMP for *Pf*TMK was 25.5 ± 0.5 μM with a k_{cat} of 3.1 ± 0.1 s^{-1}, while the K_M for dGMP was 36.2 ± 3.2 μM and the k_{cat} was measured to be 2.4 ± 0.3 s^{-1}. Our observed values are comparable to previously reported values by Kandeel et al[35], where he obtained values of 22 ± 3.6 μM and 3.2 ± 0.26 s^{-1}; and 30 ± 7.2 μM and 2.7 ± 0.44 s^{-1} for the K_M and k_{cat} values for dTMP and dGMP, respectively. ATP activity was measured at saturating dTMP concentration (1 mM), and we obtained consistent k_{cat} value and lower K_M for ATP comparing with previously reported data[28](K_M^{ATP} = 218 μM, k_{cat}^{ATP} = 3.4 s^{-1}): the measured K_M for ATP for *Pf*TMK was 100.8 ± 9.1 μM, and the k_{cat} was measured to be 3.3 ± 0.2 s^{-1}(Table 4.1) which is two-fold lower than previously reported data[35].

	K_M^{dTMP} (μM)	k_{cat}^{dTMP} (s^{-1})	K_M^{dGMP} (μM)	k_{cat}^{dGMP} (s^{-1})	K_M^{ATP} (μM)	k_{cat}^{ATP} (s^{-1})
WT	25.5 ± 0.5	3.1 ± 0.1	36 ± 3.2	2.4 ± 0.3	100.8 ± 9.1	3.3 ± 0.2
S108A	25.8 ± 1.9	3.3 ± 0.2	47.6 ± 1.7	1.3 ± 0.05	--	--
S108T	24.2 ± 1.6	3.2 ± 0.2	78.4 ± 5.7	1.3 ± 0.03	--	--
Y153F	65.9 ± 8.2	1.2 ± 0.2	96.9 ± 0.8	1.0 ± 0.07	--	--

Table 4.1 . Summary of kinetics constants for wild-type and mutant *Pf*TMK. NADH-coupled assays were carried out to examine the K_M and k_{cat} of WT and mutant PfTMK for dTMP, dGMP. dIMP and ATP (only for wild-type). Experimental results are an average of at least 2 replicate experiments. Significant differences in K_M and k_{cat} are highlighted.

4.2 Thermodynamic parameters

ITC experiments were conducted to investigate the thermodynamic parameters of dTMP, dGMP, ATP, and ADP binding to *Pf*TMK (Figure 4.1, 4.2). We used relatively high concentration of *Pf*TMK to more accurately measure the heat change during the initial injections when the substrate concentration is significantly lower than that of protein. Since *Pf*TMK is a dimeric protein, we first adopted the fitting method by Kandeel et al.[35], which is by assuming dimer concentration of the protein sample (1/2 of monomer concentration), and fitting the data to one site, or two sets of sites binding model. However, we were unable to obtain decent fitting using the previously reported fitting model, and we observed that our data was best fitted by assuming monomer concentration of the protein, and fitting to one set of site binding model.

The measured stoichiometry (n), association constant (K_a), enthalpy (ΔH°), and entropy (ΔS°), and the calculated binding affinity (K_D) and binding free energy (ΔG°) are listed in Table 4.2 with regard to the binding of dTMP and dGMP. dTMP exhibits higher binding affinity to *Pf*TMK compared with dGMP (19.87 μM ± 0.53 versus 38.67 ± 4.00 μM). The binding free energy was similar for dTMP and dGMP binding (-6.71 ± 0.09 kcal/mol versus -6.78 ± 0.28 kcal/mol), the large negative unfavorable ΔS° is compensated by the relatively large ΔH° given off during binding. The binding enthalpy and entropy of dGMP (-19.91 ± 0.26 and -13.13 ± 0.02 kcal/mol, respectively) are larger in magnitude than those of dTMP (enthalpy: -8.45 ± 1.09 kcal/mol, entropy: -1.74 ± 0.99 kcal/mol).

The larger binding enthalpy for dGMP versus dTMP can be explained by that fact that the purine ring on dGMP is larger than the pyrimidine ring in dTMP, resulting in higher heat release upon binding. The larger decrease in entropy suggests that the dGMP-*Pf*TMK complex is more rigid than the dTMP complex or binding of dTMP versus dGMP involves different solvation effects, in which the binding of relatively larger guanine base leads to displacement of additional water molecules. The more striking observation is that dGMP exhibits half site binding with consistent stoichiometry prediction of 0.5. This value suggests that in every *Pf*TMK dimer there is only one dGMP bound. This phenomenon is not observed in dTMP binding, where the stoichiometry prediction is consistently 1; we also attempted to fit the dGMP binding isotherm to a model where we set the stoichiometry to 1, and the fits were poor compared to when stoichiometry was 0.5 (Figure 4.1C).

Interestingly, we also observed that ADP (under both non-dTMP and dTMP-saturated condition) and ATP (under non-dTMP condition) appear to bind to *Pf*TMK in such a binding mode that produce little to no heat change (Figure 4.2).

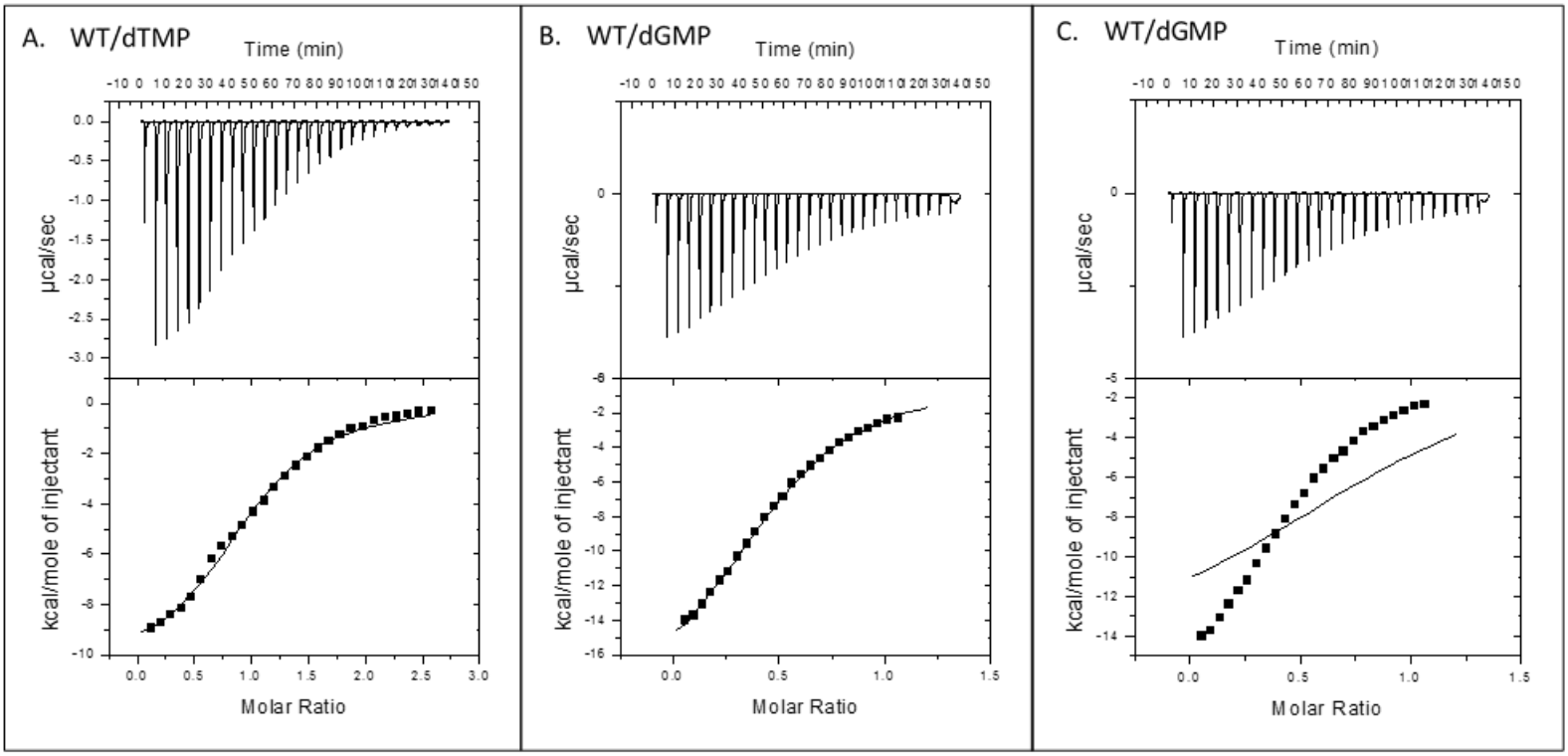


Figure 4.1 . ITC profiles of dTMP and dGMP binding to WT *Pf*TMK. Panel A: ITC profiles of dTMP binding, top panel illustrates the raw titration data, bottom panel shows integrated binding isotherms and fitted results. Panel B: ITC profiles of dGMP binding, top panel illustrates the raw titration data, bottom panel shows integrated binding isotherms and fitted results. Panel C: Fitting the same data from panel B to the same model while fixing stoichiometry to 1, fitting is very poor comparing with Panel B.

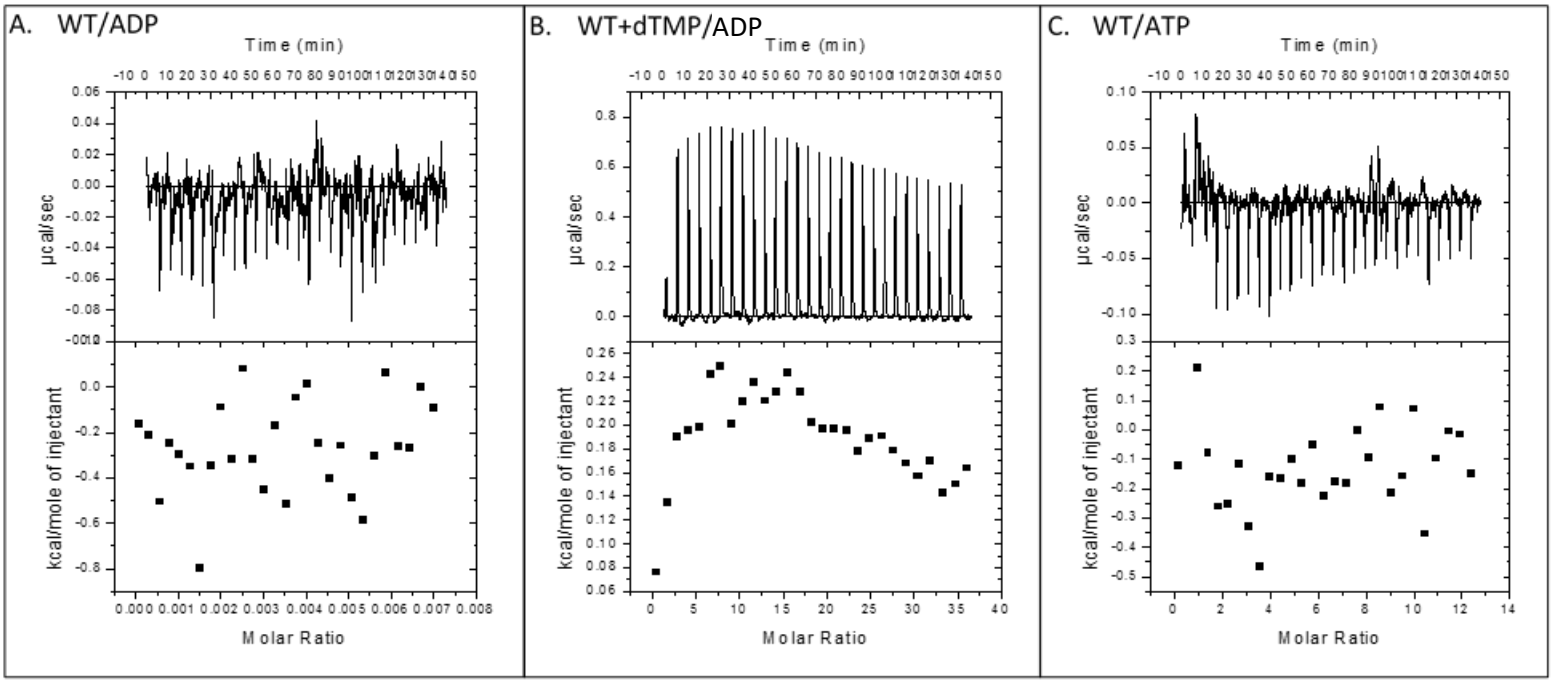


Figure 4.2 . ITC profiles of ADP and ATP binding to WT *Pf*TMK. Panel A: ITC profiles of ADP binding to PfTMK; Panel B: ITC profiles of ADP binding to dTMP-PfTMK complex; Panel C: ITC profiles of ATP binding to PfTMK. All top panels illustrate the raw titration data, bottom panel shows integrated binding isotherms and fitted results. The binding isotherms suggest that ADP and ATP may bind to PfTMK in a unique binding mode that produce little heat changes.

Substrate	Stoichiometry Prediction	K_a (μM-1)	K_D (μM)	ΔH° (kcal/mol)	TΔS° (kcal/mol)	ΔG° (kcal/mol)
dTMP	1.02± 0.042	0.05 ± 0.001	19.87 ± 0.53	-8.45 ± 1.09	-1.74 ± 0.99	-6.71 ± 0.09
dGMP	0.50± 0.009	0.026 ± 0.003	38.67 ± 4.00	-19.91 ± 0.26	-13.13 ± 0.02	-6.78 ± 0.28

Table 4.2 . Summary of thermodynamic parameters for the binding of dTMP and dGMP to *Pf*TMK. Results are an average of at least 3 replicate experiments. dGMP exhibited lower binding affinity than dTMP. Binding affinity was then obtained by equation $K_D = 1/K_a$, and the binding free energy (ΔG°) was calculated by equation ΔG° = ΔH° – TΔS°. The free binding energy is consistent between dTMP and dGMP, nonetheless, dGMP binding exhibited entropy and enthalpy with higher magnitude. dTMP titration experiments consistently yield stoichiometry of approximately 1, while dGMP titration experiments suggested half-site binding.

4.3 Binding Properties

4.3.1 Tryptophan Fluorescence Titration Assay

The binding of dTMP, dGMP, and ADP to *Pf*TMK was evaluated by tryptophan fluorescence titration experiments. Our measurement suggested a K_D of 22.8 ± 1.9 μM for dTMP, and a K_D of 42.5 ± 2.2 μM for dGMP, and a K_D of 200.4 ± 32.8 μM for ADP (Table 4.3). The K_D values for dTMP and dGMP measured by our tryptophan fluorescence titration experiments are consistent to those from our ITC experiments, and are comparable to the K_M values from our NADH coupled assay.

The rationale behind the measurement of K_D^{ADP} is to determine the optimal ADP concentration for our subsequent NMR experiments, so that the NMR samples can be saturated with ADP.

	K_D^{dTMP} (μM)	K_D^{dGMP} (μM)	K_D^{dIMP} (μM)	K_D^{ADP} (μM)
WT	22.8 ± 1.9	42.5 ± 2.2	561.6 ± 35.4	200.4 ± 32.8
S108A	29.2 ± 3.1	62.1 ± 4.3	660.5 ± 6.1	--
S108T	27.0 ± 1.6	88.5 ± 3.9	522.1 ± 41.8	--
Y153F	68.2 ± 7.5	131.9 ± 16.2	644.9 ± 51.6	--

Table 4.3 . Summary of binding affinity of dTMP, dGMP, dIMP and ADP to WT-*Pf*TMK and mutant *Pf*TMK from tryptophan fluorescence titration assay. Results are an average of at least 2 replicate experiments (each replicate experiment result is an average of the results from 3 to 5 paralleled reactions). The K_D values for dTMP and dGMP are comparable to our data collected from ITC experiments, and are comparable to the measured K_M values. Y153F mutant exhibited overall higher K_D for all substrates (K_D^{ADP} not measured for mutants). S108A exhibited similar K_D for dTMP with WT, and significantly higher K_D for both dGMP and dIMP. S108T exhibited similar K_D for dTMP and dIMP, but significantly higher K_D^{dGMP}.

4.3.2 NMR Titration Experiments

NMR titration experiments were conducted to further evaluate the K_D measurements. Series of two-dimensional HMQC methyl carbon spectra were obtained for *Pf*TMK under incremental titration of dTMP (0 to 1 mM, protein was pre-saturated with ADP), or dGMP (0 to 1 mM, protein was pre-saturated with ADP), or ADP (0 to 2 mM, protein was pre-saturated with dTMP) with saturated concentration of dTMP. NMR peaks that exhibited chemical shift perturbation upon substrate titration were used to provide an estimate of K_D values.

We observed that I11, I51, I54, I55, I85, I176, I191, and I198 exhibited significant chemical shift perturbation (CSP) upon dTMP titration (Figure 4.3). For dGMP titration. CSP is observed for I11, I54/55, I85, I176, I191, and I198 (Figure 4.4). In case of ADP, CSP is observed for I51, I152, I176, I184, I187, and I191 (Figure 4.5). The spectra were used as input to TITAN software (MATLAB) and the K_D estimations for dTMP, dGMP and ADP was measured to be to be 75 µM, 1.428 mM and 166 µM for dTMP, dGMP and ADP, respectively.

We observed significant discrepancy between our K_D^{dTMP} and K_D^{dGMP} data collected in NMR titration experiments and those from tryptophan fluorescence titration assay, this is due to the exchange process that occurred in the unliganded enzyme, which makes it difficult to extract K_D^{dTMP} and K_D^{dGMP} using CSP. On the contrary, the K_D^{ADP} estimation using CSP is reliable, as reflected by similar measured K_D values between tryptophan

fluorescence titration and NMR titration assay, because the NMR titration experiments were carried out in saturating concentration of dTMP which quenched chemical exchange in the protein. In summary, dTMP/dGMP binding significantly improved spectra quality, suggesting these two substrates play a role in stabilizing the overall protein structure.

The NMR titration experiments supported the stoichiometry obtained from our ITC data for dTMP and dGMP binding. A stoichiometry prediction of 1 : 1 (dTMP : *Pf*TMK) would suggest that both monomers in a dTMP-bound *Pf*TMK dimer will have the same conformation, and this should be reflected in NMR spectrum with well-resolved and high-intensity peaks with no peak splitting. On the contrary, a stoichiometry prediction of 0.5 : 1 (dGMP : *Pf*TMK) should result in the splitting of some peaks in NMR spectrum due to the presence of two conformations in a dGMP-bound *Pf*TMK dimer, assuming the dGMP is undergoing slow exchange process.

Peak splitting in the dGMP spectra was observed for I11 and I198 (Figure 4.6); this observation strongly suggests the presence of two different conformations within a single dGMP-bound *Pf*TMK dimer. No such splitting was seen in the dTMP saturated spectra.

The identity of the split peak was confirmed by site-directed mutagenesis, where we generated I198V mutant, obtained 2D HMQC spectrum of the mutant under saturating concentration of dGMP, and compared it with that of dGMP-saturated wild-type *Pf*TMK (Figure 4.7); we observed that I198 is indeed represented by these two peaks, which suggests the presence of two conformations. We also obtained 2D spectrum of the I198V mutant under saturating concentration of dTMP and compared it to the dTMP-

saturated WT spectrum (Figure 4.8); we observed that I198 is only reflected as one peak in the dTMP-*Pf*TMK complex.

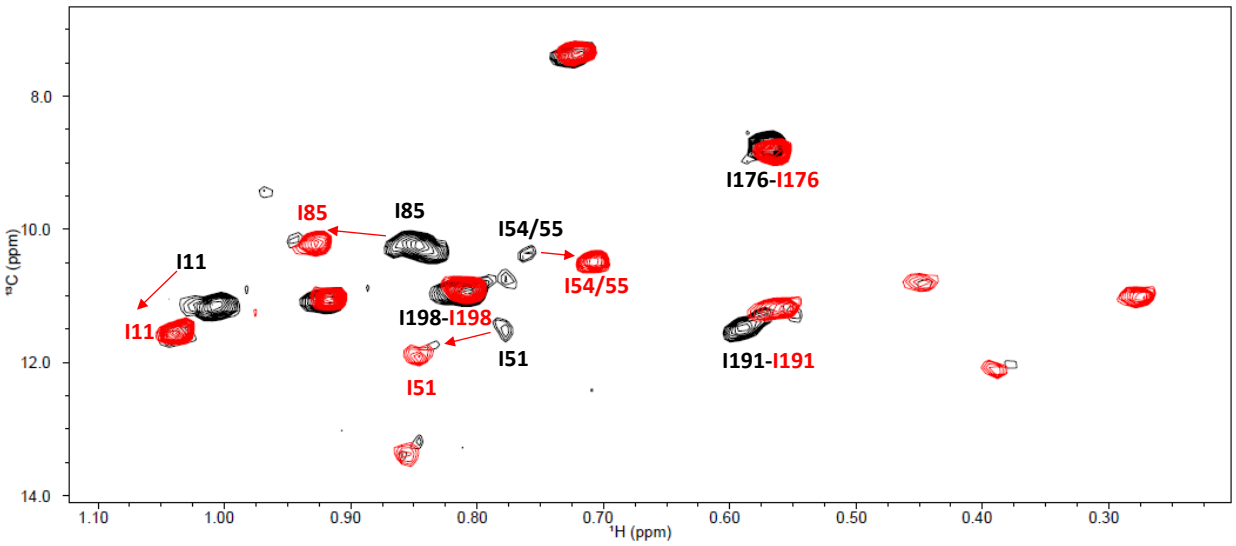


Figure 4.3 . Overlay of 2D HMQC spectrum of WT-*Pf*TMK with no dTMP (black) and spectrum of WT-*Pf*TMK with 1 mM dTMP (red). Only the isoleucine region is depicted here, peak position with non-dTMP condition is denoted in black, peak position wtih 1 mM dTMP is denoted in red. Chemical shift perturbation and peak movement is indicated by red arrows.

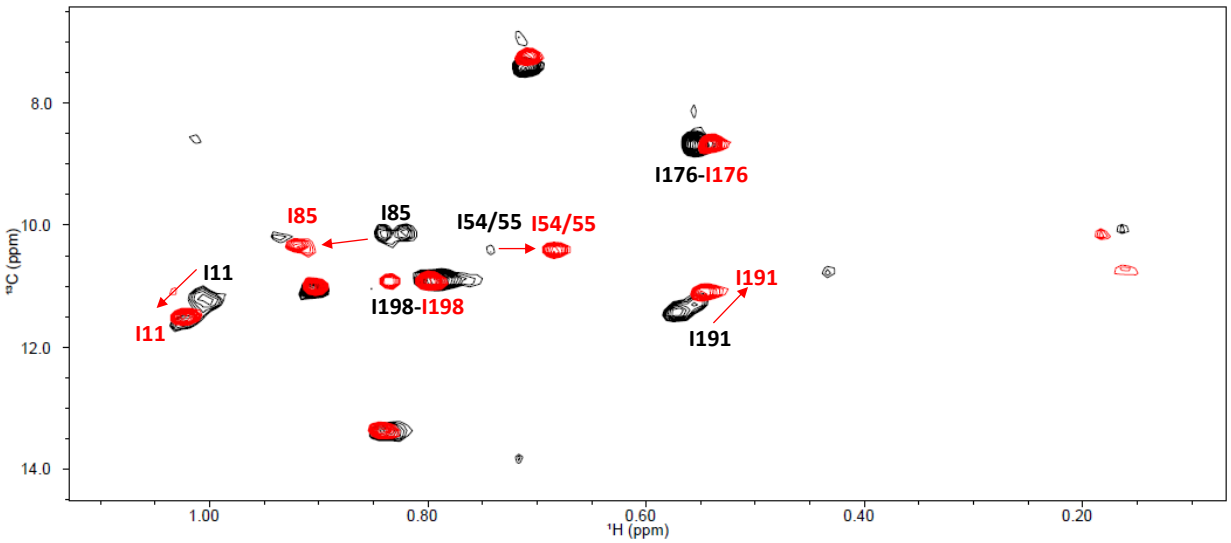


Figure 4.4 . Overlay of 2D HMQC spectrum of WT-*Pf*TMK with no dGMP (black) and spectrum of WT-*Pf*TMK with 1 mM dGMP (red). Only isoleucine region is depicted here, peak position with non-dGMP condition is denoted in black, peak position with 1 mM dGMP is denoted in red. Chemical shift perturbation and peak movement is indicated by red arrows.

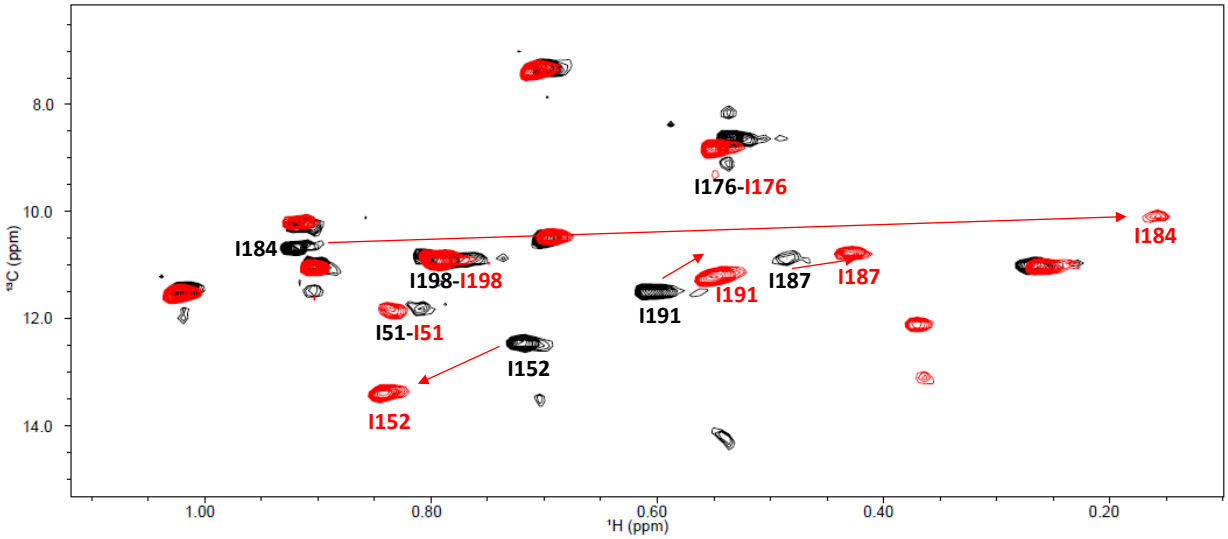


Figure 4.5 . Overlay of 2D HMQC spectrum of WT-*Pf*TMK with no ADP (black) and spectrum of WT-*Pf*TMK with 10 mM ADP (red). Only the isoleucine region is depicted here, peak position with non-ADP condition is denoted in black, peak position the 10 mM ADP is denoted in red. Chemical shift perturbation and peak movement is indicated by red arrows. I184 exhibited a very significant chemical shift perturbation under ADP titration, this is expected as I184 hovers right above the ADP binding site.

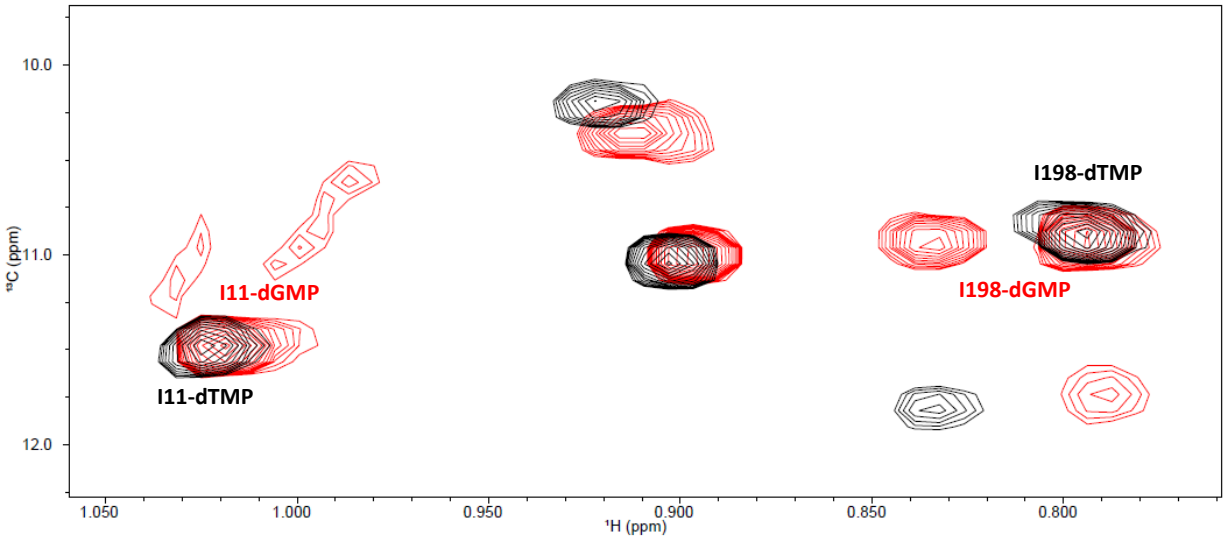


Figure 4.6 . Partial Overlay of 2D HMQC spectrum of WT-*Pf*TMK with saturating concentration of dTMP (black) and spectrum of WT-*Pf*TMK with saturating concentration of dGMP (red). dTMP-saturated PfTMK exhibits a single intense peak for residue I11 and I198, while dGMP-saturated sample exhibits peak splitting for both I11 and I198. Chemical shift difference is observed for some of the other Isoleucine residues.

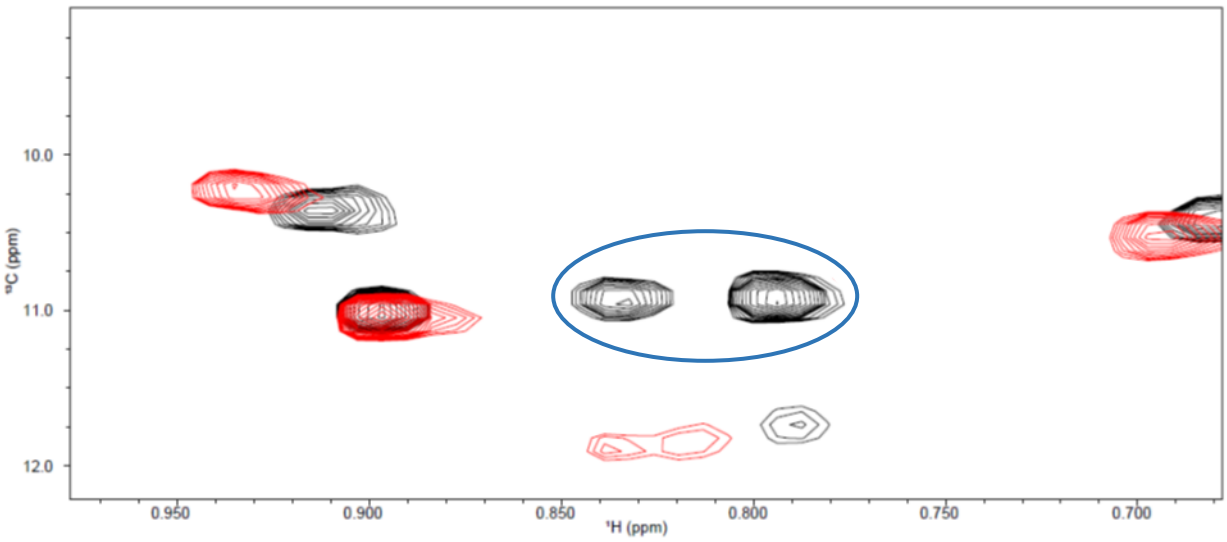


Figure 4.7 . Overlay of 2D HMQC spectrum isoleucine region of WT-*Pf*TMK (black) and spectrum of I198V mutant (red) in saturating concentration of dGMP and ADP. I198V mutant lacked two peaks in the isoleucine region (highlighted in circle), suggesting peak splitting occurs in the dGMP saturated WT PfTMK sample.

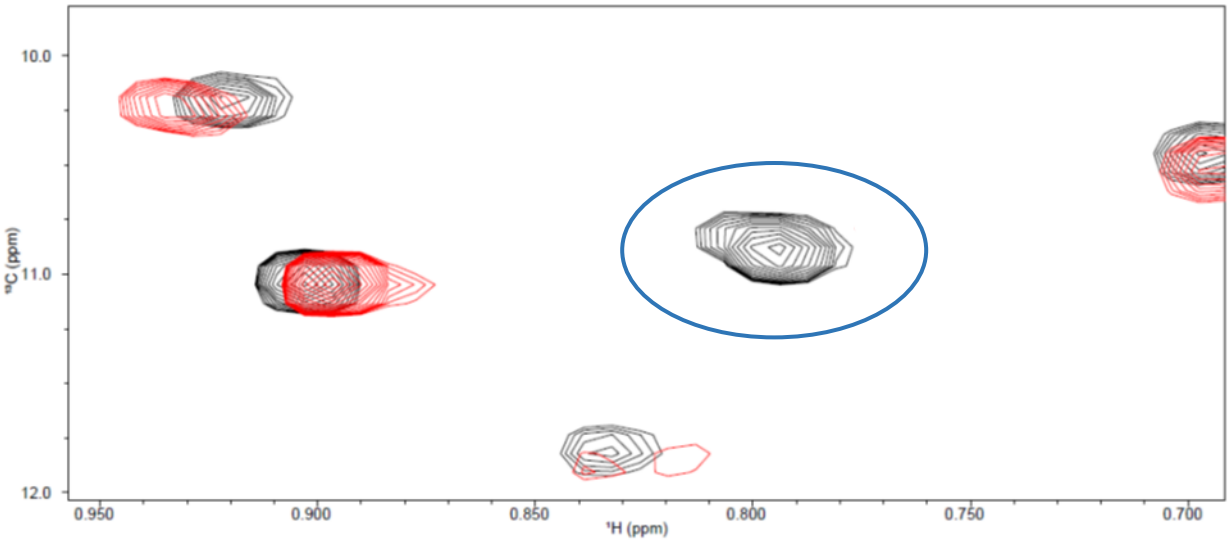


Figure 4.8 . Overlay of 2D HMQC spectrum isoleucine region of WT-*Pf*TMK (black) and spectrum of I198V mutant (red) in saturating concentration of dTMP and ADP. Note that most peaks are consistent between WT and mutant spectrum, and the I198V mutant lacked one peak in the isoleucine region (highlighted in circle).

4.4 Residues that Contribute to dGMP Binding

In a previous study by Kandeel et al.[35], several point mutations were introduced into *Pf*TMK: Y43R, Y43L, F74A, Y107F, S108T and A111K, and it was reported that the Y43R, Y43L, and S108T mutants had minor effect on both dTMP and dGMP binding, while the F74A mutant completely abolished substrate binding, and A111K significantly impacted both dTMP and dGMP binding. In a later study, Whittingham et al.[25] discussed that the effect on substrate binding by the A111K mutant could be the result of steric clash with Y153, which is predicted to interact with the base and the sugar of dTMP / dGMP. Nonetheless, none of the residues studied by Kandeel appear to play a role in the utilization of dGMP as a substrate by *Pf*TMK.

To further investigate what residues are contributing to utilization of dGMP as a substrate for *Pf*TMK, a number of point mutations were introduced into *Pf*TMK. A total of three mutants were generated: S108A, S108T, and Y153F. S108 and Y153F are both positioned in the dTMP / dGMP binding pocket of *Pf*TMK (Figure 4.9), as based on the reported crystal structure[25], it is speculated that S108 residue interacts directly with the guanine via a hydrogen bond to N1 or via a bridging water molecule with N3 of thymidine, and the -OH of Y153F forms a hydrogen bond to N2 of guanine or via a bridging water molecule to O2 of thymidine. The edge of the phenyl group of Y153 also

interacts with the furanose ring of bound substrate. In human thymidylate kinase, a threonine residue is in place of the S108 residue, and it is hypothesized that the extra methyl on the threonine residue prohibits the binding of dGMP due to spatial limitations.

A summary of the kinetic parameters for the three mutants in comparison of WT *Pf*TMK is presented in Table 4.1. Compared with WT, the S108A mutant exhibited significantly higher K_M (47.6 ± 1.7 µM vs. 36 ± 3.2 µM) and lower k_{cat} for dGMP (1.3 ± 0.05 s^{-1} vs. 2.4 ± 0.3 s^{-1}), but does not affect the overall enzyme kinetics towards dTMP. A similar trend is observed for the S108T mutant, in which the mutations had minimal effect on overall kinetics towards dTMP but resulted in significantly higher K_M (78.4 ± 5.7 µM) and lower k_{cat} (1.3 ± 0.03 s^{-1}) towards dGMP. These observations strongly suggest that S108 residue plays a critical role in the unique substrate specificity of *Pf*TMK. More importantly, the S108T mutant resulted in more significant impact on dGMP kinetics than the S108A mutant, this suggests that the steric clash is more influential than the absence of the OH substituent, and this partially explains why the human TMK is incapable of using dGMP as a substrate.

The Y153F mutation led to significantly higher K_M and lower k_{cat} for both dTMP and dGMP than WT. This indicates that the -OH substituent on the phenol is critical for the enzymatic activity towards dTMP and dGMP. This -OH group interacts with the N2 of the guanine moiety of dGMP by directly forming a hydrogen bond, and O2 to thymidine in dTMP through hydrogen bonding mediated by water molecule.

To further investigate the binding properties of *Pf*TMK for dTMP, dGMP, and dIMP, a series of tryptophan fluorescence titration assay was performed, and the K_D comparison are given in Table 4.3. All mutants shared showed similar K_D values towards dTMP, with

the exception for Y153F (68.2 ± 7.5 μM vs. 22.8 ± 1.9 μM), suggesting that only the Y153F mutant affects dTMP binding. In terms of dGMP binding, all three mutants led to significantly higher K_D, with the Y153F mutant exhibiting the largest increase. This observation is entirely consistent with our steady-states kinetics data that was presented in Table 4.1; in which we observed that S108A and S108T mutants only affected dGMP binding, and Y153F mutant affected both dTMP and dGMP binding. Between S108A and S108T mutant, we once again observed that S108T mutant led to more significant effect on dGMP binding, suggesting that the spatial limitation by the extra methyl group is more prominent than the absence of terminal -OH substituent.

The binding properties of WT and mutant *Pf*TMK for dIMP was examined through tryptophan fluorescence titration assay to determine whether the observed impact on dGMP binding is solely dependent on the interaction with the primary amine substituent on dGMP. S108A and Y153F mutants exhibited slightly higher K_D^{dIMP} compared with WT, suggesting that the interaction between S108/Y153F and dGMP is predominately through the primary amine group of guanine. Interestingly, the K_D^{dIMP} for the S108T mutants was similar to the wild-type protein, suggesting that the steric clash caused by S108T mutant was compensated for by the absence of amine in dIMP.

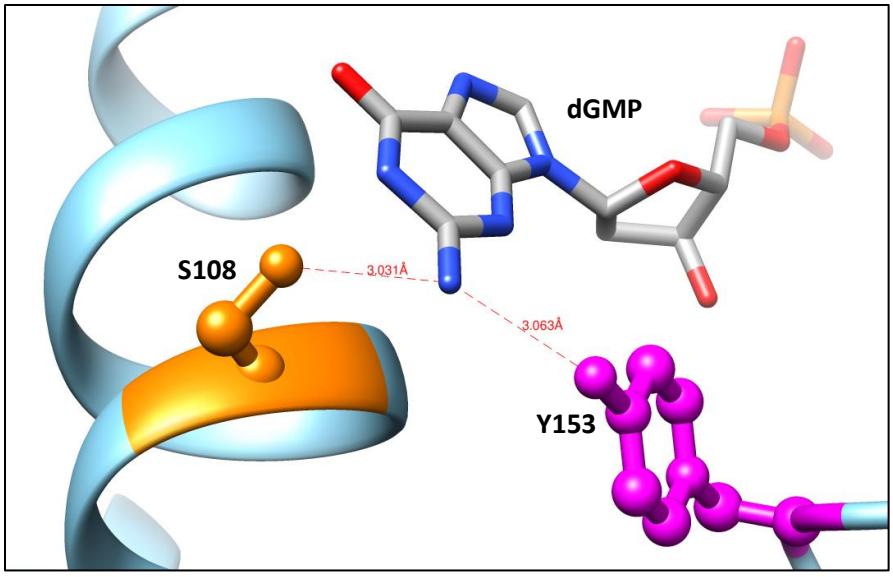


Figure 4.9 . Active site of *Pf*TMK-dGMP complex. Interaction of dGMP with S108 and Y153 residues.

4.5 Dynamic differences between dTMP and dGMP binding

NMR is a powerful tool to measure protein dynamics, which helps studying protein motions and conformation changes from ~s timescale (domain motion) to ~ps-ns timescale (molecular tumbling, loop motion and side-chain rotation)[41]. A variety of NMR methods have been developed to study the dynamic processes of protein based on the range of their timescales, such as Real-time (RT) NMR[42] (τ_{ex} > 1 s; k_{ex} < 1/s), EXchange SpectroscopY[43,44] (EXSY; τ_{ex} ≈ 10 - 5000 ms; k_{ex} ≈ 0.2 - 100; slow exchange $k_{ex} \ll \Delta\nu$), Lineshape analysis[45] (τ_{ex} ≈10 - 100 ms; k_{ex} ≈ 10 - 100/s; slow-intermediate exchange $k_{ex} \leq \Delta\nu$), Carr–Purcell Meiboom–Gill Relaxation Dispersion[46-51] (CPMG RD; τ_{ex} ≈ 0.3 - 10 ms; k_{ex} ≈ 100 – 3000/s; intermediate-fast exchange $k_{ex} \approx \Delta\nu$); Rotating Frame Relaxation Dispersion[41] (RF RD; τ_{ex} ≈ 20 - 100 μs; k_{ex} ≈ 10,000 - 50,000/s; intermediate-fast exchange $k_{ex} \geq \Delta\nu$); Nuclear Spin Relaxation[52-54] (NSR; ps - ns dynamics indirectly probed), Residual Dipolar Coupling[55,56] (RDC; ps - ms dynamics indirectly probed), and Paramagnetic Relaxation Enhancement[57] (PRE, τex ≈ 10 μs; k_{ex} ≈ 100,000 /s; fast exchange $k_{ex} \gg \Delta\nu$)).

Our goal is to elaborate on the dynamic differences between dTMP and dGMP binding to *Pf*TMK. *Pf*TMK has the unique substrate specificity of accepting dGMP as a substrate, and due to the relatively larger size of purine ring compared with pyrimidine

ring of dTMP, we speculate that dTMP-*Pf*TMK and dGMP-complex must exhibit differences in dynamics; more importantly, our ITC data suggested a unique half-site binding mechanism of dGMP binding, and this feature should be reflected in dynamic differences as well. Since we are interested in the dynamics and motion of the substrate-*Pf*TMK complex during binding/release or catalysis in the ms timescale, CPMG relaxation dispersion is the perfect tool for our experiment.

In CPMG experiments, we are looking at the exchange process in the intermediate-fast regime, in which we can observe an enhanced relaxation rate $R_{2,eff} = R_2^0 + R_{ex}$ and a broadened linewidth $\lambda=2R_{2,eff}$. The idea of the CPMG experiment is to refocus this chemical shift broadening by applying a series “spin-echo” pulse to transverse magnetization during the relaxation delay. During our CPMG experiments, we obtained 16 ^{1}H-^{13}C NMR spectra containing a fixed relaxation time of 40 ms. Each experiment contained a different number of “spin-echo” pulses (0, 1, 4, 6, 8, 10, 12, 16, 20, 24, 28, 32, 36, 40, with 8 and 16 having repeats for error estimation). We plot measured $R_{2,eff}$ values against CPMG frequency (ν_{CPMG}), and the two-dimensional data was fitted to fast-exchange model to acquire dynamic parameters[46,58].

A large number of ILV residues exhibited significant differences in effective transverse relaxation rate (R_{ex}) in the dTMP-ADP complex compared to that the dGMP-ADP complex (Table 4.4, Figure 4.10). R_{ex} is a complicated parameter that involves rate of interconversion between states (k_{ex}), and Φex, which is a parameter that combines the relative populations in each state (p_a, p_b) and the resonance frequency difference between states ($\Delta\omega$). As a result, in our analysis of data collected from CPMG experiments, we mainly focused on k_{ex}.

We observed that many methyl side-chain residues exhibited no exchange in both dTMP- and dGMP-*Pf*TMK complexes (V12, L27, V29, L42, L88, L90, I93, V132, V195, I198 and V200, colored in orange in Figure 4.10); these residues are distributed away from both dTMP / dGMP binding site and ATP binding site. We also identified some residues that show equal exchange in both dTMP- and dGMP-*Pf*TMK complexes (I11, L32, V37, I176 and I191, colored in green in Figure 4.10), these residues are positioned below the ATP binding site of *Pf*TMK.

I178 exhibited higher exchange rate in dTMP-*Pf*TMK complex than that in dGMP-*Pf*TMK complex, and it is located at relatively close distance to the bound ATP (colored in magenta in Figure 4.10). Most interestingly, we identified four residues that exhibited no exchange in dTMP-*Pf*TMK complex, but showed exchange in the dGMP-*Pf*TMK complex (I54, I55, L73 and I127, colored in cyan in Figure 4.10); these residues are all at the dimer interface of the *Pf*TMK dimer. I152, which sits on the LID domain of *Pf*TMK also next to the dimer interface, showed increased exchange rate in dGMP-*Pf*TMK complex.

Rather than a more localized dynamics difference within the active site of *Pf*TMK between dTMP and dGMP binding, we instead observed a global enzyme dynamics difference: exchange rate differences were observed in dTMP binding site, ATP binding site, and most importantly, at the dimer interface of *Pf*TMK. The dynamics difference observed at the dimer interface strongly suggests communication between the monomers in dGMP-*Pf*TMK complex, and this was not observed in the dTMP-*Pf*TMK complex. This unique inter-monomer communication in the dGMP-*Pf*TMK complex

supports our hypothesis, in that the binding of one dGMP molecule leads to conformational change that prohibits the binding of second dGMP.

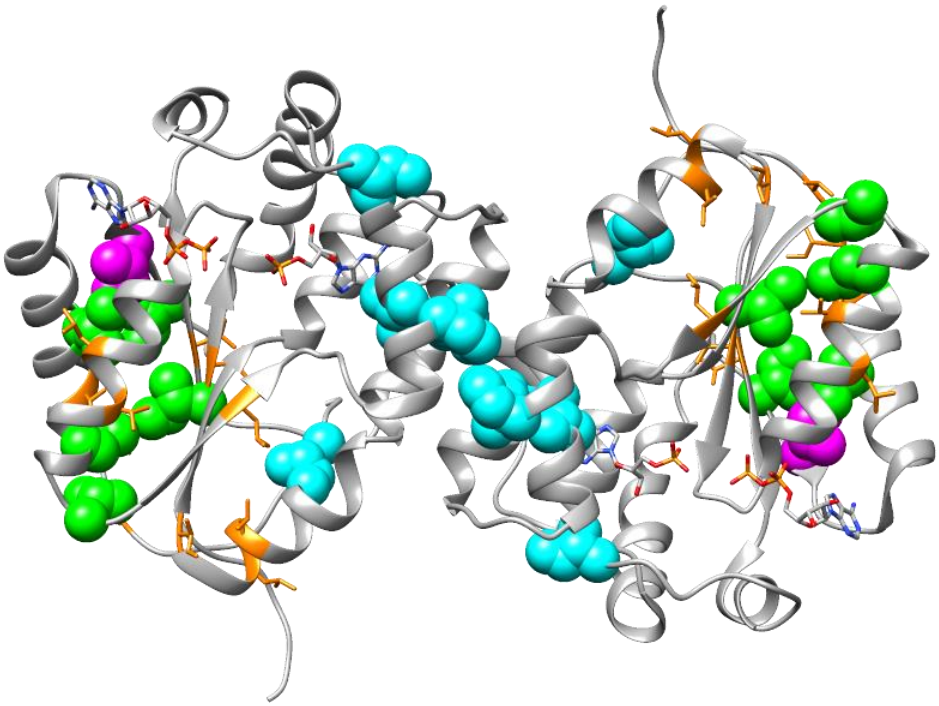


Figure 4.10. Chemical exchange in dTMP and dGMP Complexes of PfTMK. Structure of dimer with bound ADP and dTMP are shown (PDB ID 2WWF this crystal structure is used here just to illustrate the relative position of the residues of interest, not a representation of the proposed half-site binding mechanism for dGMP). ADP and dTMP are rendered in stick with carbon colored yellow. Orange indicates residues that do not show exchange, green indicates residues that show equal exchange rates for both dTMP and dGMP, cyan indicates faster exchange for dGMP and magenta indicates faster exchange for dTMP.

Residue	K_{ex}^{dTMP} (s^{-1})	K_{ex}^{dTMP} Range (s^{-1})	ϕ_{ex}^{dTMP} ($rads^{-2}s^{-2}$)	ϕ_{ex}^{dTMP} Range	K_{ex}^{dGMP} (s^{-1})	K_{ex}^{dGMP} Range (s^{-1})	ϕ_{ex}^{dGMP} ($rads^{-2}s^{-2}$)	ϕ_{ex}^{dTMP} Range
I11	275	160-420	3060	2646,3500	264	150-440	2000	1717,2677
V12	N.E.	--	N.E.	--	N.E.	--	N.E.	--
L27	N.E.	--	N.E.	--	N.E.	--	N.E.	--
V29	N.E.	--	N.E.	--	N.E.	--	N.E.	--
L32	500	240-1000	2030	1180,2523	384	130-870	1543	1011,1852
V37	150	10-330	1750	13000,2190	170	10-400	1573	21500,2615
L42	N.E.	--	N.E.	--	N.E.	--	N.E.	--
I54/55	N.E.	--	N.E.	--	1440	600-2100	12420	6830,15754
L73	N.E.	--	N.E.	--	530	460-850	7273	7256,11512
L88	N.E.	--	N.E.	--	N.E.	--	N.E.	--
L90	N.E.	--	N.E.	--	N.E.	--	N.E.	--
I93	N.E.	--	N.E.	--	N.E.	--	N.E.	--
I127	N.E.	--	N.E.	--	640	520-750	11600	9600,13300
V132	N.E.	--	N.E.	--	N.E.	--	N.E.	--
I152	570	340-700	7400	5020,9300	1200	700-1800	13600	9300,16300
I176	484	400-700	6960	6010,7560	405	310-580	3561	2941,3510
I178	608	475-1150	6980	5723,8330	376	240-540	4745	4226,4930
I191	490	350-710	3360	2600,3640	433	270-570	6010	4230,7140
V195	N.E.	--	N.E.	--	N.E.	--	N.E.	--
I198	N.E.	--	N.E.	--	N.E.	--	N.E.	--
V200	N.E.	--	N.E.	--	N.E.	--	N.E.	--

(Table description see next page)

Table 4.4: Chemical exchange in dTMP and dGMP complex. N.E. stands for no-exchange. Table shows residues that were fit to the two site fast exchange model. Residues not shown had high chi^2 values due to high-signal to noise in the spectra. Rows with no color coding indicate no observed exchange for either dTMP or dGMP complex, these are colored red in Figure 4.10. Residues that showed equal exchange rates for both dTMP and dGMP are colored green in column 1 of the table and in Figure 4.10 Those residues who showed greater exchange rates with dGMP are indicated with cyan coloring in both the table and Figure 4.10. Residues that showed a larger exchange rate with dTMP are colored magenta. N.E. stands for no-exchange.

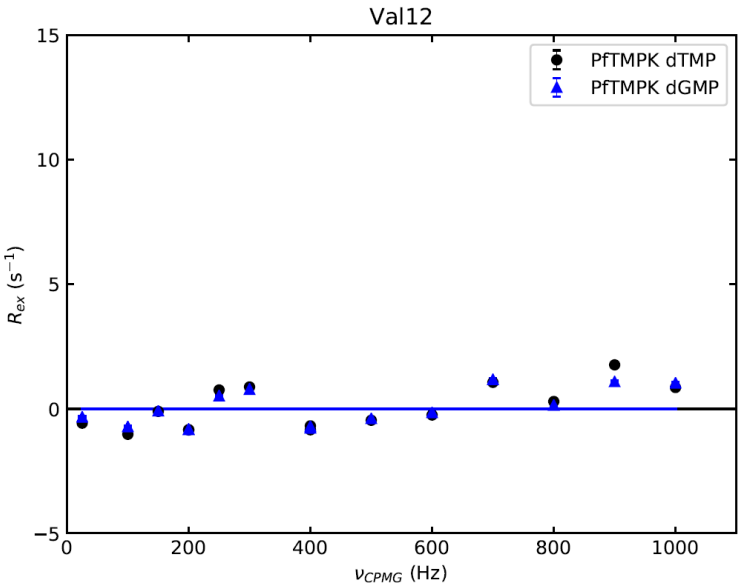


Figure 4.11 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for V12. No Exchange was observed for both complexes.

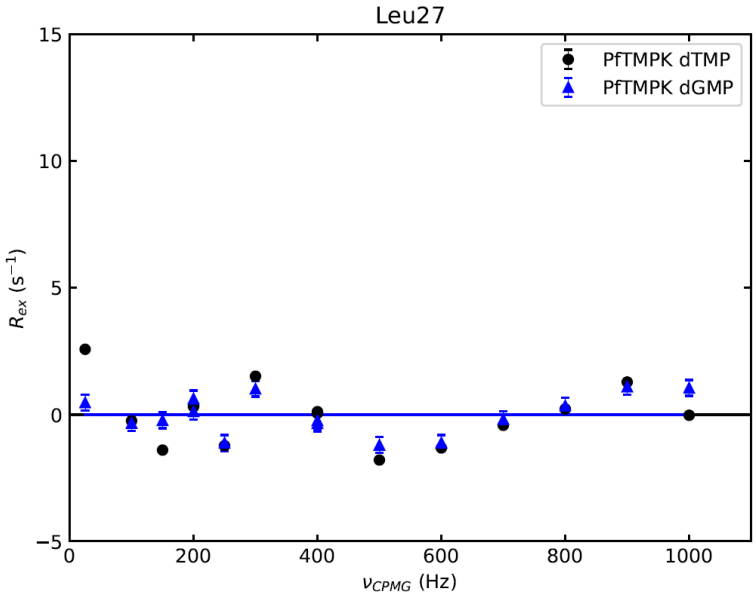


Figure 4.12 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for L27. No Exchange was observed for both complexes.

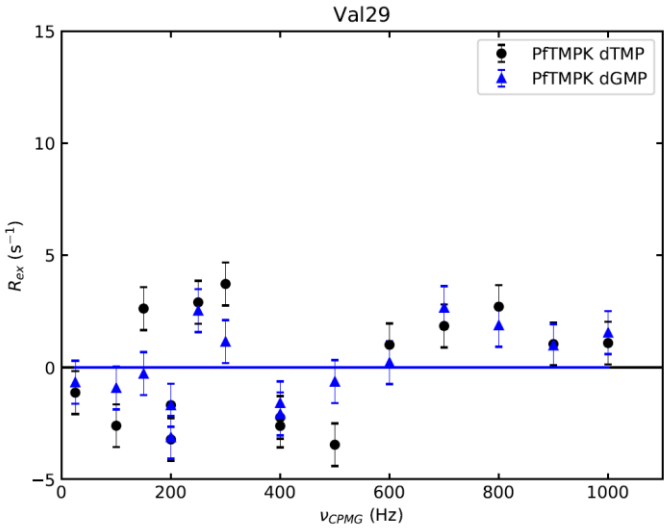


Figure 4.13 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for V29. No Exchange was observed for both complexes.

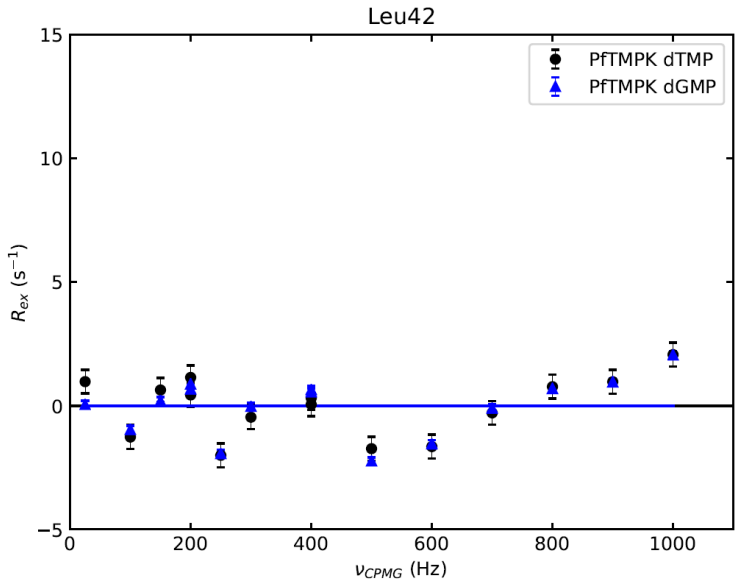


Figure 4.14 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for L42. No Exchange was observed for both complexes.

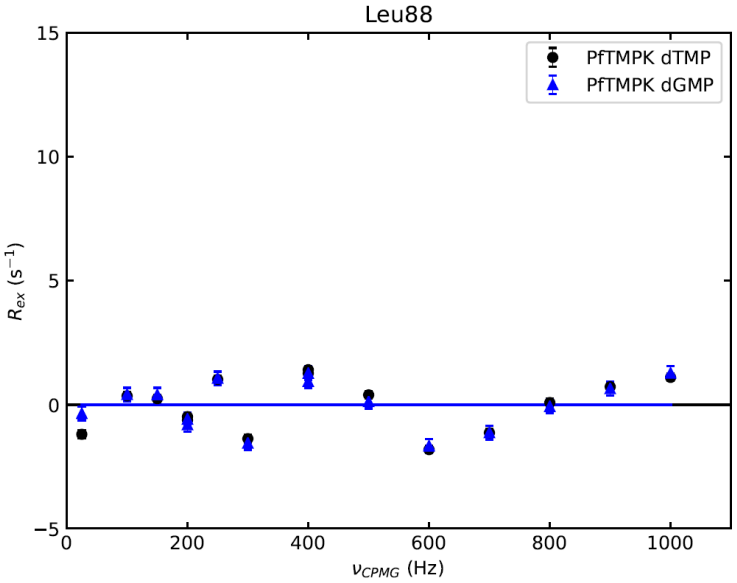


Figure 4.15 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for L88. No Exchange was observed for both complexes.

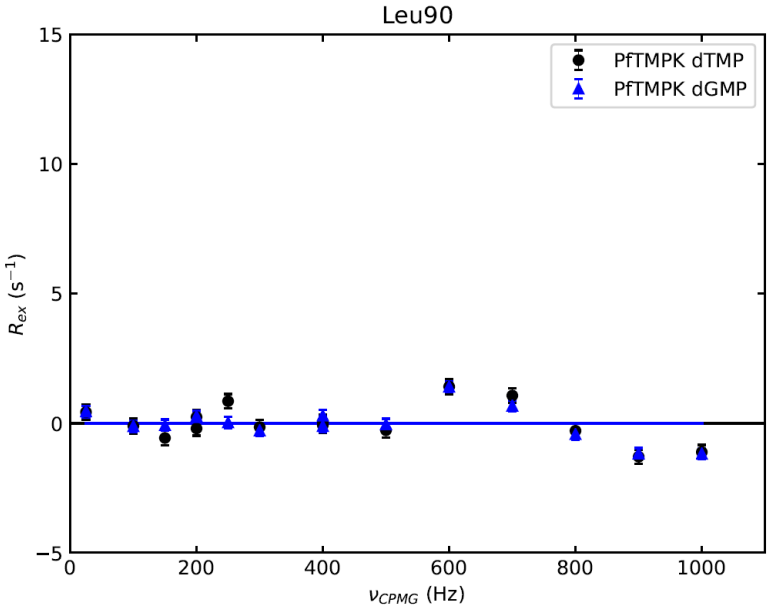


Figure 4.16 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for L90. No Exchange was observed for both complexes.

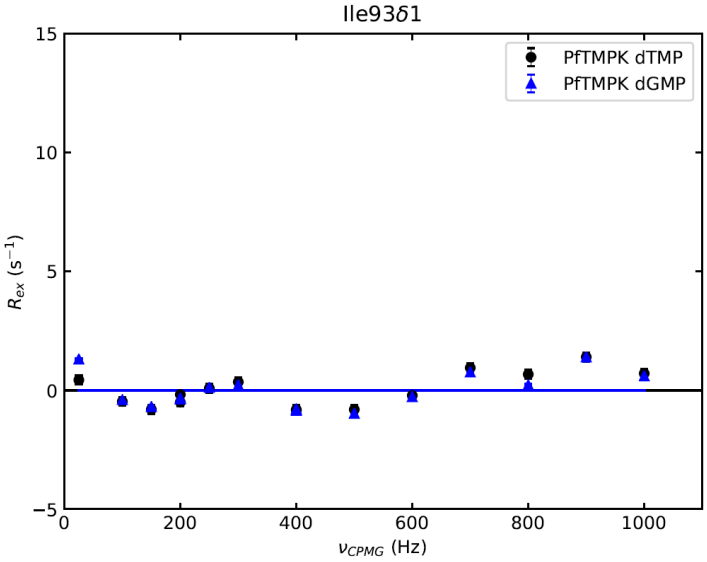


Figure 4.17 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for I93. No Exchange was observed for both complexes.

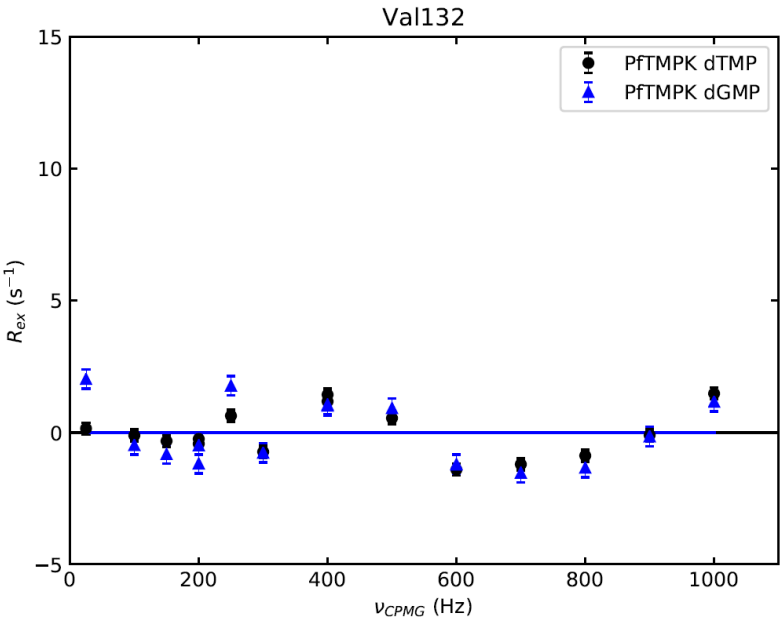


Figure 4.18 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for V132. No Exchange was observed for both complexes.

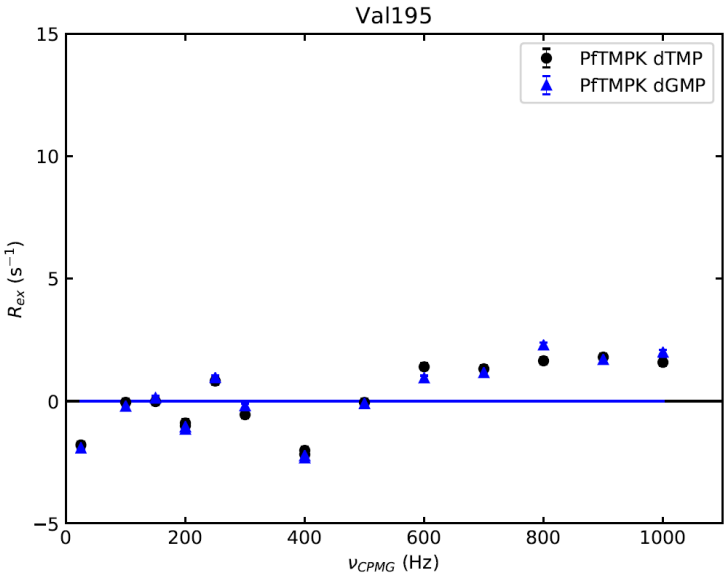


Figure 4.19 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for V195. No Exchange was observed for both complexes.

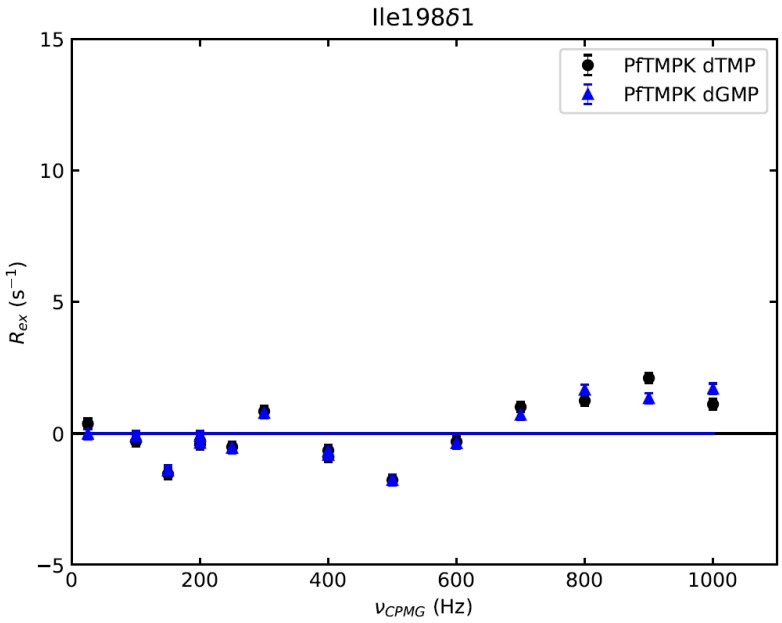


Figure 4.20 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for I198. No Exchange was observed for both complexes.

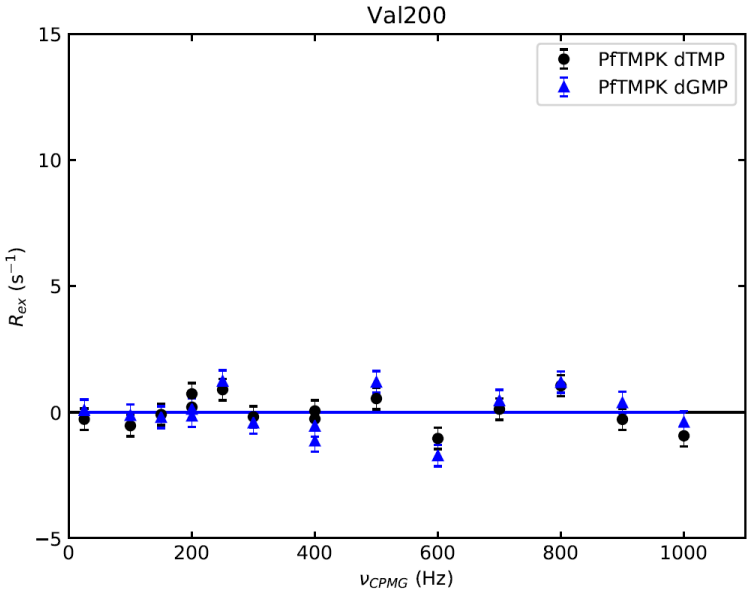


Figure 4.21 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for V200. No Exchange was observed for both complexes.

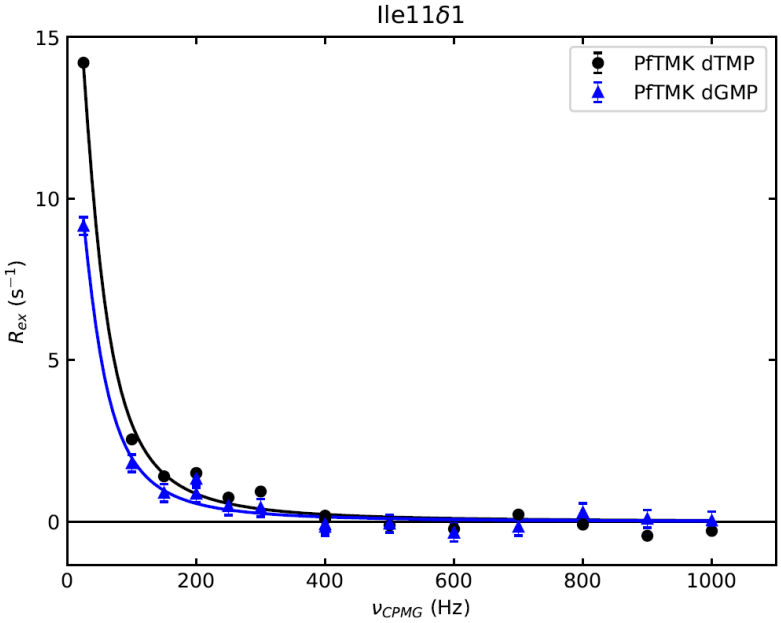


Figure 4.22 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for I11. Equal exchange was observed for dTMP- and dGMP-PfTMK complexes.

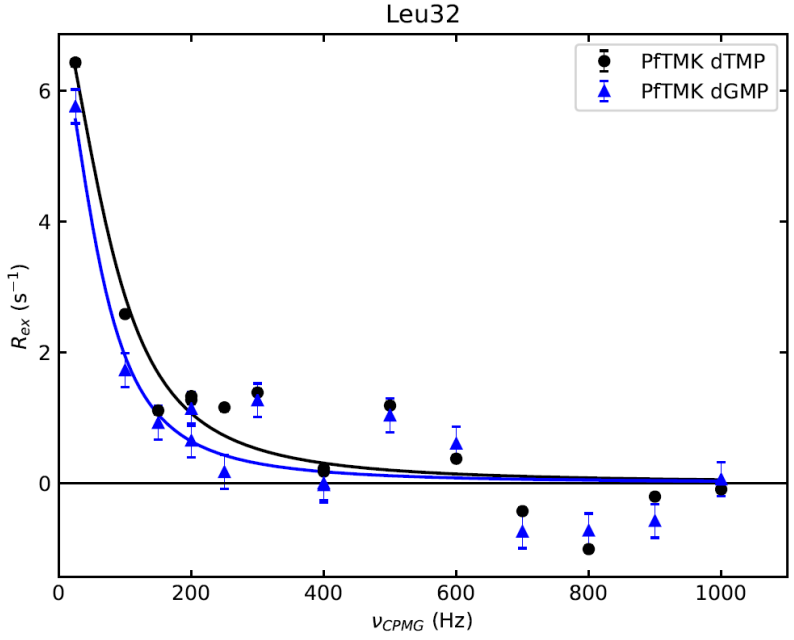


Figure 4.23 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for L32. Equal exchange was observed for dTMP- and dGMP-PfTMK complexes.

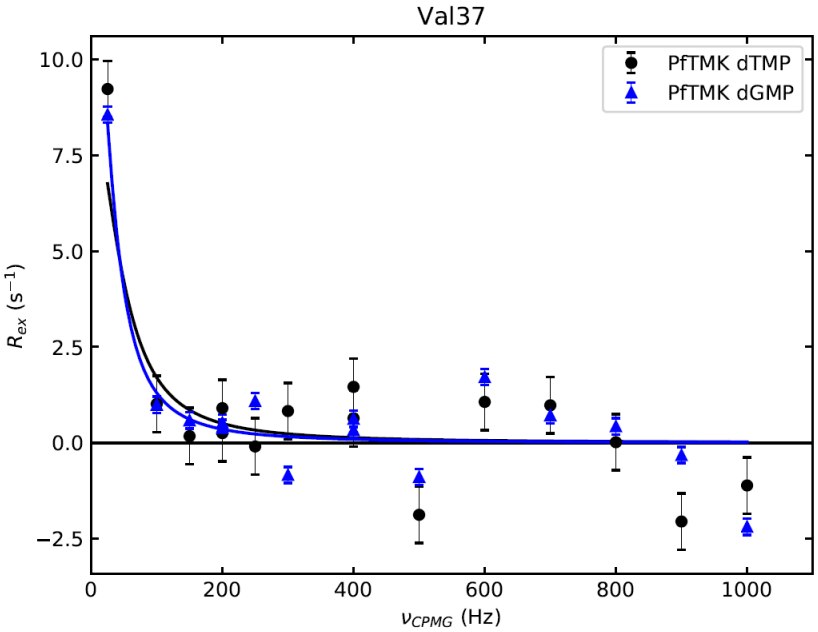


Figure 4.24 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for V37. Equal exchange was observed for dTMP- and dGMP-PfTMK complexes.

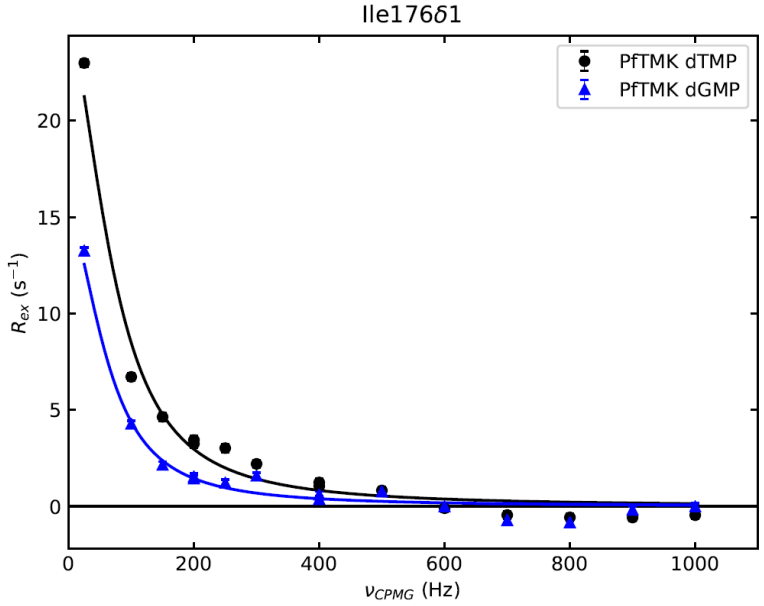


Figure 4.25 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for I176. Equal exchange was observed for dTMP- and dGMP-PfTMK complexes.

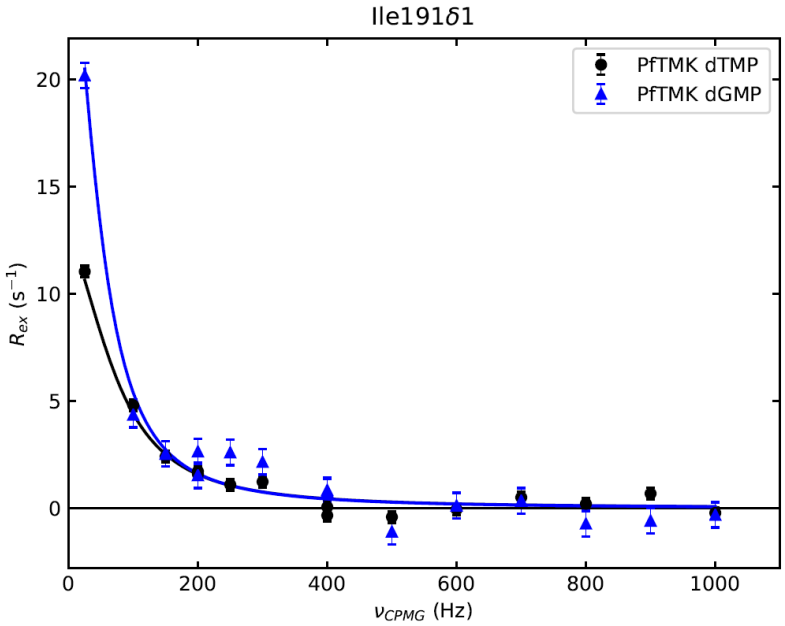


Figure 4.26 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for I191. Equal exchange was observed for dTMP- and dGMP-PfTMK complexes.

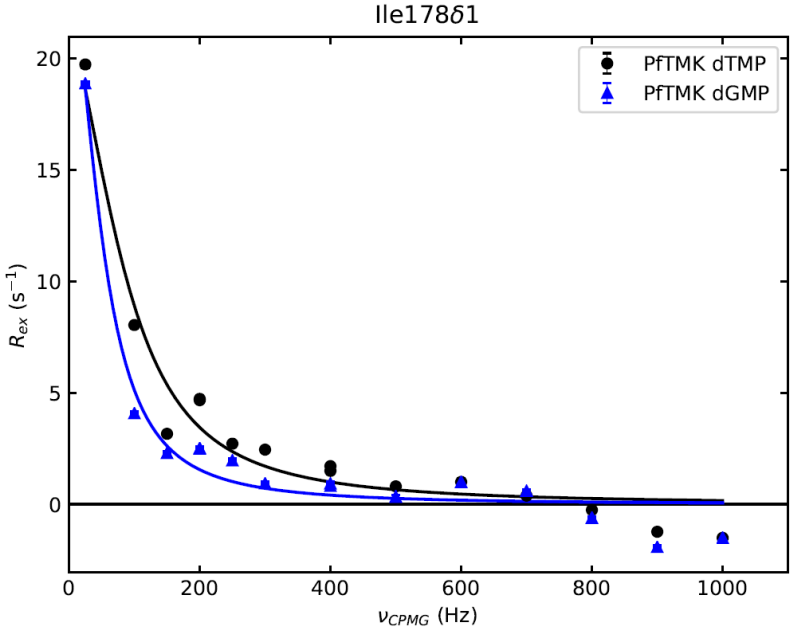


Figure 4.27 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for I178. dTMP binding showed higher k_{ex} than that in dGMP-PfTMK complex.

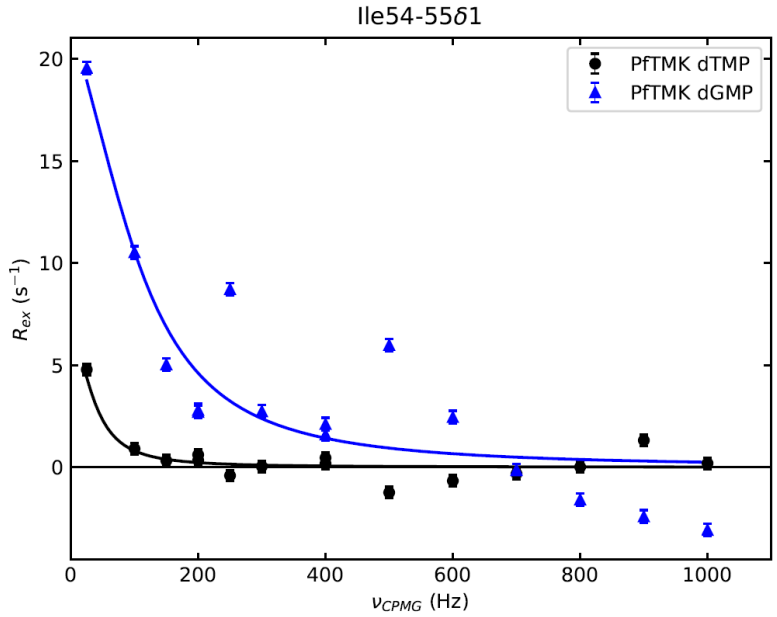


Figure 4.28 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for I54/55. I54/55 sit on the dimer interface of PfTMK, no exchange was observed in dTMP binding, while exchange occurs in dGMP binding.

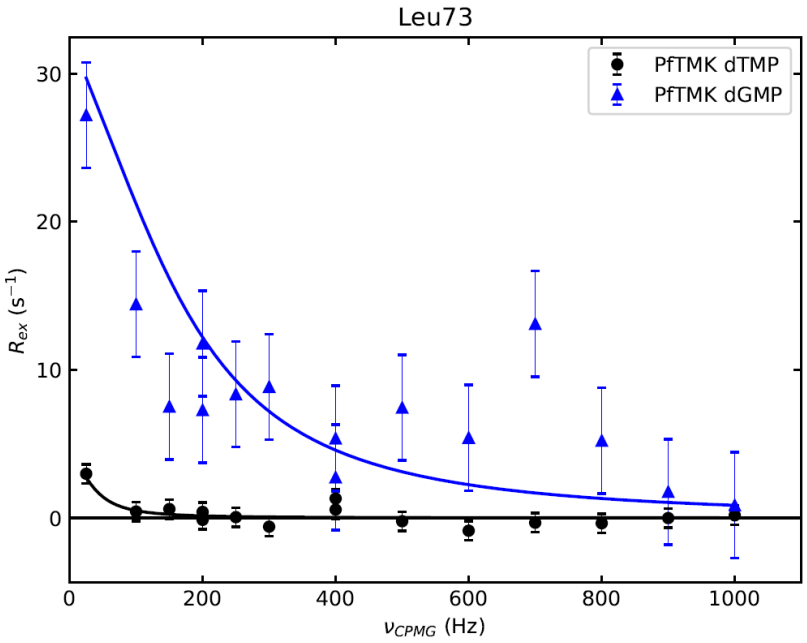


Figure 4.29 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for L73. L73 sits on the dimer interface of PfTMK, no exchange was observed in dTMP binding, while exchange occurs in dGMP binding.

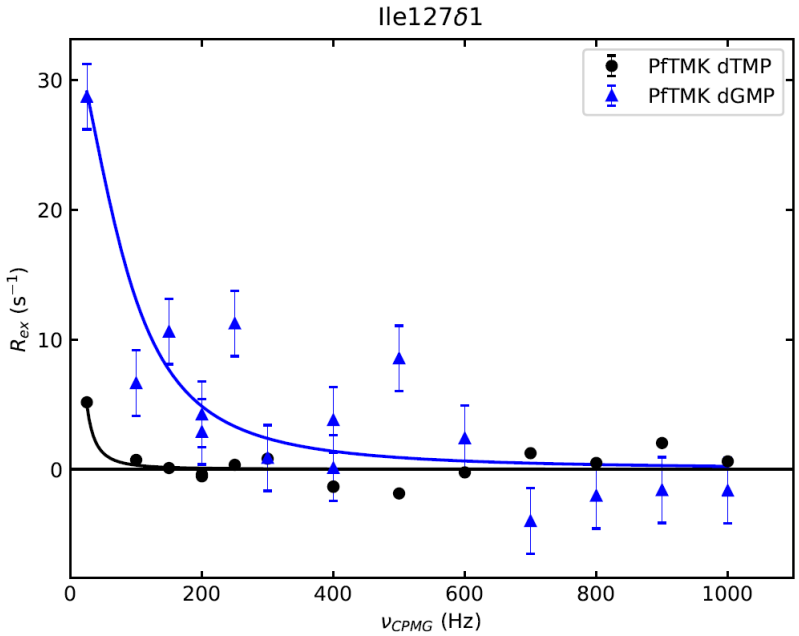


Figure 4.30 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for L127. L127 sits on the dimer interface of PfTMK, no exchange was observed in dTMP binding, while exchange occurs in dGMP binding.

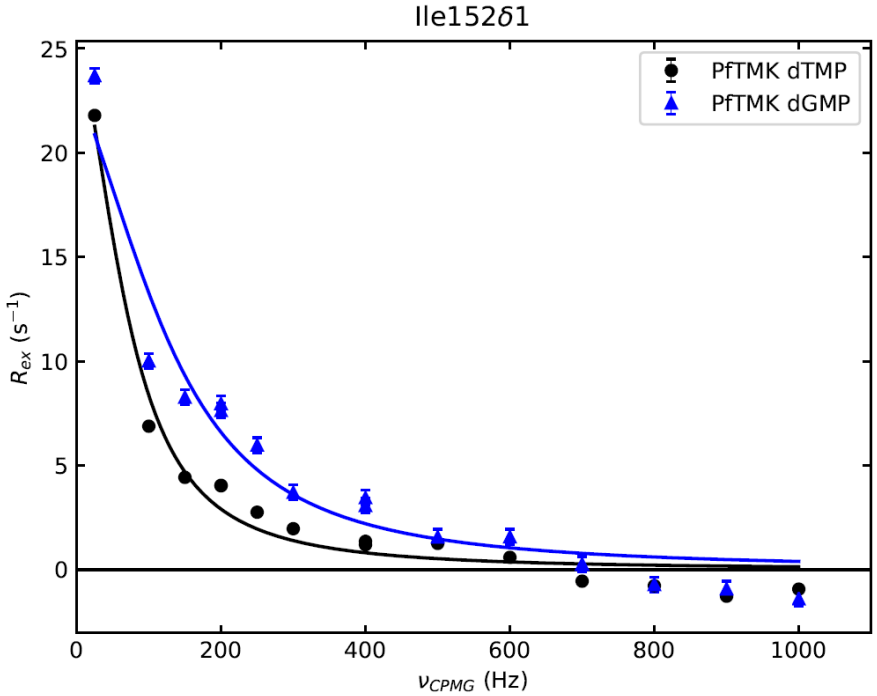


Figure 4.31 . Overlay of dTMP (black) and dGMP (blue) binding dispersion curves obtained from 800 MHz field strength for I152. L73 sits on the LID domain of PfTMK and next to dimer interface, dGMP binding exhibited significantly higher k_{ex}.

4.6 Conclusion

We have examined the kinetics, thermodynamics, and binding property of dTMP, dGMP, ADP, and ATP binding to *Pf*TMK. In addition, mutational studies and CPMG relaxation dispersion experiments were carried out to investigate which residues play a role in the unique substrate specificity of *Pf*TMK.

K_M values were obtained for dTMP, dGMP and ATP through NADH-coupled assay, and we obtained consistent K_D values for dTMP and dGMP from ITC and tryptophan fluorescence titration experiments that are comparable to the measured K_M values. dTMP and dGMP binding exhibited similar free binding energy, with dGMP binding showed much higher enthalpy and entropy changes, which can be explained by the larger size of purine ring. Our ITC data also revealed that dTMP and dGMP binding to *Pf*TMK do not share the same binding mode: dTMP binding with a 1:1 stoichiometry, while dGMP binding exhibited a unique half-site binding mode, suggesting that each *Pf*TMK homodimer is only occupied by a single dGMP molecule. This observation was supported by our NMR titration experimental data, where we observed peak splitting in the NMR spectrum obtained from dGMP-saturated sample, which was not observed in the dTMP-saturated sample.

Our CPMG relaxation dispersion experiments showed that dTMP- and dGMP-*Pf*TMK complex exhibits global difference in enzyme dynamics rather than a more localized

effect that is confined to the active site. More importantly, the fact that a significant number of residues that are located at the dimer interface exhibited different exchange rate between dTMP and dGMP binding suggest that dGMP binding leads to communication between the two monomers, which could be the result of conformational changes that lead to half-site binding.

This evidence suggests that the binding of first dGMP molecule lead to conformational change of the other monomer that prohibits binding of second dGMP molecule. Although the half-site binding mechanism was not exhibited in the previous crystal structures of dGMP-bound *Pf*TMK, in which *Pf*TMK is fully occupied by two dGMP molecules in both binding sites[25], we speculate that the formation of crystal lattice compels symmetry of the monomers, resulting in two dGMP molecules binding to the protein with reduced affinity.

Our mutational studies revealed that Y153F is involved with the binding and activity of *Pf*TMK for both dTMP and dGMP, and this agrees with the predictions based on reported crystal structure. Our data suggest that S108 residue has a more selective interaction with dGMP, and potentially contribute to the unique substrate specificity of *Pf*TMK. We also showed that substitution of the S108 residue to threonine (S108T K_D^{dGMP} = 88.5 ± 3.9 µM), which is equivalent to the human version of the enzyme, led to more significant impact on dGMP binding than removing the -OH substituent (S108A K_D^{dGMP} = 62.1 ± 4.3 µM), this partially help explaining why huTMP is unable to phosphorylate dGMP. In our K_D^{dIMP} measurements, we observed that dIMP showed significantly higher K_D for WT *Pf*TMK than that of dGMP (around 13-fold. 42.5 ± 2.2 µM vs. 561.6 ± 35.4 µM), which reflects approximately 6 KJ/mol decrease in binding energy;

while for Y153F mutant, we only observed a ~5-fold increase in K_D. This suggests that the large difference between K_D^{dIMP} and K_D^{dGMP} is a result of factors beyond the lack of the primary amine substituent, which could be differences in the base stacking effect between F74 and inosine vs. guanosine.

This unique half-site binding mode in dGMP-*Pf*TMK complex could shed some light on future design of *Pf*TMK-specific inhibitors that could potentially act as drug candidates for anti-malarial agents. Currently, there has been significant effort on designing thymidine-based *Pf*TMK inhibitors[8,9,10,11]. However, there have been limited attention on designing guanosine-based compounds targeting *Pf*TMK. A guanosine-based compound may be much more efficient in inhibiting *Pf*TMK, in that the binding of one dGMP analog could potentially act as an allosteric inhibitor that prevents any binding on the vacant monomer of a *Pf*TMK dimer, and this could potentially reduce the IC_{50} concentration of the inhibitor. Design of guanosine based *Pf*TMK specific inhibitors could consider focusing on improving upon the interaction with S108 to generate lead compounds with higher binding affinity.

Chapter 5

Summary

5.1 Summary and Conclusion

In this dissertation, we aim to address the following questions regarding known inhibitors of *Pf*TMK:

- Does compound 13b function as a competitive inhibitor against *Pf*TMK with K_i and binding affinity that is consistent with its *in-vivo* EC_{50} value?
- Can we resolve the metabolic instability of 13b while retaining its inhibition effect and binding affinity?
- Where exactly is 13b (and the synthesized derivatives) binding to *Pf*TMK, active site (competitive inhibitor) or elsewhere (allosteric inhibitor)?

We also extensively characterized the steady-state kinetics, binding properties, thermodynamic and dynamics of *Pf*TMK in hope to provide informative understanding that could guide the future development of *Pf*TMK-specific inhibitors that could potentially serve as anti-malarial drug candidate.

We were able to achieve each of these goals. We determined that although 13b exhibits impressive EC_{50} value, reflected in decent parasite growth inhibition, it is a relatively weak competitive inhibitor against *Pf*TMK. We synthesized derivatives that resolve the metabolic instability issue by substituting the chlorobenzene substituent with stable moieties like fluorobenzene and methoxybenzene; these derivatives retained the *Pf*TMK inhibition effect and binding affinity, suggesting that there would be some potential in

structurally modifying compound 13b to develop more potent inhibitors. This idea is aided by our binding site determination, which provided information on some of the compound-protein interaction that can be enhanced through structural modification. Unfortunately, because these derivatives are relatively weak inhibitors, they were not the most optimal targets for the development of drug candidates.

Through our comprehensive characterization of *Pf*TMK-substrate interaction, we obtained extensive understanding of enzyme kinetics, binding modes, binding affinity, thermodynamics, and protein dynamics. We also investigated the unique substrate specificity of *Pf*TMK, and determined a residue that potentially plays a role in conferring *Pf*TMK the unique ability of using dGMP as substrate. More importantly, we identified a unique binding mode in the *Pf*TMK-dGMP complex, in which only one dGMP molecule bind to per *Pf*TMK dimer. These studies provide information that potentially shows a more optimal direction in *Pf*TMK-specific inhibitor design, that dGMP-based analog may be more potent in *Pf*TMK inhibition and may perform better in terms of specificity. Our structural and dynamics studies also provide valuable information that guides the design of dGMP-based analog.

5.2 Future direction

Although we showed that compound 13b is relatively poor in *Pf*TMK inhibition and binding, its impressive EC50 value at 28 nM should not be overlooked. We speculate that there must be significant off-target effect that results in parasite growth inhibition, and compound 13b may target another critical enzyme in *P. falciparum* life cycle. It would be worthwhile to attempt to identify the binding target of this compound to reveal more druggable targets for future development of anti-malarial drug candidates.

One of the key elements in our study is to establish that guanine-based analogs are potentially more effective than their thymidine-based counterparts. This argument hinges on further validation of half-site binding mode of dGMP-*Pf*TMK.

In current field of *Pf*TMK inhibitor development, the majority of developed inhibitors are thymidine-based compounds. We predict that dGMP-based inhibitors should be more potent; more importantly, guanosine-based analog may perform better in binding specificity, since human version of the enzyme is unable to utilize dGMP as substrate, consequently, human TMK will be unlikely to bind to dGMP-based compounds.

5.2.1 Target identification for Cl-derivative

Compound 13b poorly binds and inhibits *Pf*TMK; nevertheless, its impressive EC50 value still suggest that it is a lead compound worth pursuing. The first step would be attempting to identify the physiological target of 13b.

One possible approach is to utilize biotin-streptavidin interaction to fish out the binding target of compound 13b. This can be accomplished by preparing Plasmodium culture and lyse the cells to harvest the lysate mixture, then expose the lysate mixture to biotinylated compound 13b, this will be followed by application of streptavidin beads to extract the binding target. The captured protein will then be identified and applied to further characterization.

One concern with this approach is where to attach the biotin labels. Although we did not generate crystal structure of *Pf*TMK in complex with compound 13b, our docking simulation data showed that the terminal chlorobenzene substituent is completely outside of the active site of the protein, and we also showed that this substituent is not required for binding or inhibition. This suggests that the terminal benzene substituent would be ideal for attaching biotin label.

Another concern is that one of the metabolites of the compound 13b consists of an epoxide. Consequently, the parasite growth inhibition could also be the result of epoxide accumulation and toxicity. This concern can simply be lifted by testing the F-derivative (compound 13c) in terms of parasite growth inhibition. Instead of the chlorobenzene

substituent, the F-derivative consists of a highly stable fluorobenzene substituent which is resistant against oxidation. If the observed EC_{50} for compound 13c is consistent with the measured K_i, that would suggest that the reported EC_{50} for the compound 13b is mostly likely a result of epoxide toxicity. On the contrary, if a compound 13c exhibited a similar EC_{50} compared to that of compound 13b, it would suggest significant off-target effect – and we could use the approach discussed in the previously section to identify the physiological target of the compound.

Alternatively, another approach would be using Drug Affinity Responsive Target Stability (DARTS) approach[60,61]. This method relies on treating aliquots of cell lysates with the compound of interest, and utilizing limited digestion with protease to break down proteins in the lysate. Binding of the compound will protect the bound protein from proteolysis, as a result, the bound protein could be identified through SDS-PAGE gel; this will be followed by mass spectrometry for the identification of the target protein.

Using this approach, we could first grow up parasite culture, harvest the cells and lyse them to collect cell lysate. The lysate will then be treated with the Cl-derivative, and this will be followed by protease treatment. The protease will then be quenched, and the post-treatment lysate will be applied to SDS PAGE gel. The bands that appear to be protected by Cl-derivative binding will be subjected to gel extraction and the obtained protein will be identified through mass spectrometry.

5.2.2 Validation of dGMP Half-site Binding Mechanism

Our NMR titration data show evidence on the half-site binding mechanism of *Pf*TMK-dGMP, and we speculate that binding of first dGMP molecule leads to conformational change in the other half of the dimer and prevent further binding, it would be worthwhile to further validate this observation by attempting to examine whether two distinct conformations exist in dGMP-bound *Pf*TMK dimer.

One possible approach would be running stopped-flow experiment coupling with time-resolved small-angle X-ray scattering to compare the structural changes of *Pf*TMK during dTMP or dGMP binding. This idea is inspired by the work by West et al[62], in which they monitored the conformational changes of ATCase from *E. coli* upon the addition of substrates.

5.2.3 Insights on Future Inhibitor Design

The vast majority of developed *Pf*TMK inhibitors are dTMP-based analogs, and only one guanosine-based inhibitor, a GZT-derivative, was reported by Cui et al. with a low EC_{50} value of 66 μM[59]. However, it is only characterized *in-vivo*, so it's difficult to determine whether it is a competitive inhibitor against *Pf*TMK. Consequently, this compound may not be the optimal lead compound to pursue. Alternatively, another starting point would be using the Kato et al[9] or Noguchi et al.[10] compound as scaffold, and replacing the thymidine component with guanine.

From our structure-activity relationship studies, we suggest that structural modifications could be made towards improving upon interaction with S108 (and possibly I152). The primary amine group of the dGMP-analog could be replaced with a more electronegative hydrogen bond accepter like fluorine to enhance its interaction with S108; alternatively, substitution to a bulkier hydrophobic group could potentially enhance hydrophobic interaction with I152 and make the active site more compact. α-guanosine-based analog should also be considered to examine if α-variants perform better than the β-counterparts like the thymidine-based compounds.

5.2.4 Fragment Library Screening Using NMR

One of the applications of NMR that has been overlooked is its usage in high throughput screening. Methyl side chain residues are mostly positioned in the core of the protein and are highly sensitive to binding events and dynamic changes due to binding events[32,33]. From our NMR titration data, we have shown that binding of substrate leads to significant chemical shift perturbations (Figure 4.3-4.5); the CSPs are not only highly informative in illustrating whether a small molecule or fragment is binding to the protein of interest, it also serves to directly probe the binding site of the bound molecule or fragment (with appropriate methyl side chain residue NMR assignments).

In a previous study by Scheich et al. [31], promising leads against thymidylate kinase from *M. tuberculosis* were identified through high throughput screening of fragment library using NMR, and the identified leads greatly supported them in the development of inhibitors against the enzyme. Henceforth, high throughput screening of fragment library using NMR will be a promising approach for us to identify leads for *Pf*TMK inhibition. Since we have already acquired a decent coverage of the methyl residue resonance assignment of *Pf*TMK, the detected fragment binding based on chemical shift perturbation can be readily distinguished between competitive or allosteric binding. The competitive-binding fragments could provide information on what specific substituents or moieties are contributing to binding and this could be applied to structural modification of existing inhibitors to improve binding affinity; the allosteric-binding fragments could

potential help in the development of allosteric inhibitors against *Pf*TMK as a new class of anti-malaria drug candidates.

We can start by screening a commercial fragment library such as the Ultimate Fragment Library (4,200 compounds) from Life Chemicals. The general procedure will be similar to all of our 2D HMQC experiment. NMR samples will be prepared in NMR buffer (1 mM HEPES, 200 mM NaCl, 25 µM EDTA, 0.02% sodium azide in D_2O) containing 10mM ADP, 10mM Mg2+, and 20 µM dTMP (slightly lower than K_M^{dTMP}). We will first acquire the non-fragment 2D spectrum, then a cocktail of 10~15 fragments will be introduced into the NMR sample, and we will examine the difference between the fragment-treated spectrum with the non-fragment spectrum to identify if chemical shift perturbation occurred due to binding. This process will be repeated for the screening of the whole library, and we could they attempt to identify the binders by reducing the number of fragments in the cocktails that led to chemical shift perturbation.

Chapter 6

Materials and Methods

6.1 Preparation of Wild-type and Isotopically Labeled PfTMK

A *Pf*TMK expression sequence that was optimized for *Escherichia coli* expression was prepared and obtained from DNA 2.0. The amino terminus encodes a His_6 tag for our subsequent purification process involving Hispur cobalt resin. The expression sequence was cloned into T7 expression plasmid pET-22b[63], and subsequently transformed into C3013 cell line. The transformant was selected by 50mg/ml ampicillin. For non-isotopically labeled *Pf*TMK production, a starter culture was prepared by inoculating a colony into 5 mL of Luria broth (LB). The starter culture was grown overnight, and was then transferred to a larger volume of LB, and was incubated at 37 ºC until the optical density at 600 nm (OD_{600}) reached 0.7. 1 mM isopropyl thiogalactoside (IPTG) was then added and the culture was induced for 3 h.

For the preparation of isotopically labeled *Pf*TMK, a starter culture was prepared by inoculating a colony into 1 mL of Studier's PG medium[64] with 50% H_2O and 50% D_2O containing 10% glucose, 2 mM Mg^{2+}, and trace metals. The starter culture was grown overnight. The D_2O level was raised to 75% the next day by adding 1 mL of D_2O Studier's PG medium (3 h incubation), and then raised to 90% by adding another 3 mL of D_2O Studier's PG medium (3 h incubation). The 90% D_2O starter culture was then transferred to a larger volume of 100% D_2O Studier's PG medium. The culture was then

grown overnight at 30 ºC until OD_{600} reached 1.4, and then precursors, ketobutyric acid-methyl 13C (50 mg/L) and α-ketoisovaleric acid (100 mg/L)[65], was added. The culture was then incubated at 30 °C for 1 h and induced with 1 mM IPTG for 18-20 h.

6.2 Protein Purification

The purification process was the same for both labeled and non-labeled *Pf*TMK. The induced cells were harvested by centrifugation and resuspended in 20 mL of wash buffer (50 mM potassium phosphate, 250 mM NaCl, 5 mM imidazole, 0.02% sodium azide, and 10% glycerol (pH 7.4)) containing 0.1 mM dTMP and 10 mM $MgSO_4$, this was followed by the addition of lysozyme (1 mg/mL) and incubation for 20 min on ice. This was followed by the addition of 1% Triton-X and sonication with a Fisher Scientific 550 sonicator. The lysate was subjected to centrifugation (19000g for 30 min), and the supernatant was collected and incubated with 1 mL of Hispur cobalt resin (Thermo Scientific) for 1 h at 4 °C. The supernatant/resin mixture was loaded into a gravity column, and was then washed with 10 column volumes of wash buffer. Lastly, the protein was eluted with 6 column volumes of elution buffer (50 mM sodium phosphate, 300 mM NaCl, and 450 mM imidazole (pH 7.4)). Buffer exchange was performed with an Amicon Ultra concentrator to remove the high concentration of imidazole in the sample. The protein sample was then loaded onto a Q-Sepharose column that was pre-equilibrated with column buffer consisting of 20 mM potassium phosphate, 0.02% sodium azide, and 1 mM EDTA (pH 8.0). The protein was then eluted with a linear salt gradient from 0 to 2 M NaCl in column buffer, and the eluate was collected on a fraction collector. The purified protein was then concentrated and stored at 4 °C, and typical yields are 50 mg of *Pf*TMK from 1 L of culture.

6.3 Steady-state Kinetics Assay

The NADH-coupled assay[66] was performed to measure the K_M of *Pf*TMK for dTMP, dGMP and ATP. In the measurement of K_M of *Pf*TMK for dTMP and dGMP, 1 mL kinase reaction mixtures was prepared containing 50 mM Tris-HCl, 50 mM KCl, 0.5 mM phosphoenolpyruvate, 1 mM ATP, 20 mM Mg^{2+}, 5 units/mL lactate dehydrogenase, 5 units/mL pyruvate kinase, 0.15 mM NADH, and a range of concentrations of dTMP or dGMP from 2 to 150 μM. The reaction was initiated by the addition of 150 nM *Pf*TMK, and the absorbance at 340 nm (A340) was measured as a function of time. For the measurement of K_M of *Pf*TMK for ATP, 1 mL kinase reaction mixtures was prepared containing 50 mM Tris-HCl, 50 mM KCl, 0.5 mM phosphoenolpyruvate, 1 mM dTMP, 20 mM Mg^{2+}, 5 units/mL lactate dehydrogenase, 5 units/mL pyruvate kinase, 0.15 mM NADH, and a range of concentrations of ATP from 2 to 150 μM. Again, the reaction was initiated by adding 150 nM of *Pf*TMK, and the A340 was measured overtime.

The K_M for each of the substrate was estimated by deducing the initial velocity at different concentrations of substrates from the A340 measurements, and the calculated initial velocities were then fitted to the equation $V = V_{max}[Substrate]/(K_M + [Substrate])$ using the Solver routine in Excel to calculate the K_M values.

The procedure for the estimation of the Ki values for Cl-, F-, and OMe-derivatives was very similar to the process above, except the synthesized compounds (100mM, 200mM

or 300mM) were added to the reaction mixture prior to the addition of enzyme. The reciprocal of the initial rates (1/v) were plot against the concentration of the synthesized compounds in a Dixon plot for the estimation of K_i values.

6.4 Tryptophan Fluorescence Titration Assay

Tryptophan fluorescence titration experiments were conducted to determine the binding properties of the inhibitors and substrates to *Pf*TMK. The general experimental condition for the Tryptophan titration assay is as follows: a reaction mixture was prepared with 300 nM *Pf*TMK, and 20mM Mg2+ in Tris buffer pH 7.4. 50 µl of the reaction mixture was pipetted into the wells on a 384-well plate. The tryptophan fluorescence data was measured with Tecan Infinite M1000 plate reader: the samples were excited at 280nm and a fluorescence scan was performed (ranging from 288nm to 500nm), and the fluorescence intensity was plotted as a function of wavelength to generate a fluorescence scan curve.

For the titration experiment, a standard tryptophan fluorescence scan was first performed. Then, 1µl of prepared compound solution (2 mM Cl-, F-, or OMe-derivative) was added to each well and another fluorescence scan was performed with the same parameters mentioned above. The same process was repeated 10 times (in order to collect fluorescence data at 40, 80, 120, 160, 200, 240, 280, 320, 360, and 400 µM of each compound) for the collection of tryptophan fluorescence intensity data.

The measurements for substrates were carried out in the same fashion, except the compound solutions were replaced by substrate solutions (from 0 to 0.4 mM for dTMP and dGMP, and 0 to 1 M for ADP).

The K_D for each of the synthesized compound was estimated using the following method: ΔF (tryptophan fluorescence change upon titration of the compound) data were plotted against the concentration of the synthesized compound, hyperbolic curve fit was then performed using equation: $\Delta F = \Delta F_{max}[C]/(K_D+[C])$, where K_D can be readily estimated from the fitted hyperbolic curve using Excel Solver.

6.5 Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were conducted to examine the thermodynamic parameters of *Pf*TMK upon dTMP, dGMP, ADP, and ATP binding. Binding of dTMP, dGMP, ADP and ATP were investigated by ITC, and all experiments were carried out on MicroCal VP-ITC microcalorimeter, and the collected data were processed with Origin 7 software with ITC addons. The reference cell was filled with ddH2O. The *Pf*TMK protein samples were dialyzed extensively overnight in ITC buffer (20 mM sodium phosphate buffer, 10 mM magnesium chloride (pH 7)) at 4 ºC. The protein concentration we used in our experiments ranged from 50 µM to 150 µM. The protein sample was first degassed at 25 ºC, and was then transferred to the 1.4 mL sample cell. The syringe was then loaded with substrate solution (ranging from 0.5 to 1 mM, degassed) prepared in ITC buffer. The ITC experiments were carried out at 30 ºC, and stirred at 350 rpm. A total of 29 injections were performed for each experiment with 300 s during between each 10 uL injection. Control experiments included titration of each of the substrate into ITC buffer. The data were fitted to one set of binding site binding model, and association constant (Ka), enthalpy of binding (ΔH°), entropy (ΔS°), and stoichiometry (n) were obtained from the fitting via Origin 7. Binding affinity was then obtained by equation $K_D = 1/Ka$, and the binding free energy (ΔG°) was calculated by equation $\Delta G° = \Delta H° – T\Delta S°$.

6.6 NMR Spectroscopy

6.6.1 2D HMQC Experiment and NMR Titration

NMR titration experiments were performed to examine the chemical shift perturbation of methyl residues in *Pf*TMK for the purpose of K_D estimation for dTMP, dGMP, and ADP, and to investigate the stoichiometry of dTMP-*Pf*TMK and dGMP-*Pf*TMK complexes.

0.3-0.4 mM wild-type isotopically labeled *Pf*TMK in NMR buffer containing 2mM ADP (or 1mM dTMP for K_D^{ADP} measurement). Two-dimensional HMQC methyl carbon spectra were obtained on either Bruker AVANCE 500 mHz or 600 mHz (proton) NMR spectrometers. Substrate was titrated into the sample to raise the substrate concentration to 0.1X. This process was repeated until the substrate concentration reached 3 mM for dTMP and dGMP, and 10 mM for ADP.

The HMQC spectra (500 MHz, 600 MHz 1H) were processed using NMRpipe[67] and input to TITAN software[68] in MATLAB for the estimation of K_D. This was achieved by selecting specific peaks on the spectrum that exhibited chemical shift perturbation during the titration, and the K_D was calculated by the two-dimensional line fitting feature from TITAN (MATLAB).

6.6.2 CPMG Relaxation Dispersion Experiment

Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion experiment was conducted to examine the local dynamics of *Pf*TMK upon dTMP and dGMP binding. NMR samples were prepared with the following condition: 0.5mM WT isotopically labeled *Pf*TMK in NMR buffer (1 mM HEPES, 200 mM NaCl, 25 µM EDTA, 0.02% sodium azide in D_2O) containing 10mM ADP, 10mM Mg2+, and 6.7 mM dTMP or 10mMdGMP. Substrate concentrations were selected to ensure that both samples shared the same fractional saturation at ~98%. Dispersion was measured at two fields, 600 MHz and 850 MHz, using Bruker AVANCE spectrometers using the relaxation dispersion experiment developed by Kay et al.[69].

Peak intensities were calculated by nmrDraw and the dispersion curves were generated by plotting the effective transverse relaxation rate (R_{ex}) against the rate of refocusing pulses (ν_{CPMG}). The dispersion curves were fitted to fast-exchange model to acquire dynamic parameters[46,58] (eq. 1):

$$R_{2,\text{eff}} = R_0^2 + \frac{\Phi_{ex}}{k_{ex}}\left[1 - \frac{4\upsilon_{\text{CPMG}}}{k_{ex}}\tanh\left(\frac{k_{ex}}{4\upsilon_{\text{CPMG}}}\right)\right] \qquad [1]$$

where $\phi ex = p_a p_b \Delta\omega^2$. The parameters p_a and p_b refer to the population of the two states, $\Delta\omega$ is the difference in the absorption frequency of the two states and k_{ex} is the

rate of exchange between the states. The ϕex and k_{ex} and were acquired by globally fitting the data from two fields using an in-house Python program.

Reference

1. Gathany J. (2014) **Malaria Drug Resistance | Worldwide Antimalarial Resistance Network**. *Wwarn.org*. N.p.
2. Ashley E, et al. (2014) **Spread of Artemisinin Resistance in Plasmodium falciparum Malaria**. *New England Journal of Medicine* 371:411-423
3. Lavie, A., and Konrad, M. (2004) **Structural requirements for efficient phosphorylation of nucleotide analogs by human thymidylate kinase.** *Mini Rev Med Chem*. 4(4):351-9.
4. Grem, JL., Danenberg, KD., Behan, K., Parr, A., Young, L., Danenberg, PV., Nguyen, D., Drake, J., Monks, A., and Allegra, CJ. (2001) **Thymidine kinase, thymidylate synthase, and dihydropyrimidine dehydrogenase profiles of cell lines of the National Cancer Institute's Anticancer Drug Screen**. *Clin Cancer Res*. 7(4):999-1009.
5. Cui, Q., Shin, WS., Luo, Y., Tian, J., Cui, H., and Yin, D. (2013) **Thymidylate kinase: an old topic brings new perspectives.** *Curr Med Chem*. 20(10):1286-305.
6. Martinez-Botella, G., Breen, J. N., Duffy, J. E., Dumas, J., Geng, B., Gowers, I. K., Green, O. M., Guler, S., Hentemann, M. F., Hernandez-Juan, F. A., Joseph-McCarthy, D., Kawatkar, S., Larsen, N. A., Lazari, O., Loch, J. T., Macritchie, J.

A., McKenzie, A. R., Newman, J. V., Olivier, N. B., Otterson, L. G., Owens, A. P., Read, J., Sheppard, D. W., and Keating, T. A. (2012) **Discovery of selective and potent inhibitors of gram-positive bacterial thymidylate kinase (TMK)**. *J. Med. Chem*. 55, 10010−10021.

7. Martinez-Botella, G., Loch, J. T., Green, O. M., Kawatkar, S. P., Olivier, N. B., Boriack-Sjodin, P. A., and Keating, T. A. (2013) **Sulfonylpiperidines as novel, antibacterial inhibitors of Gram-positive thymidylate kinase (TMK).** *Bioorg. Med. Chem. Lett*. 23, 169−173.
8. Chen, M., Sinha, K., Rule, GS., and Ly, DH. (2018**) Interaction of α-Thymidine Inhibitors with Thymidylate Kinase from Plasmodium falciparum.** *Biochemistry* 57, 19, 2868-2875
9. Kato, A., Yasuda, Y., Kitamura, Y., Kandeel, M., and Kitade, Y. (2012) **Carbocyclic thymidine derivatives efficiently inhibit Plasmodium falciparum thymidylate kinase (PfTMK)**. *Parasitol*. Int. 61, 501− 503.
10. Noguchi, Y., Yasuda, Y., Tashiro, M., Kataoka, T., Kitamura, Y., Kandeel, M., and Kitade, Y. (2013) **Synthesis of carbocyclic pyrimidine nucleosides and their inhibitory activities against Plasmodium falciparum thymidylate kinase**. *Parasitol*. Int. 62, 368−371.
11. Cui, H., Carrero-Lerida, J., Silva, a. P., Whittingham, J. L., Brannigan, J. a., Ruiz-Perez, L. M., Read, K. D., Wilson, K. S., Gonzalez-Pacanowska, D., and Gilbert, I. H. (2012) **Synthesis and evaluation of alpha-thymidine analogues as novel antimalarials.** *J. Med. Chem*. 55, 10948−10957.

12. Van Poecke, S., Munier-Lehmann, H., Helynck, O., Froeyen, M., and Van Calenbergh, S. (2011) **Synthesis and inhibitory activity of thymidine analogues targeting Mycobacterium tuberculosis thymidine monophosphate kinase**. *Bioorg. Med. Chem*. 19, 7603−7611.
13. Familiar, O., Munier-Lehmann, H., Ainsa, J. A., Camarasa, M. J., and Perez-Perez, M. J. (2010) **Design, synthesis and inhibitory activity against Mycobacterium tuberculosis thymidine monophosphate kinase of acyclic nucleoside analogues with a distal imidazoquinolinone**. *Eur. J. Med. Chem*. 45, 5910−5918.
14. Choi, J. Y., Plummer, M. S., Starr, J., Desbonnet, C. R., Soutter, H., Chang, J., Miller, J. R., Dillman, K., Miller, A. A., and Roush, W. R. (2012) **Structure guided development of novel thymidine mimetics targeting Pseudomonas aeruginosa thymidylate kinase: from hit to lead generation**. *J. Med. Chem*. 55, 852−870.
15. **Malaria Disease and Biology.** (2018) *Global Health, Division of Parasitic Diseases and Malaria.* N.p.
16. Brown, Geoff (2006). **Artemisinin and a new generation of antimalarial drugs**. *Education in Chemistry*. Vol. 43 no. 4. pp. 97–99.
17. Krieger, Johannes., Smeilus, Toni., Kaiser, Marcel., Seo, Ean-Jeong., Efferth, Thomas., Giannis, Athanassios (2018). **Total Synthesis and Biological Investigation of (−)-Artemisinin: The Antimalarial Activity of Artemisinin Is not Stereospecific**. *Angewandte Chemie International Edition*. 57 (27): 8293–8296.

18. Tilley, Leann., Straimer, Judith., Gnädig, Nina F., Ralph, Stuart A., Fidock, David A. (2016). **Artemisinin Action and Resistance in Plasmodium falciparum**. *Trends in Parasitology*. 32 (9): 682–696
19. Winzeler, E. A., M. J., Manary (2014). **Drug resistance genomics of the antimalarial drug artemisinin**. *Genome Biology*. 15 (11): 544
20. Cravo, P., Napolitano, H., Culleton, R. (2015). **How genomics is contributing to the fight against artemisinin-resistant malaria parasites**. *Acta Tropica*. 148: 1–7
21. Ostermann, N., Schlichting, I., Brundiers, R., Konrad, M., Reinstein, J., Veit, T., Goody, R. S., and Lavie, A. (2000) **Insights into the phosphoryltransfer mechanism of human thymidylate kinase gained from crystal structures of enzyme complexes along the reaction coordinate**. *Structure* 8, 629−642.
22. Lavie, A., Konrad, M., Brundiers, R., Goody, R. S., Schlichting, I., and Reinstein, J. (1998) **Crystal structure of yeast thymidylate kinase complexed with the bisubstrate inhibitor P1-(5'-adenosyl) P5-(5'-thymidyl) pentaphosphate (TP5A) at 2.0 A resolution: Implications for catalysis and AZT activation.** *Biochemistry* 37, 3677−3686.
23. Fioravanti, E., Haouz, A., Ursby, T., Munier-Lehmann, H., Delarue, M., and Bourgeois, D. (2003) **Mycobacterium tuberculosis thymidylate kinase: Structural studies of intermediates along the reaction pathway**. *J. Mol. Biol*. 327, 1077−1092.
24. Caillat, C., Topalis, D., Agrofoglio, L. a., Pochet, S., Balzarini, J., Deville-Bonne, D., and Meyer, P. (2008) **Crystal structure of poxvirus thymidylate kinase: an**

unexpected dimerization has implications for antiviral therapy. *Proc. Natl. Acad. Sci. U. S. A*. 105, 16900−16905.

25. Whittingham JL., et al (2010) **Structural basis for the efficient phosphorylation of AZT-MP (3′-azido-3′-deoxythymidine monophosphate) and dGMP by Plasmodium falciparum type I thymidylate kinase**. *Biochem. J*. 428, 499−509.
26. Sinha K., and Rule GS. (2017) **The Structure of Thymidylate Kinase from Candida albicans Reveals a Unique Structural Element.** *Biochemistry*, 2017, 56 (33), pp 4360–4370
27. Jerina, D. M., and Daly, J. W. (1974) **Arene oxides: a new aspect of drug metabolism**. *Science* 185, 573−582.
28. Rule, G.S., Hitchens, K.T. (2006) **Fundamentals of Protein NMR Spectroscopy.** *Springer*, Dordrecht, The Netherlands
29. Cavanagh, J., Fairbrother, W.J., Palmer A.G.I., Rance, M. Rance., Skelton N.J. (2007) **Protein NMR Spectroscopy**. *Elsevier*, Burlington, MA, USA
30. Levitt, M.H. (2001) **Spin Dynamics: Basics of Nuclear Magnetic Resonance**. *John Wiley & Sons Ltd*., West Sussex, England
31. Scheich, C., Puetter, V., Schade, M. (2010) **Novel Small Molecule Inhibitors of MDR Mycobacterium tuberculosis by NMR Fragment Screening of Antigen 85C.** *Journal of Medicinal Chemistry* 53 (23), 8362-8367
32. Morales, KA., Yang, Y., Cole, TR., Igumenova, TI. (2016) **Dynamic Response of the C2 Domain of Protein Kinase Cα to Ca2+ Binding**. *Biophys. J.*111 (8):1655-1667.

33. Ruschak AM., Kay, LE. (2010) **Methyl groups as probes of supra-molecular structure, dynamics and function**. *J Biomol NMR*. 46(1):75-87

34. Yampolsky, L.Y., Stoltzfus, A. (2005) **The exchangeability of amino acids in proteins.** *Genetics*, 170, 1459-1472.

35. Kandeel, M., Ando, T., Kitamura, Y., Abdel-Aziz, M., and Kitade, Y. (2009) **Mutational, inhibitory and microcalorimetric analyses of Plasmodium falciparum TMP kinase. Implications for drug discovery**. *Parasitology* 136, 11−25.

36. O. Trott, A. J. Olson. (2010) **AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading**, Journal of Computational Chemistry 31 455-461

37. Chemsketch version 14.01, Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com, 2015.

38. Michel F. Sanner. Python**: A Programming Language for Software Integration and Development**. J. Mol. Graphics Mod., 1999, Vol 17, February. pp57-61

39. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. (2009) **Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity.** J. Computational Chemistry 2009, 16: 2785-91

40. The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

41. Kleckner, IR., Foster, MP. (2011) **An introduction to NMR-based approaches for measuring protein dynamics**. *Biochim Biophys Acta*.1814(8):942-68

42. Zeeb, M., Balbach, J. (2004) **Protein folding studied by real-time NMR spectroscopy.** *Methods*. pp. 65-74
43. Palmer, AG3., Kroenke, CD., Loria, JP. (2001) **Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules** *Meth. Enzymol.* pp. 204-238
44. J. Jeener, J., Meier, B., Bachmann, P., Ernst, R. (1979) **Investigation of exchange processes by 2-dimensional NMR-spectroscopy** *J. Chem. Phys.* pp. 4546-4553
45. McConnell,H. (1958) **Reaction rates by nuclear magnetic resonance.** *J. Chem. Phys*. pp. 430-431
46. Meiboom, S., Gill, D. (2963) **Modified spin-echo method for measuring nuclear relaxation times.** *Rev. Sci. Instrum*. pp. 688-691
47. Luz, Z., Meiboom, S. (1963) **Nuclear magnetic resonance study of the protolysis of trimethylammonium ion in aqueous solution-order of the reaction with respect to solvent.** *The Journal of Chemical Physics* 39 (2), 366-370
48. Carr, H., Purcell, E. (2954) **Effects of diffusion on free precession in nuclear magnetic resonance experiments**. *Phys. Rev*. pp. 630-638
49. Loria, J., Rance, M., Palmer, A. (1999) **A relaxation-compensated Carr–Purcell–Meiboom–Gill sequence for characterizing chemical exchange by NMR spectroscopy**. *J. Amer. Chem. Soc*. pp. 2331-2332

50. Ishima, R., Torchia, DA. (2003) **Extending the range of amide proton relaxation dispersion experiments in proteins using a constant-time relaxation-compensated cpmg approach** *J. Biomol. NMR*. pp. 243-248
51. Lundström, P., Vallurupalli, P., Hansen, DF., Kay, LE. **Isotope labeling methods for studies of excited protein states by relaxation dispersion NMR spectroscopy**. *Nat. Protoc.* pp. 1641-1648
52. Kem*Pf*, JG., Loria, JP. (2003) **Protein dynamics from solution NMR: theory and applications**. *Cell Biochem. Biophys*. pp. 187-211
53. Jarymowycz, VA., Stone, MJ. (2006) **Fast time scale dynamics of protein backbones: nmr relaxation methods, applications, and functional consequences**. *Chem. Rev.* pp. 1624-1671
54. Igumenova, TI., Frederick, KK., Wand, AJ. (2006) **Characterization of the fast dynamics of protein amino acid side chains using nmr relaxation in solution.** *Chem. Rev.* pp. 1672-1699
55. Lakomek, N., Lange, OF., Walter, KFA., Farès, C., Egger, D., Lunkenheimer, P., Meiler, J., Grubmüller, H., Becker, S., Groot, BL de., Griesinger, C. (2008) **Residual dipolar couplings as a tool to study molecular recognition of ubiquitin**. *Biochem. Soc. Trans.* pp. 1433-1437
56. Tolman, JR., Ruan, K. (2006) **Nmr residual dipolar couplings as probes of biomolecular dynamics**. *Chem. Rev*. pp. 1720-1736
57. Clore, GM., Iwahara, J. (2009) **Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient**

low-population states of biological macromolecules and their complexes. *Chem. Rev.* pp. 4108-4139

58. Yuwen, T., Huang, R., Vallurupalli, P., Kay, L.E. (2019) **A Methyl-TROSY-based ^{1}H Relaxation Dispersion Experiment for Studies of Conformational Exchange in High Molecular Weight Proteins**. Angewandte Chemie, *131, 6316-6320*.

59. Cui, H., Ruda, GF., Carrero-Lerida, J., Ruiz-Perez, L. M., Gilbert, IH., Gonzalez-Pacanowska, D., and Gilbert, I. H. (2010) **Exploring new inhibitors of Plasmodium falciparum purine nucleoside phosphorylase.** *Eur J Med Chem,* 45(11):5140-9

60. Lomenic, B., Olsen, RW., Huang, J. (2011) **Identification of Direct Protein Targets of Small Molecules**. *ACS Chem. Biol*. 6134-46

61. Pai, MY., Lomenic, B., Hwang, H., Schiestl, R., McBride, W., Loo, JA., Huang, J. (2015) **Drug Affinity Responsive Target Stability (DARTS) for Small Molecule Target Identification**. *Methods Mol Biol.* 1263: 287–298

62. West, JM., Xia, J., Tsuruta, H., Guo, W., O'Day, EM., Kantrowitz., ER. (2008) **Time Evolution of the Quaternary Structure of Escherichia coli Aspartate Transcarbamoylase upon Reaction with the Natural Substrates and a Slow Tight Binding Inhibitor.** *J Mol Biol.* 384(1): 206–218

63. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J.W. (1990) **Use of T7 RNA polymerase to direct expression of cloned genes**. *Methods Enzymo*l. 185, 60−89.

64. Studier FW. (2005) **Protein production by auto-induction in high density shaking cultures**. *Protein expression and purification*, 41(1):207-234.

65. Goto NK., and Kay LE. (2000) **New developments in isotope labeling strategies for protein solution NMR spectroscopy.** *Current opinion in structural biology*, 10(5):585-592.

66. Blondin, C., Serina, L., Wiesmuller, L., Gilles, A. M., and Barzu, O. (1994) **Improved spectrophotometric assay of nucleoside monophosphate kinase activity using the pyruvate kinase/lactate dehydrogenase coupling system.** *Anal. Biochem*. 220, 219−221.

67. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., *Pf*eifer, J., Bax, A. (1995) NMRPipe: **A multidimensional spectral processing system based on UNIX pipes.** *J. Biomol. NMR*, 6, 277-293.

68. Waudby, CA., Ramos, A., Cabrita, LD., and Christodoulou, J. (2016) **Two-Dimensional NMR Lineshape Analysis**. *Scientific Reports* volume 6, 24826

69. Millet, O., Loria, J.P., Kroenke, C.D., Pons, M., Palmer III, A.G. (2000) **The static magnetic field dependence of chemical exchange linebroadening defines the NMR chemical shift time scale**. *Journal of the American Chemical Society* 122 (12), 2867-2877