A Dissertation

entitled

Structural and Biochemical Studies of Protein-Ligand Interactions: Insights for Drug Development

by

Vidhi Mishra

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the Doctor of Philosophy Degree in Chemistry

__________________________
Donald R. Ronning, Ph.D., Committee Chair

__________________________
John J. Bellizzi, Ph.D., Committee Member

__________________________
Steve Sucheck, Ph. D., Committee Member

__________________________
Constance A. Schall, Ph.D., P.E., Committee Member

__________________________
Dr. Patricia R. Komuniecki, Dean
College of Graduate Studies

The University of Toledo

December 2013
This document is copyrighted material. Under copyright law, no parts of this document may be reproduced without the expressed permission of the author.
An Abstract of

Structural and biochemical studies of protein-ligand interactions: Insights for drug development

By

Vidhi Mishra

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the Doctor of Philosophy Degree in Chemistry

The University of Toledo

December 2013

Protein ligand interactions play a key role in the majority of all biological processes. Proteins display unique binding sites that are recognized by specific ligand through distinct interactions. Characterization of these interactions provides insights that are exploitable for designing drugs. Here we study enzyme-substrate interactions in a protein encoded by Helicobacter pylori (H. pylori), which is associated with gastric and duodenal ulcers.

H. pylori MTAN (HpMTAN) catalyzes the hydrolysis of N-ribosidic bonds of at least four different adenosine based substrates; S-adenosylhomocysteine (SAH), 5’-methylthioadenosine (MTA), 5’-deoxyadenosine (5’-DOA) and 6-amino-6-deoxyfutalosine. This hydrolytic activity places MTAN at the hub of at least seven fundamental metabolic pathways: the purine salvage pathway, the methionine salvage pathway, S-adenosylmethionine (SAM)-dependent methylation pathways, polyamine biosynthesis, the production of quorum sensing molecules and menaquinone biosynthesis. Campylobacter and Helicobacter are dependent on MTAN for menaquinone synthesis, an essential metabolite for bacterial viability, making MTAN an
excellent target for the development of new treatments for Helicobacter infections. To structurally characterize the interactions between MTAN and its various substrates, complexes of an inactive mutant of the *Hp*MTAN with two known substrates, 5’-methylthioadenosine (MTA) and S-adenosylhomocysteine (SAH) were formed and crystallized. The crystal structures of mutant *Hp*MTAN complexed with SAH and MTA were solved to 1.2 and 1.6 Å, respectively. The *Hp*MTAN-SAH co-crystal structure represents the first visualization of interactions between the homocysteine moiety of SAH and the 5’-alkylthiol-binding subsite of the MTAN active site. The co-crystal structure of wild-type MTAN with products, adenine and S-ribosylhomocysteine, was determined to 1.54 Å resolution. The similarities and differences in these three structures highlight features that can be exploited to design *H. pylori* specific drugs. Additionally, a high-throughput fluorescence polarization assay was developed and optimized that will afford the identification of new drug scaffolds that bind to the *H. pylori* MTAN active site.

Additionally, we studied the I1 protein that is essential for the assembly of vaccinia virus. I1 is a telomere binding protein. We performed structural and biochemical characterizations to understand the mechanism of protein-DNA interaction.

Finally, we studied a protein involved in the folate biosynthetic pathway for drug discovery purpose against *Mycobacterium tuberculosis*.
Acknowledgements

At first, I thank my advisor, Prof. Donald Ronning for his extensive support, patience, and enthusiasm. His dedication to teach protein crystallography and explaining the details has helped me gain a better understanding, and appreciate new discoveries in the field. I would like to thank my committee members Dr. John Bellizzi and Dr. Steve Sucheck. I feel privileged that Dr. Schall has accepted to be a part of my thesis committee and thank her for the same.

I would like to thank Dr. Chi-chin Hwang for her suggestions in enzymology and various aspects of my research. I am very grateful to Dr. Ronald Viola, Dr. Max Funk, Dr. Tim Mueser, and Dr. Wendell Griffith for allowing me to use different equipments in their lab. I would like to express my heartfelt gratitude to Dr. Leif Hanson. I would like to thank my past group members Dr. Julie Boucau, Dr. Dan Lajiness, and Natalie Iacopelli for having helped me to acclimatize in the lab. The long hours I spent with Jared, Lorenza, Lucile, and Erica, in the lab has kept my graduate work lively and productive.

I would like to thank my friends and family members for their love, tremendous support, and motivation throughout my education. I am very grateful for the constant support by Nidhi Sharma. Lastly, the contribution of my husband, Santosh Kumar is of paramount significance. He has always encouraged me in science, life, and has always stood by me in times of stress. He deserves endless appreciation.
# Table of Contents

Abstract .................................................................................................................................................. iii
Acknowledgements .................................................................................................................................. v
Table of Contents ..................................................................................................................................... vi
List of Figures .......................................................................................................................................... xi
List of Abbreviations ........................................................................................................................... xiv
List of Symbols ........................................................................................................................................ xvi

1 *Helicobacter pylori* 5’-methylthio/S-adenosylhomocysteine nucleosidase:
background and significance .................................................................................................................. 1

1.1 Role of bacterial MTAN ................................................................................................................... 1
1.2 *Helicobacter pylori* MTAN as a drug target .................................................................................. 4
1.3 Active site of *HpMTAN* .................................................................................................................... 5
   1.3.1 Purine binding subsite .................................................................................................................. 5
   1.3.2 Ribose binding subsite ................................................................................................................. 6
   1.3.3 5’-alkylthio binding subsite ........................................................................................................ 8
1.4 Enzyme mechanism ........................................................................................................................... 8
1.5 Human MTAP vs. *HpMTAN* ............................................................................................................. 10
1.6 Overall structure of *HpMTAN* ......................................................................................................... 13
1.7 Ligand induced conformational changes of *HpMTAN* ................................................................. 16
Kinetic and structural insight into differential substrate binding to \textit{Hp}MTAN\ldots 18

2.1 Overview ........................................................................................................ 18

2.2 Materials and methods .................................................................................. 19

\hspace*{0.5cm} 2.2.1 Mutagenesis .................................................................................... 19

\hspace*{0.5cm} 2.2.2 Protein expression and purification .................................................... 19

\hspace*{0.5cm} 2.2.3 MTAN steady-state kinetic assays ....................................................... 22

\hspace*{0.5cm} 2.2.4 Tris inhibitory assay ............................................................................. 23

\hspace*{0.5cm} 2.2.5 Crystallization ....................................................................................... 23

2.3 Results and discussion .................................................................................. 24

\hspace*{0.5cm} 2.3.1 Kinetic characterization of wt-MTAN ................................................... 24

\hspace*{0.5cm} 2.3.2 Tris inhibitory assay ............................................................................. 26

\hspace*{0.5cm} 2.3.3 Crystallization studies .......................................................................... 27

\hspace*{0.5cm} \hspace*{0.5cm} 2.3.3.1 Interactions between homocysteine chain and 5\textsuperscript{-}alkylthio

\hspace*{0.5cm} \hspace*{0.5cm} binding subsite .................................................................................. 31

\hspace*{0.5cm} \hspace*{0.5cm} 2.3.3.2 Comparison of \textit{Hp}MTAN-D198N-SAH with \textit{Hp}MTAN-

\hspace*{0.5cm} \hspace*{0.5cm} D198-MTA ....................................................................................... 35

\hspace*{0.5cm} \hspace*{0.5cm} 2.3.3.3 Comparison of \textit{Hp}MTAN-D198N-SAH with \textit{Hp}MTAN-

\hspace*{0.5cm} \hspace*{0.5cm} SRH-ADE ......................................................................................... 37

\hspace*{0.5cm} 2.4 Conclusion .............................................................................................. 40

3 Development of high-throughput binding assay to identifying \textit{Hp}MTAN

\hspace*{0.5cm} inhibitors .................................................................................................... 41

\hspace*{0.5cm} 3.1 Rational for assay development .................................................................. 41

\hspace*{0.5cm} 3.2 Approach for the assay development ..................................................... 43
3.3 Counter-screen assay .................................................................46
3.4 Material and methods.........................................................................................47
  3.4.1 Protein preparation and adenine removal from HpMTAN active site47
  3.4.2 Fluorescence labeling of SAH ..............................................................47
  3.4.3 HPLC purification of SAH-488 ...............................................................48
  3.4.4 Characterization and concentration determination of SAH-488......48
  3.4.5 SAH-488 binding assay ..........................................................................49
  3.4.6 SAH displacement assay ...........................................................................49
  3.4.7 Assay validation .......................................................................................50
  3.4.8 Screening of compounds of NIH clinical collection.........................50
  3.4.9 Counter-screen assay ..............................................................................51
3.5 Results and discussion .........................................................................................52
  3.5.1 SAH-488 synthesis, purification and characterization .......................52
  3.5.2 SAH-488 binding constant .......................................................................54
  3.5.3 SAH inhibition constant ..........................................................................55
  3.5.4 Assay validation .......................................................................................56
  3.5.5 Screening of compounds of NIH clinical collection .........................59
  3.5.6 Hit validation and false positive hits ......................................................61
3.6 Conclusion and future work ..............................................................................62
4 Structural and biochemical analysis of II: Telomere binding protein in vaccinia virus ........................................................................................................63
  4.1 Overview .......................................................................................................63
  4.2 Structure of vaccinia virus ...........................................................................64
4.3 Replication of vaccinia virus .................................................................66
4.4 Protein parameters, expression and purification of I1 ..............................69
  4.4.1 Co-affinity purification of recombinant I1 ......................................69
  4.4.2 SDS-PAGE gel analysis.................................................................70
  4.4.3 Rationale for interest in truncated form I1 ......................................70
  4.4.4 SEC purification of truncated I1 ....................................................72
4.5 Crystallization .....................................................................................74
4.6 Truncated I1 and DNA binding assay ....................................................75
4.7 Issues with truncated I1 expressed using construct (p101-I1) ..................77
4.8 Construct design for expressing an engineered truncated I1 ..................77
  4.8.1 Mass spectrometry ........................................................................78
  4.8.2 Cloning, protein expression and purification of I1 truncated ............82
4.9 Conclusion and future studies ..............................................................84

5 Structure and functional analysis of Mycobacterial dihydronopterin triphosphate pyrophosphatase: A nudix hydrolase involve in folate biosynthesis ............86
  5.1 Tuberculosis and *Mycobacterium tuberculosis* ...................................86
  5.2 Folate biosynthetic pathway ............................................................87
  5.3 Protein expression and purification of *M.smeg* MutT3 .......................89
  5.4 MutT3 substrate specificity using ATP, GTP, CTP and UTP as a substrate...91
  5.5 Kinetic parameters using ATP as a substrate .....................................94
  5.6 Crystallization attempts .................................................................95
  5.7 Chemo-enzymatic synthesis of DHNTP ...........................................95
    5.7.1 Cloning, protein expression and purification of *E.coli* FolE ..........95
5.7.2 Chemo-enzymatic synthesis of DHNTP ........................................97
5.7.3 Purification and characterization of DHNTP .................................98
5.8 MutT3 steady-state kinetics using DHNTP as a substrate .................101
5.9 Conclusion and future work ..........................................................102
References .......................................................................................103
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Bacterial MTAN is linked to multiple pathways ........................................ 3</td>
</tr>
<tr>
<td>1-2</td>
<td>Menaquinone biosynthetic pathway in <em>Helicobacter pylori</em> .......................... 4</td>
</tr>
<tr>
<td>1-3</td>
<td>Active site of <em>Hp</em>MTAN ........................................................................... 7</td>
</tr>
<tr>
<td>1-4</td>
<td>Reaction mechanism of MTAN ...................................................................... 8</td>
</tr>
<tr>
<td>1-5</td>
<td>MTAP vs. MTAN reaction scheme .................................................................. 11</td>
</tr>
<tr>
<td>1-6</td>
<td>MTAP vs. MTAN active site .......................................................................... 12</td>
</tr>
<tr>
<td>1-7</td>
<td>MTAP active site has a truncated 5’-alkylthio binding subsite ...................... 13</td>
</tr>
<tr>
<td>1-8</td>
<td>The <em>Hp</em>MTAN monomer structure ............................................................... 14</td>
</tr>
<tr>
<td>1-9</td>
<td><em>Hp</em>MTAN exist as a homodimer ................................................................... 15</td>
</tr>
<tr>
<td>1-10</td>
<td>Ligand induced conformational change of <em>Hp</em>MTAN ................................... 17</td>
</tr>
<tr>
<td>2-1</td>
<td>SDS-PAGE gel analysis of <em>Hp</em>MTAN ......................................................... 21</td>
</tr>
<tr>
<td>2-2</td>
<td>Scheme for MTAN kinetic assay .................................................................. 22</td>
</tr>
<tr>
<td>2-3</td>
<td>Kinetic characterization of <em>Hp</em>MTAN ....................................................... 25</td>
</tr>
<tr>
<td>2-4</td>
<td>Inhibition profile of <em>Hp</em>MTAN by Tris .................................................... 26</td>
</tr>
<tr>
<td>2-5</td>
<td>An obligatory dimer of MTAN-D198N-SAH .................................................. 27</td>
</tr>
<tr>
<td>2-6</td>
<td>Crystal information of MTAN-D198N-SAH .................................................. 28</td>
</tr>
<tr>
<td>2-7</td>
<td>Crystal information of MTAN-D198N-MTA .................................................. 29</td>
</tr>
<tr>
<td>2-8</td>
<td>Crystal information of MTAN-SRH-ADE ..................................................... 30</td>
</tr>
<tr>
<td>2-9</td>
<td>Interactions between homocysteine chain and 5’-alkylthio binding subsite ..........32</td>
</tr>
<tr>
<td>2-10</td>
<td>The sequence alignment of 1000 related MTANs .................................................34</td>
</tr>
<tr>
<td>2-11</td>
<td>The comparison of SAH and MTA bound in the active site ..................................36</td>
</tr>
<tr>
<td>2-12</td>
<td>The comparison of substrate and product bound in the active site .......................39</td>
</tr>
<tr>
<td>3-1</td>
<td>Scheme of SAH-488 synthesis .................................................................................43</td>
</tr>
<tr>
<td>3-2</td>
<td>Principle of FP binding assay ..............................................................................45</td>
</tr>
<tr>
<td>3-3</td>
<td>Scheme of Amplex Red-Xanthine oxidase assay ......................................................46</td>
</tr>
<tr>
<td>3-4</td>
<td>HPLC chromatogram of SAH-488 purification ..........................................................52</td>
</tr>
<tr>
<td>3-5</td>
<td>ESI-MS of purified sample of SAH-488 .................................................................53</td>
</tr>
<tr>
<td>3-6</td>
<td>FP binding assay of SAH-488 and MTAN-D198N ....................................................54</td>
</tr>
<tr>
<td>3-7</td>
<td>SAH displacement assay ......................................................................................55</td>
</tr>
<tr>
<td>3-8</td>
<td>FP assay evaluation ............................................................................................58</td>
</tr>
<tr>
<td>3-9</td>
<td>Screening of NIH clinical collection ....................................................................60</td>
</tr>
<tr>
<td>4-1</td>
<td>Vaccinia virus .....................................................................................................65</td>
</tr>
<tr>
<td>4-2</td>
<td>Vaccinia virus DNA replication .........................................................................66</td>
</tr>
<tr>
<td>4-3</td>
<td>Vaccinia virus life cycle ....................................................................................68</td>
</tr>
<tr>
<td>4-4</td>
<td>Co-affinity chromatogram and corresponding SDS-PAGE ......................................71</td>
</tr>
<tr>
<td>4-5</td>
<td>SEC and corresponding SDS-PAGE ....................................................................73</td>
</tr>
<tr>
<td>4-6</td>
<td>Crystals of truncated I1 ......................................................................................74</td>
</tr>
<tr>
<td>4-7</td>
<td>I1/DNA binding curve .......................................................................................76</td>
</tr>
<tr>
<td>4-8</td>
<td>Peptide mass fingerprinting of truncated form of I1 ............................................78</td>
</tr>
<tr>
<td>4-9</td>
<td>Intact mass analysis ..........................................................................................80</td>
</tr>
<tr>
<td>4-10</td>
<td>Secondary structure prediction for vaccinia virus I1 .........................................81</td>
</tr>
</tbody>
</table>
4-11 The SDS PAGE gel of I1 truncated expression ..........................................................83
5-1 De novo folate biosynthesis .........................................................................................88
5-2 SDS PAGE analysis of MutT3 ......................................................................................90
5-3 Scheme of MutT3 kinetic assay ....................................................................................92
5-4 Substrate specificity of mutT3 ....................................................................................93
5-5 Kinetic characterization of mutT3 using ATP as a substrate ........................................94
5-6 SDS PAGE analysis of FolE ........................................................................................96
5-7 Chemo-enzymatic synthesis of DHNTP ....................................................................97
5-8 HPLC chromatogram using isopropanol-triethylamine-85% phosphoric acid ..........98
5-9 HPLC chromatogram using isopropanol-triethylamine-formic acid .........................99
5-10 ESI-MS of the fraction collected from HPLC ..........................................................100
5-11 Kinetic assay using DHNTP as a substrate ..............................................................101
List of Abbreviations

5’DOA.......................... 5’-deoxyadenosine
ATP............................. adenosine triphosphate
CTP.............................. cytidine triphosphate
DHFL......................... dehypoxanthine futalosine
DMSO......................... dimethylsulfoxide
DNA........................... deoxyribonucleic acid
E. coli....................... Escherididia coli
EDTA........................... ethylenediaminetetraacetic acid
ESI-MS....................... electrospray ionization mass spectrometry
GTP........................... guanosine triphosphate
HEPES....................... 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hp.............................. Helicobacter pylori
IC50.......................... half maximal inhibitory concentration
IPTG........................... isopropyl β-D-1-thiogalactopyranoside
kcat.......................... turnover number
Ki............................. inhibitory constant
KM........................... Michaelis-Menten Constant
LB............................ Luria Bertani
\textit{M. smeg} \ldots \textit{Mycobacterium smegmatis}

\textit{M. tb} \ldots \textit{Mycobacterium tuberculosis}

MALDI-MS \ldots \text{matrix-assisted laser desorption/ionization}

MTA \ldots \text{5’-methylthioadenosine}

MTAN \ldots \text{5’-methylthioadenosine/S-adenosylhomocysteine nucleosidase}

NIH \ldots \text{National Institute of Health}

OD \ldots \text{optical density}

PAGE \ldots \text{polyacrylamide gel electrophoresis}

PCR \ldots \text{polymerase chain reaction}

SAH \ldots \text{S-adenosylhomocysteine}

SAM \ldots \text{S-adenosylmethionine}

SDS \ldots \text{sodium dodecyl sulfate}

SEC \ldots \text{size exclusion chromatography}

SRH \ldots \text{S-ribosyl homocysteine}

Tris \ldots \text{tris(hydroxymethyl)aminomethane}

UTP \ldots \text{uridine triphosphate}

\text{V}_i \ldots \text{initial velocity}

\text{V}_\text{max} \ldots \text{maximum velocity}

WHO \ldots \text{World Health Organization}
List of Symbols

Å.............................Angstrom
°C .........................degree Celsius
µM..........................micromolar
kDa..........................kilodalton
mOD........................milli-optical density
mM..........................millimolar
mP...........................milipolarization
nm...........................nanometer
nM...........................nanomolar
OD..........................optical density
pH..........................potential hydrogen
rpm........................revolution per minute
w/v........................weight/volume
Chapter 1

Helicobacter pylori 5’-methylthio/S-adenosylhomocysteine nucleosidase: background and significance

1.1 Role of bacterial MTAN

The enzyme 5’-methylthio/S-adenosylhomocysteine nucleosidase (MTAN) belongs to a family of hydrolases (EC 3.2.2.9) that catalyze the hydrolysis of the N-glycosidic bond in bacteria such as Helicobacter pylori, Vibrio cholerae, Escherichia coli and Streptococcus aureus\(^1\)\(^-\)\(^4\). Bacterial MTAN catalyzes the depurination of at least three different adenosine-based substrates: S-adenosylhomocysteine (SAH), 5’-methylthioadenosine (MTA), and 5’-deoxyadenosine (5’-DOA)\(^5\)\(^-\)\(^7\) (Figure 1-1). This hydrolytic activity links MTAN to different components of the S-adenosylmethionine (SAM) metabolic pathway including the SAM utilization pathway, polyamine biosynthesis, the SAM radical pathway, purine salvage pathways, autoinducer-2 biosynthesis and the methionine salvage pathway (Figure 1-1). SAM is the primary methyl donor for the methylation of proteins, nucleic acids, carbohydrates, and lipids\(^8\). Methyltransferases utilize SAM for methyltransfer, converting SAM into SAH in the process.
SAM is also a substrate for spermidine synthase, an enzyme required for polyamine biosynthesis\(^9\). This results in the conversion of SAM into MTA. Inhibition of MTAN causes the accumulation of MTA and SAH in the bacterial cell. As the concentration of these substrates increase, MTA and SAH function as feedback inhibitors for upstream enzymes such as spermidine synthase and methyltransferases respectively, preventing polyamine biosynthesis and the methylation of biomolecules\(^8,10\). In addition, the hydrolysis of SAH and MTA is required for purine and methionine recycling\(^11,12\).

MTAN activity is also required for the synthesis of autoinducers II (4,5-dihydroxy-2, 3-pentadione), a quorum-sensing molecule\(^13\). Quorum sensing is the phenomenon by which bacterial cells communicate and co-operate with each other for biofilm formation, toxin production and other virulence factors. Inhibition of MTAN results in the disruption of biofilm formation and virulence factor production and thus represents an important chemotherapeutic target in the inhibition of quorum-sensing\(^14,15\).

Choi-Rhee and Cronan reported the role of MTAN in the SAM radical pathway\(^5\). The SAM radical enzymes such as biotin synthase utilize SAM to synthesize biotin and releases 5’-DOA radical as a side product. The highly reactive 5’-DOA radical is reduced to form 5’-DOA, that is cleaved by MTAN into adenine and 5’-deoxyribose. The inhibition of MTAN results in the accumulation of 5’-DOA in the bacterial cell. The high concentration of 5’-DOA serves as a feedback inhibitor for SAM radical enzymes\(^5\).
Figure 1-1: Bacterial MTAN is linked to multiple pathways. MTAN utilizes three different adenosine-based substrates: SAH, MTA and 5′-DOA (solid purple box). Figure adapted from reference 24.
1.2 *Helicobacter pylori* MTAN as a drug target

Recent studies provide evidence that MTAN plays an essential role in organisms such as *Campylobacter jejuni* and *Helicobacter pylori*\(^{16, 17}\). Li *et al.* first reported that these bacterial species require MTAN for the synthesis of menaquinone, an electron transporter. Menaquinone is a crucial component of the electron transport chain in bacterial respiratory systems. Three different menaquinone biosynthetic pathways have been reported; the common menaquinone pathway, the futalosine pathway and the modified futalosine pathway\(^{18}\). The last pathway utilizes MTAN for converting 6-amino-6-deoxyadeno-fualosine into de-hypoxanthine futalosine (DHFL), and is employed by the pathogenic bacteria of *Campylobacter* and *Helicobacter* species for their survival\(^{17, 18}\). This vital role of *Helicobacter pylori* MTAN (*Hp*MTAN), in addition to previous roles suggest that this enzyme is an intriguing drug target.

Figure 1-2: Menaquinone biosynthetic pathway in *Helicobacter pylori*. Figure adapted from reference 24.
1.3 Active site of HpMTAN

To analyze the structural determinants of the HpMTAN active site, the Ronning lab has performed various X-ray co-crystallization experiments of HpMTAN with different substrates and substrate analogues. These experiments result in HpMTAN in an apo form \([HpMTAN\text{-HOH} \text{ (PDB: 3NM4)}]\), and complexed with tris(hydroxymethyl)aminomethane (Tris) \([HpMTAN\text{-TRS} \text{ (PDB: 3NM4)}]\), adenine \((HpMTAN\text{-ADE})\), adenine and Tris \([HpMTAN\text{-ADE-TRS} \text{ (PDB: 3NM6)}]\) and, formycin A; nonhydrolyzable analogue of MTA-fragment \([HpMTAN\text{-FMA} \text{ (PDB: 3NM5)}]\). Based on these structural analyses, the active site of HpMTAN is divided into three subsites consistent with other bacterial MTANs: purine binding site, ribose binding site, and 5’-alkylthio binding site \(^2\) (Figure 1-3).

1.3.1 Purine binding subsite

The crystal structure of HpMTAN-FMA and HpMTAN-ADE-TRS shows that the active site residues D198, V154, and F153 form interactions with an adenine moiety \(^2\). D198 and V154 form hydrogen bond interactions and F153 forms \(\pi-\pi\) interactions with the purine ring. The Oδ1 and Oδ2 of D198 hydrogen bonds with the N6 and N7 of amino groups of adenine, respectively, with a hydrogen bond distance of 2.8 and 2.7 Å. Additional hydrogen bonding involves interactions between backbone carbonyl oxygen and amide nitrogen of V154 and N6 and N1 of adenine, respectively, with distances of 3.1 and 3.0 Å (Figure 1-3). The comparison of HpMTAN-ADE-TRS and HpMTAN-TRS crystal structures suggests that, like \(E.coli\) MTAN \(^{19, 20}\), adenine binding promotes only a slight shift of loop connecting helix \(\alpha6\) to sheet \(\beta10\). This shift moves the D198 and F153
residues towards the active site and results into F153 adenine π-π interactions. The interactions of adenine with D198 and F153 residues supports its binding in the purine binding subsite. However, these interactions are unable to promote the closing of the active site, resulting in an “open form” unless ribose or a ribose-mimic is bound in the active site. The two forms of protein; “open form” and “closed form” of this protein will be discussed later in the section 1.7.

1.3.2 Ribose binding subsite

Upon binding of the potential substrate, amino acid residues in the ribose binding subsite make multiple interactions. These interactions play two important roles; they promote the closing of the active site that will be discussed in section 1.7 and stabilize the oxocarbenium intermediate (section 1.4). The ribose moiety makes four hydrogen bonds; three direct and one mediated by the nucleophilic water molecule. The 2’-hydroxyl of ribose forms the second hydrogen bond with the amide proton of M174 and Oε1 of E175. The third hydrogen bond is formed between the 3’-hydroxyl of ribose and Oε1 of E175 (Figure 1-3). The nucleophilic water molecule is positioned in the active site via interactions with R194, E175 and E13. E13 also hydrogen bonds with the 3’-hydroxyl of ribose through the nucleophilic water (Figure 1-3).
Figure 1-3: Active site of *HpMTAN*\(^2\). (A) Formycin A (FMA) bound structure of *HpMTAN* (PDB: 3NM5). The protein is represented in green cartoon. FMA is shown in CPK color with yellow carbon. (B) Adenine and Tris bound structure (PDB: 3NM6). Adenine and Tris is shown in CPK color with white carbon. **Purine binding subsite interactions**: Adenine moiety interacts with D198 and F153 shown in (A). Adenine base also interacts with the backbone of V154 shown in (B). **Ribose binding subsite interactions**: The ribosyl moiety of FMA forms hydrogen bond with E175 and M174. The nucleophilic water forms hydrogen bonds with R194 and E13 (A). The Tris is involved in extensive hydrogen bonding with M10, V78, E175, M174, R194 and E13 (B). The nucleophilic water molecule shown in FMA bound structure is replaced by hydroxyl group of Tris and forms hydrogen bond with E13 and R194 (B).
1.3.3 5'-alkylthio binding subsite

The 5'-alkylthio binding subsite has not been thoroughly explored previously. The comparison study of bacterial MTAN and its human homologue MTAP has suggested that the 5'-alkylthio binding subsite exhibits the major structural differences between these two enzymes\(^{21}\) (section 1.5). It is apparent that the interactions between the 5'-alkylthio binding subsite and homocysteine moiety are exploitable for designing bacterial specific drugs. In this work, we explore the interactions between the 5'-alkylthio binding subsite and homocysteine moiety using X-ray crystallography and kinetic analysis of \(Hp\)MTAN.

1.4 Enzyme mechanism

Allart \textit{et al.} first described the catalytic mechanism of \textit{Escherichia coli} MTAN (\(Ec\)MTAN)\(^{22}\). This study suggests that the MTAN hydrolysis follows an \(S_N 1\)-type mechanism that involves an oxocarbenium ion intermediate. In the \(H. pylori\) strain J99 MTAN, the aspartic acid residue at position 198 (D198) serves as a general acid and donates a proton to N7 of the adenine moiety of the substrate\(^2\). To compensate for the loss of electron density, the purine base withdraws electrons from the thioribose moiety. The electron flow towards the adenine moiety develops a partial positive charge on the thioribose moiety ultimately breaking the \(N\)-glycosidic bond and forming an oxocarbenium intermediate. This intermediate is attacked at the anomeric carbon by an activated nucleophilic water to form 5’-methylthioribose (MTR) (Figure 1-4). The nucleophilic water molecule is immobilized in the active site by R194, E175 and E13.
amino acid residues and activated by E13, which acts as a general base to remove the proton to promote nucleophilic attack$^{19,20,23,24}$.

**Figure 1-4:** Reaction mechanism of MTAN. The aspartic acid (D198 in *HpMTAN* strain J99) acts as a general acid that initiates the catalytic reaction. The enzymatic reaction begins on the upper left. The bond cleavage is followed by formation of oxacarbenium ion (step 2). Nucleophilic water attacks the anomeric carbon followed by product formation (step 3). The substrate molecule (SAH) replaces product formed in active site (step 4)$^{2,24}$. 

---

9
1.5 Human MTAP vs. *Hp*MTAN

Mammalian cells do not encode MTAN. Instead, the enzyme responsible for the catabolism of MTA in mammalian cells is 5’-methylthioadenosine phosphorylase (MTAP), which is homologous to bacterial MTANs. MTAP binds MTA very similarly to that observed for MTA binding in MTAN\(^{21}\). Because of the similarities between human MTAP and bacterial MTAN, compounds that inhibit MTAN activity, also bind to MTAP with comparable affinity\(^{25, 26}\). The development of new and specific drugs are required that can selectively inhibit MTAN but not affecting the activity of MTAP.

The dissimilarity in the active sites of *Hp*MTAN and MTAP arise due to differences in the reaction mechanism, as well as preference for substrate type. The *Hp*MTAN uses a water molecule as a nucleophile to attack the oxocarbenium intermediate, while MTAP utilizes inorganic phosphate as a nucleophile\(^{21}\) (Figure 1-5). This difference is reflected in the ribose-binding pockets of the two enzymes. The ribose-binding pocket of MTAP is formed by residues H61, R60, T18, T197 and T93 and therefore has a strong positive electrostatic potential\(^{21}\) (Figure 1-6). In contrast, the *Hp*MTAN ribose-binding site is composed of E13, E174 and R194 and has a weak negative electrostatic potential (Figure 1-6).
Figure 1-5: MTAP vs. MTAN reaction scheme. (A) The reaction schemes for MTAN utilizing MTA and SAH as a substrate. The water molecule (nucleophile) is shown in orange, and the proton donated by D198 is shown in blue. (B). The reaction scheme for human MTAP enzyme. The inorganic phosphate (nucleophile) is shown in orange. The MTAP derived proton is shown in blue. Figure adapted from reference 24.
Figure 1-6: MTAP vs. MTAN active site. (A) The gray cartoon represents the backbone of MTAP. The side chains shown in gray represent amino acid residues that interact with sulfate (sulfate mimics the phosphate in the active site). Immucillin-A (ImmA: MTA based inhibitor) has CPK coloring with orange carbon (PDB: 1K27). (B) The green cartoon represents the backbone of MTAN (PDB: 3NM5). The amino acid residues (green sticks) interact with nucleophilic water and FMA (CPK color with yellow carbon).
Another difference stems from the ability of MTAN to catalyze the hydrolysis of four adenosine-based substrates (Figure 1-1 and 1-2) that differ significantly in the nature of their 5’-alkyl moieties$^{24}$. Based on the availability of substrates, we used MTA and SAH for this study (Figure 1-4). Human MTAP does not utilize SAH as a substrate and hence lacks a sufficiently large 5’-alkylthio binding subsite required for binding of the homocysteine moiety$^{21}$. In contrast, bacterial MTAN has the capability of binding and catalyzing the hydrolysis of SAH$^{24}$.

Figure 1-7: MTAP active site has a truncated 5’-alkylthio binding subsite$^{26}$. The white transparent surface represents the one molecule of MTAP backbone. The active site is highlighted by the presence of ImmA (CPK color with orange carbon). (PDB: 1K27). The SAH molecule (CPK color with green carbon) is docked in the active site showing homocysteine chain protruding out of the truncated active site. The yellow transparent surface represents the other molecule of the MTAP.
1.6 Overall structure of *Hp*MTAN

The *Hp*MTAN is 25 kDa and the central core of each monomer consists of ten mixed β-sheets flanked by six α-helices (Figure 1-8). The analytical gel filtration chromatography experiments performed by Lee *et al.* suggest that the bacterial MTAN exists as a dimer. The previously solved crystal structures of *E.coli* MTAN and *Staphylococcus aureus* MTAN have also shown the homodimer form of this enzyme.

Figure 1-8: The *Hp*MTAN monomer structure. (A) Cartoon representation (PDB: 3NM5) (B) Topology of *Hp*MTAN. All the colored cylinders represent α-helix and colored arrows represent β-sheets. The line arrow (white or black) inside the cylinder directs from N-terminal to C-terminal of α-helix.
Likewise, the crystal structures of *Hp*MTAN solved by Ronning *et al.* exist as a homodimer\(^2\). The crossover loop of one monomer interacts with the active site of other monomer and vice-versa to form the dimer (Figure 1-9). Although no evidence is available yet, it is predicted that the interaction of the two monomers via the crossover loop plays a role in the allosteric regulation of MTAN.

![Figure 1-9](Image)

**Figure 1-9:** *Hp*MTAN exist as a homodimer\(^2\). One molecule is shown in orange. Other monomer is shown in green. The active site is highlighted by the presence of ligand (Formycin A; nonhydrolysable form of MTA) bound in it. The active site is shared between the two molecules of homodimer. (PDB: 3NM5)
1.7 Ligand induced conformational changes of HpMTAN

The structural studies of bacterial MTAN from different species have established that the enzyme exists in two main conformations: open and closed\(^2,20,28\). Except for a few minor differences, all bacterial MTAN share similar structural rearrangements during active site closing. Likewise, depending on the ligand bound, HpMTAN has shown to be either in open or closed conformation\(^2\). The crystal structure of HpMTAN-HOH (the apo structure of MTAN), HpMTAN-ADE represents the open forms whereas HpMTAN-TRS, HpMTAN-ADE-TRS, and HpMTAN-FMA represent closed forms of the enzyme.

Briefly, the enzyme undergoes two major conformational changes from open to closed form. The most significant structural change occurs in the loop connecting helix \(\alpha 6\) to sheet \(\beta 10\) and contains 200-208 amino acid residues (Figure 1-10). In the open form, the absence of electron density for this loop suggests that the loop is very dynamic and the active site is more exposed to solvent. In the closed form, residues 204-208 undergo a loop to helix conformational change. The conformational changes move D198 and D209 significantly and position them into the active site. The other significance of loop to helix transition is that it covers the top of the active site making it less exposed to solvent. The structural studies performed by Ronning \textit{et al.} suggest that the closing of active site requires the binding of ligand that can make hydrogen bond with E13, E175 and R194 simultaneously\(^2\). Exemplified by Tris (Figure 1-3 B), the hydroxyl groups of Tris forms hydrogen bonding interactions with E13, E175 and R194. The bridging of these three amino acids stabilizes the closed form of active site.
Figure 1-10: Ligand induced conformational change of HpMTAN$^2$. The two forms: open (HpMTAN-HOH) and closed (HpMTAN-FMA) are superimposed. The protein molecules are shown in white cartoon except the $\alpha_6$ helices. The $\alpha_6$ helices are indicated by the green in apo form and orange in FMA bound structure. The region representing the most significant conformational ($\alpha_6$) difference is highlighted in purple. The dotted green line indicated a disordered region in the open form. The FMA molecule is shown in CPK color with yellow carbon.
Chapter 2

Kinetic and structural insight into differential substrate binding to HpMTAN

2.1 Overview

Emergence of antibiotic resistant strains of Helicobacter pylori requires the development of novel drugs. MTAN, an enzyme absent in humans but essential for Helicobacter pylori is an excellent drug target. Present inhibitors of MTAN are based on the adenosine analogue or transition state analogue of adenosine. The adenosine based inhibitors bind to bacterial MTAN and human homologue (MTAP) with comparable affinity.\(^3,26,29-31\)

Here, we performed structural and kinetic characterization of interaction of homocysteine moiety of SAH and 5’-alkylthio binding subsite of HpMTAN. To explore these interactions, we crystallize HpMTAN with SAH. The crystal structures of MTAN with SAH have never been reported before. In Ronning’s lab, previous attempt to crystallize wild-type HpMTAN (wt-HpMTAN) with SAH was unsuccessful as Tris, a buffer component, replaced the product from the active site. To allow the trapping of SAH in the active site without being converted into product we made an inactive form of HpMTAN. To inactivate the HpMTAN, the aspartic acid at position 198, which acts as a general acid, was mutated to asparagine (D198N). This mutation allowed the trapping of
SAH/MTA in the active site without hydrolyzing it. Also, we crystallized wt-\textit{Hp}MTAN with SAH in the presence of 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) instead of Tris in the crystallization buffer. This resulted in the crystallization of an enzyme-product complex, wt-\textit{Hp}MTAN-S-ribosylhomocysteine-adenine (wt-\textit{Hp}MTAN-SRH-ADE).

Additionally, we evaluate the inhibitory profile of Tris. The crystallization study has shown that the Tris has a capability to bind in the ribose binding subsite forming multiple interactions\textsuperscript{2} (section 1.7). We performed an enzymatic assay to determine the inhibitory effect of Tris on \textit{Hp}MTAN activity.

2.2 Materials and methods

2.2.1 Mutagenesis

Site directed mutagenesis was performed to make the D198N mutant using a pET-32 based plasmid containing the wild-type \textit{pfs} gene (gene encodes for MTAN) from \textit{Helicobacter pylori} (strain J99) as a template. Amino acid Aspartic acid at position 198 was replaced by asparagine by using polymerase chain reaction.

2.2.2 Protein expression and purification

The template plasmid encodes a polyhistidine tagged thioredoxin-\textit{Hp}MTAN fusion protein containing a prescission protease cut site immediately \textit{N}-terminal to the first residue of wt-\textit{Hp}MTAN. Proteolysis with prescission protease produces a polyhistidine tagged thioredoxin and untagged \textit{Hp}MTAN. Sequencing of the plasmid
resulting from the mutagenesis experiment confirmed the presence of the D918N mutation.

The plasmid containing the mutated gene was used to transform BL21 (DE3) Rosetta cells (EMD Biosciences). Cultures of LB media containing 0.1 mM chloramphenicol and 0.3 mM ampicillin were incubated at 37 °C. Cells were induced after reaching an Abs<sub>600nm</sub> = 0.6-0.8 by the addition of 0.1 mM IPTG and incubated for 18-20 hours at 16 °C. Cells were harvested by centrifugation and resuspended in buffer A (20 mM Tris, pH 7.5, 0.5 M NaCl, 5 mM β-mercaptoethanol and 25 mM imidazole). The resuspended cells were lysed by sonication and then centrifuged at 15,000 x g. The supernatant was applied to a 5 mL HisTrap column equilibrated with buffer A (GE Healthcare). Elution of recombinant proteins was performed using a linear gradient of imidazole from 25 to 250 mM over 20 column volumes.

The fractions containing the purified proteins were treated with Prescision Protease and dialyzed overnight against buffer A. This protein sample was again applied to a HisTrap column (GE Healthcare) to selectively bind the cleaved histidine tag and the protease. The fractions containing HpMTAN were then subjected to size exclusion chromatography on a Hi-Load Superdex 200 column as a polishing step (GE Healthcare). All protein samples were analyzed for purity using SDS-PAGE. The absorbance at 280 nm was used to determine protein concentration using an extinction coefficient of 3105 M<sup>-1</sup> cm<sup>-1</sup>. 
Figure 2-1: SDS-PAGE gel analysis of HpMTAN. (A) Ni affinity chromatography. The protein around 36.0 kDa corresponds to MTAN with thioredoxin tag attached to it. (B) Tag removal. After the removal of thioredoxin tag protein runs around 25.0 kDa. (C) Size exclusion chromatography. Protein is without any impurity band.
2.2.3 MTAN steady-state kinetic assays

All assays were performed in triplicate in 100 mM HEPES and 50 mM KCl buffer at pH 7.5. For determining kinetic parameters, a stock of 1 mM SAH and 1 mM MTA was prepared in the same buffer. The substrates were serially diluted to produce a final concentration range of 5 µM to 125 µM in 100 µl of assay solution. The reduction of absorbance at a wavelength of 274 nm (Figure 2-2) was monitored for 15 minutes at 37 °C using a Biotek Synergy H4 plate reader (Winooski, VT, USA). The extinction coefficient of 1.6 mM⁻¹ cm⁻¹ was used for assays with both SAH and MTA²,²⁴.

Figure 2-2: Scheme for MTAN kinetic assay²⁴. The substrates (MTA and SAH) have maximum absorbance at 274 nm. As the substrates convert into products the absorbance at 260 nm decreases.
2.2.4 Tris inhibitory assay

For determining inhibition constants ($K_i$), the assay was performed in buffer (100 mM, 50 mM KCl, pH 7.5) in triplicate. Each 100 µl of reaction consisted of 25 µM of MTA and 25 nM of MTAN. The concentration of Tris was varied from 0 to 250 mM. The decrease in absorbance was monitored at 274 nm for 15 min.

2.2.5 Crystallization

For the crystallization studies, purified protein samples were dialyzed against the crystallization buffer (20 mM Tris pH 7.5, 0.2 mM TCEP, and 1 mM EDTA) and concentrated to 16 mg/ml by ultrafiltration (Millipore). Crystals of the MTAN-D198N mutant complexed with either MTA or SAH were grown by the hanging-drop vapor diffusion method. Crystallization drops containing 1 µl of well solution, 1 µl of MTAN-D198N (16 mg/ml) and 0.5 µl of SAH/MTA (10 mM) were equilibrated with 100 µl of well solution. The well solution for producing MTAN-D198N/MTA complex crystals contained 0.2 M magnesium chloride, 0.1 M HEPES pH 7.5 and 25 % w/v PEG 3350. The well solution for producing MTAN-D198N/SAH complex crystals contained 0.05 M magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5 and 30 % PEG-MME 550. Crystals of the wt- HpMTAN/SRH/adenine complex were produced in the same conditions as the MTAN-D198N/SAH complex crystals.

X-ray diffraction experiments were performed at the LS-CAT ID-D beamline at the Advanced Photon Source, Argonne National Labs. Data were indexed, integrated and scaled using HKL2000. Molecular replacement was performed using Evolutionary
Program for Molecular Replacement (EPMR) and PDB 3NM6 as a model\textsuperscript{2}. Rigid body refinement, followed by simulated annealing, B-factors and positional refinements were carried out using PHENIX\textsuperscript{33}. The model was manually corrected using COOT\textsuperscript{34}.

2.3. Results and discussion

2.3.1 Kinetic characterization of wt-MTAN

The initial velocity at the different concentration of SAH and MTA was obtained by converting change in absorbance per minute to concentration (µM) per minute. The initial velocity values allowed us to calculate $K_M$ and $V_{max}$ using Michaelis-Menten equation (equation 2.1). The $K_M$ value for SAH and MTA were calculated to be 10 ± 1 µM and 39 ± 5 µM, respectively (Figure 2-3). The $V_{max}$ for SAH and MTA were 3 ± 0.1 µM/min and 6 ± 0.3 µM/min, respectively. The turnover number ($k_{cat}$ values) of MTAN for SAH and MTA were 1.8 ± 0.1 sec\textsuperscript{-1} and 3.8 ± 0.2 sec\textsuperscript{-1}, respectively. The lower $K_M$ value for SAH reflects the higher affinity of SAH for \textit{HpMTAN}. In the equation, $K_M$ is Michaelis constant, $S$: substrate concentration, $V_i$: rate of reaction,

$V_{max}$ : maximum rate of reaction

$$V_i = \frac{V_{max} \times [S]}{K_M + [S]} \quad \text{(Equation 2.1)}$$
Figure 2-3: Kinetic characterization of *Hp*MTAN. MTA as a substrate (upper panel) and SAH as a substrate (lower panel).
2.3.2. Tris inhibitory assay

The initial velocity of reaction in absence and in presence of varying concentration of inhibitors were used to calculate inhibitory concentration ($K_i$). The $K_i$ was calculated using equation number 2.2, adapted from reference number 13. The $K_i$ was calculated to be 25.24 mM.

\[
\frac{V'_S}{V_s} = \frac{K_M + [S]}{K_M + [S] + \frac{K_M \times [I]}{K_i}}
\]

(Equation 2.2)

Where, $K_M$: Michaelis constant, $S$: substrate concentration, $I$: Inhibitor concentration, $V'_S$: rate of reaction in absence of inhibitor, $V_s$: rate of reaction in the presence of inhibitor.

![Figure 2-4: Inhibition profile of HpMTAN by Tris](image)

Figure 2-4: Inhibition profile of HpMTAN by Tris
2.3.3. Crystallization studies

We solved three crystal structures: MTAN-D198N with SAH, MTAN-D198N and MTA, and, wt-MTAN with SRH-ADE. These crystal structures were resolved to 1.2, 1.63 and 1.50 Å respectively. These crystal structures have one monomer in the asymmetric unit, related to the other monomer by a twofold axis of crystallographic symmetry. Both protein molecules in the homodimer represent the closed form of \textit{HpMTAN}. Figure 2-5 shows the dimer of MTAN-D198-SAH.

![Diagram of MTAN-D198N-SAH dimer]

Figure 2-5: An obligatory dimer of MTAN-D198N-SAH\textsuperscript{24}. The protein crystallizes in a dimer form. One molecule of dimer is shown in green and other is shown in orange. The active sites are highlighted by the presence of SAH. The crossover loop of one protein molecule interacts with the SAH bound in the active site of other molecule.
The $F_o$-$F_c$ difference map calculations generated the strong electron densities in to the active sites, allowed us to model the respective ligands unambiguously (Figure 2-6, 2-7, and 2-8). The adenine and ribose moieties make similar interactions with adenine binding and ribose-binding subsites as discussed before in the case of HpMTAN-FMA.

<table>
<thead>
<tr>
<th>Resolution range (Å) (highest shell)</th>
<th>MTAN-D198N-SAH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P2₁2₁2₁ (122-120)</td>
</tr>
<tr>
<td>$a$, $b$, $c$ (Å)</td>
<td>80.7, 87.2, 67.2</td>
</tr>
<tr>
<td>Total reflections (unique reflections)</td>
<td>822470 (75266)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>95.0 (92.6)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>10.9</td>
</tr>
<tr>
<td>Average $I/\sigma (I)$</td>
<td>16.9</td>
</tr>
<tr>
<td>$R_{sym}$ (%) (highest shell)</td>
<td>5.0 (30.0)</td>
</tr>
<tr>
<td>Atoms/A.S.U</td>
<td>2,218</td>
</tr>
<tr>
<td>$R_{wr}$ (%)</td>
<td>15.4</td>
</tr>
<tr>
<td>$R_{free}$ (%)</td>
<td>17.5</td>
</tr>
<tr>
<td>Average B-factor protein (Å)</td>
<td>16.6</td>
</tr>
<tr>
<td>Average B-factor ligand (Å)</td>
<td>14.07</td>
</tr>
<tr>
<td>Average B-factor solvent (Å)</td>
<td>27.01</td>
</tr>
<tr>
<td>Ligand occupancy (%)</td>
<td>1000</td>
</tr>
<tr>
<td>r.m.s.d. bonds (Å)</td>
<td>0.005</td>
</tr>
<tr>
<td>r.m.s.d. angles (°)</td>
<td>1.09</td>
</tr>
<tr>
<td>Ramachandran favored (%)</td>
<td>96.9</td>
</tr>
<tr>
<td>Ramachandran disallowed (%)</td>
<td>0.0</td>
</tr>
<tr>
<td>Coordinate Error (Å)</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Figure 2-6: Crystal information of MTAN-D198N-SAH. Top: The crystal of MTAN-D198N-SAH (left). The chemical structure of SAH (right). Bottom: The diffraction and refinement statistics of MTAN-D198N-SAH (left). The SAH is modeled in the $F_o$-$F_c$ difference map present in the active site.
Figure 2-7: Crystal information of MTAN-D198N-MTA\textsuperscript{24}. Top: The crystals of MTAN-D198N-MTA (left). The chemical structure of MTA (right). Bottom: The diffraction and refinement statistics of MTAN-D198N-MTA (left). The MTA is modeled in the $F_o$-$F_c$ difference map present in the active site.
Figure 2-8: Crystal information of MTAN-SRH-ADE\textsuperscript{24}. Top: The crystal of MTAN-SRH-ADE (left). The chemical structure of SRH and adenine (right). Bottom: The diffraction and refinement statistics of MTAN-SRH-ADE (left). The adenine and SRH are modeled in the F_o-F_c difference map present in the active site.
2.3.3.1 Interactions between homocysteine chain and 5’-alkylthio binding subsite

Unlike the adenine and ribose binding subsites, both molecules of the homodimer form the 5’-alkylthio binding subsite. The homocysteine moiety of SAH extends away from the deeply buried active site of one molecule (monomer A) and makes significant interactions with the other molecule (monomer B) of the homodimer (Figure 2-5).

The amino acid residues forming the 5’-alkylthio binding site are D209 from monomer-A, F107, and H109 residues from monomer-B (Figure 2-9). The α-amino group of the homocysteine moiety forms cation-π interactions with the aromatic side chain of residue F107 from monomer B. The distance between the nitrogen of α-amino group and the aromatic ring of F107 is measured as 3.3 Å. The cation-π interactions have gained interest since in the past decade. Like other noncovalent interactions such as hydrogen bond, ionic and hydrophobic forces, cation-π interactions play an important role in protein stability and ligand recognition.

In general, cationic NH groups of lysine/arginine form electrostatic interactions with π electrons of phenylalanine/tyrosine/tryptophan. The other protein-ligand systems exemplified by acetylcholine binding protein-acetylcholine complex and γ-aminobutyric acid (GABA) receptor-GABA complex suggest the importance of this interaction in MTAN\textsuperscript{35, 36}. In the case of acetylcholine binding protein-acetylcholine complex, the quaternary amine forms multiple cation-π interactions with aromatic residues of acetylcholine receptor proteins.
The other interaction involves residue D209 forms a water-mediated hydrogen bond with the α-amino group of the homocysteine moiety. The distances from Oδ1 of D209 to the water molecule, and from the water molecule to the α-amino group were observed as 2.7 Å each. The D209 residue is located in the loop that connects the N-terminus of helix α6 to sheet β10 and remains disordered in the ligand-free active site of MTAN. The conformational change of loop to helix rearranges the active site. First, D198 is positioned so that it may initiate the catalysis. Second, D209 is positioned to enable interactions with the ligand. The third interaction observed between the α-carboxyl moiety and the Ne of H109 residue of monomer B was measured to be 2.8 Å.

Figure 2-9: Interactions between homocysteine chain and 5’-alkylthio binding subsite. The SAH (stick, CPK with white carbon) bound in to the active site of molecule (orange). The green cartoon represents the crossover loop from the other molecule.
To inspect the significance of these amino acid residues in the binding site, one thousand closely related MTAN protein sequences were aligned (Figure 2-10). This sequence alignment highlights the presence of phenylalanine at the position 107 (numbering according to HpMTAN), which was conserved in about 85% of all MTAN sequences. This high level of conservation exhibits the importance of the F107 in the 5′-alkythiobinding site and supports our observation of cation-π interaction with the homocysteine chain. It is expected that F107 may be required for making π-π interactions with the benzoate ring of 6-amino-6-deoxyfutalosine, a substrate required for menaquinone biosynthesis in HpMTAN (Figure 1-2).

The sequence comparison of 1000 MTANs sequences also highlighted the importance of polar acidic residue at position 209 in the active site. It is observed that 90% of MTANs have either aspartate or glutamate at this position suggesting the possibility of ionic interactions with the α-amino group of the homocysteine moiety of SAH. In contrast, the residue at position 109 is shown to be conserved as either a tyrosine or histidine. About 80% of the sequenced MTANs have tyrosine residues. The tyrosine residues are thought to be involved in hydrogen bonding with the α-carboxyl group of the homocysteine moiety. In the sequence alignment analysis, 10% of the 1000 MTANs belonged to the order Campylobacterales. All of these possess histidine at position 109, suggesting an ionic interaction with the α-carboxyl group of the homocysteine moiety.
Figure 2-10: The sequence alignment of 1000 related MTANs. Only the region consisting of amino acid residue from 100 to 115 (upper panel) and from 194 to 210 (lower panel) is shown. The green highlighted region corresponds to amino acids F107 and Y109 (shown as green carbon in Figure 2-9). The orange highlighted region corresponds to amino acids D198 and D209 (shown as orange carbon in Figure 2-9). Figure adopted from reference 24.
2.3.3.2 Comparison of \(H_{p}MTAN-D198N-SA\)H with \(H_{p}MTAN-D198N-MTA\)

\(H_{p}MTAN\) has a capability to bind four different substrates that differ in their 5’-alkyl functional group. To assess the role of the 5’-alkylthio binding subsite in substrate selectivity, we determined the MTAND198N-MTA structure and compared it with MTAND198N-SA. The comparative study indicates that all three binding subsites, including the 5’-alkylthio binding subsite, superimpose perfectly. The identical structural conformation suggests that the homocysteine moiety binding does not induce any conformational changes in the 5’-alkylthio binding subsite, following the lock and key model (Figure 2-11).

Moreover, a comparative study of the 5’-alkylthio binding subsite highlighted a conserved hydrogen bonding pattern, regardless of substrate type. The active site of MTAND198N-MTA structure has two water molecules, replaced by the \(\alpha\)-amine and \(\alpha\)-carboxyl moieties of homocysteine chain in the MTAND198N-SA structure. The first water molecule forms hydrogen bond with D209 via another water molecule, similar to the \(\alpha\)-amine and D209 interaction in MTAND198N-SA. The second water molecule is hydrogen bonded with H109 and the first water molecule. This creates a solvent network in to the active site. It is evident that homocysteine chain binding causes ligand and active site desolvation, resulting in a gain in entropy. However, change in enthalpy is unexpected, as the pattern and number of interactions remained unchanged.
Relation to steady state kinetics

To examine the effect of additional interactions between the homocysteine moiety of SAH and the 5’-alkythiobinding subsite, steady state kinetics were performed (section 2.3.1). The $K_M$ values for SAH and MTA were found to be 10 ±1 and 39 ±5 µM, respectively. The lower $K_M$ value for SAH was anticipated from the structural information. Three additional interactions and no major loss of entropy upon binding of the homocysteine moiety contribute to the higher affinity for SAH over MTA.

Figure 2-11: The comparison of SAH and MTA bound in the active site$^{24}$. The SAH is shown in CPK color with white carbon. Fo-Fc omit map represents MTA. The dark green protein residues belongs to MTAN-D198N-SAH and light green bound structure belongs to MTAN-D198N-MTA.
2.3.3.3 Comparison of \textit{Hp}MTAN-D198N-SA\textit{H} with \textit{Hp}MTAN-SRH-ADE

To confirm that the structural features observed in MTAND198N-SA\textit{H} complex are not a consequence of protein mutation, we determined the crystal structure of wt-\textit{Hp}MTAN with product bound ternary complex. This structure was also assessed for any conformational changes in the 5'-alkylthiobinding region that may result in substrate hydrolysis. The co-crystallization experiment of wt-\textit{Hp}MTAN with SA\textit{H} resulted in formation of the ternary enzyme-product complex, \textit{Hp}MTAN-SRH-ADE. As expected, the Fo-Fc map calculation showed two discrete density regions in the active site. This allowed us to place adenine and SRH into the respective electron density map (Figure 2-8).

To inspect the conformational changes in the enzyme active site and in the substrate following hydrolysis, we superimposed MTAND198N-SA\textit{H} and wt-MTAN-SRH-ADE (Figure 2-12). This comparative analysis highlights that the ribose moiety undergoes significant conformational changes, whereas the adenine and homocysteine moieties remain unaffected. During catalysis, ribose anomerization converts the $\beta$-ribosyl moiety of the substrate into the $\alpha$-ribosyl moiety of the product as seen in MTAND198N-SA\textit{H} and wt-MTAN-SRH-ADE, respectively. This conformational change is in accordance with the proposed catalytic mechanism, where cleavage of $N$-glycosidic bond produced an oxocarbenium ion intermediate.

Further, the ordered nucleophilic water attacks the anomeric carbon and is converted into the C1’-hydroxyl group of the $\alpha$-SRH product. Interestingly, the
interactions of E13, E175 and R194 residues with the ordered water molecule remains conserved even after hydrolysis, whereupon it forms a covalent bond with C1’ of the SRH product. To maintain the position of the nucleophilic water, ribose undergoes a change from a C4’-endo conformation to a C2’-endo conformation. It is presumed that the conformation changes in the ribosyl moiety are assisted by the conformational flexibility of the ribose binding site of the enzyme. No conformational differences were observed in homocysteine moiety or 5’-alkylthiobinding subsite, suggesting less flexibility provided by these regions.

**Relation to steady state kinetics**

The $k_{cat}$ values for SAH and MTA were calculated to be $1.8 \pm 0.1$ and $3.8 \pm 0.2 \text{ s}^{-1}$, respectively (section 2.3.1). The higher turnover rate for MTA compare to SAH can be supported by the crystallographic result of wt-MTAN-SRH-ADE. As the 5’-alkylthiobinding subsite residue makes additional interactions with the SRH product, it may slow down the release of product from the active site (Figure 2-12). This would result in a lower turnover rate.
Figure 2-12: The comparison of substrate and product bound in the active site. The SAH is shown in CPK color with white carbon. The product complex is shown in bronze carbon. The steady state kinetic indicates the lower release of product when SAH used as a substrate. This can be expected due to the extra interactions form by homocysteine moiety of SAH shown in upper panel.

\[ k_{\text{cat}} = 3.8 \pm 0.2 \text{ sec}^{-1} \]

\[ k_{\text{cat}} = 1.8 \pm 0.1 \text{ sec}^{-1} \]
2.4. Conclusion

Presented research is the first report describing interactions between 5’-alkylthiobinding subsite and the homocysteine moiety of SAH in bacterial MTAN. We successfully crystallized and solved three structures of *Hp*MTAN with its two substrates (SAH/MTA) and products (SRH and adenine). This work highlighted three specific interactions formed between the homocysteine moiety of SAH and the 5’-alkylthiobinding site. These three interactions can be exploited for designing bacterial-specific inhibitors. The comparative structural analysis enlightened that enzyme exhibit rigidity in the 5’-alkylthiobinding site, and thus contributes lesser entropic penalty on the homocysteine moiety binding. These structural evidences were further supported by observing a lower $K_m$ value of *Hp*MTAN for SAH than MTA.

To highlight the potential of targeting the homocysteine binding subsite, we compared the MTAND198N-SA and human MTAP. This structural comparison clearly exhibits that the smaller 5’-alkylthio binding of MTAP is unable to accommodate the homocysteine moiety of SAH. In conclusion, our combined results suggest that the 5’-alkylthio binding site indeed provides an excellent target for designing bacterial specific drugs. SAH based inhibitors would not only provide higher affinity for bacterial MTAN but also likely exhibit low or no affinity for MTAP. Further, observation of the Campylobacterales-specific interactions between the α-carboxyl moiety and the Ne of H109 residue could be exploited to discriminate between *Hp*MTAN and commensal bacterial MTANs.
Chapter 3

Development of high-throughput binding assay to identifying HpMTAN inhibitors

3.1 Rational for assay development

As research continues to discover novel roles of HpMTAN in different metabolic pathways, HpMTAN has gained interest as a potential drug target. Current inhibitors of MTAN are either adenosine analogues or transition-state analogues, and these inhibitors exhibit nanomolar affinity towards bacterial MTANs\textsuperscript{26, 30}. However, these inhibitors are unable to discriminate between bacterial MTAN and human MTAP; therefore MTAP is also bound with comparable affinity\textsuperscript{30, 31}. From a pharmacological point of view, adenosine-based inhibitors are not considered good drug candidates\textsuperscript{38, 39}. These inhibitors may bind non-specifically to human proteins such as MTAP, adenosine kinase and adenosine receptors, resulting in severe, toxic side effects for the patient\textsuperscript{40, 41}.

Because of this, new inhibitors that possess a non-adenosine framework and can bind specifically to HpMTAN are needed. The aim of this study is to use large diverse chemical libraries to identify a pool of non-adenosine compounds that can inhibit HpMTAN activity. The potential inhibitors can be further optimized to drug that inhibit
*Hp*MTAN selectively and cause less toxicity. Here, we developed a high-throughput binding assay that will allow rapid screening of compounds that bind to *Hp*MTAN. Current methods for inhibition study of MTAN include isothermal calorimetry\(^{28, 42}\), UV-absorption based assays\(^{24, 29}\), and radioactive labeled ligand based assays\(^{23}\). Isothermal titration calorimetry requires a large quantity of protein and is therefore not suitable for high-throughput screening. UV-absorption assays require low amount of proteins and can be performed in multi-well plate format. However, UV-absorption assays result in a high frequency of false-positive hits, as a large number of compounds in chemical libraries have aromatic moieties and thus have similar spectral properties as adenosine based substrates. The major disadvantages associated with radioisotope-based assays are that they create hazardous waste, slow and are too expensive to efficiently screen large numbers of compounds. Thus, it is imperative to develop an assay that can fulfill a majority of the characteristics suitable for high-throughput screening.

Here we developed and optimized a fluorescence polarization (FP) binding assay that is sensitive, inexpensive, require a very low amount of protein, is highly reproducible, nonhazardous and can be performed in a 384 well plate format. These features make the FP binding assay suitable for screening of thousands of compounds per day.
3.2 Approach for assay development

The development of the FP binding assay was guided by the MTAN-D198N-SAH crystal structure. The HpMTAN crystal structure shows the 5’-alkylthio binding subsite and its proximity to the protein surface is sufficiently large to accommodate a molecule of organic dye such as Alexa Fluor. We synthesized a water-soluble fluorescent probe, which we refer to as SAH-488, composed of a SAH molecule coupled with Alexa Fluor 488 dye. To synthesize SAH-488, we used an Alexa Fluor 488 sulfodichlorophenol ester (SDP), which is an amine reactive fluorescent dye that reacts with the α-amino group of SAH forming a stable amide bond. The advantageous features of the probe are: stability in aqueous medium and insensitivity to pH variation, which provide for the consistent assay results. Also, the close resemblance of SAH-488 to the substrate (SAH) offers a probe with high affinity and selectivity for MTAN.

Figure 3-1: Scheme of SAH-488 synthesis.
The presented FP binding assay requires the inactive form of enzyme, MTAN-D198N mutant to accomplish the binding of SAH-488 without hydrolyzing it. Since the α-amino group of the homocysteine moiety is not available to form either the cation-π interaction with F107 or the hydrogen-bonding interaction with D209 (section 2.3.3.1). However, it is likely that the highly conjugated ring system of the fluorescent dye may engage in π-π interactions with the F107 amino acid residue.

This assay is based on the general principle of fluorescence-polarization (FP) binding assays (Figure 3-2). In these assays, a small molecular probe (SAH-488 in this case) tumbles randomly in the reaction buffer. When polarized light travels through the reaction mixture, SAH-488 molecules absorb light and reach an excited state. The randomly tumbling SAH-488 molecules rotate before emitting light and hence the plane of the emission light differs from the excitation light, resulting in depolarized light emission. However, when bound to a large molecule, such as MTAN-D198N, the rotation of SAH-488 molecules slow down. This results in the emitted light existing in the same plane as the excitation light. The degree of polarization can be calculated using the following equation 3.1\(^2\). Where, \(I_\perp\) and \(I_\parallel\) represents the fluorescence intensity in the parallel plane and perpendicular plane (with respect to the plane of excitation light), respectively.

\[
P = \frac{I_\perp - I_\parallel}{I_\perp + I_\parallel} \quad \text{(Equation 3.1)}
\]

In the equation, the free state of SAH-488 results in a low polarization value and as the concentration of SAH-488 bound to MTAN-D198N increases, the value of
polarization increases. To perform the FP binding assay, we used a constant concentration of SAH-488 while increasing the concentration of MTAN-D198N until saturation is achieved. For inhibitor screening, compounds that bind to the active site of MTAN-D198N will compete with SAH-488 and displace it from the active site. The displacement of SAH-488 will result in low fluorescence polarization. Therefore, the binding of a compound competitive for SAH-488 can be observed by measuring a decrease in polarization.

Figure 3-2: Principle of FP binding assay. (A) The SAH molecule is shown as gray stick model. The green star represents the dye. In free state it rotates faster and emit light in different plane as excitation light. (B) The SAH-488 bound in to the active site of MTAN-D198N. The binding slows down the rotation of SAH-488 resulting in increased polarization of light.
3.3 Counter-screen assay

The follow-up assay was utilized to validate hits (compounds that cause low polarization value in FP competitive binding assay) from the chemical libraries. To confirm enzymatic inhibition by compounds, an Amplex Red-Xanthine oxidase assay was employed\textsuperscript{44}. The assay involves the oxidation of Amplex Red to resorufin which is a fluorescent molecule having excitation and emission wavelength of 571 and 585 nm. The assay involves the following steps; MTAN catalyzes the de-purination of SAH releasing adenine as a product. Xanthine oxidase oxidizes adenine converting it to uric acid and hydrogen peroxide (\(\text{H}_2\text{O}_2\)). Horseradish peroxidase utilizes \(\text{H}_2\text{O}_2\) to oxidize Amplex Red to resorufin.

Figure 3-3: Scheme of Amplex Red-Xanthine oxidase assay.
3.4 Material and methods

3.4.1 Protein preparation and adenine removal from *Hp*MTAN active site

The MTAN-D198N was prepared as discussed in the section 3.2.2. In addition, the protein was dialyzed against activated charcoal for two days and two nights in 20 mM HEPES (pH 7.5). This dialysis was performed to remove adenine from the active site of the enzyme. Adenine is present in bacterial cell environment and binds to MTAN during protein expression in the cell. Adenine binds to MTAN so tightly that it remains bound throughout all protein purification steps, resulting in an elevated 260/280 absorbance ratio of purified protein.

3.4.2 Fluorescence labeling of SAH

To synthesize the SAH-488 probe, 0.6 mg of Alexa Fluor 488 5-SDP ester (Life technology) was dissolved in 15 µl of DMSO immediately before the labeling reaction. 0.7 mg of SAH (Cayman Chemical) was dissolved in 285 µl of 0.1 M sodium bicarbonate buffer (pH 8.3). The solutions were combined, placed on a low-speed oscillating shaker and incubated for 6 hours at room temperature. The reaction tube was kept covered with aluminum foil to avoid photobleaching of the Alexa Fluor. To ensure the thorough mixing of reactants, the reaction tube was vortexed gently every 30 minutes for first two hours of reaction.
3.4.3 HPLC purification of SAH-488

Labeled SAH was purified by high-performance liquid chromatography (Waters 2487) with a reverse phase analytical column (BDS HYPERSIL C18, 4.6 X 150 mm). The mobile phase A and B consist of 1% TFA in water and acetonitrile, respectively. The SAH-488 was separated from the reaction mixture by running a linear gradient of buffer B at the flow rate of 1 mL/min. The peaks were monitored using dual wavelengths of 260 and 488. The λ\text{max} of SAH and Alexa Fluor 488 are 260 nm, and 488 nm, respectively. Overall two major peaks were observed: the first peak was observed at 260 nm but not at 488 nm, which is the assigned absorption for SAH. The second peak has strong absorbance at 488 nm with some absorbance at 260 nm (Figure 3-4).

3.4.4 Characterization and concentration determination of SAH-488

The second peak containing SAH-488 was collected, concentrated using a SpeedVac and subjected to high-resolution nanospray mass spectrometry with the assistance of Dr. Jingshu Guo (Dr. Wendell Griffith’s lab). The concentration of SAH-488 was determined by measuring the absorbance at 488 nm using a Genesys 6 UV-Vis spectrophotometer (Thermo Electron Corporation). The Beer-Lambert equation was used to calculate the concentration using an extinction coefficient of 71,000 cm\(^{-1}\) M\(^{-1}\). The crude reaction mixture and purified SAH-488 samples were stored at -20 °C until needed (Figure 3-5).
3.4.5 SAH-488 binding assay

All FP experiments were carried out in black, 384-well microplates (Corning) using a Synergy H4 hybrid multi-mode plate reader (BioTek). The excitation and emission wavelengths used to measure the fluorescent polarization were 485 nm and 528 nm, respectively. The binding experiment was performed in triplicate in reaction buffer (20 mM HEPES, 50 mM KCl, pH 7.5) at room temperature. The stock of 50 nM SAH-488 was prepared in this reaction buffer. In each 50 µl of reaction, 10 µl of SAH-488 stock solution was added to reach final concentration of 10 nM SAH-488. The concentration of MTAN-D198N was varied from 0 nM to 750 nM. During all FP assays, one reaction well containing just buffer was used a blank. The fluorescence polarization was measured after 15 min of incubation at room temperature (Figure 3-6).

3.4.6 SAH displacement assay

The assay was performed in triplicate using constant concentrations of MTAN-D198N and SAH-488 at 70 nM and 10 nM, respectively. The concentration of SAH was varied from 0 µM to 2.5 µM. The appropriate amount of buffer was added to make up the 50 µl reaction volume. The concentration of MTAN-D198N used was guided by the saturation curve from SAH-488 binding assays. The concentration of MTAN-D198N was chosen to be sure of having 50-80 % of protein–probe complex formation (Figure 3-7).
3.4.7 Assay validation

To assess reproducibility and consistency of this assay the Z’ factor (define in section 3.5.4) was calculated in two different days. Each day eighteen reactions were set up, nine reactions as positive controls and nine reactions as negative controls. For positive controls, the 50 µl reactions contained 70 nM of MTAN-D198N, 10 nM of SAH-488 and 1% DMSO. For negative controls, each 50 µl reactions contained 10 nM of SAH-488 and 1% DMSO. The appropriate amount of buffer was added to reach the 50 µl reaction volume. The fluorescence polarization was measured after 15 min of incubation at room temperature. The observed values were used to calculate the mean and standard deviation of the Z’ factor.

3.4.8 Screening of compounds of NIH clinical collection

To screen compounds from the NIH clinical collection, each 50 µl sample reaction was assembled containing 70 nM of MTAN-D198N, 10 nM of SAH-488 and 100 µM of tested compound. The concentration of compounds provided in the NIH clinical collection are 10 mM each dissolved in 100% of DMSO. Adding 0.5 µl of compound to each 50 µl sample reaction to reach the final concentration of 100 µM. The assay is suitable to screen available 446 compounds in one day but to evaluate the interday variation in signal, the assay was performed in two separate days.
3.4.9 Counter-screen assay

We used a commercially available Xanthine Oxidase Amplex Red kit (Molecular probes) to monitor the production of adenine. The stock solutions of following solutions were prepared as given in the kit manual: 10 mM Amplex red reagent (AR), 100 U Horseradish Peroxide (HRP) and 10 U xanthine oxidase. The aliquots for Amplex Red regent and Horseradish Peroxide were stored in -20 °C until required. All reactions were performed using 20 mM HEPES pH 7.5, 50 mM KCl, pH 7.5 at room temperature.

To make sure that the MTAN is a rate limiting enzyme we optimize the MTAN concentration for the assay, the 25 µl reactions were set up to contain 12 µM SAH, 50 µM AR reagent, 0.2 U HRP, 300 U xanthine oxidase. The concentrations of MTAN were varied from 1.25 nM to 5 nM. For inhibition experiments, the 25 µl reactions were set up to contain 12 µM SAH, 50 µM AR reagent, 0.2 U HRP, 300 mU xanthine oxidase, 2.5 nM MTAN. The drug concentrations were varied from 10 to 100 µM.
3.5 Results and discussion

3.5.1 SAH-488 synthesis, purification and characterization

Two peaks were observed, one for SAH at 260 nm and a second peak representing SAH-488 at 488 nm in addition absorbance at 260 nm. This can be explained by the fact that SAH was used in excess of Alexa dye during synthesis. The two peaks were sufficiently separated to afford purification of the desired product. The second peak eluted in approximately 20% acetonitrile (Figure 3-4). When characterizing the product, ESI-MS data indicated that the sample contained a single solute exhibiting a 901.05 m/z, which is equivalent to the mass of SAH-488 shown in the Figure 3-5.

Figure 3-4: HPLC chromatogram of SAH-488 purification. The blue line represents the absorbance at 488. The red line represents the absorbance at 260. The green line is a gradient of mobile phase B.
Figure 3-5: ESI-MS of purified sample of SAH-488. The corresponding molecule of SAH-488 is shown in inset.
3.5.2. SAH-488 binding constant

In FP binding assays the concentration of SAH-488 was fixed at 10 nM and the MTAN-D198N concentration was increased until the binding reached saturation (Figure 3-6). The fluorescence polarization values (mP) for each concentration of MTAN-D198N were measured and used to calculate (using equation 3.2) the dissociation constant ($K_d$). The data was fitted to give $K_d$ value of 51.24 ± 4.11 nM using GraphPad Prism 4.0.

\[
K_d = \frac{[R][L]}{[RL]} \quad \text{(Equation 3.2)}
\]

In the equation, $R$ represents the concentration of free receptor (MTAN-D198N) and $L$ represents free ligand concentration (SAH-488). The $RL$ represents the complex of protein and ligand (MTAN-D198N and SAH-488).

![Figure 3-6: FP binding assay of SAH-488 and MTAN-D198N.](image)

Figure 3-6: FP binding assay of SAH-488 and MTAN-D198N.
3.5.3. SAH inhibition constant

To determine the IC$_{50}$ for SAH, dose dependent experiment was carried out. The concentration of SAH-488 and MTAN-D198N complex was kept constant. The concentration of untagged SAH was increased gradually. As the concentration of SAH increased, the decrease in FP values was observed. This suggests an SAH concentration dependent displacement of SAH-488 from the active site. The data [log (SAH) versus FP values for corresponding concentration of SAH] was fitted to nonlinear regression using GraphPad prism 4.0. The IC$_{50}$ for SAH was calculated to be 54.76 nM.


3.5.4 Assay validation

The quality of our developed assay was evaluated by calculating the statistical parameter $Z'$ in the fashion proposed by Zhang et al. To calculate the $Z'$ factor\textsuperscript{45, 46}, the experiment was performed using positive controls and negative controls. The positive control reactions produce a signal that is statistically different (higher or lower depending on the assay type) from the negative controls (background). In general, the positive control for any FP binding assay contains a constant concentration of a labeled ligand and the target protein. The complex of the labeled ligand and protein produce a significantly higher signal than negative control (Figure 3-8).

The negative control provides the background value that is associated with the labeled ligand, buffer type or pH of reaction. For FP binding assays the negative control can be designed based on one of two possibilities\textsuperscript{45, 46}. For one option, the negative control consists of the free labeled ligand without any protein present in the reaction. Another option possibility is to use a complex of the labeled ligand bound to the protein, but in the presence of a known strong inhibitor. Since the strong inhibitor will displace the labeled ligand from protein binding site, there will be a significantly lower polarization value. This type of negative control is appropriate if complete displacement of the labeled ligand occurs. The data collected from positive controls (signal) and negative controls (background) are used to calculate the signal-to-background ratio of a particular assay.
The statistical parameter, “Z’ factor” can be utilized to calculate the signal-to-background ratio\textsuperscript{45}. This parameter takes an account of mean and standard deviation values associated with the positive and negative control datasets. To calculate the Z’ factor the following equation is used\textsuperscript{45}:

\[
Z' = 1 - \frac{3\sigma_+ + 3\sigma_-}{|\mu_+ - \mu_-|} \quad \text{(Equation 3.3)}
\]

In the equation, $\sigma_+$ and $\sigma_-$ represent the standard deviation for the positive control and negative controls, respectively. Meanwhile, $\mu_+$ and $\mu_-$ represent the mean values for the positive and negative controls, respectively.

The Z’ values within a range of 0.5 to 1 represent an assay with good to excellent quality\textsuperscript{45}. The assays with a Z’ value closer to unity are higher quality, more reliable, optimized, and are considered useful for screening libraries. The Z’ values below 0.5 represent a low quality assay or an assay that requires further optimization. Here, to evaluate the quality and reproducibility of our described FP assay, we performed the assay across two different days to determine the inter-day signal variation. Each day eighteen reactions, nine positive controls (10 nM SAH-488 and 70 nM MTAN-D198N) and nine negative controls (10 nM SAH-488 in absence of MTAN-D198N) were performed. As the majority of chemical libraries are stored in DMSO, we added an equivalent concentration of DMSO to each control reaction (typically 1% DMSO). The Z’ factor for the FP binding assay was calculated using equation no. 3.3. The inter-day Z’ factor for the developed assay was calculated to be 0.86.
Figure 3-8: FP assay evaluation. The red boxes represent the polarization values from positive control reactions. The green boxes represent the polarization values from negative control reactions. The three times standard deviation is shown by the black line. The difference between the three times standard deviation of negative control and positive control represent the “signal window”[45].
3.5.5 Screening of the NIH clinical collection

Another statistical parameter, Z factor\textsuperscript{45}, was calculated to assess the affect of different compounds on the described assay to determine the suitability of the assay for high-throughput screening. The compounds from chemical libraries may cause a high level of signal variability, resulting in a Z-factor lower than 0.5. A Z-factor below 0.5 indicates that the assay is not suitable for high-throughput screening without further optimization. To optimize an assay for screening chemical libraries, careful analysis of the intrinsic properties of chemicals in the library is necessary.

We screened 446 compounds from the NIH clinical collection. Negative controls (10 nM SAH-488 in absence of MTAN-D198N) were used for each plate. As shown in Figure 3-8, the negative controls provide lower polarization values. The mean and standard deviations for the negative control dataset were calculated. The higher polarization values are associated with the “sample reaction”\textsuperscript{45}, which consists of 10 nM SAH-488, 70 nM MTAN-D198N and 100 µM of a compound from the NIH clinical collection. A majority of the compounds are biologically inactive and do not displace SAH-488 from the active site of MTAN-D198N, thus producing polarization values equivalent to those of the positive controls. However, a few sample reactions have a lower polarization that lies within the “signal window” or “separation band” zone\textsuperscript{45, 46}. The compounds associated with these lower polarizations are considered hits and will be evaluated by secondary screening methods.
At this point, the mean and standard deviations of the “sample reactions” were calculated. The Z-factor of our developed FP assay for screening the NIH clinical collection was calculated using following equation\textsuperscript{45}:

\[ Z = 1 - \frac{3\sigma_s + 3\sigma_c}{|\mu_s - \mu_c|} \]  

(Equation 3.4)

Where \( \mu_s \) and \( \mu_c \) represent the mean values for the sample reactions and negative controls, respectively, While \( \sigma_s \) and \( \sigma_c \) represent the standard deviations for the sample reactions and negative controls, respectively. The Z-factor was calculated to be 0.62, validating the compatibility of our developed FP assay for high-throughput screening of an available chemical library.

Figure 3-9: Screening of NIH clinical collection. The blue and green diamonds represent the polarization values from sample reaction and control reactions respectively. The three times standard deviation is shown by black line. The difference between the three times standard deviation of sample and control represent the “signal window” or “separation band”\textsuperscript{45,46}. 

60
3.5.6 Hit validation and false positive hits

Out of 446 compounds, eight compounds were found to be in the “separation band”\textsuperscript{45} representing hits. These compounds include epirubicin, idarubicin, doxorubicin, telithromycin, epigallocatechin, topotecan, and isoquercitrin. First three compounds; epirubicin, idarubicin and doxorubicin are fluorescent red compounds; belong to anthracycline family; have excitation and emission wavelength of 488 and 560 nm. Since these three compounds absorb at the same wavelength as SAH-488, they clearly cause interference in the polarization assay.

The compounds epigallocatechin, topotecan isoquercitrin and telithromycin were subjected to Amplex Red-Xanthine Oxidase to evaluate the inhibition activity. None of these compounds were found to be inhibiting the MTAN activity. It has been observed that colored or fluorescent compounds causes interference in fluorescence based assays. However, the compound very close to three times standard deviation of sample mean (3\*\(\sigma\) sample) was evaluated by drug dose dependent FP assay. As expected from the structure it is found to be replacing the substrate from the active site. The xanthine oxidase based assay cannot be performed on this compound, as the compound is hydrolysable by MTAN.
3.6 Conclusion and future work

Although this assay is of high quality and reproducibility, it can still be improved further. One area of weakness lies in the limited size of the library that was screened. The NIH Clinical Collection contains 446 compounds, while other libraries can contain upwards of 10,000 compounds. If a larger collection is screened, then the potential number of hits will obviously increase. Also, the NIH collection includes a relatively large ratio of compounds with the same wavelength as the Alexa Fluor 488. This results in a greater percentage of false positives, as these compounds will exhibit similar absorption/emission data as the Alexa Fluor in this assay. Therefore, a larger screening library would be advantageous in correcting this abundance of false positives, as a larger collection would dilute the number of compounds with similar wavelengths, leading to a smaller signal to noise ratio. Another option for eliminating these false positives would be to choose a difference fluorescent probe for the MTAN. For example, if we select a probe with a longer wavelength such as shifting to red we would circumvent the false positive hits that were related to the sharing of a similar wavelength with the green Alexa Fluor.
Chapter 4

Structural and biochemical analysis of I1: Telomere binding protein in vaccinia virus

4.1 Overview

Vaccinia virus is a member of family Poxviridea, and the virus is used as a vaccine for smallpox. The use of this vaccine facilitated worldwide eradication of smallpox that was certified by the World Health Organization in 1980\textsuperscript{17}. The research interest in vaccinia virus has resurfaced due to the risk of the use of smallpox as a bioterrorism agent\textsuperscript{48, 49}. Smallpox virus is an ideal biological weapon as it is highly infectious, transmissible by aerosol route, and can result in lethal infection once it has spread, as there is no modern treatment for the Smallpox disease. This vaccine is employed rarely in today’s world, resulting in a high susceptibility amongst the general populace. In addition to its usefulness as a vaccine, for the past two decades, the recombinant vaccinia virus has been used as a vector for expressing foreign antigens in mammalian cells and studying their immunological response\textsuperscript{50}. Extensive research to use genetically engineered vaccinia viruses for cancer and HIV therapeutics are underway\textsuperscript{51}. 
4.2 Structure of vaccinia virus

Vaccinia virus is a large, enveloped, double stranded DNA (dsDNA) virus with a complex structure. The brick-shaped viruses have an outer envelope-covering dumbbell shaped core. The core contains the dsDNA and proteins associated with replication, transcription and assembly of virion. Because vaccinia virus replicates in cytoplasm of infected cells it encode most of the protein required for replication and transcription. Theses proteins are encoded by a large linear dsDNA genome of 192 kilobase pairs. The two strands of DNA are covalently linked at both termini giving an appearance of one strand of DNA. These termini, also known as telomeres, consist of a 104-nucleotide-long hairpin structure. The four unpaired nucleotides are present in each termini forming a loop of hairpin structure. Eighty-eight AT rich nucleotides are engaged in base pairing, forming a duplex. The rest of the twelve nucleotides are unpaired, extra helical: two nucleotides on one strand and ten nucleotides on the other strand (Figure 4-1).
Figure 4-1: Vaccinia virus. (A) The structure of enveloped virus with core composed of dsDNA and viral proteins. (B) The genome with highlighted viral telomere (hairpin structure); the genome is consists of 192 kbp, while the telomere is made up of 104 nucleotides; 88 nucleotide involved in base pairing, four nucleotides forms loop region and twelve nucleotides are extra helical nucleotides; ten (shown in red circle) on one side and two (shown in yellow triangles on other side). Adapted from reference 52.
4.3 Replication of vaccinia virus

Viral telomeres are associated with DNA replication\textsuperscript{53,54}. The nick generated near the terminal end exposes the 3’ hydroxyl that serves as a primer 3’end. This initiates the DNA replication by viral DNA polymerases (Figure 4-2). In the course of researching these proteins that are involved in telomere binding and replication, two proteins, I1 and I6, were identified\textsuperscript{55}. The biochemical and genetic analyses of these newly discovered proteins have suggested a preferable binding of I1 and I6 to the viral hairpin of the vaccinia virus\textsuperscript{55,56}. Klemperer \textit{et.al} reported that I1 plays a role in the late phase of the viral replication life cycle and is essential for the assembly of the virion\textsuperscript{56}.

![Figure 4-2: Vaccinia virus DNA replication; nick generates in the proximity of telomere loop, DNA polymerase (shown in blue oval solid) binds and polymerizes viral DNA from 3’ to 5’ direction. Adapted from reference 53,54.](image-url)
In general, the life cycle of enveloped viruses involve the following steps: 1) Infection; viral envelope fuses with host cell membrane releasing the nucleic acid containing protein shell (capsid) inside the cell, 2) Uncoating; degradation of capsid occur releasing the nucleic acid in the cell, 3) Replication; the synthesis of multiple copies of viral nucleic acid accompanied by the protein synthesis, 4) Assembly; arrangement of viral protein occurs around viral nucleic acid, producing many copies of matured virion, 5) Shedding: the matured virus leaves the cell and spreads throughout the environment\textsuperscript{57}. The vaccinia virus follows a similar life cycle in two of these specific stages relevant to the presented work; immature virion with nucleoid (IVN) and the mature virion (MV)\textsuperscript{56}. These stages are associated with assembly of virion in host cell (Figure 4-3).

IVN is characterized as a spherical particle with an undefined core containing a dense, granular nucleoid. Alternately, MV is characterized as a brick shaped, well-defined dumbbell shaped core containing dsDNA\textsuperscript{56}. IVN is a noninfectious form of the virion, while MV is an infectious form. Morphogenesis from IVN to MV depends on several late stage virion proteins. Biochemical and electron microscopic analyses have suggested that the I1 protein expresses in the late stage of virion life cycle and plays a critical role in its assembly. It has been revealed that in the absence of I1 protein expression, the progression from IVN to MV is arrested.

To summarize, I1 is a telomere-binding protein that is essential for the progression from the noninfectious to infectious virion. Further, to elucidate the mechanism of I1-telomere binding and effect of inhibitor on binding, we performed structural and biochemical studies.
Figure 4-3: Vaccinia virus life cycle. (1) Infection; enveloped extracellular virus (EEV) enter inside the host cell. (2) Uncoating; releases viral protein and nucleic acid in host cell cytoplasm. (3) Replication; viral DNA multiplies making copies of viral genome. (4) Assembly; genome packaging into late viral proteins starts. In the presence of viral proteins including I1, noninfectious IVN develops into infectious MV.
4.4 Protein parameters, expression and purification of I1

The I1 protein parameters; 318 amino acid (including 6XHis tag), 36.7 kDa molecular weight, 8.70 pI and 16390 M⁻¹ cm⁻¹ were calculated using ProtParam tool (ExPASY website). The recombinant plasmid p101 containing the I1 gene (C-terminal His tag) used to transform E. coli Rosetta™ by a previous student in Dr. Ronning’s lab. The transformed bacterial cells were cultured in Luria broth (LB) broth containing 0.1 mM chloramphenicol and 0.3 mM ampicillin. The cells were grown at 37 °C until cells density reached to Abs₆₀₀nm of 0.6-08. The culture was induced with 0.1 mM IPTG at 16 °C for 24 hours. The cells were harvested and the pellets were resuspended in buffer A (20 mM HEPES pH 7.5, 0.5 M NaCl, 5 mM β-mercaptoethanol, and 5 mM imidazole). The cells were stored at -80 °C until needed.

4.4.1 Co-affinity purification of recombinant I1

The cell lysis of induced cells was performed by adding 1 mM lysozyme and 0.1 mM of DNaseI and kept on ice for 20 minutes. The cell suspension was then subjected to sonication and centrifuged at 15000g for 30 minutes at 4 °C. The supernatant was further filtered to remove any residual cell debris. The filtered supernatant was loaded on 5 ml HisTrap FF (GE Healthcare), Co-affinity column pre-equilibrated with buffer A (20 mM HEPES pH 7.5, 0.5 M NaCl, 5 mM β-mercaptoethanol, and 5 mM imidazole). The recombinant protein was eluted using the gradient of buffer B (20 mM HEPES pH 7.5, 0.5 M NaCl, 5 mM β-mercaptoethanol, and 150 mM imidazole).
4.4.2 Sodium dodecyl sulfate-polyacrylamide gel electrophoretic analysis

The peak fractions from the affinity chromatography were analyzed by running Sodium dodecyl sulfate-polyacrylamide gel electrophoretic (SDS PAGE). Three major bands of different molecular weight were observed. The band around 35 kDa corresponds to full length I1 protein. The second band around 70 kDa corresponds to a dimeric form of full length if I1. The third band observed near 28 kDa is a truncated form of I1 that is a result of the proteolytic degradation of the I1 protein in the E. coli host cell. Since the recombinant I1 protein has a C-terminal His tag, it is evident from the gel that there is a loss of roughly 7 kDa of N-terminus peptide (Figure 4-4). It may very well be expected that the N-terminus is dynamic and exposed to proteolytic enzymes in E.coli cells. The idea of a flexible N-terminus is in accordance with crystallization experiments using full length and truncated forms of I1 protein discussed below.

4.4.3 Rationale for interest in truncated form I1

The crystallization experiments using full-length protein failed to produce crystals. However, the truncated form of protein has a potential to crystallize, which can be exploited for the structural studies. However one may suspect that the N-terminally truncated protein might lose its ability to bind DNA. To investigate the DNA binding ability of N-terminally truncated protein, an electrophoresis mobility shift assay (EMSA) was performed in Ronning lab. The EMSA analysis suggests that the truncated form of protein retains its ability to bind DNA and can be used in crystallization and binding studies.
Figure 4-4: Co-affinity chromatogram and corresponding SDS-PAGE: The bands at around 37.0 and 25.0 kDa correspond to full length and truncated form of I1 respectively.
4.4.4 SEC purification of truncated I1

I1 was further purified using SEC. The fractions containing only truncated form and mixture of truncated form and full length of I1 were pooled and dialyzed overnight against size exclusion buffer (20 mM HEPES pH 7.5, 250 mM NaCl, 0.5 mM TCEP). The protein was concentrated to reach the volume of 3-5 ml and injected on to a pre-equilibrated Hi-Load Superdex 200 column (GE Healthcare). The peak fractions were analyzed using SDS PAGE (Figure 4-5). The first peak corresponds to the full length of protein and thus elute before truncated form. The fractions containing truncated I1 (apparent molecular weight 25.0 kDa) were pooled and used for crystallization and binding experiments.
Figure 4-5: SEC and corresponding SDS-PAGE. The bands at around 37.0 and 25.0 kDa correspond to full length and truncated form of I1 respectively.
4.5 Crystallization

The truncated I1 was concentrated to 23 mg/ml using ultrafiltration (Millipore). The protein concentrations were determined by measuring the absorbance at 280 nm and the extinction coefficient used was 16390 M$^{-1}$ cm$^{-1}$. The initial screening experiment was performed using Index HT (Hampton Research) and the hanging drop vapor diffusion method. The drop contained 1 µl of well solution, 1 µl of truncated I1 (23 mg/ml). The crystallization trays were incubated at 20 °C. Only one condition (0.2 M Magnesium chloride hexahydrate, 0.1 M Bis-Tris pH 6.5 and 25% w/v polyethylene glycol 3350) found to produce promising crystals (Figure 4-6). The crystals were screened for the X-ray diffraction. The crystals diffracted to 6-7 Å resolutions. To improve the quality of crystals, two-dimensional screens (varying the concentration of MgCl$_2$ and PEG 3350) were carried out. These experiments produced better appearance crystals (Figure 4-6) and will be subjected to X-ray diffraction studies.

Figure 4-6: Crystals of truncated I1. (left) Initial hit found in condition. (right) Crystals found 2-D screening
4.6 Truncated I1 and DNA binding assay

To determine the affinity of I1 for dsDNA, we performed the DNA binding experiment. The fluorescein labeled ssDNA sequence and unlabelled complementary strands were purchased from Integrated DNA technologies. To anneal the DNA, the 20 µl (500 µM) of each labeled and unlabelled strands were dissolved in 80 µl of deionized water separately, to reach the final concentration of 100 µM of stock ssDNA. These two samples of labeled and unlabelled ssDNA were mixed and incubated at 95 °C for 20 minutes. The heat block was turned off and the ssDNA mixture was allowed to cool to room temperature. The tagged dsDNA was stored in -20 °C.

To perform the FP assay, the stock of 100 nM of tagged dsDNA was prepared using buffer 20 mM HEPES pH 7.5. The assay was performed in black, 384-well plates and FP was measured using a Biotek Synergy H4 Plate reader. Each 50 µl reaction consists of 5 µl of 100 nM tagged dsDNA to reach a concentration of 10 nM. The I1 concentration was varied from 2 nM to 500 nM and appropriate amount of buffer was added to make up the 50 µl of reaction. After 15 minute of incubation the, FP was measured using excitation and emission wavelength of 485 nm and 528 nm. The data was fit and analyzed using Prism (GraphPad software). The binding constant $K_d$ was determined to be 9.8 ± 3.2 nM (Figure 4-7).
Figure 4-7: I1/DNA binding curve. The assay was performed in triplicate.
4.7 Issues with truncated I1 expressed using construct (p101-I1)

We observed inconsistent results in protein expression and crystallization experiments when truncated I1 resulting from degradation in *E. coli* cells was used. The ratio of full-length protein to truncated protein varied from batch to batch protein preparation. Also, based on the batch of protein, use of identical crystallization condition produces different crystal morphology or does not produce any crystals. We hypothesize that these inconsistencies are associated with differential proteolysis by proteolytic enzymes in *E.coli* cells. The difference in proteolytic activity may generate different forms of truncation.

4.8 Construct design for expressing an engineered truncated I1

To avoid variation, we designed an expression construct (plasmid-I1\text{truncated}) for expressing only truncated version of I1 will be referred as I1\text{truncated}. From this truncated construct, we expected the following implications: 1) the yield of truncated form of protein will be significantly increased as the *E. coli* will use its nutrient to overexpress truncated form only. 2) the overexpressed truncated I1 protein will be more homogenous and hence expected to give more consistent crystallization results. 3) In comparison to the original construct, it is more feasible to use this construct for a selenomethionine crystallization experiments. To determine the N-terminal truncation site on I1 protein, mass spectrometry analysis of full length and truncated form were carried out.
**4.8.1 Mass spectrometry**

On a SDS-PAGE gel the bands that corresponded to the truncated form of I1 were cut out and subjected to trypsin digestion and analysed on MALDI-TOF/TOF. The peptide mass fingerprinting resulted in 34 % coverage of the protein sequence highlighted in Figure 4-8. The 100 amino acid residues at the N-terminus were not observed. This result suggests that the truncated form consists of amino acids residues from E101 to C-terminal end. However this does not negate the presence any residue before E101. It is possible that high frequency of lysine residues at positions 90, 93, 99, and 100 results in small peptides not observed in the MS experiment.

![Protein sequence coverage: 34%](image)

**Figure 4-8:** Peptide mass fingerprinting of truncated form of I1. The result showing 34 % of sequence coverage. The covered sequences are highlighted in red. The amino acid residues from E101 to C-terminal are highlighted with yellow background.
The protein samples containing both forms were subjected to intact mass analysis in three different experiments using MALDI (Matrix-assisted laser desorption/ionization). Two peaks were observed, the major peak corresponds to truncated form while the minor peak corresponds to full lengths of I1. The major peak in three different experiments corresponds to m/z 28145.600, 28138.722, and 28157.064. The minor peak in three different experiments corresponds to mass 36760.158, 36744.876, and 36769.766. The peak differences from these three experiments calculated were 8614.558, 8606.154, and 8612.702. The differences in mass suggest that a 8.0 kDa segment is absent in truncated form. We were unable to calculate the exact mass difference between the full length and truncated form because low sensitivity of instrument used for resolving the accurate intact mass. As highlighted in Figure 4-9, the experiment indicates that the truncation site is located at around amino acid residue I76.

Furthermore, we performed a secondary structure prediction for vaccinia virus I1 protein using Jpred software\textsuperscript{58}. (Figure 4-10). As shown in Figure 4-9, the secondary structure prediction indicated that residue I76 located with in an $\alpha$- helix. Since I76 is predicted to be found in a helix, it is less likely to be the truncation site. However, the helix predicted to contain I76 is preceded by a predicted loop that may represent the site of truncation. Taken together, the mass spectrometry data suggest that when designing a plasmid to express a truncated I1 protein, the amino acid residues around I76 can be chosen as the N-terminus. The secondary structure prediction guided us to start at G73.
Figure 4-9: Intact mass analysis. The peaks from three experiments are shown in blue, purple and green. The mass difference between full length (small peaks) and truncated form is highlighted in red. Overall the mass difference was found to be approximately 8.0 kDa. (Shown in upper panel). The peptide (expected missing peptide from truncated form) corresponding to approximately 8.0 kDa is highlighted using yellow background (Shown in lower panel).
Figure 4-10: Secondary structure prediction for vaccinia virus I1$^{58}$. The “H” indicates that the amino acids are involved in α-helix while “E” indicates that amino acid residues are involved in β-sheet formation. The black line represents the loop region of protein. The amino acid residue I76 is located in α-helix (highlighted by yellow box). We chose G (two amino acid upstream from I76) as the N-terminus end for I1$_{truncated}$ gene.
4.8.2 Cloning, protein expression and purification of I1\textsubscript{truncated}

The I1 gene was amplified by polymerase chain reaction (PCR) using following primer: Forward primer: 5' - CAC CAT GGG AGG AAT CCT TAT CAG TCT TAT TAA TAG T -3' and Reverse primer: 5'- TTG GAT CCT CAG TGG TGG TGG TGG TGG TGT TCA GCA TTA CTT GAT ATA GTA ATA TTA GGC -3'

The amplified gene was inserted into a pET-28-based plasmid. The restriction sites used were NcoI and BamHI. The sequence was confirmed by nucleotide sequencing performed by Eurofins MWG operon. The BL21 \textit{E. coli} cells were transformed with pET-28-I1\textsubscript{truncated} and cultured in LB media at 37 °C for 5-6 hours. Once reaching an \textit{Abs\textsubscript{600}} of 0.6, the culture was induced at 16 °C for 24 hours. The protein purification was performed similar to the full-length protein. The protein was found to be overexpressing well and present in soluble fraction of cell lysate (Figure 4-11). Unfortunately, we observe complete protein precipitation immediately following the first step of protein purification.
Figure 4-11: The SDS PAGE gel of I_{truncated} expression. In the second lane the band around 25.0 KDa represents the soluble expression of I_{truncated}. 
4.9 Conclusion and future studies

We performed crystallization and biochemical studies to characterize the I1 protein from vaccinia virus. The expression of p101-I1 in *E.coli* host results into the full length as well as truncated form of I1. Since full-length protein does not crystallize, we chose the truncated form of I1 protein as this form tended to crystallize easily. I1 truncated form retains its ability to bind DNA and can be used to study I1 DNA interaction. The inconsistency in protein expression and crystallization were observed when using truncated I1 that may be associated with the differential proteolysis in *E.coli* host.

To avoid the inconsistency in truncated I1 protein, we constructed a recombinant plasmid for the expression of truncated form of I1 only. The peptide mass finger printing, intact mass analysis and secondary structure analysis lead to the design of gene I1\textsubscript{truncated}. We successfully designed a construct that over expresses a soluble form of I1\textsubscript{truncated} protein. However, the protein precipitated out just after the first step of purification.

We hypothesize that soluble form of I1\textsubscript{truncated} protein is associated with the unknown binding partner in the solution. After first step of purification the I1\textsubscript{truncated} protein loses that binding partner that leads to destabilization of the protein in solution. On the basis of one peptide finger printing experiment of truncated I1 protein we think that the binding partner could be a small peptide from the N-terminal domain of the full-length protein. The said peptide finger printing of I1 truncated form indicates the presence of small peptide that belongs to the N-terminal region of full length of protein. Considering that, in future co-expression of I1\textsubscript{truncated} with the expected binding partner
peptide will be performed. Once the method for soluble purified $I1_{\text{truncated}}$ is achieved the protein will be subjected to crystallization and DNA binding studies.
Chapter 5

Structure and functional analysis of Mycobacterial dihydroleopterin triphosphate pyrophosphatase: A nudix hydrolase involved in folate biosynthesis

5.1 Tuberculosis and Mycobacterium tuberculosis

Tuberculosis (TB) affects one third of the world population. Every year, 9 million persons develop an active infection. In 2011, the World Health Organization (WHO) reported 1.4 million deaths\(^5\). TB is mostly caused by *Mycobacterium tuberculosis* (*M. tb*). Treatment for TB usually involves combination of several drugs (first line drugs such as isoniazid, rifampicin, pyrazinamide, or ethambutol and second-line drugs including aminoglycosides or thioamides for instance) and last between 3 to 9 months. Due to non-compliance in patients treated for TB, multidrug-resistant (MDR-TB) and extensive drug-resistant tuberculosis (XDR-TB) have emerged. In the first case, the bacteria are resistant to at least two first-line drugs; while in the second, first-line drugs as well as several of the second-line drugs are inefficient.

More recently, totally drug-resistant tuberculosis (TDR-TB) has also been discovered\(^6\). Additionally, the co-infection with Human Immunodeficiency Virus (HIV) arises as a problem as most of the protease inhibitors used to treat HIV are not compatible
with the drugs involved in the treatment of TB. All together, these data stress the need of studying *Mycobacterium tuberculosis* (*M. tb*) and finding new treatments for tuberculosis.

### 5.2 Folate biosynthetic pathway

The enzymes that are linked to one carbon metabolism utilize folic acid. One carbon metabolism is essential for DNA synthesis, RNA modification and amino acid biosynthetic pathways\(^{61}\). Because of this, interference of folate-dependent pathways inhibits cell growth and replication. Antagonists of folate or antifolates have been used as therapeutic agents for the treatment of many diseases such as cancer, malaria and bacterial infections\(^{62-65}\). Conventional antifolate drugs such as sulfonamides and trimethoprim primarily inhibit enzymes that are involved in *de novo* synthesis of folate\(^{66}\). Because humans and other mammals lack the enzymes necessary for *de novo* synthesis of folate, inhibiting this pathway in bacteria has been commonly targeted for developing antibacterial drugs. While early emergence of antibiotic resistance to antifolate antibiotics and side effects initially moved researchers toward the development of alternative targets, interest in antifolate compounds has resurfaced.

The second pathway in folate biosynthesis involves the removal of pyrophosphate from dihydronopterin triphosphate (DHNTP) to convert it into dihydronopterin monophosphate (DHNP)\(^{67}\). In *Mycobacterial* species, it is suspected that enzyme encoded by Rv0413 catalyzes the hydrolysis of pyrophosphate of DHNTP, belongs to nudix hydrolase family and called mutT3\(^{68}\). In general, the substrates of nudix hydrolases family has common nucleoside diphosphates moiety to variable moiety X\(^{69}\). Most studied nudix hydrolases belong to subgroup mutT, enzymes that act as antimutator, removes
deleterious metabolites or mutated nucleosides from the cell. The members of nudix hydrolases share the common motif that forms the active site of enzyme and is responsible for catalysis and binding of substrate\textsuperscript{69}. Here we aim to study the structural and kinetic properties of mutT3 for the purpose of drug development.

Figure 5-1: De novo folate biosynthesis. MutT3 play role in the second step of the pathway.
5.3 Protein expression and purification of *M.smeg* MutT3

The *M.smeg* MutT3 protein parameters; 165 amino acids (including C-terminal, 6XHis tag), 17.9 kDa molecular weight, 6.33 pI and 34490 M⁻¹ cm⁻¹ were calculated using ProtParam tool (ExPASY website). The recombinant plasmid pVN832 (pET11c based, *M.smeg* mutT3 gene, C-terminal His tag, provided by Dr. Liem Nugyen from Case Western Reserve University) used to transform BL-21 cells. The transformed bacterial cells were cultured in Luria broth (LB) broth containing 0.3 mM ampicillin. The cells were grown at 37 °C until cells density reached to Abs₆₀₀ nm of 0.6-08. The culture was induced with 0.1 mM IPTG at 16 °C for 24 hours.

The cells were harvested and the pellets were resuspended in binding buffer (20 mM Tris pH 7.5, NaCl 500 mM, BME 5 mM, MgCl₂ 10 mM, glycerol 10% and imidazole 25 mM). Protein was purified using Ni-affinity chromatography using elution buffer (20 mM Tris pH 7.5, NaCl 500 mM, BME 5 mM, MgCl₂ 10 mM, glycerol 10% and imidazole 500 mM). The protein was further purified using size exclusion chromatography using buffer (20 mM Tris 7.5, 10 mM MgCl₂, 250 mM NaCl, 0.5 mM TCEP and 10% glycerol).
Figure 5-2: SDS PAGE analysis of MutT3. (A) After Ni affinity chromatography. The intense band at 18.0 kDa corresponds to mutT3 protein. (B) After size exclusion chromatography. Only a single band was observed that corresponds to mutT3 protein.
5.4 MutT3 substrate specificity using ATP, GTP, CTP and UTP as a substrate

To assess the nucleoside hydrolytic activity we perform a kinetic assay using four different deoxyribonucleotides (dNTPs); UTP, CTP, GTP and CTP. Although, DHNTP is a real substrate of mutT3 but due to the unavailability of the substrate we use different dNTPs. The assay was performed using malachite green method\(^{70}\). For each enzymatic reaction, 100 µl of final reaction volume was setup.

All assays were performed using 20 mM HEPES and 50 mM KCl buffer at pH 7.5. Each reaction containing 100 µM of dNTP, 2 mM of MgCl\(_2\), 2 U of inorganic pyrophosphatase and 400 nM of mutT3 were incubated at 37 °C. The reactions were incubated for 10 minute, 20 minute and 30 minute and quenched by adding malachite green solution (Figure 5-3). After 15 minute the absorbance was measured at 630 nM using a BiotekSynergy H4 plate reader (Winooski, VT, USA).

As shown in Figure 5-4 the product release with time was observed only with ATP as substrate. No increased was observed when GTP, CTP or UTP was used as a substrate. These experiments suggest that the mutT3 exhibit pyrophosphate hydrolytic activity. Also, it indicates that mutT3 shows specificity for ATP as a substrate. This result is in accordance with the substrate specificity shown by \textit{E.coli} nudix hydrolase\(^{69}\).
Figure 5-3: Scheme of MutT3 kinetic assay.
Figure 5-4: Substrate specificity of mutT3. The reactions for each nucleotide were quenched in 10, 20 and 30 minutes. The blue dots represent the absorbance from ATP hydrolysis by mutT3. No increase of absorbance was observed for GTP, CTP or UTP.
5.5 Kinetic parameters using ATP as a substrate

All assays were performed in triplicate in 20 mM HEPES and 50 mM KCl buffer at pH 7.5. Each reaction contains 2 mM of MgCl$_2$, 2 U of inorganic pyrophosphatase and 400 nM of mutT3. The concentration of ATP was varied from 50 to 1000 µM. There were six setup of reactions performed. Each setup contains ATP concentration from 50 to 1000 µM. These six reactions were quenched by adding malachite green after 10, 20, 30, 40, 50 and 60 minutes of incubation. The absorbance was measured at 630 nm. The rate of increase in absorbance was calculated for each concentration of ATP. The data was used to calculate the $K_M$ using equation 2.1. The $K_M$ value for ATP was calculated to be 155.6 ± 28 µM.

![Figure 5-5: Kinetic characterization of mutT3 using ATP as a substrate. The initial velocity is shown in terms of absorbance (mOD).](image)
5.6 Crystallization attempts

For the crystallization studies, purified protein samples were concentrated up to 4 mg/ml by a 3 kDa cut off Amicon system (Millipore). The crystal plates were set up using vapour diffusion method and sitting drop method in microbatch plates using the index screen (Hampton research.) The crystallization trays were incubated at 20 °C and 4 °C. Various attempts of crystallization was performed using mutT3 with and without ATP and AMPCPP. Unfortunately, no protein crystals were ever observed.

5.7 Chemo-enzymatic synthesis of DHNTP

To calculate the kinetic parameters using DHNTP as a substrate, we performed chemoenzymatic synthesis. To synthesis DHNTP we use the upstream enzyme FolE. In the presence of FolE, GTP is converted into DHNTP. We cloned, expressed, and purified recombinant E.coli FolE using affinity and size exclusion chromatography.

5.7.1 Cloning, protein expression and purification of E.coli FolE

The E.coli FolE gene was cloned into pET32 based plasmid and the construct was used to transform T7 express cells. The transformed bacterial cells were cultured in Luria broth at 37 °C until cells density reached to Abs$_{600\ nm}$ of 0.6-08. The culture was induced with 0.1 mM IPTG at 16 °C for 18-24 hours. The cells were harvested and the pellets were resuspended in binding buffer (20 mM Tris pH 7.5, NaCl 500 mM, BME 5 mM, and imidazole 25 mM. The protein was purified using Ni affinity chromatography followed by size exclusion chromatography using 20 mM Tris 7.5, 250 mM NaCl, 0.5
mM TCEP and 10% glycerol. The fraction containing purified protein (24.8 kDa) was collected and used for chemoenzymatic reaction.

Figure 5-6: SDS PAGE analysis of FolE. (A) After Ni affinity chromatography. The intense band at 25.0 kDa corresponds to FolE. (B) After size exclusion chromatography. The higher bands are the oligomeric form of FolE.
5.7.2 Chemo-enzymatic synthesis of DHNTP

To synthesize DHNTP, the reaction was performed in buffer 20mM MES 6.5 at 37 °C. The 1 mL reaction mixture contained 1 mM of GTP and 1 mM of FolE. The 50 µL reaction was observed for the DHNTP formation. Since DHNTP absorbs at 330 nm the reaction was followed until reached to the equilibrium (Figure 5-7). Once the reaction reached equilibrium the reaction mixture was passed through a 3.0 kDa cut off Amicon ultrafiltration device (Millipore). The filtrate containing a substrate-product mixture was purified using HPLC.

Figure 5-7: Chemo-enzymatic synthesis of DHNTP. GTP converts in to DHNTP in presence of FolE. The formation of DHNTP was followed by monitoring the absorbance at 330 nm.
5.7.3 Purification and characterization of DHNTP

The DHNTP was purified by high-performance liquid chromatography (Waters 2487) with a reverse phase analytical column (BDS HYPERSIL C18, 4.6 X 150 mm). The mobile phase isopropanol-triethylamine-85% phosphoric acid and water in the ratio of 8:10:3:979(v/v), pH 7 was used at flow rate of 1 ml/min. The peaks were monitored using dual wavelengths of 260 and 330. The $\lambda_{max}$ of GTP and DHNTP is 260 nm, and 330 nm, respectively. Overall two major peaks were observed: the first peak was observed at 330 nm but not at 260 nm. The second peak has strong absorbance at 260 nm with no absorbance at 330 nm assigned as GTP (Figure 5-8).

Figure 5-8: HPLC chromatogram using isopropanol-triethylamine-85% phosphoric acid as a mobile phase. The first peak in channel 2 corresponds to DHNTP.
However, the purified sample of DHNTP contained a high concentration of phosphoric acid. The high concentration of phosphoric acid causes a high background in kinetic assays, since we use a malachite green assay that detect the release of phosphate in the reaction. To avoid the high concentration of phosphate in the sample, we tried different mobile phases for the HPLC purification. Ultimately, we used isopropanol-triethylamine-formic acid pH 7.5 to separate DHNTP from GTP (Figure 5-9). The separation was not as good as with isopropanol-triethylamine-85% phosphoric acid. However the fractions containing DHNTP were collected with care and analyzed by ESI-MS.

Figure 5-9: HPLC chromatogram using isopropanol-triethylamine-formic acid as a mobile phase. The first peak in channel 2 corresponds to DHNTP. In upper channel the major peak corresponds to GTP. Other smaller peaks are the hydrolyzed of GTP.
Figure 5-10: ESI-MS of the fraction collected from HPLC.
5.8 MutT3 steady-state kinetics using DHNTP as a substrate

The kinetic assay was performed as previously described. A malachite green assay was used to detect the release of phosphate. The 50 µl reaction contains 2 mM of MgCl₂, 2 U of inorganic pyrophosphatase and 400 nM of mutT3. The concentration of DHNTP varied from 5 to 100 µM. The reaction was quenched after 30 min by adding malachite green solution. There were no significant differences observed between the control and enzyme containing reactions.

Figure 5-11: Kinetic assay using DHNTP as a substrate. The blue dots represent the control reactions. The red dots represent the reaction in presence of mutT3. The reaction was quenched after 30 min. and the absorbance was measured for each reaction containing different concentration of DHNTP.
5.9 Conclusion and future work

Here we expressed and purified *M. smegmatis* dihydroneopterin pyrophosphatase (mutT3) for the kinetic and structural characterization. We found that mutT3 utilizes ATP as a substrate but does not utilize GTP, CTP and UTP. We calculated the kinetic parameters for ATP as a substrate. To calculate the kinetic parameters for its hypothesized biological substrate, we synthesized DHNTP chemoenzymatically. However, no significant catalytic activity was observed when using DHNTP as a substrate. It is likely that the purified DHNTP sample has a high content of free phosphate that causes the low signal to background ratio. Different assay methods could be used to detect product formation such as HPLC analysis or use of a pyrophosphate sensor. Additionally, NMR studies could be performed to detect any dihydroneopterin pyrophosphatase activity. However, no strong evidence has ever shown that the mutT3 has a dihydroneopterin pyrophosphatase activity.
References


Receptor Alpha and Thymidylate Synthase Protein Expression in Patients with Non-Small-Cell Lung Cancer Treated with Pemetrexed, *Journal of Thoracic Oncology* 2013, 8, 19-30.


