2009

Structural analysis of 5'-Methylthioadenosine/S-Adenosylhomocysteine nucleosidase from Helicobacter pylori for the purpose of drug development

Natalie Marie Iacopelli

The University of Toledo

Follow this and additional works at: http://utdr.utoledo.edu/theses-dissertations

Recommended Citation

http://utdr.utoledo.edu/theses-dissertations/1078

This Thesis is brought to you for free and open access by The University of Toledo Digital Repository. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of The University of Toledo Digital Repository. For more information, please see the repository's About page.
A Thesis

Entitled

Structural Analysis of 5’-Methylthioadenosine/S-Adenosylhomocysteine Nucleosidase from *Helicobacter pylori* for the Purpose of Drug Development

By

Natalie Marie Iacopelli

Submitted as a partial fulfillment of the requirements for

The Master of Science degree in Chemistry

______________________________
Advisor: Donald Ronning

______________________________
Committee Member: Max Funk

______________________________
Committee Member: Timothy Mueser

______________________________
College of Graduate Studies

The University of Toledo

August 2009
An Abstract of

Structural Analysis of 5’-Methylthioadenosine/S-Adenosylhomocysteine Nucleosidase from *Helicobacter pylori* for the Purpose of Drug Development

Natalie Marie Iacopelli

Submitted as a partial fulfillment of the requirements for

The Master of Science in Chemistry

The University of Toledo

August 2009

5’-Methylthioadenosine (MTA)/S-Adenosylhomocysteine (SAH) nucleosidase is an essential enzyme for all bacteria. It is involved in many cellular processes including: polyamine synthesis, quorum sensing, methionine metabolism, and biological methylation.\(^1\)\(^-\)\(^4\) Because of the essential nature of this enzyme, MTA/SAH nucleosidase is a viable broad-spectrum antibiotic target. The gene encoding for MTA/SAH nucleosidase from *Helicobacter pylori* was cloned, expressed, and purified using immobilized metal affinity and size exclusion chromatography. The purified MTA/SAH nucleosidase was crystallized in native form then subjected to X-ray diffraction studies and structurally analyzed using macromolecular X-ray crystallography. MTA/SAH nucleosidase was also crystallized in two additional states: one bound to a substrate analog (Formycin-A), and one with adenine. Analysis of the three structures supports previously theorized catalytic mechanisms and models.\(^5\)\(^,\)\(^6\) Also, the presence of a Tris molecule within the active site of the native and adenine-bound structures provides insight into the development of a Tris-based enzyme inhibitor.
I would like to dedicate this thesis in loving memory of my grandparents, Joseph Iacopelli, Natale Palazzolo, and Lucy Palazzolo (Aimar), as well as my grandmother Anne Iacopelli (Pagano). So much of me has come from you.
Acknowledgements

The completion of my Master in Science would not have been possible without the help of some very important people. First, my father and mother, Joseph and Maria Iacopelli have encouraged me from a young child to now an adult in every adventure of my life. Beyond that, they have given me a strength to withstand any storm. Thank you.

Dr. Donald Ronning has been very influential in my research as my advisor and guide through this maze we call graduate school. I have never met a person with such an abundance of patience and willingness to explain every fine detail. Thank you for your wisdom and not laughing too hard when I made all of my “not thinking” mistakes.

I would like to extend my gratitude to my committee members Dr. Max Funk and Dr. Tim Mueser for their guidance. Also, thank you Dr. Funk for getting me back in the pool and introducing me to the great people of the Laurel Hills swim club.

I could have not survived the many long days in the lab had it not been for my group members Julie Boucau, Dan Lajiness, and Jonathan Collins. Thank you for all the help when I was first learning all of the lab techniques, and letting me steal lanes on a gel. I would also like to thank the undergraduate researchers in the lab Kushal Patel, and Alex Fitzhum. I also have to include my next door neighbor group members Lindsey Easthon, Sean Seaver, and Deepa Singh. Thank you Lindsey for always providing a great one liner to get me through the day and Sean for helping me out so much with the X-ray crystallography programs. I would also like to thank Sean for peer pressuring me into many a morning workout and my first half marathon.

Thank you to my family, especially my sister, Liza Iacopelli, and brother, Ross Iacopelli, to the instrumentation center, especially Dr. Leif Hanson for the guidance on
X-ray crystallography, to the University of Toledo Chemistry department, especially Char Hansen for keeping me ahead of all the paper work, to friends, including Tamam Baiz, Christa Graham, and Nina Potente, and to Andrew Behrle, you have been my rock and North star throughout this whole process.
Table of Contents

An Abstract of ........................................................................................................................................... ii
Acknowledgements ................................................................................................................................. iv
List of Equations .......................................................................................................................................... ix
List of Figures ................................................................................................................................................ x
List of Schemes ........................................................................................................................................... xii
List of Tables .............................................................................................................................................. xiii
List of Abbreviations ............................................................................................................................... xiv
Chapter 1 .................................................................................................................................................... 1
  Introduction .................................................................................................................................................. 1
    1.1. Helicobacter pylori .............................................................................................................................. 1
    1.2. MTA/SAH Nucleosidase as a Target ................................................................................................. 3
    1.3. SAM Metabolism .............................................................................................................................. 6
      1.3 A. Polyamines ...................................................................................................................................... 6
      1.3 B. Biological Methylation ................................................................................................................. 7
      1.3 C. Autoinducers ................................................................................................................................. 9
      1.3 D. Methionine .................................................................................................................................... 11
    1.4. MTA/SAH Nucleosidase .................................................................................................................... 13
    1.5. MTA/SAH Nucleosidase Homology ................................................................................................. 15
Chapter 2 .................................................................................................................................................... 20
  General Methods: Expression and Purification of a Protein of Interest .............................................. 20
    2.1. Introduction ........................................................................................................................................ 20
    2.2. Molecular Biology Techniques: Vector Construct ......................................................................... 21
      2.2 A. Polymerase Chain Reaction (PCR) ............................................................................................. 21
      2.2 B. Gel Electrophoresis and Gel Extraction ...................................................................................... 22
      2.2 C. Restriction Digest ....................................................................................................................... 23
      2.2 D. Ligation ......................................................................................................................................... 23
    2.3. Molecular Biology Techniques: Transformation ........................................................................... 24
      2.3 A. Transformation ............................................................................................................................. 25
      2.3 B. Mini-prep: ...................................................................................................................................... 26
    2.4. Expression of Protein of Interest ................................................................................................. 27
    2.5. Analytical Techniques: Protein Purification .................................................................................. 28
      2.5 A. Harvesting of E. coli .................................................................................................................... 28
      2.5 B. Lysis of E. coli ............................................................................................................................. 28
      2.5 C. Immobilized Metal Affinity Liquid Chromatography ................................................................... 29
      2.5 D. Immobilized Metal Affinity Liquid Chromatography (Tag Removal) ................................... 29
      2.5 E. Size Exclusion Chromatography (SEC) ..................................................................................... 30
    2.6. Summary .......................................................................................................................................... 31
Chapter 3 .................................................................................................................................................... 32
Methods: Expression and Purification of MTA/SAH Nucleosidase from *H. pylori*. 32

3.1. Introduction................................................................................................................. 32

3.2. Vector Construct: pET32 (Novagen®) + pfs .......................................................... 33
  3.2 A. Amplification of pfs using polymerase chain reaction (PCR) ......................... 33
  3.2 B. Restriction digest of pfs and pET32 (Novagen®) ............................................. 35
  3.2 C. Ligation of pfs and pET32 .................................................................................. 36
  3.2 D. Transformation of XL-10 with pET32 + pfs ..................................................... 36
  3.2 E. Miniprep and verification of gene insertion ...................................................... 37

3.3. Transformation of T7 express® + pROS with pET32 + pfs ........................................ 38

3.4. Expression and solubility test of the fusion protein SAH/MTA nucleosidase (MTAN) and Trx-Tag............................................................................................................. 39

3.5. Protein Purification .................................................................................................... 41
  3.5 A. Culture and Harvest of T7® + pROS + (pET32 + pfs) .................................... 41
  3.5 B. Lysis of T7® + pROS + (pET32 + pfs) .............................................................. 42
  3.5 C. Immobilized Metal Affinity Chromatography (IMAC) ..................................... 42
  3.5 D. IMAC Tag Removal ........................................................................................... 45
  3.5 E. Size Exclusion Chromatography (SEC) ............................................................ 46

3.6. Dialysis of pure MTAN.............................................................................................. 48

3.7. Summary.................................................................................................................... 49

Chapter 4 ............................................................................................................................ 51

General Methods: Macromolecular X-ray Crystallography ........................................... 51

4.1. Introduction............................................................................................................... 51

4.2. Macromolecular X-ray crystallography theory ....................................................... 52

4.3. Protein Crystallization ............................................................................................. 54

4.4. Diffraction experiment and data integration ........................................................... 56

4.5. Data scaling............................................................................................................... 57

4.6. Resolution, I/σ, and completeness .......................................................................... 57

4.7. Molecular replacement ............................................................................................ 58

4.8. Model refinement: simulated annealing, manual manipulation, minimization, and B-factor refinement .................................................................................................................. 58

4.9. Summary.................................................................................................................... 60

Chapter 5 ............................................................................................................................ 61

Methods: Macromolecular X-ray Crystallography of MTA/SAH Nucleosidase from *H. pylori*. .............................. 61

5.1. Introduction............................................................................................................... 61

5.2. Crystallization of MTAN ......................................................................................... 62

5.3. Crystallization of MTAN with Formycin-A (FMA) and SAH .............................. 64

5.4. X-Ray diffraction ..................................................................................................... 65
  5.4 A. X-ray diffraction of native MTAN ................................................................. 65
  5.4 B. X-ray diffraction of MTAN pre-incubated with 10 mM FMA .................... 67
  5.4 C. X-ray diffraction of MTAN co-crystallized with 2 mM SAH .................... 68

5.5. Molecular replacement and refinement .................................................................... 72

Chapter 6 ............................................................................................................................ 75

Results ................................................................................................................................ 75

vii
6.1. Introduction.................................................................................................................. 75
6.2 A. Secondary structure of MTA/SAH nucleosidase native ........................................ 75
6.2 B. Secondary structure of MTA/SAH nucleosidase with FMA ................................. 77
6.2 C. Secondary structure of MTA/SAH nucleosidase with SAH ................................. 82

Chapter 7 .................................................................................................................................. 84

Discussion .................................................................................................................................. 84
7.1. Introduction ...................................................................................................................... 84
7.2. MTA/SAH nucleosidase active site and catalysis ......................................................... 84
7.3. Drug Development ......................................................................................................... 95
7.4. MTA/SAH nucleosidase as a target for a broad-spectrum antibiotic ......................... 97
7.5. Future work ..................................................................................................................... 98

References ................................................................................................................................. 100

Appendices ................................................................................................................................. 105
Appendix I: *pfs* sequencing from *Helicobacter pylori*. Chromatogram of *pfs* sequencing ................................................................................................................................. 106
Appendix II: ProtParam: of MTA/SAH Nucelosidase from *Helicobacter pylori*45 107
Appendix III: MTA/SAH Nucleosidase Sequence Alignment of Helicobacter pylori, Mycobacterium tuberculosis, and Staphylococcus aureus47 108
Appendix IV: *pfs* sequence from *Mycobacterium tuberculosis* 109
Appendix V: HisTrapChe (Ni²⁺) chromatogram of MTAN from *M. tuberculosis*, and SDS-PAGE of fractions F3-F5. ................................................................. 112
Appendix VI: HisTrapChe (Ni²⁺) chromatogram Tag Removal of fractions F3-F6 from chelating column (Appendix V) ................................................................. 113
Appendix VII: Anion exchange chromatogram of flow through and fractions F1 and F2 from tag removal chromatography (Appendix VI), and SDS Page of fractions H11-I7. ..................................................................................... 114
Appendix VIII: Crystal hits from Hampton Research Index Screen ............................... 115
List of Equations

Equation 3.1: Beer-Lambert Law ................................................................. 49
Equation 4.1: Bragg’s Law ........................................................................ 53
Equation 4.2: The square root of the intensity of a reflection is equal to structure factor, which contains information of the amplitude of the diffracted X-ray wave. 54
Equation 4.3: Calculating an electron density map...................................... 59


### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td><em>Helicobacter pylori</em> infection</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>SAM</td>
<td>5</td>
</tr>
<tr>
<td>1.3</td>
<td>SAM metabolic pathways</td>
<td>5</td>
</tr>
<tr>
<td>1.4</td>
<td>Sequence alignment of MTA/SAH nucleosidase from <em>S. aureus</em>, <em>E. coli</em>, and <em>H. pylori.</em></td>
<td>16</td>
</tr>
<tr>
<td>1.5</td>
<td><em>E. coli</em> MTA/SAH nucleosidase</td>
<td>17</td>
</tr>
<tr>
<td>1.6</td>
<td><em>S. aureus</em> MTA/SAH nucleosidase</td>
<td>17</td>
</tr>
<tr>
<td>1.7</td>
<td>MTA/SAH nucleosidase active site from <em>E. coli</em> and <em>S. aureus</em>.</td>
<td>18</td>
</tr>
<tr>
<td>1.8</td>
<td>MTA/SAH nucleosidase active site from <em>E. coli</em> and <em>S. aureus</em>.</td>
<td>19</td>
</tr>
<tr>
<td>2.1</td>
<td>Restriction Enzyme Cut Site (NciI)</td>
<td>24</td>
</tr>
<tr>
<td>2.2</td>
<td><em>E. coli</em>’s Heat Shock of Discontinuous Membrane</td>
<td>25</td>
</tr>
<tr>
<td>3.1</td>
<td>Agarose Gel of the test digestion of pET32 + pfs (SNU refers to pfs)</td>
<td>38</td>
</tr>
<tr>
<td>3.2</td>
<td>Expression and solubility test of MTAN</td>
<td>40</td>
</tr>
<tr>
<td>3.3</td>
<td>SDS-PAGE gel of fractions B1-B4 of HiTrap chelating column of MTAN</td>
<td>44</td>
</tr>
<tr>
<td>3.4</td>
<td>Chromatogram of HiTrap chelating (Ni²⁺) column of T7® + pROS + pET32 + pfs lysate</td>
<td>44</td>
</tr>
<tr>
<td>3.5</td>
<td>Chromatogram of HiTrap chelating (Ni²⁺) tag removal column of fractions B1-B4 from the initial HiTrap chelating column</td>
<td>45</td>
</tr>
<tr>
<td>3.6</td>
<td>SDS-PAGE gel of flow through and fraction A1 from HiTrap chelating (Ni²⁺) tag removal column</td>
<td>46</td>
</tr>
<tr>
<td>3.7</td>
<td>SEC of concentrated MTAN from tag removal</td>
<td>47</td>
</tr>
<tr>
<td>3.8</td>
<td>SDS-PAGE gel of concentrated MTAN from tag removal</td>
<td>48</td>
</tr>
<tr>
<td>5.1</td>
<td>MTAN (12.19 mg/mL) crystal from crystallization condition F3 of Crystal Lite Screen (Qiagen®) 0.1 M Tris pH 8.5, 4% weight/vol PEG 8000</td>
<td>63</td>
</tr>
<tr>
<td>5.2</td>
<td>MTAN (12.19 mg/mL) crystals from coarse gradient crystallization condition 0.1 M Tris pH 8.5, 16% weight/vol PEG 8000</td>
<td>64</td>
</tr>
<tr>
<td>5.3</td>
<td>Formycin-A (FMA)</td>
<td>65</td>
</tr>
<tr>
<td>5.4</td>
<td>X-ray diffraction pattern of MTAN native crystal from crystal condition 16% weight/vol PEG 8000 and 0.1M Tris.HCl pH 8.5.</td>
<td>67</td>
</tr>
<tr>
<td>5.5</td>
<td>The diffraction pattern of MTAN with pre-incubated 1 mM FMA</td>
<td>70</td>
</tr>
<tr>
<td>5.6</td>
<td>The diffraction pattern of MTAN co-crystallized with 2 mM SAH</td>
<td>71</td>
</tr>
</tbody>
</table>
List of Schemes

Scheme 1.1: (A) Spermidine Synthesis (B) Spermine Synthesis ........................................ 7
Scheme 1.2: Biological Methylation from SAM ................................................................. 9
Scheme 1.3: Autoinducer I Synthesis .................................................................................. 10
Scheme 1.4: Autoinducer II Synthesis ............................................................................... 11
Scheme 1.5: Methionine Synthesis .................................................................................... 12
Scheme 1.6: (A) SAH Degradation (B) MTA Degradation ................................................. 14
Scheme 1.7: MTA/SAH Nucleosidase ................................................................................. 15
Scheme 7.1: MTA/SAH nucleosidase mechanism from H. pylori ....................................... 85
Scheme 7.2: Mechanism of MTA phosphorylase ............................................................... 97
Scheme 7.3. Conversion of adenine to 2,8-dihydroxyadenine by xanthine oxidase .. 97
List of Tables

Table 3.1: PCR mixture for the amplification of pfs. .............................................. 34
Table 3.2: PCR parameters for the amplification of pfs. ........................................ 34
Table 3.3: Restriction digest reaction of pfs and pET32 ......................................... 35
Table 3.4: Ligation Reaction of pET32 + pfs ......................................................... 35
Table 3.5 Test restriction digest of pET32+pfs plasmid ......................................... 36
Table 4.1: Unit cell dimensions of the seven crystal systems ................................. 52
Table 5.1: Optimization of 0.1 M Tris.HCl pH 8.5, 4% weight/vol PEG 8000 ....... 62
Table 5.2: Diffraction and refinement statistics of MTAN native, MTAN with FMA and MTAN with SAH 49, 50 ................................................................. 69
Table 6.1: Hydrogen binding of FMA within the active site of MTA/SAH nucleosidase. 53 .......................................................... 79
Table 6.2: Hydrogen bonding of adenine and Tris within the active site of MTA/SAH nucleosidase. 53 .......................................................... 82
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acyl Carrier Protein</td>
<td>ACP</td>
</tr>
<tr>
<td>Acyl homoserine lactone</td>
<td>AHL</td>
</tr>
<tr>
<td>Advance Photon Source</td>
<td>APS</td>
</tr>
<tr>
<td>β-mercaptoethanol</td>
<td>BME</td>
</tr>
<tr>
<td>Calf intestinal phosphatase</td>
<td>CIP</td>
</tr>
<tr>
<td>Crystallography and NMR System</td>
<td>CNSsolve 1.2</td>
</tr>
<tr>
<td>Deionized H$_2$O</td>
<td>dH$_2$O</td>
</tr>
<tr>
<td>Deoxyribonucleic acid</td>
<td>DNA</td>
</tr>
<tr>
<td>Deoxynucleotide triosphosphates</td>
<td>dNTP</td>
</tr>
<tr>
<td>4,5-dihydroxy-2,3-pentadione</td>
<td>DPD</td>
</tr>
<tr>
<td>Formycin-A</td>
<td>FMA</td>
</tr>
<tr>
<td>Immobilize metal affinity chromatography</td>
<td>IMAC</td>
</tr>
<tr>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
<td>IPTG</td>
</tr>
<tr>
<td>Luria-Bertani</td>
<td>LB</td>
</tr>
<tr>
<td>Life Science Collaborative Access Team</td>
<td>LS-CAT</td>
</tr>
<tr>
<td>5’-methylthioadenosine</td>
<td>MTA</td>
</tr>
<tr>
<td>MTA/SAH nucleosidase</td>
<td>MTAN</td>
</tr>
<tr>
<td>5’-methylthioribose</td>
<td>MTR</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>MW</td>
</tr>
<tr>
<td>Nuclear Magnetic resonance</td>
<td>NMR</td>
</tr>
<tr>
<td>Pellet</td>
<td>P</td>
</tr>
</tbody>
</table>
Polymerase chain reaction  PCR
Polyethylene glycol  PEG
pRosetta  pROS
Release Factor 1  RF1
Release Factor 2  RF2
Reverse Osmosis  RO
Supernatant  S
S-Adenosylhomocysteine  SAH
S-Adenosylhomocysteine Nucleosidase  SAHN
S-Adenosylmethionine  SAM
Sodium dodecyl sulfate poly-acrylamide gel electrophoresis  SDS-PAGE
Size exclusion chromatography  SEC
S-Ribosylhomocysteine  SRH
MTA/SAH nucleosidase  SNu
Tetrahydrofolate  THF
Thioredoxin poly-histidine tag  Trx-Tag
Chapter 1

Introduction

1.1. *Helicobacter pylori*

Robin Warren and Barry Marshall first described the *Helicobacter pylori* in the early 1980’s as a Gram-negative bacterium that inhabits the gastro-duodenal tract of the gastro-intestines. Warren and Marshall were also able to provide experimental evidence, using upper gastro-intestinal endoscopy, supporting *H. pylori* as the causative agent of peptic ulcers. This discovery revolutionized the treatment of peptic ulcers to include a regimen of broad-spectrum antibiotics. Furthermore, *H. pylori* has since been strongly coupled to gastric cancer and is categorized as a class I carcinogen by the World Health Organization. For their discovery of *H. pylori*, Warren and Marshall were awarded the Nobel Prize in medicine/physiology in 2005.
*H. pylori* infection is described as a chronic and transmissible gastro-duodenal disease. Transmission of the bacteria is fecal-oral, in which a person becomes infected by consuming infected water or food, or by direct oral contact with someone already infected.\(^9\) After the bacteria are ingested they migrate to the gastric mucosa. The gastric mucosa is a protective barrier for the gastric lining acting as a buffer from the gastric acid (pH 2.0-3.0) that aids in digestion.\(^11\) Once in the gastric mucosa the bacteria begin to divide and proliferate. At this time *H. pylori* begins to degrade the mucosal layer, using an enzyme called mucinase, allowing gastric acid and pepsin, a protease, to reach the underlining tissue.\(^12\) The immune system, recognizing the bacteria as an antigen, produces and releases T-cells to combat the infection. However, T-cells are not able to penetrate the gastric mucosa layer. The T-cells die and release superoxide radicals as a last effort to combat *H. pylori*, which causes inflammation. The combination of the inflammation, gastric acid, and pepsin results in the formation of gastro-duodenum ulcers (Figure 1.1).\(^11,13\)

As previously mentioned, the treatment for the infection of *H. pylori* is with broad-spectrum antibiotics including, clarithromycin, tetracycline, and amoxicillin.\(^8\) However, the use of antibiotics for the treatment of bacterial infections has led to the ever-growing problem of antibiotic resistance. The bacteria acquire antibiotic resistance by point mutation or transfer of antibiotic resistance plasmids between the bacteria. The mechanism of resistance acts in three ways: producing an enzyme that degrades the antibiotic, alteration of the microbes target molecule that binds the antibiotic, or the development of a pathway to pump the antibiotic out of the cell.\(^14\) Once a strain of *H. pylori* has resistance against a particular antibiotic the treatment with that antibiotic is
ineffective. Also, of the three approved antibiotics for *H. pylori* infection, resistant strains have been identified.\textsuperscript{15, 16} Thus, it is important to develop novel antibiotics to continue to treat *H. pylori* infections.

1.2. MTA/SAH Nucleosidase as a Target

A potential drug target for the development of novel antibiotics is 5'-Methylthioadenosine (MTA)/S-Adenosylhomocysteine (SAH) nucleosidase. SAH/MTA nucleosidase is involved in the metabolic pathways of S-Adenosylmethionine (SAM) (Figure 1.2). The metabolism of SAM leads to the production of polyamines, autoinducers, methionine, and methylated biomolecules.\textsuperscript{1-4} A description of each of these pathways and the importance of each product on the cell cycle is described in Section 3.1.

SAM metabolic pathways involve a variety of enzymes, all of which produce either 5'-Methylthioadenosine or S-Adenosylhomocysteine as a product, which can exhibit feedback inhibition of other enzymes involved in SAM metabolism (Figure 1.3).\textsuperscript{4, 17-21} Feedback inhibition is a process in which the product of the enzymatic reaction can act as an inhibitor of the forward reaction. Thus, it is vital that SAH and MTA are further degraded to prevent the inhibition of SAM-dependent or S-Adenosylmethylthiopropamine-dependent enzymes. SAH/MTA nucleosidase fulfills this role and degrades both SAH and MTA. SAH/MTA nucleosidase is also a good drug target because mammalian cells use two separate and different enzymes for the degradation of SAH and MTA, SAH hydrolase and MTA phosphorylase, respectively.\textsuperscript{22-23}
Figure 1.1: *Helicobacter pylori* infection.\textsuperscript{11-13}
Figure 1.2: SAM

Figure 1.3: SAM metabolic pathways. The X stands for inhibition of the reaction. (5’-Methylthioadenosine nucleosidase (MTAN) and S-Adenosylhomocysteine (SAHN))
1.3. SAM Metabolism

Four biological pathways metabolize SAM. These pathways include polyamine, methionine, and autoinducer synthesis as well as biological methylation. All of the SAM metabolic pathways produce SAH or MTA, and exhibit feedback inhibition from SAH or MTA. The following is a description of each of the SAM metabolic pathways.

1.3 A. Polyamines

Polyamines can be described as organic compounds with two or more primary amino groups. There are four types of polyamines that are found in biology and include: cadavrine, putrescine, spermine, and spermidine. Spermine and spermidine are both synthesized from SAM. SAM is the amino propyl donor in the formation of both spermine and spermidine. The first step in both the synthesis of spermine and spermidine is the decarboxylation of the methionine substituent of SAM by SAM decarboxylase. In the next step, the decarboxylated SAM donates an amino propyl group to either putrescine or spermidine to generate spermidine or spermine respectively (Scheme 1.1).

The roles of polyamines include growth stimulation, oxidative stress survival, and low pH survival. Also, polyamines provide stabilization and increase the temperature necessary to disrupt the hydrogen bonding of duplex deoxyribonucleic acids (DNA). This property comes from electrostatic interactions between the anionic DNA and cationic polyamines.
1.3 B. Biological Methylation

Biological methylation has effects on many biomolecule interactions such as binding, processing, and activation. A few examples of the significance of biological methylation include DNA and protein methylation. DNA methylation in bacterial cells...
acts as a defense mechanism against harmful foreign DNA inserted into the cell by bacteriophages. The bacterial host has DNA methyltransferases that label the bacteria’s DNA as self. The foreign DNA is not methylated, is recognized and degraded by restriction endonucleases. Restriction endonucleases are enzymes that bind and cleave specific DNA sequences, protecting the cell from bacteriophages. DNA methylation also assists in DNA single strand mismatch repair that occurs during DNA replication. Before DNA replication the template/parent strand is methylated and as the nascent/daughter strand is being synthesized it lacks methylation. A mismatch in the nascent strand is corrected by proofreading exonucleases that have the ability to identify the nascent strand based on the absence of methylation. The exonucleases excise a fragment of the nascent strand with the mismatch and DNA polymerase correctly fills in the missing fragment.

The effects of methylation on protein’s form and function are very dynamic. The methylation of proteins can act as an on/off switch or for the proper interactions between protein and other biomolecules. For example, GTPase proteins are generally associated with the plasma membrane. This association occurs by the addition of unsaturated hydrocarbons to C-terminal thiol groups of GTPase. The unsaturated hydrocarbon becomes integral to the plasma membrane and anchors the GTPase protein to the membrane. Proteolysis of the protein occurs after the thiol group with the attached unsaturated hydrocarbon. After proteolysis, the freed C-terminal carboxylate group is methylated. The methylation of the carboxylate lowers the energy for the association/insertion of the GTPase protein into the membrane. Another example of protein methylation function is the methylation of release factor 1 (RF1) and release
factor 2 (RF2) in the termination of protein synthesis. The release factors recognize the stop codon of mRNA and hydrolyze the full-length protein from tRNA disconnecting the nascent polypeptide from tRNA. Methylation of δ-N-glutamine of a glycine-glycine-glutamine (GGQ) motif of RF1 and RF2 helps position the water molecule for hydrolysis of the protein from tRNA.²¹

Biomolecules are methylated primarily by methyltransferases, and the methyl donor is generally SAM.²⁴,³¹ The methylation reaction is a simple $S_{N2}$ reaction. The first step of the mechanism is the binding of the biomolecule and SAM to the methyltransferase. The formal charge on the sulfur atom of SAM is negative, which polarizes the methyl group (CH$_3$δ$^+$). The transfer of the methyl group to the biomolecule is equivalent to the transfer of CH$_3$ to an electron rich nucleophile of an $S_{N2}$ reaction. The thermodynamics of the transfer reaction is $-17$ kcal/mol, indicating that the methyl transfer is the driving force of the reaction. The products released are a methylated biomolecule and SAH (Scheme 1.2).²⁸

![Scheme 1.2: Biological Methylation from SAM](image)

1.3 C. Autoinducers

Autoinducers are vital biomolecules for the prokaryotic process known as quorum sensing. Quorum sensing is a prokaryotic biological process which utilizes autoinducers to signal to the culture its density and environment.³²,³³ The individual bacterium gene
expression then responds accordingly. This communication allows the unicellular bacterium to act as a multi-cellular organism. Processes that are generally under the control of autoinducers are those that cannot be undertaken by one cell and include: virulence, antibiotic production, biofilm formation, and sporulation.\textsuperscript{32, 34}

There are two types of autoinducers, autoinducer I and II, and both are generated by SAM degradation.\textsuperscript{3} Autoinducer I is generated from SAM and is catalyzed by acyl homoserine lactone synthase (AHL synthase).\textsuperscript{35} The first step is an intramolecular lactonization of the carboxyl oxygen reacting with C2” of SAM. The acyl chain is added onto the amino group, which is donated by an acyl carrier protein (ACP). The resulting products are a homoserine lactone (autoinducer I) and MTA (Scheme 1.3). Autoinducer I is then used for intra-species communication.\textsuperscript{3, 31, 35}

\begin{center}
\textbf{Scheme 1.3: Autoinducer I Synthesis}\textsuperscript{31}
\end{center}

Autoinducer II is generated not directly from SAM, but from one of its degradation products, SAH.\textsuperscript{36} SAH is most likely a product of a biological methylation reaction. SAH is first metabolized by SAH nucleosidase to produce S-ribosylhomocysteine (SRH) and adenine as a side product. SRH is degraded by LuxS removing the homocysteine chain from SRH as a side product and opening the ribose ring to produce 4,5-dihydroxy-2,3-pentadione (DPD). DPD is further metabolized by
LuxP to generate furanosyl boron diester (autoinducer II) (Scheme 1.4). Autoinducer II is then used for inter-species communication.\textsuperscript{17,37}

Scheme 1.4: Autoinducer II Synthesis\textsuperscript{17}

The manipulation of bacterial quorum sensing to promote or inhibit bacterial growth is well noted in nature. Two examples of such manipulation are Delisha pulchra, an Australian macroalgae, and Staphylococcus aureus. Delisha pulchra produces furanones similar to an acyl homoserine lactone (autoinducer I), which inhibits harmful bacteria’s intracellular communication.\textsuperscript{38} Staphylococcus aureus autoinducers increase the expression virulence factors within its own culture and decreases the expression of virulence factors in competing strains of Staphylococcus aureus to out compete those strains.\textsuperscript{39}

1.3 D. Methionine

The production of both SAM and methionine are within the same cellular cycle and are dependent on each other. The first step of the methionine cycle is SAM
degradation to SAH by a methyltransferase. Next, SAH is broken down into adenine and SRH by SAH nucleosidase. LuxS further degrades SRH into DPD and homocysteine. The final step in methionine synthesis is the addition of a methyl group to the sulfur of homocysteine. SAM is then regenerated from methionine and ATP by SAM synthetase (Scheme 1.5).⁴  

![Scheme 1.5: Methionine Synthesis](image)

Scheme 1.5: Methionine Synthesis⁴
1.4. MTA/SAH Nucleosidase

MTA/SAH nucleosidase, as mentioned, is specific to MTA and SAH, and is biologically active as a dimer. Each monomer of the dimer possesses one active site.\(^1\) The nucleosidase from *E. coli* is active over a broad pH range from 4.5-9.5, and favors an acidic pH for improved catalysis. In general, most enzymatic studies have been performed at a pH of 7.0. Also, ideal catalytic temperature is between 37-45 °C for the *E. coli* nucleosidase.

Binding of MTA/SAH to the active site proceeds first with the binding of MTA/SAH followed by orientation of a nucleophilic water. Once MTA/SAH is bound to the active site the enzyme changes from an “open” conformation to a “closed” conformation inducing a C4’ endo sugar pucker on the ribose.\(^2\) Once in the “closed” conformation all water is excluded from the active site except for the nucleophilic water. The nucleophilic water is orientated within the active site by hydrogen bonding interactions with polar residues within the active site. Catalysis occurs, breaking the N-glycosidic bond between N9 and C1’ and producing adenine and either SRH (SAH substrate) or 5’-Methylthioribose (MTR) (MTA substrate) (Scheme 1.6).\(^2,^3\) SRH and MTR do not exhibit any feedback inhibition, and the reaction is not reversible.
Aspartic acid and its conjugate base drive the MTA/SAH nucleosidase reaction. The first step of MTA/SAH nucleosidase reaction is protonation at the N7 position of the adenine base by aspartic acid (Scheme 1.7). The electrons in the $\pi$ bond between N7 and C8 are delocalized within the base resulting in a partial negative charge on N9. In addition, the protonation of N7 is the rate-limiting step of the reaction. After protonation the reaction is in a theorized transition state. The theorized transition state has an extension of the N-glycosidic bond between C1’ and N9. The extension of the C1’-N9 is the result of the formation of a partial negative charge on N9 and a partial positive charge on C1’. The partial positive charge of C1’ is resonance stabilized by the lone pair of electrons of the ribose oxygen resulting in the formation of an oxocarbenium ion. An enzyme oriented nucleophilic water will perform an $S_N1$ attack on the oxocarbenium ion.
cleaving the glycosidic bond between C1’ and N9. N9 acts as a Lewis base and donates a pair of electrons to form a nitrogen-hydrogen bond with a hydrogen from the water molecule now incorporated in the thioribose. The electrons from the oxygen-hydrogen bond of the water are then displaced onto the oxygen. The final step is the removal of a proton on N7 by the conjugate base of aspartic acid regenerating the enzymatic acid.\textsuperscript{44}

![Diagram of MTA/SAH Nucleosidase]

Scheme 1.7: MTA/SAH Nucleosidase\textsuperscript{44}

1.5 MTA/SAH Nucleosidase Homology

MTA/SAH nucleosidase is not unique to \textit{H. pylori}, and is found in many prokaryotes including \textit{Mycobacterium tuberculosis}, \textit{Staphylococcus aureus}, and \textit{Escherichia coli}. MTA/SAH nucleosidase in each of these organisms has a similar primary amino acid sequence (Figure 1.4). The structure of the nucleosidase from \textit{S. aureus} and \textit{E. coli} has been previously solved using macromolecular X-ray crystallography. The two models show that the protein forms a homodimer. Each monomer is composed of two sets of $\alpha$ helices, which are separated by a $\beta$ sheet that
extends the length of the protein. Also, present from residues 97-122 in the *E. coli* model and residues 96-121 in the *S. aureus* is a loop of random coil from monomer A that crosses over and interacts with the active site of monomer B and vice versa. This crossover loop may be indicative of allosteric regulation of catalysis between the two active sites (Figure 1.5 and 1.6).

Figure 1.4: Sequence alignment of MTA/SAH nucleosidase from *S. aureus*, *E. coli*, and *H. pylori*. Residues that are in bold and underline text are conserved active site residues among the organisms.
The active site in both the *S. aureus* and *E. coli* structure is subdivided into three sections: adenine, ribose, and alkylthio binding site. The adenine-binding site is a deep pocket composed of both β sheets and α helices. The adenine is bound to the enzyme by electrostatic and π-stacking interactions. Contributing residues for the electrostatic interactions between the adenine base include backbone hydrogen bonding with Ile 152/151 and the residue carboxyl group of aspartic acid 197/196 (*E. coli* / *S. aureus*).
Also, phenylalanine 153/152 will participate in van der Waals interactions with the adenine base (Figure 1.7).

The ribose-binding site interacts with the ribose through electrostatic and van der Waals interactions. The ribose sugar participates in hydrogen bonding between the side chain carboxyl group of glutamic acid 174/173 and with backbone hydrogen bonding methionine 173/172. Also, methionines 9/8 and 173/172 participate in van der Waals interactions with the ribose ring (Figure 1.7).

Figure 1.7: MTA/SAH nucleosidase active site from *E. coli* (the second number of the amino acid is the sequence number for the *S. aureus*). Amino acids M 173, E 174, R 193 compose the ribose-binding site. Amino acids D 197, S 196, I 150, and F 151 compose the adenine-binding site. Formycin A is a substrate analog.

The alkylthio binding site is composed of mostly hydrophobic residues forming a hydrophobic pocket for the binding of the alkylthio chain. These residues include isoleucine 50/49 and phenylalanines 151/150 and 207/206. Other hydrophobic residues
that compose the hydrophobic pocket come from the neighboring residues crossover loop and include valine 102 (alanine 101 \textit{S. aureus}), phenylalanine 105/104, and proline 113/112 (Figure 1.8).

![Figure 1.8: MTA/SAH nucleosidase active site from \textit{E. coli} (the second number of the amino acid is the sequence number for the \textit{S. aureus}). Amino acids F 107 (B), V 102, P 113, I 50, M 9, F 151, and F 207 compose the alkylthio-binding site. Formycin A is a substrate analog.]

The objective of this thesis was to determine the structure of MTA/SAH nucleosidase from \textit{H. pylori} in three forms: native, with a substrate analog bound, and with substrate bound, as well as the development of an assay for the screening of MTA/SAH inhibitors designed based on information from the crystal structures.
Chapter 2

General Methods: Expression and Purification of a Protein of Interest

2.1. Introduction

The techniques that were used to obtain purified protein follow a similar methodology. This process includes molecular biology and separation techniques. The molecular biology techniques are used to genetically engineer *E. coli* to produce the protein of interest. The separation techniques employed liquid chromatography in order to separate the protein of interest from the rest of the bacterial proteins. Summarized here are the general molecular biology and separation techniques used in the production and purification of MTA/SAH nucleosidase from *H. pylori*. 
2.2. Molecular Biology Techniques: Vector Construct

A vector, or plasmid, is a circular piece of DNA ranging from 3-5 kbp that can be incorporated into bacteria and its genes expressed. This ability of bacteria to be able to absorb DNA from its environment and produce proteins encoded by that DNA is commonly used in biochemical labs to synthesize the protein of interest. The molecular biology techniques that are used to generate a vector for cloning are polymerase chain reaction, gel electrophoresis, restriction digest, and ligation. The following is further description of each step of vector construction.

2.2 A. Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) amplifies the gene of interest using the enzymatic activity of DNA polymerase. The reaction requires a template containing the gene to be amplified, primers, DNA polymerase, and deoxyribonucleotide triosphosphates (dNTPs). The gene to be amplified acts as a template for DNA polymerase and can come from two sources, either genomic DNA or from a previously cloned vector containing the gene of interest. Two complimentary primers, one for each strand of DNA, are needed in order to form a duplex DNA with the template. The duplex formed is necessary for the binding of DNA polymerase. The primers are engineered to have restriction enzymes recognition sites for later molecular cloning steps, and tend to be 18-30 base pairs in length. The role of DNA polymerase is to incorporate the correct deoxyribonucleotide triosphosphates dNTPs into the gene fragment that is being amplified.
There are three steps involved in PCR: melting, annealing, and elongation. The melting step is heating of the reaction to disrupt the hydrogen bonding of the duplex DNA, and separates the two strands of DNA. The annealing step is the cooling of the reaction to approximately 2 °C below the melting temperature of the template-primer duplex, which promotes template-primer interactions. The melting temperature is a calculated temperature at which the primer-template duplex is disrupted and promotes the formation of single stranded DNA. Since the annealing temperature is based off the melting temperature of the template-primer duplex this step can vary from reaction to reaction. The elongation step is the elongation of the primer using the template as a guide, to synthesize the gene of interest in the 5’ to 3’ direction. The temperature during the elongation stage is based off the type of DNA polymerase being used and in general the temperature for the following experiments was 68 °C. This process of melting, annealing, and elongation is commonly repeated 35 times.

2.2 B. Gel Electrophoresis and Gel Extraction

The PCR product is separated from the primers and the template using gel electrophoresis. The DNA migrates through the gel based on its size. Larger DNA fragments will move slower than smaller DNA fragments. The gel is run with a DNA ladder that has a specific set of DNA fragments of a known size. The amplified gene product molecular weight is also known and can be identified based on its relative position to the molecular weight ladder. The amplified PCR product is extracted from the gel using QIAquick® Gel Extraction Kit (Qiagen®).
2.2 C. Restriction Digest

As previously noted the primers for PCR are engineered to have restriction enzyme cut sites. A restriction enzyme is a bacterial enzyme that recognizes small specific DNA sequences, called cut sites, and cleaves the DNA. Restrictions enzymes generally leave an overhang of base pairs after cleavage. The overhang of base pairs is referred to as “sticky ends” that will favor hydrogen bonding with a complimentary “sticky end” (Figure 2.1). A restriction digest involves one or two restriction enzymes to cut the ends of the gene of interest and the plasmid that the gene is to be ligated into. After digestion, the gene of interest and plasmid should have complimentary “sticky ends” that will favor hydrogen-bonding interactions. The products from the digestion are separated by gel electrophoresis and gel extracted (Section 2.3).

2.2 D. Ligation

The ligation reaction brings together two fragments of DNA and ligates them together. The reaction contains the digested gene of interest, plasmid, and DNA ligase. The “sticky ends” of the gene of interest and plasmid participate in hydrogen bonding which helps align the two fragments to be pasted/ligated together by DNA ligase. However, “sticky ends” aren’t necessary for ligation to occur. Blunt ends can also be ligated. Also, vital to a ligation reaction is a phosphate group, which is on the end of the cleaved DNA. The vector now has the gene of interest inserted into the desired plasmid.
2.3. Molecular Biology Techniques: Transformation

Once the plasmid has been engineered it has to be replicated. Replication of the plasmid is achieved by utilizing *Escherichia coli*. Generally, the *E. coli* used in the laboratory are chemically competent. Chemically competent *E. coli* have an increased likelihood of accepting DNA from their environment. After the *E. coli* has been transformed with the plasmid, the bacteria will begin to replicate the plasmid. The replicated plasmid is then harvested from the cells and purified. During this process the bacterial cell is lysed with a basic solution containing detergent. The plasmid is purified by binding to an anion exchange resin. The following sections are a more detailed description of transformation and plasmid purification.
2.3 A. Transformation

Transformation refers to the ability of bacteria to absorb plasmids from the environment. The first step of transformation is ice incubation of chemically competent *E. coli* and the ligated plasmid. The bacteria and plasmid mixture is then heat shocked at 42 °C for 30 seconds. The heat shock step opens up the *E. coli*’s discontinuous membrane. Once the membrane is open, plasmids have the opportunity to be absorbed into the cell. The membrane is reconstituted by immediate ice incubation after heat shock.

After transformation the bacteria are incubated long enough to assure the bacteria have divided three times. *E. coli* divide approximately every 20 minutes so the incubation time is generally one hour at optimal growing temperature (37 °C). The reason for the one-hour outgrowth is all plasmids used in the lab have a gene coding for a particular antibiotic resistance. The resistance gene allows for the selection of bacteria that contain the transformed plasmid. Selection is achieved by culturing the bacteria on plates that have Luria-Bertani broth (LB), agar, and a specific antibiotic corresponding to the antibiotic resistant gene of
the plasmid. Bacteria that have the plasmid will form colonies, and those that do not have the plasmid will not form colonies. However, immediately after transformation the antibiotic resistant gene of the plasmid has not had time to begin expression. For this reason the cells are incubated at optimal growing temperature to allow for the expression of the antibiotic gene before plating. After the one-hour incubation at 37 °C the bacteria are plated on LB, agar, and antibiotic (e.g. ampicillin) plates. The bacteria are then incubated 12-14 hours, at 37 °C.

2.3 B. Mini-prep:

Colonies on the LB, agar, and antibiotic plate are picked and added to LB and antibiotic media. The bacteria are placed in an incubator set at 37 °C and shaken. The culture is grown until the cell density is 0.6 - 0.8 absorbance at 600 nm. The culture is centrifuged at 4000 rpm for 20 minutes to pellet the bacteria. The following steps of the mini-prep are per the documented protocol for the QIAquick® Spin Miniprep Kit (Qiagen®). The mini-prep lyses the cells and extracts plasmid DNA. After mini-prep the plasmid is pure and ready for further characterization. Further characterization of the plasmid includes a test restriction digest to confirm gene insertion and sequencing of the gene of interest to confirm that the insert is the gene of interest. The test restriction digest is similar to the original restriction digest, except the sample volume is smaller and the ligated plasmid is only digested. Gel electrophoresis is then used to confirm an insertion. If insertion did occur into the plasmid then a band at the molecular weight of the gene of interest should separate from the plasmid on the gel. The University of Michigan DNA sequencing core for the presented research did the gene of interest sequencing. Once, the
plasmid has been verified to have the proper gene insert it is transformed into an engineered \textit{E. coli} strain that allows the expression of protein. These strains of \textit{E. coli} are referred to as expression strains. The procedure for transformation of the plasmid into the expression strain is the same as transformation previously described (section 3.1).

\textbf{2.4. Expression of Protein of Interest}

The expression vector within an expression strain of \textit{E. coli} is now able to express the protein of interest. The expression strain is cultured in LB and a selecting antibiotic media at 37 °C and shaken until it has a cell density of 0.6 – 0.8 absorbance at 600 nm. Up to this point expression of the gene of interest has been inhibited by \textit{lac} repressor, which binds upstream of the gene of interest on a sequence known as the lac operator. The lac operator is the regulator region of the \textit{lac} operon. Once lactose binds to the lac operating region the \textit{lac} repressor is released and expression induced. \textit{lacI} is induced experimentally using isopropyl \textbeta-D-1-thiogalactopyranoside (IPTG), which is a lactose gratuitous inhibitor. IPTG binds to the \textit{lac} repressor, releasing it, and inducing expression. IPTG is a gratuitous inducer so it will not be degraded within the cell and continue to induce expression. The incubation temperature is generally dropped at the time of induction to aid in the solubility of the protein (20-16 °C). The culture is generally induced for a period of 16 – 20 hours.
2.5. Analytical Techniques: Protein Purification

The induced cells are harvested from the media, resuspended in a buffer, lysed, and the protein of interest purified using liquid chromatographic techniques. The liquid chromatographic techniques used are immobilized metal affinity and size exclusion chromatography. The immobilized metal affinity chromatography separates the protein of interest from the bacterial proteins and the size exclusion chromatography is to remove any aggregates from the protein of interest. The following is a more in depth description of each step of the purification process.

2.5 A. Harvesting of E. coli

Following the induction of the protein of interest the cells are harvested using Millipore Pellicon 2 Mini filter (0.45 µm). The filter separates the bacterial cells from the liquid media. However, some of the media still remains after the filtration. The cells are completely separated from the media by centrifuging at 4000 rpm for 30 minutes. The supernatant is decanted and the pellet re-suspended in a low imidazole buffer. Further details of the low imidazole buffer will be discussed in section 5.3

2.5 B. Lysis of E. coli

E. coli is lysed with lysozyme and sonication. DNase I is also added to the bacteria to degrade DNA. Lysozyme is an enzyme that degrades the polysaccharides on the cell membrane. Sonication utilizes ultrasound to vibrate and disrupt the cell
membrane lysing the cells. After lysis the cell fragments are pelleted by centrifugation. The supernatant is removed from the pellet and should contain the protein of interest.

2.5 C. Immobilized Metal Affinity Liquid Chromatography

The protein expressed is either a fusion protein with a poly-histidine tag or has an engineered poly-histidine tag. A poly-histidine tag is generally six to ten histidines located at either the N-terminus or the C-terminus of the protein. A fusion protein consists of the protein of interest and a “tag” protein that has a poly-histidine tag at the N-terminus. An engineered poly-histidine tag can either be on the N-terminus or C-terminus of the protein. A histidine tag is helpful in the purification of protein because the secondary amine of the imidazole ring of histidine will chelate Ni\(^{2+}\). Accordingly, the gene of interest with either a fusion tag protein or an engineered poly-histidine tag will chelate to the immobilized Ni\(^{2+}\) of the column, while the other bacterial proteins flow through. The protein is eluted off the column by increasing the concentration of imidazole over time. Imidazole is similar to the aromatic ring of histidine, and as the imidazole concentration increases it will compete for binding to Ni\(^{2+}\) eluting off the protein of interest.

2.5 D. Immobilized Metal Affinity Liquid Chromatography (Tag Removal)

This step is only taken if the protein of interest is expressed as a tag fusion protein or has a poly-histidine tag with an engineered protease cut site. Between the tag protein and the protein of interest there is a relatively small amino acid sequence that codes for a
cut site of a specific protease. This cut site is not found within the tag or the protein of interest so upon incubation with the protease only the specific engineered cut site will be cut. The protease is generally incubated with the fusion protein at the same time as the fusion protein is being dialyzed into a low imidazole buffer. The length of this step is dependent on the activity of the protease and the estimated time it needs to fully cleave the tag from the protein of interest. The dialysis into the low imidazole buffer is to allow for a second Ni\(^{2+}\) column to be used to separate the protein of interest from the tag protein. The poly-histidine tag will bind to the Ni\(^{2+}\) column. The protease should also bind to the column because it also has a poly-histidine tag. The flow through should have the protein of interest because it lacks a poly-histidine tag.

2.5 E. Size Exclusion Chromatography (SEC)

Size exclusion chromatography (SEC) is used to separate the protein of interest from aggregated protein and protein that has been degraded during the purification process. A size exclusion column has a stationary phase of polymer beads packed within the column. Relatively small protein molecules will migrate slowly through the column because they are able to fit into the internal volume of the polymer beads. Relatively larger protein molecules will migrate faster because they are not able to fit into the internal volume of the polymer beads. This allows for the separation of protein molecules based on size. Thus, aggregates should elute first, followed by the protein of interest, and then any truncated protein molecules. After the final purification step of SEC other polishing steps may be needed to further purify the protein (i.e. ion exchange
However, in the case of MTA/SAH nucleosidase from *H. pylori* the protein was deemed pure after SEC.

The molecular weight of the protein is estimated using sodium dodecyl sulfate poly-acrylamide gel electrophoresis (SDS-PAGE). Fractions from SEC containing the protein of interest are run on a SDS-PAGE gel with a molecular weight ladder. The molecular weight ladder, similar to the ladder used in the agarose gel electrophoresis, has a set of known molecular weight proteins. The molecular weight of the protein of interest can be calculated using the relative position of the ladder.

### 2.6. Summary

The steps to obtained purified protein include the genetic engineering of a plasmid and cloning of that plasmid into an expression cell line of *E. coli*. The *E. coli* are cultured and induced to express protein using IPTG. Following induction the *E. coli* is harvested from the media and lysed. The supernatant from the lysis has the protein of interest and is purified from the bacterial proteins using immobilized metal affinity chromatography and size exclusion chromatography. After the protein is determined to be in its most pure state it is generally dialyzed into either a crystal buffer for crystallization experiments or a catalysis buffer for catalytic analysis.
Chapter 3

Methods: Expression and Purification of MTA/SAH

Nucleosidase from *H. pylori*

3.1. Introduction

MTA/SAH nucleosidase of *H. pylori* is the only enzyme that degrades both MTA and SAH. It is a viable drug target because of its involvement in many cellular pathways such as polyamine, methionine, and autoinducer synthesis, as well as biological methylation. Further evidence supporting MTA/SAH as a good drug target is that the mammalian cell uses two separate and different enzymes for the degradation of MTA (i.e. MTA phosphorylase) and SAH (i.e. SAH hydrolase).

The following is a detailed description of the molecular cloning, expression, and purification of MTA/SAH nucleosidase. The MTA/SAH nucleosidase (MTAN) gene
from *H. pylori, pfs*, was amplified using PCR and ligated into pET32 vector (Novagen). T7 express® + pRosetta (pROS) *E. coli* cells were transformed with pET32 + pfs. T7 express® + pROS + pET32 + pfs were cultured and MTAN expression induced with IPTG. T7 express® + pROS + pET32 + pfs were harvested and lysed, and MTAN extracted and purified using liquid chromatographic techniques.

### 3.2. Vector Construct: pET32 (Novagen®) + pfs

The following methods were used by Dr. Donald Ronning for the development of the pET32 + pfs vector. *pfs* was amplified with polymerase chain reaction (PCR). The PCR product and pET32 were restriction digested with restriction enzymes NcoI and BamHI. pET32 and *pfs* were ligated together using DNA ligase and transformed into XL-10 *E. coli* cells. Colonies from the transformation were cultured and the transformed plasmid was purified. Test digestion and gene sequencing confirmed the *pfs* insertion into pET32. The following is a more detailed description of the vector construct methods.

#### 3.2 A. Amplification of *pfs* using polymerase chain reaction (PCR)

The *pfs* gene was amplified from genomic *H. pylori* DNA as the template using a polymerase chain reaction (PCR). Primers for the *pfs* amplification were for the forward primer 5’-CACCATGGGCAAAAAATTGGCATTTTAGGGGC-3’ and for the reverse primer 5’-CCGGATCCCTAAAGCTCATCCACCATGCTTT-3’. The forward primer binds the beginning of gene and elongation proceeds forward. The reverse primer binds the end of the gene and elongation proceeds backwards, or reverse. The components of
the PCR reaction are listed in Table 3.1. \textit{Pfx50} (Invitrogen™) was the DNA polymerase used in the PCR and formamide was included to decrease non-specific interactions between the primers and template as well as increase amplification. Also, the 10x \textit{Pfx50} buffer was provided with the \textit{Pfx50} from Invitrogen™. The PCR parameters had a melting temperature of 95 °C, annealing temperature of 53 °C, and an elongation temperature of 68 °C (Table 3.2). The 4 °C hold temperature is to keep the reaction cold until gel electrophoresis. The PCR was run on an Eppendorf Mastercycler personal thermocycler.

<table>
<thead>
<tr>
<th>Volume</th>
<th>PCR Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5 µL</td>
<td>Genomic DNA (\textit{H. pylori})</td>
</tr>
<tr>
<td>1.0 µL</td>
<td>5’ primer</td>
</tr>
<tr>
<td>1.0 µL</td>
<td>3’ primer</td>
</tr>
<tr>
<td>37.5 µL</td>
<td>Deionized H₂O (dH₂O)</td>
</tr>
<tr>
<td>5.0 µL</td>
<td>10x \textit{Pfx50} buffer</td>
</tr>
<tr>
<td>1.0 µL</td>
<td>Deoxynucleotide triosphosphates (dNTPs)</td>
</tr>
<tr>
<td></td>
<td>(10mM)</td>
</tr>
<tr>
<td>1.0 µL</td>
<td>Formamide</td>
</tr>
<tr>
<td>1.0 µL</td>
<td>\textit{Pfx50} (DNA polymerase 5 U / µL)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Replication</th>
<th>Temperature</th>
<th>Duration</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1x</td>
<td>95 °C</td>
<td>5 min.</td>
<td>Melting temp.</td>
</tr>
<tr>
<td>35x</td>
<td>95 °C</td>
<td>30 sec.</td>
<td>Melting temp.</td>
</tr>
<tr>
<td></td>
<td>53 °C</td>
<td>45 sec.</td>
<td>Annealing temp.</td>
</tr>
<tr>
<td></td>
<td>68 °C</td>
<td>2 min.</td>
<td>Elongation temp.</td>
</tr>
<tr>
<td>1x</td>
<td>68 °C</td>
<td>4 min.</td>
<td>Elongation temp.</td>
</tr>
<tr>
<td>1x</td>
<td>4 °C</td>
<td>hold</td>
<td></td>
</tr>
</tbody>
</table>

The PCR product was separated from the genomic DNA and primers with gel electrophoresis. To the PCR product 10 µL of 6X DNA loading dye was added and loaded into the wells of a 1% agarose gel. The gel electrophoresis parameters were 135
V for 30 min. The PCR was successful and the product was extracted per the documented protocol of QIAquick\textsuperscript{®}: Gel Extraction Kit (Qiagen\textsuperscript{®}).

### 3.2 B. Restriction digest of pfs and pET32 (Novagen\textsuperscript{®})

The amplified pfs and pET32 (Novagen\textsuperscript{®}) were restriction digested with NcoI and BamHI from New England Biolabs. The restriction digest reactions for both the pET32 and pfs are in Table 3.3. Also, added to the pET32 reaction was calf intestinal phosphatase (CIP). The role of CIP is to cleave off the phosphates at the end of linear DNA fragments. The cleaving of the phosphate ends of pET32 will decrease the likelihood that upon ligation pET32 will self ligate because it will not have an available phosphate for the ligation reaction. The reaction was incubated at 37° C for 30 min. The digested fragments were separate using gel electrophoresis. To each restriction digest 8 µL of 6X DNA loading dye was added and loaded into the wells of a 1% agarose gel. The gel electrophoresis parameters were 135 V for 30 min. The digest gene of interest and the plasmid was gel extracted per the documented protocol of QIAquick\textsuperscript{®}: Gel Extraction Kit (Qiagen\textsuperscript{®}).

<table>
<thead>
<tr>
<th>pfs Restriction Digest Reaction</th>
<th>Volume</th>
<th>Restriction Digest Components</th>
<th>pET32 Restriction Digest Reaction</th>
<th>Volume</th>
<th>Restriction Digest Components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.0 µL</td>
<td>10X Buffer #3 (NEB)</td>
<td></td>
<td>4.0 µL</td>
<td>Buffer #3 10X (NEB)</td>
</tr>
<tr>
<td></td>
<td>4.0 µL</td>
<td>dH\textsubscript{2}O</td>
<td></td>
<td>18.5 µL</td>
<td>dH\textsubscript{2}O</td>
</tr>
<tr>
<td></td>
<td>1.0 µL</td>
<td>NcoI</td>
<td></td>
<td>1.0 µL</td>
<td>NcoI</td>
</tr>
<tr>
<td></td>
<td>1.0 µL</td>
<td>BamHI</td>
<td></td>
<td>1.0 µL</td>
<td>BamHI</td>
</tr>
<tr>
<td></td>
<td>0.5 µL</td>
<td>BSA</td>
<td></td>
<td>0.5 µL</td>
<td>BSA</td>
</tr>
<tr>
<td></td>
<td>30.0 µL</td>
<td>pfs</td>
<td></td>
<td>15.0 µL</td>
<td>pET32</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td></td>
<td>0.5 µL</td>
<td>CIP</td>
</tr>
</tbody>
</table>
3.2 C. Ligation of \textit{pfs} and pET32

The restriction digest of \textit{pfs} and pET32 were ligated using DNA ligase (Roche). The ligation reaction is in Table 3.4. The reaction was incubated at room temperature for 5 minutes. The ligated product of pET32 + \textit{pfs} creates a fusion protein of MTA/SAH nucleosidase and a poly-histidine tagged thioredoxin (Trx-Tag) protein. The fusion of MTA/SAH nucleosidase with Trx-Tag will be used later for protein purification.

Table 3.4: Ligation Reaction of pET32 + \textit{pfs}

<table>
<thead>
<tr>
<th>Volume (µL)</th>
<th>Ligation Reaction Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>5x DNA dilution buffer (Roche)</td>
</tr>
<tr>
<td>5.0</td>
<td>2x DNA dilution buffer (Roche)</td>
</tr>
<tr>
<td>0.5</td>
<td>Cleaved pET32</td>
</tr>
<tr>
<td>3.5</td>
<td>Cleaved \textit{pfs}</td>
</tr>
<tr>
<td>0.5</td>
<td>Rapid DNA ligase (Roche)</td>
</tr>
</tbody>
</table>

3.2 D. Transformation of XL-10 with pET32 + \textit{pfs}

XL-10 \textit{E. coli} cells were transformed with the ligated pET32 + \textit{pfs}. One µL of ligated pET32 + \textit{pfs} was added to 20 µL of XL-10 cells and incubated on ice for 5 minutes. XL-10 + (pET32 + \textit{pfs}) was heat shocked at 42 °C for 30 seconds and immediately placed back in an ice bath. XL-10 + (pET32 + \textit{pfs}) were incubated on ice for 2 minutes. To the XL-10 + (pET32 + \textit{pfs}) mixture 200 µL of Luria-Bertani (LB) broth was added and incubated at 37° C for one hour. The transformation of XL-10 + (pET32 + \textit{pfs}) was added to a LB, 10% agar, and 0.269 mM ampicillin plate and spread around using a sterile glass stick. The plate was incubated at 37 °C overnight.
3.2 E. Miniprep and verification of gene insertion

Colonies of XL-10 + (pET32 + pfs) grew overnight. Of those colonies, two colonies were picked and each was added to a separate 15 mL conical tube containing 5 mL LB and 0.269 mM ampicillin. Colonies were cultured to a cell density between 0.6-0.8 absorbance units at 600 nm. The bacteria were harvested from the media by centrifugation at 4000 rpm for 20 minutes at 4 °C. The harvested cells were miniprepped with QIAprep: Miniprep using a microcentrifuge (Qiagen®) per the documented protocol to extract the plasma DNA from the XL-10 cells.

The pET32 + pfs plasmid was verified for insertion by a test restriction digest (Table 3.5) and DNA sequencing. The test restriction digest was incubated at 37 °C for 30 minutes and analyzed for insertion using gel electrophoresis. To the test restriction digest reaction 3.0 µL of 6x DNA loading dye was added to the reaction. The gel was a 1% agarose gel and the parameters for the gel electrophoresis were 135 V for 30 minutes (Figure 3.1). The DNA sequencing was done by University of Michigan DNA sequencing core and results from the sequencing confirmed the inserted gene was pfs (Appendix I).

Table 3.5 Test restriction digest of pET32+pfs plasmid

<table>
<thead>
<tr>
<th>Volume</th>
<th>Test Restriction Digest Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5 µL</td>
<td>10X Buffer #3 (NEB)</td>
</tr>
<tr>
<td>4.35 µL</td>
<td>dH2O</td>
</tr>
<tr>
<td>0.5 µL</td>
<td>NcoI</td>
</tr>
<tr>
<td>0.5 µL</td>
<td>BamHI</td>
</tr>
<tr>
<td>0.15 µL</td>
<td>BSA</td>
</tr>
<tr>
<td>8.0 µL</td>
<td>pET32</td>
</tr>
</tbody>
</table>
3.3. Transformation of T7 express® + pROS with pET32 + pfs

T7 express® E. coli cells were the expression strain used for the expression of pfs. The T7 express® strain had been previously transformed with a pROS plasmid which codes for tRNA with the anti-codons AUA, AGG, AGA, CUA, CCC, or GGA. These tRNA anti-codons are deemed to be rare anti-codons in E. coli. Since the pfs is from H. pylori the availability of these rare anti-codons may affect expression. pROS overcomes this potential problem by expressing those rare anti-codons. Also, pROS has a chloramphenicol resistant gene to select for the pROS plasmid. The transformation of T7® + pROS with pET32 + pfs follows the same protocol as the transformation of XL-10 with pET32 + pfs (Section 2.4). The only modification to the protocol was the plate has the addition of 99 µM chloramphenicol (i.e. LB, 10% agar, 0.269 mM ampicillin, and 99 µM chloramphenicol). Colonies of T7® + pROS with pET32 + pfs grew overnight.
3.4. Expression and solubility test of the fusion protein SAH/MTA nucleosidase (MTAN) and Trx-Tag

One colony from the transformation of T7® + pROS with pET32 + pfs was picked and added to 5 mL LB with 0.269 mM ampicillin, and 99 µM chloramphenicol. The culture was incubated at 37 ºC until it reached an optical density of 0.6-0.8 absorbance units at a wavelength 600 nm. A glycerol stock of 200 µL of T7® + pROS + pET32 + pfs and approximately 50 µL of glycerol was added to a microcentrifuge tube. The glycerol stock was vortexed to mix the cells and glycerol and placed in the -80 ºC freezer. The remaining 4.8 mL of T7® + pROS + (pET32 + pfs) cells were induced with 4.8 µL of 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG). The cells were incubated at 37 ºC for 3 hours, and harvested by centrifugation at 4000 rpm for 20 minutes at 4 ºC. The supernatant was decanted and the pellet resuspended in a 200 µL buffer of 20 mM sodium phosphate pH 7.5, 25 mM imidazole, 5 mM β-Mercaptoethanol (BME), and 0.5 M NaCl and transferred to a microcentrifuge tube. From the resuspended pellet 15 µL were removed and added to a microcentrifuge tube and labeled pellet. The remaining resuspended cells were lysed with 1 µL of 1000x lysozyme stock. Also, added to the resuspended cells was 1 µL of 1000x stock of DNase I. The cells were incubated with the lysozyme and DNase I for 20 minutes at room temperature.

Lysis of the cells was done by three freeze/thaw cycles using liquid nitrogen. The microcentrifuge tube of resuspended cells was placed in the liquid nitrogen until frozen and thawed in a 42 ºC water bath. This cycle was repeated a total of three times. The lysed cells were pelleted to separate the cell components from the soluble protein. The supernatant was extracted from the pellet using a pipette, placed in a microcentrifuge
tube, and the pellet discarded. From the extracted cells 15 µL were added to a microcentrifuge tube and the sample was labeled supernatant. To both the pellet and supernatant sample 5 µL of NuPAGE dye was added and incubated at 72 °C for 10 minutes. Both samples were checked for MTAN expression by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

The conditions for SDS-PAGE were 185 V for 35 min. (Figure 3.2). The gel was added to Commassie blue stain. The gel and stain were heated using a microwave for 30 sec. and shaken for 10 min. at room temperature. The gel was transferred from the staining solution to a destaining solution, which was 10% vol/vol in water acetic acid. The gel was heated by a microwave for 30 sec. and shaken for 10 minutes at room temperature. The pellet and the supernatant had expression of MTAN indicating that MTAN was being expressed and was soluble in the 20 mM Na₂HPO₄ pH 7.5, 25 mM imidazole, 5 mM BME, and 0.5 M NaCl buffer.

![Figure 3.2: Expression and solubility test of MTAN (P = pellet, S = supernatant, and Trx-SNu = fusion protein MTAN and Trx-Tag)]
3.5. Protein Purification

Protein purification uses a variety of analytical techniques to separate MTAN from *E. coli* proteins that are also soluble in the chosen buffer. The steps for the purification of MTAN were growth of a 4 L culture of T7® + pROS + (pET32 + pfs), induction, harvest, and lysis of T7® + pROS + (pET32 + pfs), immobilized metal affinity chromatography (IMAC), and size exclusion chromatography (SEC) of lysate. The following is a more detailed description of each step of protein purification.

3.5 A. Culture and Harvest of T7® + pROS + (pET32 + pfs)

The T7® + pROS + (pET32 + pfs) glycerol stock from the expression solubility test was added to 100 mL LB with 0.269 mM ampicillin, and 99 µM chloramphenicol. The cells were incubated at 37 °C until they reached a cell density of 0.6-0.8 absorbance units at 600.00 nm. From the 100 mL culture 50 mL was transferred to 4 L of LB with 0.269 mM ampicillin, and 99 µM chloramphenicol. The remaining 50 mL was utilized for two glycerol stocks (25 mL of T7® + pROS + (pET32 + pfs) and 10 mL glycerol). The cells were incubated at 37 °C until they reached a cell density of 0.6-0.8 absorbance at 600 nm, and were induced with 1 mM IPTG. The incubation temperature was dropped to 16 °C and incubated between 16-20 hours.

The induced The T7® + pROS + (pET32 + pfs) were harvested by filtration and centrifugation. The culture was filtered by a Millipore Pellicon 2 Mini filter (0.45 µm) until the volume of the culture was approximately 800 mL. An additional 400 mL of cells were rinsed from the filter with water. The total volume of cells, 1200 mL, was fully separated from the media by centrifuging at 4000 rpm for 20 min at 4 °C. The supernatant was decanted and the pellet resuspended in 25 mL of 20 mM Na₂HPO₄ pH
7.5, 25 mM imidazole, 5 mM BME, and 0.5 M NaCl buffer. The resuspended cells were added to a 50 mL conical tube and placed in the -80 °C freezer until lysis and purification.

3.5 B. Lysis of T7® + pROS + (pET32 + pfs)

Harvested cells of T7® + pROS + (pET32 + pfs) were thawed and transferred to a 50 mL beaker. To the thawed cells 10 µL of 1000x lysozyme and 1000x DNase I was added. Also, 15 µL of 12 M BME was added to the thawed cells. Cells were placed on ice and lysed using sonication instead of freeze thaw cycles. The sonicator was a Misonix 3000 sonicator and the program was a cycle of 30 sec. sonication, at a power level of 6, and 2 min. off. This cycle was repeated five times. The sonicated cells were incubated on ice for 20 min., transferred to a 50 mL conical tube, and centrifuged at 11,000 rpm for 30 min. at 4 °C. The supernatant, or lysate, was poured off into a 50 mL beaker and the pellet discarded. The lysate was filtered using a Millex® -GP (33mm) Millipore syringe filters (0.22 micron cutoff) to remove any precipitates.

3.5 C. Immobilized Metal Affinity Chromatography (IMAC)

The first liquid chromatography step of protein purification was IMAC. The column used for IMAC was a 5 mL GE Healthcare HiTrap Chelating column and run on an AKTA FPLC. The HiTrap Chelating column immobilized metal was Ni²⁺. Unicorn software was used to run a designed automated purification protocol and record information regarding the gradient and absorbance of the chromatographic protocol.
The HiTrap Chelating column was first charged with 5 mL of 50 mM nickel (II) sulfate. Excess nickel (II) sulfate was washed from the column with 5 mL dH₂O. The column was placed on the AKTA FPLC. Running 20 mM Na₂HPO₄ pH 7.5, 25 mM imidazole, 5 mM BME, and 0.5 M NaCl buffer over the column until there was no change in absorbance over time equilibrated the column. The syringe-filtered lysate was manually loaded onto the column. Flow through from the column was collected. The column was washed with 10 column volumes of the loading buffer (20 mM Na₂HPO₄ pH 7.5, 25 mM imidazole, 5 mM BME, and 0.5 M NaCl), and the protein eluted off the column by a gradient of 0% - 100% elution buffer (20 mM Na₂HPO₄ pH 7.5, 0.25 M imidazole, 5 mM BME, and 0.5 M NaCl) over 20 column volumes. To assure that all of the protein was eluted from the column 2 column volumes of the elution buffer were run over the column. Fractions from the elution buffer were collected throughout the gradient and elution buffer rinse (Figure 3.4). Fractions B1-B4 were run on a SDS-PAGE gel (15 µL fraction and 5 µL of NuPAGE dye) and stained with Comassie blue using the same procedure in section 4 (Figure 3.3).

Fractions B1-B4 were dialyzed into the loading buffer and incubated with PreScission Protease to cleave the Trx-Tag. The Fisher dialysis tubing (seamless-cellulose: width 32 mm, diameter 20.4 mm, 12,000 Da molecular weight (MW) cutoff) was hydrated for 5 minutes in reverse osmosis (RO) water. One side of the dialysis tubing was clamped and fractions B1-B4 with 1 mL PreScission Protease were added to dialysis tubing and the other side clamped. The dialysis tubing was added to 500 mL of loading buffer and dialyzed overnight at 4 °C. After the dialysis the imidazole
concentration decreases to 25 mM and the fusion protein has been cleaved to form the Trx-Tag and MTAN.

Figure 3.3: SDS-PAGE gel of fractions B1-B4 of HiTrap chelating column of MTAN

Figure 3.4: Chromatogram of HiTrap chelating (Ni\(^{2+}\)) column of T7\(^{\oplus}\) + pROS + pET32 + pfs lysate. The blue line is the absorbance of the lysate and the green line is the gradient of elution buffer.
3.5 D. IMAC Tag Removal

The second IMAC column follows the same protocol as the first (section 5.3), but instead of loading the lysate the dialyzed fractions B1-B4 were loaded onto the column (Figure 3.5). Also, the column had already been charged for the first column so the column was not recharged again for the second column. The flow through and the first fraction (A1) were run on a SDS-PAGE gel (15μL of sample and 5 μL NuPAGE Dye) and stained with Commassie blue using the same procedure as in Section 4 (Figure 3.6).

![Figure 3.5: Chromatogram of HiTrap chelating (Ni²⁺) tag removal column of fractions B1-B4 from the initial HiTrap chelating column (Figure 3.3). The blue line is the absorbance and the green line is the gradient of elution buffer.](image-url)
3.5 E. Size Exclusion Chromatography (SEC)

Flow through and fraction A1 from the IMAC tag removal were pooled and concentrated by (NH$_4$)$_2$SO$_4$ precipitation. (NH$_4$)$_2$SO$_4$ was added to the pooled protein to reach a concentration of 2.40 M (375 mg/mL). The protein was incubated on ice until all of the (NH$_4$)$_2$SO$_4$ was dissolved. The protein plus (NH$_4$)$_2$SO$_4$ were centrifuged at 11,000 rpm for 20 minutes at 4 °C. The supernatant was decanted and the pellet resuspended in 4 mL Size Exclusion Chromatography (SEC) buffer (20 mM Na$_2$HPO$_4$ pH 7.5, 0.3 M NaCl, and 5 mM BME). The resuspended pellet was injected onto an Amersham Biosciences Superdex HiLoad 16/60 column for SEC (Figure 3.7). Fractions 19-29 were
run on SDS-PAGE gel (15 µL of fraction and 5 µL of NuPAGE dye) and stained with Comassie blue using the same procedure as in section 4 (Figure 3.8).

Figure 3.7: SEC of concentrated MTAN from tag removal. The blue line is UV absorbance at 280 nm and the red line is conductivity.
3.6. Dialysis of pure MTAN

Fractions 29-35 were pooled and an absorbance at 280 nm taken to determine the concentration. A GENESYS 6 UV-Vis Spectrophotometer from Thermo Scientific was used for determining the absorbance at 280 nm. The blank used for determining the absorbance was the SEC buffer. The absorbance at 280 nm was 0.416. The extinction coefficient was calculated using a program called ProtParam. This program uses the primary amino acid sequence to calculate a theorized extinction coefficient. MTAN from \textit{H. pylori} had a theorized extinction coefficient of 3105 cm$^{-1}$ M$^{-1}$ and a 0.1% absorbance
equaling 0.124 (Appendix II) Using the Beer-Lambert law (Equation 3.1) the concentration of MTAN after SEC was 3.35 mg/mL. The total volume of the protein was 17 mL, which indicates a total yield of 57.0 mg for the purification. The protein was concentrated to 5 mL using Amicon® Ultra – 15 Centrifugal Filter Devices with a MW cutoff of 10,000 Da. The filter was rinsed separately with 15 mL dH₂O and 15 mL of SEC buffer by centrifuging at 4000 rpm for 10 minutes at 4°C. The pooled protein was added to the centrifugal filter and concentrated by centrifuging at 4000 rpm for 10 minutes at 4 °C to a volume of 5 mL. The protein was removed from the centrifugal filter by a pipette and 1 mL was dialyzed into catalysis buffer and the rest into a crystal buffer. The catalysis buffer had 100 mM HEPES pH 7.5 and 50 mM KCl. The crystal buffer had 20 mM Tris pH 7.5, 0.2 mM TCEP, and 1 mM EDTA. The absorbance at 280 nm of the catalysis buffer dialyzed protein was 0.927 for a concentration of 7.48 mg/mL and for the crystal buffer dialyzed protein 1.512 for a concentration of 12.19 mg/mL. Further catalytic and structural analysis MTAN will be discussed in the following chapters.

\[ A = \varepsilon bc \]

\( A = \) absorbance  
\( \varepsilon = \) molar absorptivity  
\( b = \) path length  
\( c = \) concentration

*Equation 3.1: Beer-Lambert Law*

### 3.7 Summary

The *pfs* gene has been cloned into pET32 (Novagen®) plasmid. The ligation of the *pfs* into pET32 generates a fusion protein of MTAN + Trx-Tag. The Trx-Tag has a poly-histidine tag that will bind to the Ni²⁺ IMAC column. The initial IMAC column separated the protein of interest from the rest of the bacterial proteins. The tag has an
engineered cleavage site for PreScission Protease, which was used to cleave the tag from MTAN. A second Ni^{2+} IMAC column was used to separate MTAN from the cleaved tag. A polishing step for the purification of MTAN was SEC. The pure protein was dialyzed into crystallization and catalysis buffer for further catalytic and crystallization analysis.
Chapter 4

General Methods: Macromolecular X-ray Crystallography

4.1. Introduction

Protein interactions with biomolecules and organic compounds rely substantially on the tertiary and quaternary structure of the protein. However, the ability to experimentally solve the tertiary and quaternary structure of a protein is quite difficult. Two techniques that are currently being used to solve protein structures are \( ^1H \) Nuclear Magnetic Resonance (NMR) and macromolecular X-ray crystallography. The technique that was used in the elucidation of MTA/SAH nucleosidase (MTAN) protein structure was macromolecular X-ray crystallography. The following is a description of the macromolecular X-ray crystallography theory, crystallization techniques, and general methods on solving a protein structure with X-ray crystallography.
4.2. Macromolecular X-ray crystallography theory

The principle to which all X-ray crystallography is based off is that crystals will diffract an incident X-ray beam to produce a diffraction pattern. Data obtained from the diffraction pattern can then be used to solve for the atomic positions of the molecular structure of the crystal.

A crystal is a three dimensional molecular array of repeating units. There are seven crystal systems that fit the criteria for a repeating structural motif (Table 4.1). The crystal systems are furthered characterized into 14 Bravais lattices. The Bravais lattices are used to describe the unit cell. The unit cell is the smallest repeat unit of the crystal with the highest symmetry elements. Also, Bravais lattice points are where protein molecules are arranged within the crystal. The protein molecules on the lattice points compose the asymmetric unit. The asymmetric unit is composed of the backbone structure of the protein and has none of the protein atoms systematically related, and is the fundamental unit for the construction of the unit cell.48

<table>
<thead>
<tr>
<th>Crystal System</th>
<th>Unit Cell Edges</th>
<th>Unit Cell Angles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cubic</td>
<td>a = b = c</td>
<td>α = β = γ = 90°</td>
</tr>
<tr>
<td>Hexagonal</td>
<td>a = b ≠ c</td>
<td>α = β = 90°, γ = 120°</td>
</tr>
<tr>
<td>Rhombohedral</td>
<td>a = b = c</td>
<td>α = β = γ ≠ 90°, 120°</td>
</tr>
<tr>
<td>Tetragonal</td>
<td>a = b ≠ c</td>
<td>α = β = γ = 90°</td>
</tr>
<tr>
<td>Orthorhombic</td>
<td>a ≠ b = c</td>
<td>α = β = γ = 90°</td>
</tr>
<tr>
<td>Monoclinic</td>
<td>a ≠ b ≠ c</td>
<td>α = γ = 90°, β ≠ 90°, 120°</td>
</tr>
<tr>
<td>Triclinic</td>
<td>a ≠ b ≠ c</td>
<td>α ≠ β ≠ γ ≠ 90°, 120°</td>
</tr>
</tbody>
</table>

In addition to Bravais lattice symmetry there is point symmetry. Point symmetry are symmetry elements that cannot be translated. The space group of a protein crystal describes the combination of the Bravais lattice and point symmetry.48
The symmetry elements combine to make planes that will reflect X-rays in accordance with Bragg’s law (Equation 4.1). When \( n \) is an integer of the wavelength constructive interference is observed and a spot, or reflection, on the diffraction pattern from the constructive interference is observed. The constructive interference waves give rise to the diffraction pattern, and a reciprocal lattice of the crystal lattice. When \( n \) is equal to zero or not an integer, the absence of a reflection on the diffraction pattern is observed. The presence and absence of reflections from diffraction varies from space group to space group. Also, the analysis of the diffraction pattern for the systematic absences allows for the determination of the space group.\(^{48}\)

\[ n\lambda = 2d_{hkl}\sin\theta_{hkl} \]

\( n \) = integer  
\( \lambda \) = wavelength  
\( d_{hkl} \) = spacing between lattice planes  
\( \theta_{hkl} \) = angle of incidence beam on the lattice plane

*Equation 4.1: Bragg’s Law*

A reflection on the diffraction pattern is from a diffracted electromagnetic X-ray wave of light. The electrons of atoms arranged in a crystalline manner generate the diffraction pattern. Atomic positions are then calculated from the amplitude and phase of the resultant wave from the X-ray. However, the diffraction pattern only holds information on the amplitude of the wave and not the phase (Equation 4.2). This problem is known as the “phase problem” in macromolecular X-ray crystallography and can be overcome either by molecular replacement, anomalous scattering, or isomorphous replacement. The technique used to phase the data for MTAN was molecular replacement. Molecular replacement uses a previously solved homologous structure as a reference phase and solves for the experimental phase relative to that reference phase.
The only difficulty with molecular replacement is obtaining an appropriate reference structure. Ideally the reference model will have > 30% sequence identity to the unknown structure. Once the phase and amplitude are both known of the resultant waves from the diffraction of atoms within the crystal structure the atomic positions can be solved.\(^4\)

\[
\sqrt{I} = |F_{hkl}|
\]

\(I = \text{intensity of the reflection}
\)

\(F_{hkl} = \text{structure factor}
\]

*Equation 4.2: The square root of the intensity of a reflection is equal to structure factor, which contains information of the amplitude of the diffracted X-ray wave.*

### 4.3. Protein Crystallization

The crystallization of a protein for analysis with macromolecular X-ray crystallography is the first step for structure analysis. The process of crystallization is the increase of protein concentration to the levels of super saturation, nucleation of a crystal, and then growth of that crystal. The different types of crystallization techniques include batch, vapor diffusion, and dialysis. The technique used for the crystallization for MTAN was vapor diffusion. Vapor diffusion slowly reaches super saturation, which is followed by a nucleation event and a rapid decrease in protein concentration. The two most common techniques for vapor diffusion are sitting and hanging drop. A sitting drop has the protein and a crystallization solution mixed together in a sitting drop next to a well solution. The well solution has the crystallization solution at higher concentration of precipitants than the drop. Water vapor diffuses from the sitting drop to the well solution increasing precipitant concentration in the drop and promoting nucleation. The hanging drop is similar to the sitting drop, but instead sitting on the edge of the well the drop is suspended over the well on a cover slide.\(^4\)
The crystallization solution for each protein varies from protein to protein. Indeed, there is no systematic method for determining the best crystallization solution to promote nucleation and crystal growth. Instead, sparse matrix screening is used to identify crystallization solutions. Crystallization conditions screening solutions are known conditions that have produced crystals for other proteins. These conditions can vary in terms of salts, volatile organic compounds, polymers, and pH. Other factors that can affect nucleation and crystal growth include drop size, temperature, and drop shape.\textsuperscript{48}

After sparse matrix screening any conditions that produce microcrystals or crystals are optimized. The goal of optimization is to decrease nucleation and increase crystal growth. The optimization includes the variation of the factors affecting crystallization such as the concentration of the crystallization condition, drop size, and temperature.\textsuperscript{48}

Once the protein has crystallized it is ready for analysis by macromolecular X-ray crystallography. Generally, before X-ray analysis is to proceed the crystal has to be flash cooled by placement in a cryo stream. There are two reasons for freezing a crystal for X-ray analysis. The first reason is freezing the crystal will prevent some of the radiation damage from the X-ray. The other is to decrease atomic vibration for X-ray analysis. Since the goal of macromolecular X-ray crystallography is to solve for the atomic positions of the protein it is best that the movement of the atom is decreased as much as possible. A problem resulting from flash cooling is damage to the crystal. This problem is overcome through the use of cryoprotectants. Examples of cryoprotectants include polyethylene glycol (PEG) 400, glycerol, or ethylene glycol. However, cryoprotectants can also disrupt the crystal lattice, so screening for the best cryoprotectants is necessary.\textsuperscript{48}
4.4. Diffraction experiment and data integration

Once the crystal is frozen in the cryo stream it is time for data collection. The crystal is exposed to an X-ray source and a detector records the diffracted waves of light. The diffracted waves that appear on the diffraction pattern are generated by constructive interference and intersect with the Ewald Sphere. The Ewald sphere is a sphere with a radius of $1/\lambda$. When a reciprocal lattice point intersects with the Ewald sphere it is fulfilling the Bragg equation for constructive interference and a corresponding reflection appears on the diffraction pattern. Since the Ewald sphere is three-dimensional and the detector is set in two dimensions the crystal or detector must rotate to detect all reciprocal lattice point that satisfy the Bragg’s law. The most common form of rotation is rotating the crystal while the detector remains stationary. Rotation needed to collect enough data to solve the structure is dependent on space group and needs to be calculated for each data set. Also, a crystal may diffract but the resolution of the diffraction is critical. Resolution is reported as a distance and in order to solve most protein crystal structures need to have a diffraction pattern with a resolution of 3.0 Å or better.  

After the first initial frames of diffraction, the data is indexed to solve for the unit cell type and dimensions. After the unit cell is known the dimensions for the unit cell are refined along with detector distance, crystal position, and mosaicity. Mosaicity is representative of the crystal packing of a protein crystal. Generally the desired value for mosaicity is from 0.3-0.5. After indexing and determining the crystal system the amount of rotation needed to solve the structure is determined and the data set collected.

Integration of the full data set calculates the intensity of a spot over the background. The square root of the intensity determines the structure factor for each
atom, which is the amplitude of the wave and the first component needed to solve for atomic positions. Also, during integration mosaicity and crystal position are recorded.48

4.5. Data scaling

The diffraction pattern is composed of diffracted X-rays that participate in constructive interference and is the reciprocal lattice of the crystal lattice. Such as the crystal lattice has symmetry related elements so does the reciprocal lattice. Data scaling is a process that merges the symmetry related reflections of the reciprocal lattice together based on a selected space group. The quality of merging the diffraction spots is reported as $R_{\text{merge}}$. A good $R_{\text{merge}}$ lies between 0.03-0.10. For this reason, scaling of the data involves selecting multiple higher symmetry elements of the initial crystal system and testing each to see which gives the best $R_{\text{merge}}$ value. Data scaling also reports the extent of the resolution, $I/\sigma$, and completeness

4.6. Resolution, $I/\sigma$, and completeness

A resolution of 3.0 Å or less is typically sufficient to determine amino acid positions. However, the $I/\sigma$, intensity of a reflection over the background, at the selected resolution needs to be considered. An accepted $I/\sigma$ value is much debated in the literature, but in general the greater value for $I/\sigma$ at higher resolution indicates a well diffracting crystal. Also, completeness of the collected data reports the amount of observed reflections relative to the amount of reflections that should be observed of a specified space group. The closer the completeness value is to 100% the better. Furthermore, resolution, $I/\sigma$, and completeness need to be considered to determine the quality of diffraction, and if the data needs to be truncated.
4.7. Molecular replacement

After the data has been scaled, solving for the phase can proceed. A technique that can be used to solve for the phase is molecular replacement. The process involves using the intensities of the experimental data and merging it with the phases of a structure that has already been solved. A requirement for using molecular replacement is that the two structures must be homologous and have a 30% primary amino acid sequence identity. An electron density map is generated from the experimental intensities and the phase from the homologous structure. The first step in the molecular replacement is fitting of the known homologous structure into the electron experimental density map. The sequence of the homologous structure is mutated to have the primary amino acid sequence of the experimental structure. The last step of the refinement is generally a rigid body refinement, which large subunits of the protein are moved to fit the electron density. After the rigid body refinement two statistical R factors are reported (i.e. $R_{\text{work}}$ and $R_{\text{free}}$). The R factor reports the quality of the model structure. However, the ability to artificially lower the $R_{\text{work}}$ is problematic and leads to incorrect model building. This problem is overcome by the $R_{\text{free}}$, which sequesters 10% of the reflections from $|F_{\text{hkl}}|_{\text{obs}}$ in the R factor calculation. This prevents an artificial minimization between $|F_{\text{hkl}}|_{\text{obs}}$ and $|F_{\text{hkl}}|_{\text{calc}}$. The goal for the rest of the model building is to lower both the R and $R_{\text{free}}$ between the values of 0.15 and 0.25.\(^{48}\)

4.8. Model refinement: simulated annealing, manual manipulation, minimization, and B-factor refinement

Model building follows a two-step cycle that involves the refinement of the structure by computational algorithms and manual manipulation. The model is manually
manipulated into the electron density and then selected algorithms are used to refine the new positions. Simulated annealing is one of the computational refinements. Simulated annealing theoretically heats the model to break the covalent bonds between the atoms removing the atoms from local minima. The algorithm then begins to decrease the temperature factor, bringing the atoms into lowest energy state, and more favorable positions. After the simulated annealing two-electron density maps are generated (Equation 4.3). The molecule that was just refined with simulated annealing is manually manipulated. For manual manipulation the crystallographer examines each residue’s position relative to the map and positions a residue into the electron density if needed.\(^\text{49}\)

\[
\begin{align*}
A. & \quad 2 |F_{hkl}^{\text{obs}} - F_{hkl}^{\text{calc}}| \\
B. & \quad |F_{hkl}^{\text{obs}} - F_{hkl}^{\text{calc}}|
\end{align*}
\]

\[|F_{hkl}^{\text{obs}}| = \text{observe structure factors} \quad |F_{hkl}^{\text{calc}}| = \text{calculated structure factors}\]

\textit{Equation 4.3: Calculating an electron density map}

After simulated annealing refinement the model is refined using a minimization algorithm. The minimization algorithm rotates the bonds angles of an atom to obtain the lowest energy state. Again, this refinement is followed by manual manipulation. Another refinement is B-factor refinement, which is refining the thermal B-factor of each atom. The B-factor reports the thermal vibration of an atom and a desired value is between 20.0 and 30.0.\(^\text{49}\)

Once the structure of the native protein is determined and ligands that are present are added to the structure. Ligand positions are often refined using minimization. Also, a B-factor refinement is used to refine atoms of the ligand. The last step of refinement is to
add water molecules to the model. Water molecules compose the solvation shell of the protein and need to be accounted for a complete X-ray structure.\textsuperscript{49}

A model is verified to be a good structure for the protein generally by three criteria and they include R factors, Ramachandran plot, and comparison to previously solved structures. A $R$ and $R_{\text{free}}$ value between 0.15 and 0.25 is indicative of a good model. A Ramachandran plot is a plot of $\phi$ and $\psi$ angles of the protein’s backbone and has accepted regions based on $\phi$ and $\psi$ angles of documented proteins. A good model has a majority of the residues’ backbones in accepted regions. The last verification method is comparison of the structure to homologous structures. A shortcoming of this method is if there are not previously solved structures of the protein of interest.\textsuperscript{49}

\textbf{4.9. Summary}

Macromolecular X-ray crystallography is a very powerful technique for solving protein structures. The foundation of the technique is that protein crystals will diffract X-rays and from that diffraction pattern atomic positions can be determined. One problem of macromolecular X-ray crystallography is the “phase problem”. This can be overcome using molecular replacement. After the amplitude and the phase is known of the diffracted X-rays an electron density map can be generated and the residues of the protein fit to the electron density. The model is refined several times in an effort to have the best fit between the observed and calculated data. The quality of the model is reported in terms of $R$ and $R_{\text{free}}$, and is further verified by a Ramachandran plot, and comparison to homologous structures.
Chapter 5

Methods: Macromolecular X-ray Crystallography of MTA/SAH Nucleosidase from *H. pylori*

5.1. Introduction

Structural analysis of MTA/SAH nucleosidase (MTAN) active site provides insight into the chemical interactions between the enzyme and the substrate, and, with this information, structure-based inhibitor design can proceed. The goal of this thesis was to solve the protein structure of MTAN in its native state as well as with a substrate and a substrate analog using macromolecular X-ray crystallography for structural based inhibitor design. The following is a description of crystallization, X-ray diffraction, and model building of MTAN methods.
5.2. Crystallization of MTAN

Purified protein of MTAN at a concentration of 12.19 mg/mL in a buffer of 20 mM Tris pH 8.5, 0.2 mM TCEP, and 1 mM EDTA was screened for crystallization against the Classics Lite Screen (Qiagen®). The technique used for crystallization was hanging drop vapor diffusion. The volume of crystal solution in the well was 1.0 mL and the drop volume consisted of 2 µL of MTAN (12.19 mg/mL) and 2 µL of crystallization solution. The plates were incubated at room temperature over several days. Condition F3 (0.1 M Tris.HCl pH 8.5, 4% weight/vol polyethylene glycol (PEG) 8000) from the Classics Lite Screen produced crystals after eight days (Figure 5.1). That crystal condition was optimized by a coarse gradient (2%-20% weight/vol) of PEG 8000 keeping the concentration of the Tris.HCl pH 8.5 constant (Table 5.1).

Table 5.1: Optimization of 0.1 M Tris.HCl pH 8.5, 4% weight/vol PEG 8000 using a coarse gradient of the PEG 8000 from 20%-2% weight/vol.

<table>
<thead>
<tr>
<th>Expansion Conditions</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 M Tris.HCl pH 8.5, 4% PEG 8000</td>
<td></td>
</tr>
<tr>
<td>20% PEG 8000, 0.1 M Tris.HCl pH 8.5</td>
<td></td>
</tr>
<tr>
<td>16% PEG 8000, 0.1 M Tris.HCl pH 8.5</td>
<td></td>
</tr>
<tr>
<td>12% PEG 8000, 0.1 M Tris.HCl pH 8.5</td>
<td></td>
</tr>
<tr>
<td>8% PEG 8000, 0.1 M Tris.HCl pH 8.5</td>
<td></td>
</tr>
<tr>
<td>4% PEG 8000, 0.1 M Tris.HCl pH 8.5</td>
<td></td>
</tr>
<tr>
<td>2% PEG 8000, 0.1 M Tris.HCl pH 8.5</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.1: MTAN (12.19 mg/mL) crystal from crystallization condition F3 of Crystal Lite Screen (Qiagen®) 0.1 M Tris pH 8.5, 4% weight/vol PEG 8000

The coarse gradient produced crystals after 24 hours with PEG 8000 concentrations between 8%-20% weight/vol (Figure 5.2). The coarse gradient was further optimized with a fine gradient of 10%-20% weight/vol (2% increments) PEG 8000 keeping the 0.1 M Tris.HCl pH 8.5 constant.
Figure 5.2: MTAN (12.19 mg/mL) crystals from coarse gradient crystallization condition
0.1 M Tris pH 8.5, 16% weight/vol PEG 8000

The fine gradient of 10%-20% weight/vol PEG 8000 and 0.1 M Tris.HCl pH 8.5 was able
to reproduce MTAN crystals after a 24-hour period, and was determined to be the
optimum crystallization condition of MTAN (12.19 mg/mL).

5.3. Crystallization of MTAN with Formycin-A (FMA) and SAH

Substrate and substrate analogs were used in an attempt to elucidate the MTAN
structure as a substrate complex. Formycin-A (FMA) is a non-hydrolyzable substrate
analog due to a lack of a N-glycosidic bond between C1’ and N9. Instead, there is a C1’-
C9 bond (Figure 5.3). Also, FMA is a known inhibitor of MTAN nucleosidase.
However, FMA is not a good drug because it also inhibits host enzymes causing adverse effects.

![Formycin-A (FMA)](image)

The goal of using FMA was to lock the MTAN in the enzyme substrate complex. SAH was used to test catalytic activity of MTAN in the crystal structure. If the SAH was present in the structure then the crystallized MTAN wasn’t catalytically active. However, if the adenine or S-ribosylhomocysteine were present in the structure then the crystallized MTAN was catalytically active.

FMA and SAH were co-crystallized with MTAN using two different methods. One method of co-crystallization had a pre-incubation step that involved adding 1 mM FMA or 2 mM SAH to MTAN and incubating on ice for 30 minutes. The crystal drop was composed of 2.0 µL of MTAN, pre-incubated with 1 mM FMA or 2 mM SAH, and 2.0 µL of crystal condition (i.e. well solution). The other co-crystallization method used to incorporate FMA and SAH to MTAN was adding 0.4 µL of 10 mM FMA or 0.4 µL of 20 mM SAH to the drop along with 1.6 µL of well solution and 2.0 µL of MTAN (12.19 mg/mL). Both methods produced MTAN crystals.

5.4. X-Ray diffraction

X-ray diffraction data were collected at the Argonne National Laboratory Advanced Photon Source (APS) at the Life Science Collaborative Access Team (LS-
CAT) beam line. The cryoprotectant for the crystals was 20% ethylene glycol added directly to the drop. The X-ray wavelength was 1.07813 and the detector was a MAR-CCD 225. The detector distance was 180.00 mm. All crystals were initially screened for diffraction and five test frames 0.0°-5.0° with a 1.0° oscillation were taken and the diffraction data indexed using the program HKL2000 to determine the crystal system. A full data set for MTAN, MTAN with FMA, and MTAN with SAH was collected after indexing. The following is a description of the data collection, indexing, and integrating of each MTAN structure.

5.4 A. X-ray diffraction of native MTAN

The MTAN native crystal from crystal condition 16% weight/vol PEG 8000 and 0.1M Tris.HCl pH 8.5 diffracted to a resolution of 1.58 Å (Figure 5.4). The indexed cell of the MTAN native crystal was primitive hexagonal (P3). The indexed unit cell dimensions are in Table 5.3. A full data set was collected from 0°-180° with a wedge of 1.0°. The data were integrated and scaled to P3_1_21 using HKL2000. The average R_merge for the scaling was 0.06. The number of observed reflections was 1,413,985 and the number of unique reflections was 136,114. The average redundancy of reflections was 19.8 and the mosaicity was 0.507. The resolution range was from 50.00-1.58 Å and an average I/σ of 9.4 (Table 5.3).
Figure: 5.4: X-ray diffraction pattern of MTAN native crystal from crystal condition 16% weight/vol PEG 8000 and 0.1M Tris.HCl pH 8.5. The crystal diffracted to a resolution of 1.58 Å. The pictured diffraction is at a 90° rotation.

5.4 B. X-ray diffraction of MTAN pre-incubated with 10 mM FMA

The MTAN with FMA crystal was pre-incubated on ice with 1 mM FMA for 30 minutes prior to crystallization. The crystal condition for MTAN with 1 mM FMA was 16% weight/vol PEG 8000 and 0.1M Tris.HCl pH 8.5. The crystal diffracted to a resolution of 1.58 Å. The indexed cell was primitive hexagonal (P3) (Figure 5.5). The
indexed unit cell dimensions are in Table 5.3. Similar to the MTAN native data set, data were collected from 0°-180° with a wedge of 1.0°. The data were integrated and scaled to P3₁2₁ using HKL2000. The average R<sub>merge</sub> for the scaling was 0.100. The number of observed reflections was 431,547 and the number of unique reflections was 66,677. The average redundancy of reflections was 12.8 and the mosaicity was 0.380. The resolution of diffraction ranged from 50.00-1.58 Å, but the I/σ for the reflections dropped off significantly at a resolution greater than 2.00 Å so the data were truncated at 2.00 Å. At a resolution of 2.00 Å the average I/σ was 12.3 (Table 5.3).

5.4 C. X-ray diffraction of MTAN co-crystallized with 2 mM SAH

The MTAN with SAH crystal was co-crystallized together with SAH acting as an additive in crystallization. The drop for crystallization was 0.4 µL of 20 mM SAH, 1.6 µL of crystallization condition, and 2.0 µL of MTAN. The crystallization condition was 16% weight/vol PEG 8000 and 0.1M Tris.HCl pH 8.5. The crystal diffracted to a resolution of 1.58 Å. The indexed cell was primitive hexagonal (P3) (Figure 5.6). The indexed unit cell dimensions are in Table 5.3.
Table 5.2: Diffraction and refinement statistics of MTAN native, MTAN with FMA and MTAN with SAH\textsuperscript{49, 50}

<table>
<thead>
<tr>
<th>Diffraction statistics</th>
<th>MTAN Native</th>
<th>MTAN+FMA</th>
<th>MTAN+SAH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P3(_1)21</td>
<td>P3(_1)21</td>
<td>P3(_2)21</td>
</tr>
<tr>
<td>(a)</td>
<td>81.337</td>
<td>81.665</td>
<td>81.259</td>
</tr>
<tr>
<td>(b)</td>
<td>81.337</td>
<td>81.665</td>
<td>81.259</td>
</tr>
<tr>
<td>(c)</td>
<td>135.551</td>
<td>134.411</td>
<td>67.522</td>
</tr>
<tr>
<td>(\alpha=\beta=90^\circ, \gamma=120^\circ)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of measured reflections</td>
<td>1,413,985</td>
<td>431,547</td>
<td>607,295</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>136,114</td>
<td>66,677</td>
<td>35,186</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>50.00-1.58</td>
<td>50.00-2.00</td>
<td>50.00-1.58</td>
</tr>
<tr>
<td>(R_{merge}) (%)</td>
<td>6.0</td>
<td>10.0</td>
<td>5.4</td>
</tr>
<tr>
<td>Redundancy</td>
<td>19.58</td>
<td>12.3</td>
<td>10.8</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.5 (95.3)</td>
<td>98.5 (97.4)</td>
<td>98.6 (94.7)</td>
</tr>
<tr>
<td>Average I/(I)</td>
<td>9.4</td>
<td>3.9</td>
<td>5.32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refinement statistics</th>
<th>MTAN Native</th>
<th>MTAN+FMA</th>
<th>MTAN+SAH</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of atoms</td>
<td>3462</td>
<td>3530</td>
<td>1750</td>
</tr>
<tr>
<td>No. of water molecules</td>
<td>260</td>
<td>184</td>
<td>75</td>
</tr>
<tr>
<td>No. of ligand atoms</td>
<td>8</td>
<td>17</td>
<td>38</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>50.0-1.6</td>
<td>50.0-2.0</td>
<td>50.0-1.6</td>
</tr>
<tr>
<td>(R_{cryst}) (%)</td>
<td>19.42</td>
<td>19.68</td>
<td>21.40</td>
</tr>
<tr>
<td>(R_{free}) (%)</td>
<td>21.67</td>
<td>21.67</td>
<td>22.88</td>
</tr>
<tr>
<td>Average B-factor (Å(^2))</td>
<td>58.80</td>
<td>53.07</td>
<td>39.23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R.M.S. deviation</th>
<th>MTAN Native</th>
<th>MTAN+FMA</th>
<th>MTAN+SAH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonds (Å)</td>
<td>0.0043</td>
<td>0.0051</td>
<td>0.0045</td>
</tr>
<tr>
<td>Angles (°)</td>
<td>1.1781</td>
<td>1.1658</td>
<td>1.2178</td>
</tr>
<tr>
<td>% Residues in allowed regions of Ramachandran plot</td>
<td>99.33</td>
<td>99.34</td>
<td>99.11</td>
</tr>
</tbody>
</table>
Figure 5.5: The diffraction pattern of MTAN with pre-incubated 1 mM FMA. The crystal condition was 16% weight/vol PEG 8000 and 0.1M Tris.HCl pH 8.5. The crystal diffracted to a resolution of 1.58 Å. The pictured diffraction is at 90° rotation.

The data set for MTAN with 2 mM SAH was from 0°-180° with a wedge of 1°. The data was integrated and scaled to P321 using HKL2000. The average R_{merge} for the scaling was 0.054. The number of observed reflections was 607,195 and the number of unique reflections was 35,186. The average redundancy of reflections was 10.8 and the
mosaicty was 0.255. The resolution range was from 50.00-1.58 Å with an average I/σ of 5.323 (Table 5.3).

Figure 5.6: The diffraction pattern of MTAN co-crystallized with 2 mM SAH. The crystal condition was 16% weight/vol PEG 8000 and 0.1M Tris.HCl pH 8.5. The crystal diffracted to a resolution of 1.58 Å. The pictured diffraction is at 120° rotation.
5.5. Molecular replacement and refinement

The molecular replacement of MTAN native was done with Phaser from the ccp4i suite. The homologous structure used for molecular replacement was MTA/SAH nucleosidase with FMA from *Staphylococcus aureus* (3bl6.pdb). The sequence identity of MTAN from *S. aureus* and *H. pylori* was 30%. A dimer was searched for in the asymmetric unit during molecular replacement. The *S. aureus* model was rigid body refined using Refmac from the ccp4i suite. The *S. aureus* primary sequence was mutated to have the same sequence of MTAN from *H. pylori*.

The mutated structure was further refined using simulated annealing from Crystallography and NMR System (CNSsolve 1.2). The molecular dynamics was torsional, and the annealing schedule was slow cool. For the first simulated annealing refinement the starting temperature was 5500 K, and starting temperatures thereafter were 2500 K. The temperature drop was done in 50 K increments and the final steps were at a 300 K. Following the simulated annealing a minimization refinement was conducted and it was carried out for the default 100 cycles. After each simulated annealing refinement a 2Fo-Fc and Fo-Fc electron density map were generated and the refined coordinates for the structure were manually manipulated into the electron density using COOT software. The cycle of simulated annealing refinement and manual manipulation proceeded until the model was deemed to be well refined for the addition of ligands.

In the active site of the MTAN structure there was an electron density that could not be accounted for by any amino acid residue. Reviewing the crystallization condition it was thought that the electron density belonged to a molecule of Tris. One molecule of Tris was modeled into the structure as a ligand and its position, as the well as the
protein’s, was refined with minimization refinement from CNSsolve 1.2.\textsuperscript{49} 2Fo-Fc and Fo-Fc electron density maps were generated after each refinement and the ligand’s and protein’s coordinates were manually manipulated into the electron density. The cycle of minimization refinement and manual manipulation was continued until the structure was again determined to be well enough refined for B-factor refinement and the addition of water.\textsuperscript{52}

B-factor refinement is a refinement of the each individual atom’s thermal parameter. Since it is not positional, manual manipulation does not need to follow B-factor refinement. The B-factor refinement was also done using CNSsolve 1.2. Water molecules were added using the water pick program from CNSsolve 1.2 based on the criteria for having an I/\sigma equal to or greater than 3.0.\textsuperscript{49} A minimization refinement was done after water molecules were added to the structure as well. A 2Fo-Fc and Fo-Fc map was generated and the water molecules were manually manipulated into the electron density, or deleted if there was no surrounding electron density.\textsuperscript{52} The water pick program was run twice. Some water molecules that had been placed in the structure were exchanged for ethylene glycol. Indication that the picked water molecules could be ethylene glycol was the close proximity of two water molecules relative to each other. The ethylene glycol coordinates were refined in the same manner as the Tris ligand.

MTAN with FMA structures was solved isomorphously after the initial molecular replacement and refinement of MTAN native. The isomorphous replacement of the ligand bound structure entailed the building of the electron density map from the ligand structure data and using the refined MTAN structure as the model to manipulate within the ligand’s electron density. The refinement of the ligand structure was similar to the
native structure in terms of the general methods. The only modification to the methods for FMA structure was the addition of the FMA in place of a Tris molecule. Refinement of the MTAN with FMA structure then proceeded similarly as the native.

MTAN with SAH structure was solved using molecular replacement. The homologous structure used for molecular replacement was the refined MTAN native structure. The molecular replacement proceeded in a similar manner as for the native structure (Section 5.5). Modifications for the MTAN with SAH structure was searching for a monomer in the asymmetric unit for molecular replacement, and not mutating the primary amino acid sequence. The refinement of the ligand structures was similar to the native structure. One modification to the native structure model building and refinement of the added ligand was the addition of a molecule of adenine as well as the Tris to the active site. Refinement of the MTAN with SAH then proceeded similarly as the native.
Chapter 6

Results

6.1. Introduction

Macromolecular X-ray crystallography was used to solve for the protein structure of native MTA/SAH nucleosidase and MTA/SAH nucleosidase with FMA or SAH. The following are the modeled results of the MTA/SAH nucleosidase structures.

6.2 A. Secondary structure of MTA/SAH nucleosidase native

The native structure of the MTA/SAH nucleosidase has a dimer in the asymmetric unit. The structure has both α and β structure. There is a central 7 strand β sheet that is a mixture of parallel and anti-parallel β strands (β1-7). Flanking the central β sheet are α helices. Two of the α helices (α4 and α5) are orientated on the inner face of the β sheet and form the dimer interface. The outer face of the β sheet has three α helices. The
central outer α helix (α2) contributes to formation of the active site. The other two α helices are on either side of the α2 and interact with the central β sheet and α2. Above the active site there are three short parallel β strands (β8-10) that form the roof of the active site. Also, above β8-10 is an α helix. A structural feature that is important for catalysis is a crossover loop of random coil from the opposing monomer that contributes to the formation of the active site upon the binding of the substrate (Figure 6.1)

Figure 6.1: MTA/SAH nucleosidase native secondary structure model.53
At the dimer interface the largest interaction is between $\alpha_4^a$ and $\alpha_4^b$ (a and b refers to monomer a or b). The major interaction between $\alpha_4^a$ and $\alpha_4^b$ is hydrophobic with the charged residues pointing away from the interface (Figure 6.2). However, there are polar residues at the dimer interface that contribute hydrogen bonds through ordered water molecules. Another contributing interaction to the dimer interface is between $\alpha_4^a$ and $\alpha_5^b$. The interaction is again hydrophobic. Also, contributing to the $\alpha_4^a$ and $\alpha_4^b$ interaction is $\pi-\pi$ stacking between H 53 of chain a and F 178 of chain b.

6.2 B. MTA/SAH nucleosidase with FMA

The MTA/SAH nucleosidase with FMA also crystallized as a dimer in the asymmetric unit. The secondary structure of MTA/SAH nucleosidase with FMA is very similar to the native structure. However, in the FMA structure $\alpha_2$ is extended by two helical turns and shifts between 4.00-12.00 Å (Figure 6.3).

The active site of MTA/SAH nucleosidase with FMA has a closed active site with FMA present in both monomers of the dimer (Figure 6.4 and 6.5). The active site of MTA/SAH nucleosidase is divided into three sub-sites: adenine, ribose, and alkylthio. The hydrogen bonding of each of these sub-sites is described in Table 6.1.
Figure 6.2: Dimer interface of MTA/SAH nucleosidase native structure. Chain a is pink and chain b is green.
Figure 6.3: Overlay of MTA/SAH nucleosidase native (green) and FMA (blue). In the FMA structure $a_2$ is extended by two helical turns and displaced between 4.00-12.00°.

Table 6.1: Hydrogen binding of FMA within the active site of MTA/SAH nucleosidase.

<table>
<thead>
<tr>
<th>Residue</th>
<th>Binding Site</th>
<th>Distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E13 O$^\gamma$</td>
<td>WAT 86</td>
<td>2.40</td>
</tr>
<tr>
<td>V 154 O$^\beta$</td>
<td>Adenine NH$_2$ (C6)</td>
<td>3.08</td>
</tr>
<tr>
<td>V 154 N$^a$</td>
<td>Adenine N1</td>
<td>2.98</td>
</tr>
<tr>
<td>M 174 N$^a$</td>
<td>Ribose OH (C2’)</td>
<td>2.81</td>
</tr>
<tr>
<td>E 175 O$^\gamma$</td>
<td>Ribose OH (C2’)</td>
<td>2.54</td>
</tr>
<tr>
<td>E 175 O$^\alpha$</td>
<td>Ribose OH (C3’)</td>
<td>2.68</td>
</tr>
<tr>
<td>R 194 N$^\delta$</td>
<td>WAT 86</td>
<td>3.10</td>
</tr>
<tr>
<td>S 197 OH$^b$</td>
<td>D 198 O$^\gamma$</td>
<td>2.85</td>
</tr>
<tr>
<td>D 198 O$^\gamma$</td>
<td>Adenine N7</td>
<td>2.70</td>
</tr>
<tr>
<td>D 198 O$^\delta$</td>
<td>Adenine N (C6)</td>
<td>2.82</td>
</tr>
<tr>
<td>WAT 86</td>
<td>Ribose OH (C3’)</td>
<td>2.73</td>
</tr>
<tr>
<td>WAT 86</td>
<td>Ribose O5’</td>
<td>3.00</td>
</tr>
</tbody>
</table>
Figure 6.4: Hydrogen bonding schematic of MTA/SAH nucleosidase with FMA. Dashed lines indicate hydrogen bond distance in Å between protein residues and ligand.\(^{53}\)

Other interactions that contribute to the adenine-binding site are \(\pi-\pi\) stacking of the adenine base with F 153. F 153 also participates in \(\pi-\pi\) stacking with F 107 (B).

There is also an opportunity for \(\pi-\pi\) stacking between F 107 (B) and F 208. The alkylthio-binding site is a hydrophobic pocket for the homocysteine and methyl chain of SAH and MTA respectively. This alkylthio binding site in the FMA structure is a hydrophobic pocket composed of \(\alpha_2\), the crossover loop, and a \(\beta\) sheet.
Figure 6.5 (A): MTA/SAH nucleosidase with FMA bound to the active site. (A) is all residues contributing to the binding of FMA. (B) is all hydrogen bonding residues and in a different orientation than (A).

Figure 6.5 (B): MTA/SAH nucleosidase with FMA bound to the active site. (A) is all residues contributing to the binding of FMA. (B) is all hydrogen bonding residues and in a different orientation than (A).
6.2 C. MTA/SAH nucleosidase with SAH

MTA/SAH nucleosidase with SAH had a monomer in the asymmetric unit, and α2 is extended and shifted as in the FMA structure. The other secondary structure features of the MTA/SAH nucleosidase with SAH are similar to the native structure. The active site is assumed to have catalyzed the hydrolysis of the SAH because only an adenine is found in the active site. Also, present in the active site is a Tris molecule from the crystallization buffer (Figure 6.6 and 6.7). The hydrogen bonding of the active site residues with adenine and Tris are described in Table 6.2. Another interaction of the active site with the adenine is π-π stacking with F 153 similar to the FMA structure. Amino acid F 153 also interacts with F 107 (B) via π-π stacking.

Table 6.2: Hydrogen bonding of adenine and Tris within the active site of MTA/SAH nucleosidase.53

<table>
<thead>
<tr>
<th>Residue</th>
<th>Substituent</th>
<th>Distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M 10 S</td>
<td>Tris OH (1)</td>
<td>2.90</td>
</tr>
<tr>
<td>E13 Oγ</td>
<td>Tris OH (1)</td>
<td>2.90</td>
</tr>
<tr>
<td>V 78 Oβ</td>
<td>Tris (NH₂)</td>
<td>2.94</td>
</tr>
<tr>
<td>V 154 Oβ</td>
<td>Adenine NH₂ (C6)</td>
<td>2.88</td>
</tr>
<tr>
<td>V 154 Nα</td>
<td>Adenine N1</td>
<td>3.00</td>
</tr>
<tr>
<td>M 174 Oβ</td>
<td>Tris OH (3)</td>
<td>2.90</td>
</tr>
<tr>
<td>E 175 Oβ</td>
<td>Tris OH (3)</td>
<td>2.50</td>
</tr>
<tr>
<td>E 175 Oα</td>
<td>Tris OH (3)</td>
<td>3.10</td>
</tr>
<tr>
<td>R 194 Nα</td>
<td>Tris OH (2)</td>
<td>2.90</td>
</tr>
<tr>
<td>S 197 OH</td>
<td>D 198 Oγ</td>
<td>2.90</td>
</tr>
<tr>
<td>D 198 Oγ</td>
<td>Adenine N7</td>
<td>2.60</td>
</tr>
<tr>
<td>D 198 Oα</td>
<td>Adenine N (C6)</td>
<td>3.03</td>
</tr>
<tr>
<td>Adenine N9</td>
<td>Tris NH₂</td>
<td>2.70</td>
</tr>
</tbody>
</table>
Figure 6.6: MTA/SAH nucleosidase with adenine and Tris in the active site. Not pictured is the alkylthio binding site which is composed of F 153, F 209, and F 107 (B).\textsuperscript{53}

Figure 6.7: Hydrogen bonding schematic of MTA/SAH nucleosidase with adenine and Tris in the active site. Dashed lines indicate hydrogen bond distances in Å between protein residues and ligand.\textsuperscript{53}
Chapter 7

Discussion

7.1. Introduction

MTA/SAH nucleosidase from \textit{H. pylori} was crystallized in its native form, with FMA bound, and with adenine bound. The following is a detailed interpretation of the MTA/SAH nucleosidase modeled results.

7.2. MTA/SAH nucleosidase active site and catalysis

The active site of MTA/SAH nucleosidase with FMA has a well positioned aspartic acid (D 196) for the protonation N7 of the adenine ring at an atomic distance of 2.7 Å. This supports the theorized mechanism for acid protonated catalysis (Scheme 7.1), and previously solved MTA/SAH nucleosidase active site models. The active site has an induced conformational change upon substrate binding. This is evident by the
displacement and extension of α2 of the FMA and adenine structure. Since the reaction is similar to an $S_{N}1$ reaction an α and β ribose should be produced in equal molar ratios. However, the nucleophilic water (WAT 86) is positioned behind the ribose sugar in the modeled structure of FMA. This provides evidence that the ribose product from MTA/SAH nucleosidase is an α ribose.

The MTA/SAH nucleosidase that was co-crystallized with SAH was found to have only adenine and a Tris molecule in the active site. This observation indicates that the MTA/SAH nucleosidase is catalytically active in the crystal. Also, the presence of adenine in the structure of the MTA/SAH nucleosidase with SAH supports the hypothesis that adenine is not released from the active site of MTA/SAH nucleosidase until the binding of another substrate.

A Tris molecule is also found in the native structure. This indicates that binding of Tris to the active site is not dependent on the presence of adenine. However, adenine
or substrate analog appears to be needed for the modeling of the residues 196-203 (α2). In the native structure the electron density for these residues is not well defined and cannot be modeled into the structure. An explanation for this observation may be that residues 196-203 of α2 are not ordered until the binding of the substrate/product to the enzyme.

7.3. Comparison of MTA/SAH Nucleosidase with FMA bound and with Adenine bound

Similar active site interactions are found in both the nucleosidase with FMA bound and with adenine bound include, residues V 154, S 197, D 198, and R 194 hydrogen bonding interactions as well as van der Waals interaction with residues F 153, F 208 and F 107 (B) (Figure 7.1). Amino acids R 194 and E 175 in the adenine structure are hydrogen bonding to hydroxyl groups of Tris in the adenine structure instead of the ribose ring as in the FMA structure. Another difference between the FMA and adenine active sites is the opportunity for hydrogen bonding between V 78 backbone and the amine group of Tris, as well as hydrogen bonding between E13 and M 10 residues with hydroxyl groups of Tris in the adenine structure. Hydrogen bonding distances between the adenine base and V154 and D198 also varies from the adenine and FMA structure. V 154 Oβ hydrogen bond distance to adenine NH2 (C6) is shortened by 0.20 Å and D 198 Oγ hydrogen bond to adenine N7 is shortened by 0.10 Å. The overall increase and shortening of hydrogen bonding opportunities between the residues in the active site and the molecule of Tris in the adenine structure supports the building of an inhibitor based of a Tris molecule.
Figure 7.1: A side-by-side comparison of MTAN with adenine and Tris bound (A) and MTAN with FMA bound (B).

7.4. Comparison of MTA/SAH Nucleosidase from *H. pylori* to *S. aureus* and *E. coli*

As previously mentioned, MTA/SAH nucleosidase is conserved among many bacteria including *S. aureus* and *E. coli* (Figure 7.2). Comparison of the *S. aureus* and *E. coli* with FMA bound active sites to the *H. pylori* with FMA bound to the active site reveals a conservation of residues that are essential for positioning of the substrate and catalysis (Figure 7.3). All three models have three phenylalanines composing the hydrophobic alkylthio-binding site. These phenylalanines in the *H. pylori* structure are F153, 208, and 107 (B). Also, present in all three structures are the hydrogen bonding interactions between M 174 backbone and the hydroxyl of C2’ as well as hydrogen bonding between the residue carboxyl group of E 175 and hydroxyl of C2’ and C3’. The nucleophilic water in each structure is positioned in the active site by hydrogen bonding.
interactions between R 194 and with the hydroxyl groups of C2’ and C3’ of the substrate (residue numbers are in reference to the *H. pylori* sequence) (Figure 7.4 and 7.5).

Figure 7.2: Superposition of *E. coli* (red), *S. aureus* (green), and *H. pylori* (purple) with FMA bound structures.

The *H. pylori* and *E. coli* structures with FMA bound both have a more closed and less solvent exposed active site in comparison to the *S. aureus* structure. Also, unlike the *H. pylori* and *E. coli* models there are two water molecules in the active site of the *S. aureus* in comparison to one water in the *H. pylori* and *E. coli* models. The *E. coli* and *S. aureus* models have the side chain of S 76/75, respectively, participating in hydrogen
bonding interactions within the active site, but the in the *H. pylori* at the sequence aligned position for S 76/75 the residue is mutated to a V 78. The S 76/75 to V 78 mutation leads to a decrease in hydrogen bonding opportunities for the substrate within the active site. Another difference between the active sites of *S. aureus* and *E. coli* to *H. pylori* is V 154 in the *H. pylori* structure is mutated to an I 152/151 in the *E. coli*/*S. aureus* respectively. However, this change from V 154 to I 152/151 should not have a great effect on the binding of the substrate since the substrate is hydrogen bonding to the backbone of V/I.

Both the *S. aureus* and the *E. coli* structure have been concluded to have more hydrogen bonding interactions than the presented *H. pylori* active site. However, the *S. aureus* and *E. coli* structure have an extra polar uncharged residue S 75/76 respectively, in the active site that is positioned well for hydrogen bonding interactions with the substrate. Furthermore, the *S. aureus* model has more hydrogen bonding interactions because it has an additional water molecule in the active site. Since the active sites of *S. aureus* and *E. coli* have more hydrogen bonding opportunities a possible way to select for *H. pylori* would be to have a more hydrophobic inhibitor. For example, the S 76/75 to V 78 mutation could be exploited by introducing a hydrophobic group at position 8 of the adenine base.
Figure 7.3: Sequence alignment of MTA/SAH nucleosidase from *S. aureus*, *E. coli*, and *H. pylori*
Figure 7.4: MTA/SAH nucleosidase with FMA bound in the active site from E. coli (red), S. aureus (green), and H. pylori (purple). A and B are two different orientations.
Figure 7.5: (A) MTA/SAH nucleosidase from *H. pylori*
Figure 7.5: (B) MTA/SAH nucleosidase from *E. coli*
7.5. Drug Development

MTA/SAH nucleosidase with adenine and Tris bound in the active site provides an insight for inhibitor design. A possible inhibitor could be built from the Tris molecule using it as a scaffolding. The MTA/SAH nucleosidase with FMA structure could be used to design other non-hydrolysable inhibitors as well. Also, given the position of F 153, 209, and 107 (B) and their location to the alkylthio-binding site in both the adenine and FMA structure it could be advantageous to include aromatic residues in place of the alkylthio chain when designing an inhibitor.
A problem with developing an inhibitor specific for MTA/SAH nucleosidase is indentifying a compound that does not interfere with any mammalian cellular processes. For example, FMA is a well known inhibitor for MTA/SAH nucleosidase, but it has negative effects on mammalian cellular processes. Another difficulty is MTA/SAH nucleosidase mechanism proceeds in a similar manner as MTA phosphorylase (Scheme 7.2). MTA phosphorylase produces adenine and 5’-Methylthioribose-1-phosphate. Comparison of the two mechanisms demonstrates that the major differences are the N7 of the adenine ring is protonated with N 243 and a coordinated phosphate acts as the nucleophile instead of a water molecule in the case of MTA phosphorylase. Lee et al. has compared MTA/SAH nucleosidase and MTA phosphorylase structures and has suggested drug development specific to MTA/SAH nucleosidase. These designs include variation at the C8 position of the adenine ring with electronegative groups, introduction of an electropositive group on OH 2’ or OH 3’ of the ribose ring, and replacing the alkylthio chain of SAH with bulky groups.

An enzymatic oxidation assay of adenine has been well documented throughout the literature and can be used to test the inhibition of MTA/SAH nucleosidase from designed drug inhibitors.\textsuperscript{54, 55} The assay utilizes xanthine oxidase, which will react with the free adenine base from the MTA/SAH nucleosidase reaction and produce 2,8-dihydroxyadenine (Scheme 7.3). The production of 2,8-dihydroxyadenine is monitored by UV-VIS absorption at 293 nm.
Scheme 7.2: Mechanism of MTA phosphorylase

Scheme 7.3. Conversion of adenine to 2,8-dihydroxyadenine by xanthine oxidase

7.6. MTA/SAH nucleosidase as a target for a broad-spectrum antibiotic

MTA/SAH nucleosidase is found in many other bacteria such as *Mycobacterium tuberculosis* and *Staphylococcus aureus*, as well as *Helicobacter pylori*. All three of these enzymes have a similar primary amino acid sequence and have been proposed to have similar catalytic mechanisms (Appendix III). Thus, the development of an
antibiotic specific to MTA/SAH nucleosidase of *H. pylori* could be effective on other bacterial MTA/SAH nucleosidase.

Additional work that has been done into the investigation of the development of broad-spectrum antibiotics is the expression, purification and crystallization of MTA/SAH nucleosidase from *Mycobacterium tuberculosis*. The *pfs* gene coding for MTA/SAH nucleosidase was codon optimized and synthesized from GeneArt® (Appendix IV). The *pfs* gene also had engineered cleavage sites for NcoI and BamHI restriction enzymes. It was ligated into pET32 (Novagen) and expressed in T7®+pROS cells. Purification protocol proceeded in a similar fashion as for MTA/SAH nucleosidase from *H. pylori*. Modifications of the protocol included alteration of the buffers used for Ni²⁺ affinity and size exclusion chromatography and the addition of an anion exchange column before size exclusion chromatography (Appendix V-VII). MTA/SAH nucleosidase from *M. tuberculosis* was screened for crystallization against an index screen from Hampton Research (Appendix VIII). Crystals of MTA/SAH nucleosidase from *M. tuberculosis* are currently being optimized for analysis with macromolecular X-ray crystallography. Also, the purified MTA/SAH nucleosidase from *M. tuberculosis* does have catalytic activity, which was verified by the xanthine oxidase assay.

### 7.7. Future work

Further analysis of MTA/SAH nucleosidase includes testing the affinity of the Tris molecule in the SAH co-crystallized structure, structure based design of inhibitors, screening inhibitors for binding to MTA/SAH nucleosidase, and structural analysis of MTA/SAH nucleosidase from *M. tuberculosis*. An experiment that has already been performed to test for the affinity of Tris to the active site is a competition experiment.
with D-ribose. MTA/SAH nucleosidase was pre-incubated with 2 mM SAH for 30 min. followed by incubation with 100 mM D-ribose for 30 min. with both incubations on ice. MTA/SAH nucleosidase with 2 mM SAH and 100 mM D-ribose was crystallized in a screen of 0.1 Tris pH 8.5 and 12%-18% PEG 8000, and analyzed at Argonne National Lab Advanced Photon Source (APS), Life Science Collaborative Access Team (LS-CAT) beam line. Diffraction data is currently being analyzed.

Other aspects to investigate would be inhibitor structure based design based off the adenine with Tris structure or FMA structure. Also, the screening of inhibitors for binding interactions with MTA/SAH nucleosidase can be performed using macromolecular X-ray crystallography.
References


Appendices

**Appendix I:** *pfs* sequencing from *Helicobacter pylori*

**Appendix II:** ProtParam: of MTA/SAH Nucleosidase from *Helicobacter pylori*

**Appendix III:** MTA/SAH Nucleosidase Sequence Alignment of *Helicobacter pylori*, *Mycobacterium tuberculosis*, and *Staphylococcus aureus*

**Appendix IV:** *pfs* sequence from *Mycobacterium tuberculosis*

**Appendix V:** Ni\(^{2+}\) affinity chromatography (chelating) of MTA/SAH Nucleosidase from *Mycobacterium tuberculosis*

**Appendix VI:** Ni\(^{2+}\) affinity chromatography (tag removal) of MTA/SAH Nucleosidase from *Mycobacterium tuberculosis*

**Appendix VII:** Anion exchange chromatography of MTA/SAH Nucleosidase from *Mycobacterium tuberculosis*

**Appendix VIII:** MTA/SAH nucleosidase crystal hits from Hampton Research Index Screen
Appendix I: *pfs* sequencing from *Helicobacter pylori*. Chromatogram of *pfs* sequencing. The highlighted blue region is the engineered NcoI cut site.
Appendix II: ProtParam: of MTA/SAH Nucleosidase from Helicobacter pylori\textsuperscript{a5}

Amino Acid Sequence

MQKIGILGAM REEITPILEL FGVDFFEIPL GGNVFHKGVY HNKIEIVAYS KIGKXHSTLT TTSMILAFGV
QKVLFSGVGAG SLVKDLKIND LLVATQLVQH DVLPSAFDHP LGFIPESAIF IETSGLNAL AKKANEQH
ALKEGVASG DQFVHSKERK EFLVSEFKAS AVEMEGASVA FVCQKFGVPC CVLRSISDNA DEKAGMSFDE
FLEKSAHTSA KFLKSMVDE

Number of amino acids: 230
Molecular weight: 25019.9
Theoretical pI: 5.49

Amino acid composition:

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala (A)</td>
<td>20</td>
<td>8.7%</td>
</tr>
<tr>
<td>Arg (R)</td>
<td>3</td>
<td>1.3%</td>
</tr>
<tr>
<td>Asn (N)</td>
<td>6</td>
<td>2.6%</td>
</tr>
<tr>
<td>Asp (D)</td>
<td>11</td>
<td>4.8%</td>
</tr>
<tr>
<td>Cys (C)</td>
<td>3</td>
<td>1.3%</td>
</tr>
<tr>
<td>Gln (Q)</td>
<td>7</td>
<td>3.0%</td>
</tr>
<tr>
<td>Glu (E)</td>
<td>19</td>
<td>8.3%</td>
</tr>
<tr>
<td>Gly (G)</td>
<td>17</td>
<td>7.4%</td>
</tr>
<tr>
<td>His (H)</td>
<td>8</td>
<td>3.5%</td>
</tr>
<tr>
<td>Ile (I)</td>
<td>17</td>
<td>7.4%</td>
</tr>
<tr>
<td>Leu (L)</td>
<td>22</td>
<td>9.6%</td>
</tr>
<tr>
<td>Lys (K)</td>
<td>19</td>
<td>8.3%</td>
</tr>
<tr>
<td>Met (M)</td>
<td>6</td>
<td>2.6%</td>
</tr>
<tr>
<td>Phe (F)</td>
<td>16</td>
<td>7.0%</td>
</tr>
<tr>
<td>Pro (P)</td>
<td>5</td>
<td>2.2%</td>
</tr>
<tr>
<td>Ser (S)</td>
<td>20</td>
<td>8.7%</td>
</tr>
<tr>
<td>Thr (T)</td>
<td>8</td>
<td>3.5%</td>
</tr>
<tr>
<td>Trp (W)</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>Tyr (Y)</td>
<td>2</td>
<td>0.9%</td>
</tr>
<tr>
<td>Val (V)</td>
<td>21</td>
<td>9.1%</td>
</tr>
</tbody>
</table>

Total number of negatively charged residues (Asp + Glu): 30
Total number of positively charged residues (Arg + Lys): 22

Atomic composition:

<table>
<thead>
<tr>
<th>Element</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbon</td>
<td>C 1129</td>
</tr>
<tr>
<td>Hydrogen</td>
<td>H 1793</td>
</tr>
<tr>
<td>Nitrogen</td>
<td>N 287</td>
</tr>
<tr>
<td>Oxygen</td>
<td>O 334</td>
</tr>
<tr>
<td>Sulfur</td>
<td>S 9</td>
</tr>
</tbody>
</table>

Formula: \( C_{1129}H_{1793}N_{287}O_{334}S_9 \)

Total number of atoms: 3552

Extinction coefficients:

This protein does not contain any Trp residues. Experience shows that this could result in more than 10% error in the computed extinction coefficient.

Extinction coefficients are in units of \( M^{-1} \text{ cm}^{-1} \), at 280 nm measured in water.

Ext. coefficient 3105

Abs 0.1% (=1 g/l) 0.124, assuming ALL Cys residues appear as half cystines
**Appendix III: MTA/SAH Nucleosidase Sequence Alignment of Helicobacter pylori, Mycobacterium tuberculosis, and Staphylococcus aureus**

<table>
<thead>
<tr>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H. pylori</strong></td>
<td>-MQKIGILGANREEITPILELFGVDFEEIPLGGNVHKVHNYEIVAYSKIGKVSHTL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S. aureus</strong></td>
<td>---MIIGIAMEEET-ILKNKLTQLSEISVAVHVFYETGILKDRETYTQSGIGKVNAAI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M. tuberc.</strong></td>
<td>MAVTVGICAIPOELA-YLRGVLVDARKQQVAQLFDFSGQLDAHRVLAAAGMGKVNTGL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
</tr>
<tr>
<td>70</td>
<td>80</td>
<td>90</td>
<td>100</td>
<td>110</td>
<td>120</td>
</tr>
<tr>
<td><strong>H. pylori</strong></td>
<td>TTTSMILAFGVQKVLFSGVGSLVYKDLKNDDLLVATQLVQHVDLSAEFDHPLGFIFESA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S. aureus</strong></td>
<td>STTLLINKFKEIINTGSGAGLESLNVGDLISDDVYHDADATAFGYEYQIPQMPV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M. tuberc.</strong></td>
<td>TATLLADRFGCRTIVFTGVAGGLDPELCIGIDIVADRVQHPGLFLPPERLBPVQFCHIP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
</tr>
<tr>
<td>130</td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
</tr>
<tr>
<td><strong>H. pylori</strong></td>
<td>FIETSGSLNALAKKIANE----------QHIALKEGVIASGQFVHSK</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S. aureus</strong></td>
<td>AFQSSKPIIEKSVQVQQ----------QQLTAKVLVSQGDSFEGSV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M. tuberc.</strong></td>
<td>FIEPTERGLYPQVDPAVDRVKHDRLDGFTALPLSTAAAGGRQPRIYYGTILTDQXLCHCE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
</tr>
<tr>
<td>190</td>
<td>200</td>
<td>210</td>
<td>220</td>
<td>230</td>
<td>240</td>
</tr>
<tr>
<td><strong>H. pylori</strong></td>
<td>ERKFLVSEFE-ASAVEMSESAVAFVCQF RGPCVLRISSDVADEKACISMDFEPLX</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S. aureus</strong></td>
<td>EQRQKIKAFFNAMAVE-MAKTAIAQTCYQONPVVRAYVAVSLANGEAEISFEPFLEAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M. tuberc.</strong></td>
<td>RTRNRLHHELG-GMAVEMSAVAEIASC.FDIPWLVLIRALSDLAGADSGVFNFVRGEVA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
</tr>
<tr>
<td>250</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>H. pylori</strong></td>
<td>HTSAKFLKSMVDEL---</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S. aureus</strong></td>
<td>VSSSQTVEALVSQ---</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M. tuberc.</strong></td>
<td>ASSARVLRLLPVLTC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>:<em>:</em></td>
<td>:<em>:</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix IV: \textit{pfs} sequence from \textit{Mycobacterium tuberculosis}

\begin{verbatim}
60TACCCCATGGCTACCCCTTTGCTCTGATCTGCAACTT

TGACCCCTTTGGTATGATGGTGGGTTGCTGACGCTGCTT

1  ---------------+-----------------------------

CAGCAGGAACTGGCAATGCTGCTGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG

CAGCAGGAACTGGCAATGCTGCTGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG

61  ---------------+-----------------------------

CAGCAGGAACTGGCAATGCTGCTGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG

CAGCAGGAACTGGCAATGCTGCTGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG

121  ---------------+-----------------------------

AAGGAGAATGCTACCCCTGCTGCAGCCTTGCCGGCCCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTG

181  ---------------+-----------------------------

ACCATTTTTTTTACCCCTGCTGCAGCCTTGCCGGCCCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTTGCTT

241  ---------------+-----------------------------

1
\end{verbatim}
GTTATTGCCGACGTGTGTTGAGCAGCATGATTCTGCTGACCAGATGACGACCTGCG

CGTATGAGGCTGGCTATATTCCGGTTAATGAAACGGAGCAGCAGCTGTGGGTTATCCGTT

GATCCGAGCATTTATGTCTCTGTATATTACATCGTCTGGATGTTTTACCTGGCAGCGCTG

421

TCTACAGCAGCCCGTGTTGGCTCGACCGCTCACGTAATTTATATGCACCATTGTGAC

481

GGTGATCATGATCTCAGTGATCTGAACTACCCGAATCCTGCTGTCACTGACATTCGCG

541

ATGCGAGTTGAAAATCGAAAGGTGCTGCAAGTTGCTGCCAGCCTTGTGATATTCG

601
TGGCTGGTTATTCGACTGACGCGCATCTGCCAGCGAGATAGCGGTGTTTGAAT
TGGCTGGTTATTCGACTGACGCGCATCTGCCAGCGAGATAGCGGTGTTTGAAT

661  -----------------+-----------------+-----------------+-----------------+

CGTTTGTGGGTTGAACTGCAAGAACAGGGCAGCTGTTCTGCTGCTGCTGCTGTT
CGTTTGTGGGTTGAACTGCAAGAACAGGGCAGCTGTTCTGCTGCTGCTGCTGTT

721  -----------------+-----------------+-----------------+-----------------+

CTGACCGCATGTTGAGGATCCGAGCTC
CTGACCGCATGTTGAGGATCCGAGCTC

781  -----------------+-----------------+-----------------+-----------------+
Appendix V: HisTrapChe (Ni$^{2+}$) chromatogram of MTAN from *M. tuberculosis*, and SDS-PAGE of fractions F3-F5.

Loading Buffer: 20 mM Na$_2$HPO$_4$, 0.5 M NaCl, 10% Glycerol, 5 mM BME, 25 mM Imidazole
Elution Buffer: 20 mM Na$_2$HPO$_4$, 0.5 M NaCl, 10% Glycerol, 5 mM BME, 0.25 M Imidazole
Appendix VI: HisTrapChe (Ni\(^{2+}\)) chromatogram Tag Removal of fractions F3-F6 from chelating column (Appendix V)

Loading Buffer: 20 mM Na\(_2\)HPO\(_4\), 0.5 M NaCl, 10% Glycerol, 5 mM BME, 25 mM Imidazole

Elution Buffer: 20 mM Na\(_2\)HPO\(_4\), 0.5 M NaCl, 10% Glycerol, 5 mM BME, 0.25 M Imidazole

MTAN (M. tuberculosis)
Appendix VII: Anion exchange chromatogram of flow through and fractions F1 and F2 from tag removal chromatography (Appendix VI), and SDS Page of fractions H11-I7.

Loading Buffer: 20 mM Tris pH 8.5, 5 mM BME
Elution Buffer: 20 mM Tris pH 8.5, 5 mM BME, 1 M NaCl

MTAN (M. tuberculosis)
Appendix VIII: Crystal hits from Hampton Research Index Screen

0.1 M Succinic acid pH 7.0, 15% PEG 3350

0.15 M DL-Malic acid pH 7.0, 20% PEG 3350