Enzymatic and structural characterization of proteins linked to Mycobacterium tuberculosis pathogenicity

Julie Boucau

Recommended Citation
http://utdr.toledo.edu/theses-dissertations/1171

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

This Dissertation 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 Dissertation
Entitled

Enzymatic and Structural Characterization of Proteins Linked to *Mycobacterium tuberculosis* Pathogenicity

By
Julie Boucau

Submitted as partial fulfillment of the requirements for
the Doctor in Philosophy in Chemistry

________________________________________
Advisor: Dr. Donald R. Ronning

________________________________________
College of Graduate Studies

The University of Toledo
December 2008
An Abstract of

Enzymatic and Structural Characterization of
Proteins Linked to Mycobacterium tuberculosis Pathogenicity

By

Julie Boucau

Submitted as partial fulfillment of the requirements for
the Doctor of Philosophy in Chemistry

The University of Toledo

December 2008

Tuberculosis (TB) is the leading cause of death by any treatable infectious disease.\(^1\) An estimated 2 billion people are infected by the etiological agent Mycobacterium tuberculosis worldwide.\(^2\) In 2006, 9.2 million new cases were reported and 1.7 million deaths from TB occurred globally.\(^1\) The strong resurgence of the disease is due to HIV coinfection and the appearance of drug-resistant strains of the mycobacterium. In addition to the typical bacterial drug-resistance mechanisms,\(^3\) M. tb possesses a physical barrier that promotes resistance due to its thick hydrophobic cell envelope. The importance of the cell envelope was identified early on and its structure
has been well characterized. Molecules involved in the cell wall maintenance are of great interest in tuberculosis research. The antigen 85 complex (Ag85), responsible for the attachment of fatty acid moieties to carbohydrate components of the cell wall, is a validated drug target. The existing radiometric assay developed for testing the activity of Ag85 is time consuming, thus not readily adapted to the high throughput format required in the drug discovery process. The first part of this work describes the development of a colorimetric enzymatic assay adapted to drug screening. The coupled assay uses a glycoconjugate, \( p \)-nitrophenyl 6-\( O \)-octanoyl-\( \beta \)-D-glucopyranoside, as a substrate. Ag85 transfers the octyl moiety to an acceptor and \( \beta \)-glucosidase releases the chromophore, \( p \)-nitrophenol. After optimization of the assay parameters, the assay was used to screen libraries of synthetic and natural compounds. In addition to its utility for screening applications, the assay was also used for the kinetic characterization of the Ag85 family of enzymes.

Another interesting feature of \( M. \ tb \) is its ability to interfere with the host immune response. For example, the maturation of the phagosome is arrested in the early stages after phagocytosis of the mycobacterium and this process allows the intracellular survival of the pathogen. One protein promoting intracellular survival was identified and called the enhanced intracellular survival (Eis) protein. The pathological effects of Eis have been characterized, however little is known about the exact function of the protein. This work presents crystallization studies of Eis. The putative enzyme was successfully crystallized; however the diffraction achieved did not result in exploitable data.
Acknowledgments

First and foremost, I would like to thank my advisor Dr. Donald Ronning for his help and guidance throughout my adventures in protein biochemistry. I am especially grateful because he took a chance on me and gave me the opportunity to join his laboratory late in my graduate career. I really enjoyed our many conversations about science, the restaurants report cards and life in general (the rule of 25!!!).

I would also like to thank my doctoral committee: Dr. Funk, Dr. Hacker and Dr. Sucheck for their valuable suggestions. I also need to thank Dr. Aditya Sanki from the laboratory of Dr. Sucheck for synthesizing some of the compounds used in this work.

I wish to express my gratitude to Dr. Kirchhoff, Dr. Pinkerton for allowing me to change advisor and Dr. Lind for helping me transition between laboratories.

I would like to thank the chemistry department for funding my teaching assistantship and the DeArce Memorial Endowment Fund for funding my research assistantship over the summer.

I would like to thank the staff of the department of chemistry at the University of Toledo for their help in their domain of expertise.

All the people I have worked with over the years made my doctoral experience fun, stressful at times but more enjoyable overall. Many thanks to all the friends I made over the years, especially the “French connection” and the students in Dr. Mason’s laboratory for providing hours of entertainment.

I would like to thank my family and especially my parents for supporting my decision to study abroad.
# Table of Contents

Abstract ......................................................................................................................... i
Acknowledgments .......................................................................................................... iv
Table of Contents ......................................................................................................... v
List of Figures ................................................................................................................ x
List of Tables ................................................................................................................ xii
List of abbreviations .................................................................................................... xiii
Chapter One .................................................................................................................... 1
Pathogenicity and Virulence of Mycobacterium tuberculosis ................................. 1
  1. Introduction ............................................................................................................. 1
  2. Mycobacterium tuberculosis .................................................................................. 2
    2.1. Mycobacterium tuberculosis is the etiological agent of tuberculosis .......... 2
    2.2. Tuberculosis treatment and drug resistance ................................................. 3
    2.3. The mycobacterial cell envelope .................................................................. 7
      2.3.1. Description of the peptidoglycan ......................................................... 9
      2.3.2. Description of the arabinogalactan ..................................................... 9
      2.3.3. Description of the free glycolipids ..................................................... 10
      2.3.4. Description of the capsule ................................................................. 11
      2.3.5. Bacterial cell envelope .................................................................... 13
  3. Pathogenesis of Mycobacterium tuberculosis ..................................................... 13
    3.1. General description of the disease process .................................................. 13
    3.2. Molecular interaction of *M. tb* and the host cell ....................................... 15
    3.3. Importance of the bacterial cell envelope in relation to pathogenesis .......... 16
  4. Overview and significance of the work ................................................................ 17
Chapter Two ................................................................................................................... 19
Experimental Methods ................................................................................................. 19
  1. Introduction .......................................................................................................... 19
  2. Molecular biology techniques ............................................................................. 19
    2.1. Polymerase chain reaction ......................................................................... 21
    2.1.1. PCR amplification ............................................................................ 21
2.1.2. Mutagenesis PCR ................................................................. 23
2.2. Cloning ............................................................................. 23
2.2.1. Restriction enzymes digestion ...................................... 24
2.2.2. Ligation ......................................................................... 25
2.3. Gel electrophoresis ............................................................. 26
2.4. Transformation ................................................................. 27
2.5. DNA Sequencing ............................................................... 28
3. Protein purification .............................................................. 29
3.1. Protein expression ............................................................. 29
3.1.1. Transformation ............................................................. 30
3.1.2. Inducing protein expression through the lac operon .... 31
3.2. Harvesting the cells ......................................................... 32
3.3. Protein purification ............................................................. 32
3.3.1. Cell Lysis ..................................................................... 32
3.3.2. Chromatography .......................................................... 33
3.3.2.1. Metal affinity chromatography ................................ 33
3.3.2.2. Ion exchange chromatography ............................... 35
3.3.2.3. Size Exclusion chromatography also called Gel filtration .......... 36
3.3.3. Sodium dodecyl sulfate polyacrylamide gel electrophoresis .... 36
3.3.4. Salting out as a concentration method ......................... 36
3.3.5. Dialysis ....................................................................... 37
4. Spectrophotometric applications .......................................... 37
4.1. Introduction to the Beer-Lambert law .............................. 38
4.2. Monitoring the protein purification process using spectrophotometry ............. 39
4.3. Calculation of protein concentration ................................ 39
4.4. Purity determination of protein samples ......................... 39
5. Macromolecular crystallography .......................................... 40
5.1. Introduction to macromolecular crystallography ............... 40
5.2. Crystallization methods .................................................. 40
5.2.1. Hanging drop vapor diffusion crystallization method .......... 40
5.2.2. Microbatch crystallization method .................................................. 41
5.3. Experimental method ................................................................. 41

Chapter Three .................................................................................. 42

Development of a High-Throughput Assay to Test the Activity of Antigen 85  42
1. Introduction to the drug discovery process ....................................... 42
2. The antigen 85 Complex ................................................................. 44
   2.1. Introduction to the antigen 85 complex ....................................... 44
   2.2. The antigen 85 complex is a validated drug target ..................... 49
2.3. Existing assay ............................................................................. 51
3. Development of the high-throughput assay ..................................... 52
   3.1. Design .................................................................................... 52
   3.2. Synthesis of the substrate ......................................................... 54
   3.3. Cloning of *fbpC2*, expression and purification of Ag85C ........ 55
      3.3.1. Cloning of *fbpC2* .................................................................. 55
      3.3.2. Expression and purification of Ag85C .................................. 55
   3.4. Analysis of the data using enzyme kinetics ............................... 60
   3.5. Preliminary experiments .......................................................... 63
      3.5.1. Wavelength determination .................................................. 63
      3.5.2. Determination of the molar extinction coefficient of *p*-nitrophenolate .... 64
   3.6. Experimental methods ................................................................ 65
4. Enzymology Results ...................................................................... 66
   4.1. Positive and negative control experiments for Antigen 85C ......... 66
      4.1.1. Proof-of-principle ................................................................. 66
      4.1.2. Turbidity issues .................................................................... 67
      4.1.3. Rate-determining step ........................................................ 68
      4.1.4. Optimal assay parameters .................................................. 71
   4.2. Kinetic parameters of Ag85C ..................................................... 71
      4.2.1. Introduction .......................................................................... 71
      4.2.2. Results ................................................................................ 72
5. Conclusion on the enzymology ....................................................... 74
6. Future work ........................................................................................................ 75

Chapter Four ........................................................................................................ 76

Screening Applications of the Antigen 85 Assay .................................................... 76

1. Introduction to screening applications .............................................................. 76
1.1. Validation of the assay for high throughput screening application ............... 76
1.2. Secondary control assay ............................................................................. 78
1.3. Type of libraries screened ........................................................................... 78

2. Synthetic compounds ...................................................................................... 79
2.1. Introduction ................................................................................................. 79
2.2. Screening of a library of methyl 5-S-alkyl-5-thio-D-arabinofuranosides ........ 80
2.2.1. Experimental method .......................................................................... 81
2.2.2. Results .................................................................................................. 81
2.3. Screening of arabinofuranoside and trehalose esters library ....................... 82
2.3.1. Experimental methods ........................................................................ 84
2.3.2. Results .................................................................................................. 85
2.4. National Institute of Health Clinical Collection .......................................... 86
2.4.1. Experimental methods ........................................................................ 86
2.4.2. Results .................................................................................................. 87

3. Natural compounds .......................................................................................... 89
3.1. Introduction ................................................................................................. 89
3.2. Screening of natural extracts ..................................................................... 89
3.2.1. Experimental method ........................................................................ 91
3.2.2. Results .................................................................................................. 91
3.2. Conclusion .................................................................................................. 92
3.3. Future work ............................................................................................... 92

Chapter Five ......................................................................................................... 94

Study of the acyl acceptor specificity of the antigen 85 complex ......................... 94
through kinetic and structural characterization .................................................... 94
1. Introduction to the antigen 85 complex substrate paradox ............................ 94
1.1. Experimental methods ............................................................................. 97
1.2. Kinetic characterization of Ag85A and B using D-glucose as the acyl acceptor... 107
1.3. Acyl acceptor specificity

| 1.3.1. Kinetic parameters determination for arabinofuranosyl arabinoside as the acyl acceptor | 109 |
| 1.3.2. Kinetic parameters determination for trehalose as the acyl acceptor | 110 |
1.4. Acyl acceptor specificity

| 1.5. Conclusion | 114 |

2. Structural study of the Antigen 85 complex in relation to its substrate specificity. 115

| 2.1. Structural study of modified Ag85A and Ag85B | 116 |
| 2.2. Kinetic study of Ag85A and Ag85B mutants | 118 |
3. Conclusion

| 4. Future work | 121 |

Chapter Six: Crystallization Studies of the Enhanced Intracellular Survival Protein. 122

| 1. Introduction | 122 |
| 2. Cloning, protein expression and purification of Eis. | 124 |
| 2.1. Cloning of the eis gene | 125 |
| 2.2. Protein expression | 127 |
| 2.3. Protein purification | 129 |
| 3. Crystallization Studies | 134 |
| 4. Conclusion | 136 |
| 5. Future work | 136 |

References 137
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Treatment guidelines for drug–susceptible tuberculosis</td>
<td>6</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Schematic diagram of the <em>M. tb</em> cell wall.</td>
<td>8</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Structure of the monomer of peptidoglycan</td>
<td>9</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Terminal arabinosyl unit of the mycolyl-arabinogalactan</td>
<td>10</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Structure of mycolic acids</td>
<td>12</td>
</tr>
<tr>
<td>Figure 6</td>
<td>General overview of the experimental procedure</td>
<td>20</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Drug Discovery Process</td>
<td>43</td>
</tr>
<tr>
<td>Figure 8</td>
<td>General ping pong mechanism</td>
<td>46</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Mechanism of Ag85C</td>
<td>48</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Mycolyltransferase assay developed by Belisle and coworkers</td>
<td>51</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Substrate used in the assay</td>
<td>53</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Coupled assay</td>
<td>54</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Synthesis of the acyl donor</td>
<td>54</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Metal affinity chromatogram and the corresponding SDS-PAGE</td>
<td>57</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Anion exchange chromatogram and the corresponding SDS-PAGE</td>
<td>58</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Size exclusion chromatogram and the corresponding SDS-PAGE gel</td>
<td>59</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Wavelength scan of <em>p</em>-nitrophenolate in the reaction buffer</td>
<td>64</td>
</tr>
<tr>
<td>Figure 18</td>
<td>Standard curve of <em>p</em>-nitrophenolate</td>
<td>65</td>
</tr>
<tr>
<td>Figure 19</td>
<td>Positive and negative control experiments</td>
<td>66</td>
</tr>
<tr>
<td>Figure 20</td>
<td>DMSO concentration course</td>
<td>68</td>
</tr>
<tr>
<td>Figure 21</td>
<td>Ag85C concentration course</td>
<td>69</td>
</tr>
<tr>
<td>Figure 22</td>
<td>Linear regression plot of initial velocity versus Ag85C concentration</td>
<td>70</td>
</tr>
<tr>
<td>Figure 23</td>
<td><em>β</em>-glucosidase concentration course</td>
<td>70</td>
</tr>
<tr>
<td>Figure 24</td>
<td>Absorbance versus time graphs for different substrate concentrations</td>
<td>70</td>
</tr>
<tr>
<td>Figure 25</td>
<td>Michaelis-Menten Fit of Vi values versus substrate concentration</td>
<td>73</td>
</tr>
<tr>
<td>Figure 26</td>
<td>Control reaction for <em>β</em>-glucosidase</td>
<td>73</td>
</tr>
<tr>
<td>Figure 27</td>
<td>Published libraries of Ag85 inhibitors</td>
<td>78</td>
</tr>
<tr>
<td>Figure 28</td>
<td>Library of methyl 5-S-alkyl-5-thio-D-arabinofuranosides</td>
<td>80</td>
</tr>
<tr>
<td>Figure 29</td>
<td>Absorbance versus time curve for C4α at different concentrations</td>
<td>81</td>
</tr>
<tr>
<td>Figure 30</td>
<td>Structure of the mechanism-based inhibitors</td>
<td>83</td>
</tr>
<tr>
<td>Figure 31</td>
<td>Percent activity of Ag85C in presence of methyl ester</td>
<td>85</td>
</tr>
<tr>
<td>Figure 32</td>
<td>Pictures of the disk diffusion assay plates</td>
<td>88</td>
</tr>
<tr>
<td>Figure 33</td>
<td>Structure of Clofazimine and Ebselen</td>
<td>88</td>
</tr>
<tr>
<td>Figure 34</td>
<td>Screening procedure for the natural extracts</td>
<td>90</td>
</tr>
<tr>
<td>Figure 35</td>
<td>Acyl acceptors used in the assay</td>
<td>97</td>
</tr>
<tr>
<td>Figure 36</td>
<td>Ag85A Nickel affinity chromatogram and the corresponding SDS-PAGE gel</td>
<td>100</td>
</tr>
</tbody>
</table>
Figure 37  Ag85B Nickel affinity chromatogram and the corresponding SDS-PAGE gel ................................................................. 101
Figure 38  Ag85A anion exchange chromatogram and the corresponding SDS-PAGE gel ............................................................. 102
Figure 39  Ag85B anion exchange chromatogram and the corresponding SDS-PAGE gel ............................................................. 103
Figure 40  Ag85A size exclusion chromatogram and the corresponding SDS-PAGE gel ................................................................. 104
Figure 41  Ag85B size exclusion chromatogram and the corresponding SDS-PAGE gel ................................................................. 105
Figure 42  Test reaction with 100 nM of each enzyme ................................................................. 106
Figure 43  Michaelis-Menten fit when D-glucose is used as an acceptor ................................................................. 108
Figure 44  Michaelis-Menten fit when AA is used as an acceptor ................................................................. 109
Figure 45  Michaelis-Menten fit when trehalose is used as an acceptor ................................................................. 111
Figure 46  Comparison of the Michaelis Menten fit obtained for Ag85A, B and C ................................................................. 113
Figure 47  DEP modification of Ag85C ................................................................. 115
Figure 48  Comparison of the 210-221 mobile loop of Ag85C and Ag85C-DEP ................................................................. 116
Figure 49  Picture of Ag85B-DEP crystals ................................................................. 117
Figure 50  Picture of Ag85A-DEP crystals ................................................................. 118
Figure 51  Sequence alignment of Ag85A, C and B from amino acid 209 to 228 ................................................................. 119
Figure 52  Sequence alignment of the eis gene plasmid with the R69L mutation ................................................................. 126
Figure 53  N-terminus sequence of pET28-Eis ................................................................. 128
Figure 54  C-terminus sequence of pET28-Eis ................................................................. 129
Figure 55  Nickel affinity chromatogram and corresponding SDS-PAGE ................................................................. 131
Figure 56  Tag Removal Chromatogram and corresponding SDS-PAGE ................................................................. 132
Figure 57  Size Exclusion Chromatogram and corresponding SDS-PAGE ................................................................. 133
Figure 58  Initial hits visualized by optical microscopy ................................................................. 134
Figure 59  Initial crystal hit visualized by optical microscopy ................................................................. 135
Figure 60  Crystals with acetyl-CoA ................................................................. 136
List of Tables

Table 1  Commonly used drugs for tuberculosis treatment .................................. 4
Table 2  Optimal reaction conditions .................................................................... 71
Table 3  Kinetic parameters with D-glucose as the acyl acceptor ......................... 108
Table 4  Kinetic parameters with arabinofuranosyl arabinoside as the acyl acceptor 110
Table 5  Kinetic parameters when trehalose is used as an acceptor ...................... 111
Table 6  Kinetic Parameters of Ag85 with different acceptors ............................. 112
Table 7  Enzymatic activity of the Ag85A and Ag85B mutants ............................. 119
Table 8  Predicted parameters of Eis ...................................................................... 124
List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>% v/v</td>
<td>percent of the total volume in solution</td>
</tr>
<tr>
<td>6xHis</td>
<td>poly histidine tag</td>
</tr>
<tr>
<td>AA</td>
<td>methyl-2-O-β-D-arabinosylfuranosyl-α-D-arabinofuranoside</td>
</tr>
<tr>
<td>AChE</td>
<td>acetylcholine esterase</td>
</tr>
<tr>
<td>ADT</td>
<td>6-azido-6-deoxy-α,α'-trehalose</td>
</tr>
<tr>
<td>AG</td>
<td>arabinogalactan</td>
</tr>
<tr>
<td>Ag85</td>
<td>Antigen 85 complex</td>
</tr>
<tr>
<td>Ag85A</td>
<td>Antigen 85A</td>
</tr>
<tr>
<td>Ag85B</td>
<td>Antigen 85B</td>
</tr>
<tr>
<td>Ag85C</td>
<td>Antigen 85C</td>
</tr>
<tr>
<td>AMK</td>
<td>amikacin</td>
</tr>
<tr>
<td>C4</td>
<td>5-S-butyl-5-thio-D-arabinofuranoside</td>
</tr>
<tr>
<td>C6</td>
<td>5-S-hexyl-5-thio-D-arabinofuranoside</td>
</tr>
<tr>
<td>C8</td>
<td>5-S-octyl-5-thio-D-arabinofuranoside</td>
</tr>
<tr>
<td>CCD</td>
<td>charge-coupled device</td>
</tr>
<tr>
<td>CIP</td>
<td>calf intestinal alkaline phosphatase</td>
</tr>
<tr>
<td>CP</td>
<td>ciprofloxacin</td>
</tr>
<tr>
<td>CPM</td>
<td>capreomycin</td>
</tr>
<tr>
<td>DEP</td>
<td>diethyl phosphate</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTPs</td>
<td>deoxyribonucleotides</td>
</tr>
<tr>
<td>DOTS</td>
<td>Directly Observed Therapy Short Course</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>DZI</td>
<td>diameter of zone of inhibition</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene diamine tetraacetate</td>
</tr>
<tr>
<td>Eis</td>
<td>Enhanced intracellular survival protein</td>
</tr>
<tr>
<td>EMB</td>
<td>ethambutol</td>
</tr>
<tr>
<td>ETH</td>
<td>ethionamide</td>
</tr>
<tr>
<td>Fbp</td>
<td>fibronectin-binding protein</td>
</tr>
<tr>
<td>FPLC</td>
<td>fast protein liquid chromatography</td>
</tr>
<tr>
<td>GL</td>
<td>glycolipids</td>
</tr>
<tr>
<td>HIV</td>
<td>Human Immunodeficiency Virus</td>
</tr>
<tr>
<td>HTS</td>
<td>high throughput screening</td>
</tr>
<tr>
<td>IC₅₀</td>
<td>concentration required to achieve 50% inhibition</td>
</tr>
<tr>
<td>IL-10</td>
<td>interleukin-10</td>
</tr>
<tr>
<td>IL-4</td>
<td>interleukin-4</td>
</tr>
<tr>
<td>INF-γ</td>
<td>interferon-γ</td>
</tr>
<tr>
<td>INH</td>
<td>isoniazid</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactoside</td>
</tr>
<tr>
<td>ITC</td>
<td>isothermal calorimetry</td>
</tr>
<tr>
<td>KAN</td>
<td>kanamycin</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$</td>
<td>catalytic efficiency</td>
</tr>
<tr>
<td>$k_{\text{cat}}$</td>
<td>turn-over number</td>
</tr>
<tr>
<td>$K_i$</td>
<td>inhibitory constant</td>
</tr>
<tr>
<td>$K_M$</td>
<td>Michaelis-Menten constant</td>
</tr>
</tbody>
</table>
LB: Luria-Bertani
*M. smegmatis*: *Mycobacterium smegmatis*
*M. tb*: *Mycobacterium tuberculosis*
mAGP: mycolyl arabinogalactan peptidoglycan
MIC: minimum inhibitory concentration
MLSMR: Molecular Libraries Small Molecule Repository
mRNA: messenger ribonucleic acid
NCC: NIH Clinical Collection
NCE: new chemical entity
NF-κB: nuclear factor κB
NIH: National Institute of Health
NO: nitric oxide
OD: optical density
OL: outer layer
PAS: *p*-aminosalicylic acid
PCR: polymerase chain reaction
PG: peptidoglycan
PKS-TE: polyketide synthesis thioesterase
PZA: pyrazinamide
RBT: rifabutin
RIF: rifampin
RNA: ribonucleic acid
SDS: sodium dodecyl sulfate
SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis
STM: streptomycin
TAC: thiacetazone
TB: terrific broth
TB: tuberculosis
TBE: tris/borate/ethylene diamine tetraacetate
TDM: trehalose dimycolate
TE: thioesterase
TFF: tangential flow filtration
TLR: toll-like receptors
TMM: trehalose monomycolate
TNF-α: tumor necrosis factor α
V_i: initial velocity
V_max: maximum velocity
WHO: World Health Organization
Chapter One

Pathogenicity and Virulence of Mycobacterium tuberculosis

1. Introduction

With 1.7 million lives claimed in 2006, Tuberculosis (TB) remains one of the deadliest treatable infectious diseases. 9.2 million cases were reported in 2006 worldwide. An estimated 14 million people are infected every year by *Mycobacterium tuberculosis* (*M. tb*), the bacterium responsible for TB. The disease was considered under control in the second half of the twentieth century after the discovery of antitubercular drugs, but it recently revived as a global health emergency. New challenges in treatment are arising due to the occurrence of TB and human immunodeficiency virus (HIV) coinfection and the emergence of drug resistant forms of the etiological agent. In conjunction with poor control programs, those challenges are leading to the resurgence of the disease. The international scare prompted by an American tourist infected with a drug resistant form of TB in 2007 was a powerful reminder that TB is no longer history but is active and still spreading. Therefore research aimed at developing new methods, in addition to identifying and characterizing new drug targets is of prime importance. This chapter introduces the causative agent of TB with an emphasis on its cellular envelope, the causes of drug resistance and a brief presentation of mycobacterial pathogenesis.
2. Mycobacterium tuberculosis

2.1. Mycobacterium tuberculosis is the etiological agent of TB

The infectious nature of TB was identified by Robert Koch in 1882. His work characterized the etiological agent of TB as a bacterium: *Mycobacterium tuberculosis*, also called tubercle bacillus or Koch’s bacillus.\(^7\) Robert Koch was awarded the Nobel Prize in Physiology or Medicine in 1905 for his pioneering work in bacteriology and, especially, his investigations and discoveries in relation to tuberculosis.

*M. tb* belongs to the Mycobacterium genus, Mycobacteriaceae family, Corinebacterinae suborder, Actinomycetales order, Actinobacteria phylum in the Bacteria kingdom. Besides *M. tb*, a few other species have been identified to cause TB. *Mycobacterium africanum, bovis, canetti, caprae, microti* and *tuberculosis* are known as the Mycobacterium tuberculosis complex.\(^8\) Other species from that genus are pathogenic bacteria such as *Mycobacterium leprae* that causes leprosy, *Mycobacterium avium* complex and *Mycobacterium kansasii* that cause pulmonary diseases distinct from TB.\(^9\)

Mycobacteria are aerobic bacteria often described as Gram-positive bacteria due to their lack of an outer membrane. All species are characterized by a thick hydrophobic waxy cell wall that is believed to confer the bacteria most of its resistance and survival properties.\(^10, 11\) Among those, *M. tb* is a slow-dividing rod-shaped bacterium with a division time of 100 to 500 hours making it cumbersome to culture. The study of *M. tb* progressed during the twentieth century with scientific advances, and the complete sequencing of the genome of the H37Rv laboratory strain in 1998 led to tremendous research progress.\(^12, 13\) However the recent emergence of drug-resistant strains of the bacterium emphasizes the need to continue the study of *M. tb*. 
2.2. Tuberculosis treatment and drug resistance

The bacterial nature of the disease prompted the early use of antibiotics to treat TB. For example, streptomycin was used to treat TB shortly after its discovery. Unfortunately the bacterium rapidly developed resistance and the same phenomenon was observed after the introduction of new drugs. As a consequence the use of combination therapy was adopted. The main drugs prescribed to treat TB are presented in Table 1. They are divided into two categories: first–line and second–line drugs. The first line drugs, isoniazid (INH), rifampin (RIF), pyrazinamide (PZA) and ethambutol (EMB), are more effective and usually prescribed when the \textit{M. tb} strain is identified as drug-susceptible. Because of the development of drug resistance, through the typical bacterial artillery against antibiotics such as efflux pumps, drug-modifying enzymes and mutations, second-line drugs are necessary for treatment. The most common second-line therapy includes cycloserine, ethionamide (ETH), streptomycin (STM), amikacin/kanamycin (AMK/KAN), capreomycin (CPM), \textit{p}-aminosalicylic acid (PAS) and ciprofloxacin (CP).

Three classes of resistant strains have been identified: single-drug resistant, multidrug resistant (MDR) and extensive drug resistant (XDR). MDR strains are resistant to two or more of the first-line drugs, such as INH and RIF. XDR strains are resistant to first-line drugs and among second-line drugs any fluoroquinolones and at least one of the three injectable drugs (AMK, KAN and CPM). While MDR-TB is treatable, XDR-TB often results in the death of the patient because of its rapid progression.
<table>
<thead>
<tr>
<th>Classification</th>
<th>Drug (abbreviation)</th>
<th>Mechanism of action</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>First-line</td>
<td>Isoniazid (INH)</td>
<td>Synthesis of mycolic acids</td>
<td>ketoenoylreductase InhA</td>
</tr>
<tr>
<td></td>
<td>Rifampin (RIF)</td>
<td>mRNA synthesis</td>
<td>β-DNA-dependent RNA polymerase</td>
</tr>
<tr>
<td></td>
<td>Rifabutin (RBT)</td>
<td>mRNA synthesis</td>
<td>β-DNA-dependent RNA polymerase</td>
</tr>
<tr>
<td></td>
<td>Pyrazinamide (PZA)</td>
<td>Fatty acid synthesis</td>
<td>Fatty acid synthetase I FasI</td>
</tr>
<tr>
<td></td>
<td>Ethambutol (EMB)</td>
<td>Cell wall synthesis</td>
<td>Arabinosyl transferase</td>
</tr>
<tr>
<td>Second-line</td>
<td>Cycloserine</td>
<td>Peptidoglycan synthesis</td>
<td>Alanine racemase D-alanine ligase</td>
</tr>
<tr>
<td></td>
<td>Ethionamide (ETH)</td>
<td>Mycolic acid synthesis</td>
<td>unknown</td>
</tr>
<tr>
<td></td>
<td>Streptomycin (STM)</td>
<td>Protein synthesis</td>
<td>16S ribosomal RNA</td>
</tr>
<tr>
<td></td>
<td>Aminoglycosides</td>
<td>Protein synthesis</td>
<td>30S ribosomal subunit</td>
</tr>
<tr>
<td></td>
<td>Amikacin/Kanamycin</td>
<td>Protein synthesis</td>
<td>30S ribosomal subunit</td>
</tr>
<tr>
<td></td>
<td>(AMK/KAN)</td>
<td></td>
<td>unknown</td>
</tr>
<tr>
<td></td>
<td>Polypeptides</td>
<td>Protein synthesis</td>
<td>unknown</td>
</tr>
<tr>
<td></td>
<td>Capreomycin (CPM)</td>
<td></td>
<td>unknown</td>
</tr>
<tr>
<td></td>
<td><em>p</em>-aminosalicylic acid (PAS)</td>
<td>Iron uptake</td>
<td>Thymidylate synthase thyA</td>
</tr>
<tr>
<td></td>
<td>Fluoroquinolones</td>
<td>Cell division</td>
<td>DNA gyrase</td>
</tr>
<tr>
<td></td>
<td>Ciprofloxacin (CP)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Commonly used drugs for tuberculosis treatment
The primary cause of drug resistance is patient non-compliance. The drug regimen is usually heavy and long. Depending on the diagnosis, the treatment can last up to 18 months and requires the administration of up to four drugs as shown in Figure 1.\textsuperscript{18}

The typical symptoms are a chronic cough with blood-tinged sputum, fever, night sweats and weight loss. The diagnosis usually relies on chest X-ray, a tuberculin skin test, blood test and microbiological culture of bodily fluids. The treatment algorithm for an at-risk patient with drug-susceptible TB is presented in Figure 1. When the initial patient evaluation shows typical symptoms along with abnormal chest X-ray and positive tuberculin test, there are two possible options. In case of low suspicion of active tuberculosis, no treatment is prescribed at first. In case of high suspicion of active TB, a combination of four drugs (INH, RIF, EMB and PZA) is prescribed for two months. After two months, a repeat evaluation is performed and at that time the culture results are known. For the low suspicion cases, even if TB is not confirmed, the patient is treated with one of three drugs for three to nine months. For the high suspicion cases, depending on the test results, a drug regimen of INH and RIF combined is prescribed for up to seven months. For drug-resistant TB, the treatment algorithm is even more complex.

Whenever the treatment is not completed, the wrong treatment has been prescribed, or the dosage and/or the length are erroneous, drug resistance may arise. The strategies of drug resistance are well known. Bacteria can inactivate or modify the drug as well as alter or overexpress the target. They can also reduce drug accumulation through barrier mechanisms or the use of efflux pumps and by decreasing drug permeability.\textsuperscript{16} In addition to developed resistance mechanisms, \textit{M. tb} possesses an intrinsic resistance due to its cell envelope.
Figure 1  Treatment guidelines for drug–susceptible tuberculosis  
(Adapted from reference 18)
The development by the World Health Organization (WHO) of the Directly Observed Therapy Short Course (DOTS) helped the implementation of guidelines to monitor patient compliance. The five-point strategy required government commitment to sustainable TB control, diagnosis through sputum smear microscopy, provision of standardized short-course chemotherapy treatment, a functioning drug supply system and a recording and reporting system to assess treatment results. DOTS helped reduce the global burden of TB.\textsuperscript{3} Initiatives such as Stop TB, a network of international organizations interested in eliminating TB, have been created to expand the DOTS strategy.\textsuperscript{3} So far, the efforts to limit emergence of new drug-resistant mycobacteria and the spread of existing MDR and XDR strains seem successful. It is nevertheless necessary to continue research efforts on \textit{M. tb}. The need to achieve a better understanding of the mycobacterium and the disease is fundamental for new drug discovery and eradication of TB.

2.3. The mycobacterial cell envelope

The mycobacterial cell envelope is a dynamic structure, and it is hypothesized that mycobacteria adapt the composition of their cell wall depending on the environment and their growth phase.\textsuperscript{10} However, a general model of the cell wall is nevertheless widely accepted. The envelope, shown in Figure 2, is made of four major components in addition to the plasma membrane: the peptidoglycan (PG) and the arabinogalactan (AG), a covalently-linked complex of mycolic acids, a layer of free glycolipids (GL) and an outer layer (OL) sometimes called the capsule. The PG, anchored in the PM, is covalently linked to the AG through a diglycosyl phosphoryl bridge. The mycolyl-arabinogalactan-peptidoglycan complex (mAGP) results from esterification at the non-
reducing ends with mycolic acids. The GL interact with the mycolyl moiety of the mAGP. Finally the OL is a mixture of polysaccharides and proteins.\textsuperscript{19}

Figure 2  Schematic diagram of the \textit{M. tb} cell wall.
2.3.1. Description of the peptidoglycan (PG)

PG is composed of chains of glycan formed from alternating units of N-acetylglucosamine and muramic acid moieties that are glycolated or acetylated. The L-alanyl-D-isoglutaminyl-meso-diaminopimelyl-D-alanine tetrapeptide chains are attached to the muramic acid residues as shown in Figure 3. This tetrapeptide is responsible for the cross-linking of PG.\textsuperscript{19}

![Figure 3](image3.png)

Figure 3 Structure of the monomer of peptidoglycan

2.3.2. Description of the arabinogalactan (AG)

AG is composed of D-arabinofuranosyl and D-galactofuranosyl residues. The homogalactan polymer is made of thirty linear $\beta$-D-galactofuranosyl residues
alternatively linked through their C-5 and C-6. Two or three arabinan chains are attached to the C-5 of 6-linked galactofuranosyl residues. The galactan component is bound to the C-6 of PG N-glycolylmuramic acid by a diglycosyl phosphate linkage. The homoarabinan chains are composed of linear α-D-arabinofuranosyl residues with branching produced by 3,5-linked α-D-arabinofuranosyl units substituted at both positions by α-D-arabinofuranosyl residues. The non-reducing termini of the arabinan chains consist of pentaarabinosyl units. Within a particular arabinan chain, two-thirds of the α-D-arabinofuranosyl residues are esterified by a group of four mycolic acids, as shown in Figure 4.11,20

![Figure 4](image)

**Figure 4**  Terminal arabinosyl unit of the mycolyl-arabinogalactan

(Adapted from references 11, 20)

### 2.3.3. Description of the free glycolipids (GL)

Free glycolipids such as trehalose dimycolate (TDM), also called cord factor, different acylated forms of glucose and trehalose, phthiocerol dimycocerosates, phenolic
glycolipids and sulfolipids interact with the mAGP forming a hydrophobic bilayer around the cell.\textsuperscript{20}

Mycolic acids are long chain $\alpha$-branched-$\beta$-hydroxyl fatty acids and contain 60 to 90 carbon atoms. They possess a meromycolate moiety with carbon chain lengths of up to 56 atoms and a long saturated $\alpha$-branch of 20 to 24 carbon atoms. The mycolic acids contain various organic functional groups such as $cis$ and $trans$ cyclopropane rings, $cis$ and $trans$ double bonds, and keto, methoxy, epoxy and wax-ester functional groups in addition to the $\beta$-hydroxy acid unit. They are divided into two main classes: $\alpha$-mycolic acids and oxygenated mycolic acids, as shown on Figure 5.\textsuperscript{11,20}

2.3.4. Description of the capsule

The outer layer of the envelope is a mixture of proteins and polysaccharides. The capsule was first identified when it accumulated around unstirred $\textit{M. tb}$ in cultures; several of its components were then characterized in the bacterial host.\textsuperscript{10,11} The main polysaccharide components of the capsule are a glucan, an arabinomannan and a mannan. The peptide component of the capsule is not as well characterized since it depends greatly on the isolation conditions. Different enzymes have been identified such as superoxide dismutase, glutamine synthetase and thioredoxin.\textsuperscript{21} The last major component of the cell envelope is a mixture of glycolipids which spans the cell wall and is anchored to the plasma membrane. It is composed of phosphatidyl-$\textit{myo}$-inositol mannosides, lipomannans and lipoarabinomannans and several proteins such as porins, that allow nutrients to cross the cell envelope.\textsuperscript{11,20}
Figure 5  Structure of mycolic acids

(Adapted from reference \textsuperscript{11, 20})
2.3.5. Bacterial cell envelope

The mycobacterial cell wall is a complex structure essential for mycobacterial survival for multiple reasons. First, the lipid bilayer formed by the free GL and the mycolate moiety of mAGP creates a passive barrier impermeable to water-soluble substances. By forming a close packed layer around the cell, the mycolic acids limit entry of hydrophilic molecules deep into the cell wall. Second, mycolic acids have been linked to a plethora of major biological effects in the human host. They offer protection against environmental stress and reactive species as well as a natural resistance against drugs, and participate in mycobacterial pathogenesis.8

3. Pathogenesis of Mycobacterium tuberculosis

To develop new efficient drugs, it is important to understand the disease process. One of the characteristics of M. tb infection is the ability of the mycobacteria to exist in two different states: active and latent infection. Once it has infected its host, the bacterium does not always promote acute disease; it can also persist in an inactive state for long periods of time. In contrast to the active state, the dormant state of the bacteria is asymptomatic and non-transmissible. Because of this, infected individuals with a healthy immune system have a 5-10% lifetime chance to develop active TB.22

3.1. General description of the disease process

TB is transmitted by exposure to M. tb present in aerosol droplets produced by a cough or a sneeze of an infected patient. The bacteria enter the upper respiratory tract and end up in the terminal air spaces of the lung. The infection goes through different stages.23 In the early events of the infection, the mycobacteria are phagocytosed by
resident macrophages. The phagocytosis is initiated by bacterial contact with macrophage receptors. Upon entry in the macrophage, *M. tb* is encased in an endocytic vacuole called the phagosome. Once a foreign molecule has been internalized, the phagosome usually goes through a maturation process involving acidification, fusion with a lysosome and production of reactive oxygen and nitrogen species. However in the case of *M. tb*, the maturation is arrested in its early stage. The involved mechanisms are not known, but more and more experimental observations are revealing a very complex process. Once established in the phagosome, the mycobacteria can progress through the different stages of infection which are replication, dissemination, establishment, maintenance of latency and reactivation. After the arrest of phagosome maturation, the mycobacteria replicate within the macrophage. The infected cells then migrate into different tissues during the dissemination stage. Stimulation of the alveolar macrophages causes their implantation in the lung epithelium. In the later events of the disease, infected macrophages release chemokines that attract inactivated monocytes, lymphocytes and neutrophils. These recruited immune cells and the infected macrophages form the granuloma, that usually contains the spread of the bacteria. As cellular immunity develops, macrophages loaded with bacilli are killed, which results in the formation of the caseous center of the granuloma, surrounded by a cellular zone of fibroblasts, lymphocytes, and blood-derived monocytes. Some organisms are not killed and remain in the granuloma in a dormant state. If and when the immune system of the host weakens, the granuloma center liquefies and serves as a rich medium in which the revived bacteria can replicate uncontrollably. At this point the viable mycobacteria, now
active, can spread. The exact conditions for reactivation of the bacteria are not well understood.

3.2. Molecular interaction of M. tb and the host cell

Numerous mycobacterial and host molecules are involved in the different stages of the infection process. Only the observations about phagocytosis and macrophage maturation, directly related to this research, are described here. Some of the macrophage receptors involved in phagocytosis have been identified. These include the CD36 family of scavenger receptors, the type 3 complement receptor and mannose receptors. The mycobacterial mannose receptor agonists are the mannan components of the capsule. The type 3 complement receptors interact with the C3 protein from the complement that opsonizes mycobacteria; the complement receptors also recognize different secreted proteins including the antigen 85 complex.

For the modulation of phagosome maturation, some of the bacterial effectors have been identified, but the mycobacterium also uses some host molecules to its advantage. The main characteristics of the maturation arrest are the prevention of acidification of the phagosome, the preclusion of phagosome-lysosome fusion and the suppression of reactive oxygen and nitrogen intermediate production. For example the exclusion of H⁺-ATPases from the vacuole has been linked to the absence of phagosomal acidification, but the exact mechanism remains unknown. The prevention of phagosome-lysosome fusion requires inhibition of vesicle migration in the host cell. A small GTPase, named Rab5 protein, is involved with early endosome trafficking. Several studies show an accumulation and segregation of Rab5 in the phagosome vacuole. As a result of its subcellular localization, Rab5 is no longer able to interact with the membrane of the
phagosome and promote endosomal trafficking. At the same time, mycobacteria modulate cellular trafficking using their own factors such as the proteins SapM and PknG and lipids like TDM and mannose lipoarabinomannan.\textsuperscript{16, 28} The inhibition of Ca\textsuperscript{2+} signaling events has also been observed in relation to phagocytosis maturation. Down-regulation of calmodulin-dependent signal transduction and inhibition of sphingokinase prevent the accumulation of Ca\textsuperscript{2+}. This prevents the respiratory burst and formation of oxygen reactive species, nitric oxide (NO) and cytokine production, protecting the mycobacteria against the usual antibacterial host defense mechanisms.\textsuperscript{15} Finally, other proteins causing the intracellular survival phenotype were identified, but their mechanism of action is still unknown. One such protein is Eis, the enhanced intracellular survival protein.

3.3. Importance of the bacterial cell envelope in relation to pathogenesis

The thick and hydrophobic mycobacterial envelope that was detailed in a previous section, offers a passive form of protection by acting as a barrier. Several of its components are involved with different stages of mycobacterial pathogenesis, including interaction with the host immune system. For instance, Toll-like receptor (TLR) agonists include a variety of cell wall products. In a healthy host, TLR proteins are part of a signaling cascade that leads to the activation and the nuclear translocation of nuclear factor κB (NF-κB) which is involved in the expression of many immune response genes such as those encoding cytokines, tumor necrosis factor α (TNF-α) and interleukins.\textsuperscript{16} TNF-α is a major determinant of immunity involved in NF-κB activation, protein-kinase pathway activation and apoptosis induction. Another TLR-specific response is the
expression and secretion of interferon-\(\gamma\) (INF-\(\gamma\)) by activated macrophages. INF-\(\gamma\) is an early gene product that causes potent autocrine and paracrine effects on the targeted cell such as macrophage activation. It is linked with antigen presentation, leukocyte-endothelium cell interactions, cell growth and apoptosis, reactive oxygen and nitrogen intermediates production and phagosome-lysosome fusion. In the case of an infected macrophage, several components of the mycobacterial capsule interfere with the normal immune response.\(^{21}\) The lipoarabinomannan inhibits INF-\(\gamma\) activation of macrophages, induces the release of TNF-\(\alpha\) from macrophages and scavenges oxygen radicals. There are also reports that phenolic glycolipids induce non-specific inhibition of lymphoproliferation. The consequence of inhibiting lymphoproliferation is a change in the pattern of cytokine production and the failure to activate macrophages. The arabinomannnan has been shown to exhibit immunosuppressive properties.\(^{9}\)

Many host-pathogen interactions and molecular processes involved in \(M. \text{tb}\) pathogenesis are still unknown. Nonetheless, recent advances in the characterization of the mycobacterium suggest that tuberculosis remains a treatable disease.

4. Overview and significance of the work

The knowledge of the agent causing tuberculosis is getting more detailed. The complete sequencing of the mycobacterial genome provided valuable clues on possible new drug targets.\(^{19,20}\) The collection of experimental evidence regarding the interaction of \(M. \text{tb}\) with its cellular host is growing, yielding a better understanding of mechanistic and structural features essential for survival of the pathogen. The importance of the cell envelope for mycobacterial survival was presented in this chapter. As a result, cell wall components and any molecules linked with their maintenance are possible drug targets.\(^{29}\)
This approach was indeed successful with the development of several antitubercular drugs targeting the cell wall, until some bacterial strains developed resistance to the drugs. Also, among the processes involved in pathogenesis, the intracellular survival of mycobacteria in macrophages is important, however it is poorly understood. This work focuses on the study and characterization of proteins involved in the cell wall maintenance, the antigen 85 complex (Ag85) and the intracellular survival of the pathogen, enhanced intracellular survival protein (Eis).

The present chapter gave a partial introduction to tuberculosis and *Mycobacterium tuberculosis*. Chapter two will describe the general experimental methods used during this work. Chapter three will present the development of an assay testing the activity of the antigen 85 complex while chapter four will introduce the application of the assay as a screening methodology for drug development and chapter five will present the enzymology studies of the Ag85 complex. Finally chapter six will describe the progress made on crystallization studies of the enhanced intracellular survival protein, Eis.
Chapter Two

Experimental Methods

1. Introduction

The same general strategy was used for the different proteins studied. The genes of interest were cloned from genomic deoxyribonucleic acid (DNA) using traditional molecular biology techniques and the recombinant proteins were purified through a standard fractionation process. A general overview of the typical procedure is given in Figure 6. The pure proteins were then used for biochemical and structural characterization.

This chapter presents the theoretical aspect of the methods and techniques employed for this research. Practical details will be given in the different chapters when needed. First the molecular biology techniques will be detailed followed by the protein expression and purification methodologies.

2. Molecular biology techniques

Molecular biology techniques include polymerase chain reaction (PCR), molecular cloning, gel electrophoresis and DNA sequencing. This section describes the different molecular biology techniques used in the laboratory.
Protein Purification

PCR
- Traditional Cloning
- Gateway® Cloning
- TOPO®-TA cloning

Cloning/expression vector
- Sequencing
- Transformation into competent cells
- Bacterial culture/ expression induced with IPTG
- Metal affinity chromatography
  - N-terminal Tag Tag Removal
  - C-terminal Tag
- Ion Exchange Chromatography (if needed)
- Size Exclusion Chromatography

Molecular Biology

Figure 6 General overview of the experimental procedure
2.1. Polymerase chain reaction

The polymerase chain reaction (PCR) was used to amplify the genes of interest and carry out mutagenesis experiments. In the first case, this technique is used to obtain a large quantity of DNA. In the second case, it permits the modification of the sequence of nucleotides of the gene of interest and ultimately can be used for the synthesis of a mutated DNA product.

PCR uses the property of double-stranded DNA to melt at high temperatures. The PCR reaction proceeds in three steps: a denaturation step, an annealing step and an elongation step. Those three steps, representing one PCR cycle, are usually repeated several times. During the denaturation step, the reaction mixture is heated to 95 °C to disrupt hydrogen bonds in the double stranded template; producing single stranded DNA. The annealing step, typically carried out 2 or 3 °C below the predicted melting temperatures of the primers, promotes the annealing of the primers to their complimentary sequences in the template. Finally, the elongation step is run at 68 °C, which is the optimal activity temperature for the heat-stable Pfx50 polymerase used for all PCR reactions (Invitrogen). In this step, the DNA polymerase extends the primer by adding nucleoside triphosphates in the 5'- to 3'- direction.

2.1.1. PCR amplification

The genes of interest were amplified by PCR. The PCR primers were designed to include nucleotide sequences complementary to the 5'- and 3'- termini of the gene of interest and the specific recognition sequences of the desired restriction enzymes to be used for molecular cloning.
A representative PCR reaction was prepared with:

- 40.5 µL of sterile water,
- 1 µL of each primer (IDT) (0.1 mM stock)
- 1 µL of template DNA (200 ng/µL)
- 5 µL of 10X reaction buffer (Invitrogen)
- 1 µL of PfX50 DNA polymerase (Invitrogen)
- 0.5 µL of deoxynucleotides (dNTPs) mix (25 mM)

In the case of PCR using genomic DNA as a template, only 39.5 µL of sterile water and 1 µL of formamide or dimethyl sulfoxide (DMSO) were used instead of 40.5 µL. Formamide prevents non-specific binding of the primers to the genomic DNA.

A typical PCR program had an initial melting step of 5 min at 95 ºC. The reaction then proceeded with a 30 s denaturation step at 95 ºC, followed by a 45 s annealing step and a 1.5 to 2 min per 1000 basepairs of amplified DNA elongation step at 68 ºC depending on the length of the gene. The PCR cycle was repeated 35 times. A polishing step (5 min at 68 ºC) was generally added at the end of the reaction.

All PCR reactions were run using a Mastercyler personal thermal cycler from Eppendorf, GmBh. The outcome of the PCR experiment is determined by gel electrophoresis *vide infra*. After completion of the PCR program, the reaction mixtures were subjected to agarose gel electrophoresis *vide infra*. When the desired product was observed, it was cut out of the gel and extracted using a QIAGEN gel extraction kit and stored at -20 ºC.
2.1.2. Mutagenesis PCR

The mutagenesis primers were designed to be identical to the template sequence except for the desired base changes. Typically those oligonucleotides were longer than primers used for PCR amplification in order to overcome the energetic penalty caused by the mismatched base pairs. The mutagenesis reactions were typically run at 54 °C unless specified otherwise. A typical PCR program had an initiation step of 5 min at 95 °C, followed by a 30 s denaturation step at 95 °C, a 45 s annealing step and a 1 min per 1000 basepairs of DNA in the plasmid to be mutated elongation step at 68 °C depending on the length of the template. This PCR cycle was repeated 16 cycles. If needed, the program was modified to include a touchdown annealing step. The composition of the mutagenesis PCR experiments was identical to that of PCR amplification reactions described earlier. Once the reaction was complete, the reaction mixture was submitted to digestion with DpnI enzyme, purchased from New England Biolabs Inc., for one hour at 37 °C. This enzyme is a restriction endonuclease that cleaves methylated DNA. DpnI cuts 5'-GATC-3' sequences after the methylated adenine. This digestion step permits the degradation of the methylated template DNA molecules that do not bear the mutation. Only the mutated DNA molecules are left in the solution. After digestion, 1 µL of the digestion reaction is used to transform competent cells 

2.2. Cloning

Once the gene of interest has been amplified, it is digested with restriction endonucleases then inserted into a bacterial plasmid using DNA ligase.
2.2.1. Restriction enzymes / digestion

Restriction digest experiments use Type II restriction endonuclease enzymes that recognize specific palindromic sequences of DNA. Typically, the recognized palindromes are 4 to 6 bases long. Some enzymes cleave the two strands of DNA at staggered positions producing complimentary single stranded fragments sometimes called sticky ends; the alternative would be blunt ends. The use of sticky ends generally permits higher cloning success rate since the digested molecules can base pair with other fragments obtained with the same restriction enzyme.

Double restriction digest experiments were elected in order to allow directional cloning; moreover, if possible, only enzymes producing sticky ends were used. This strategy usually maximizes the chances of cloning success. For each reaction, it was necessary to check which buffer had to be used and if sequential digestion reactions were needed. The experimental procedures were performed as recommended by New England Biolabs® guidelines. In the plasmid digestion reaction, calf intestinal alkaline phosphatase (CIP) was added. The removal of the 5' phosphate group from the plasmid molecules prevents self-ligation of the plasmid during the ligation experiment.

The PCR product and the desired plasmids extracted from the agarose gel were submitted to restriction digestion using the appropriate restriction enzymes. Typical digestion reactions were as follows:

- 17 µL of PCR product
- 2 µL of the appropriate 10X buffer
- 0.5 µL of each restriction enzyme
or

- 16.5 µL of plasmid
- 2 µL of the appropriate 10X buffer
- 0.5 µL of each restriction enzyme
- 0.5 µL of CIP

All the components were mixed in 1.5 mL centrifuge tubes and incubated for 30 to 60 min at 37 °C. The reaction components were separated by gel electrophoresis in a 1% agarose gel. The desired products were cut out of the gel, extracted and stored at -20 °C until further use.

2.2.2. Ligation

During a ligation reaction, an enzyme called DNA ligase is used to catalyze the formation of a phosphodiester bond between the 3'- hydroxyl end of one oligonucleotide and the 5'-phosphate end of another. When using restriction enzymes producing sticky ends, the complimentary single stranded overhangs from the digested PCR product and plasmid can base pair. This increases the likelihood that the ligation reaction is successful. The ratios of plasmid to PCR product used were estimated from the agarose gels. All the ligation reactions were performed using the Rapid DNA Ligation Kit from Roche Applied Science.

A typical reaction combined 4 µL of PCR product and plasmid, 0.5 µL of T4 DNA ligase with a buffer system supplied in the Rapid DNA Ligation Kit. The reactions
were carried out at room temperature for 5-15 min. The reaction mixtures were then used to transform bacteria or stored at -20 °C until needed.

2.3. Gel electrophoresis

Gel electrophoresis achieves the separation of nucleic acids according to their size. Electrophoresis is typically carried out in a gel like matrix such as agarose, a loose mesh of carbohydrate polymers. It relies on the property of a charged molecule to move in an electric field. The velocity of the charged molecules correlates with the overall charge density, the size and the shape of the molecule. In the case of nucleic acids, the charge density and the shape of the molecules are considered constant for homogeneous samples. Therefore only the size of the molecules dictates their migration velocity in the matrix: the smaller molecules move faster, travelling farther. By using a mixture of DNA molecules of known size or a DNA ladder as a reference, it is possible to evaluate the size of the samples run on a gel. Typically 1% agarose gels were used, prepared with of 0.6 g of SeaKem® LE agarose purchased from Lonza and 60 mL of 1 X Tris/borate/ethylene diamine tetraacetate (TBE) buffer and 0.5 µL of the DNA gel stain, SYBR® Safe dye (10,000X) purchased from Invitrogen™. The dye fluoresces when it intercalates into DNA; therefore if it is added to the gel mixture, any DNA molecule larger than 50 base pairs will appear on the gel. The 1% agarose gels are usually run for 35 min at 135 V in a BioRad MiniSub® Cell GT apparatus. The products were visualized on a Safe Imager blue-light transilluminator using an amber filter unit or viewing glasses. The bands containing the desired products were cut out and the DNA was extracted using a
QIAquick® gel extraction kit from QIAGEN. The DNA samples were stored at -20 ºC until needed.

2.4. Transformation

Once the ligation experiment has been performed, larger quantities of the cloned vector are obtained by transformation into competent bacterial cells. Competency of the cells means that the bacterial cells have been modified to maximize their efficiency in taking up relaxed circular DNA. The efficiency of the transformation process is limited; it is estimated to be about 0.1%. However, one transformed cell can multiply and form a colony with millions of bacterial cells containing the desired vector. The only requirement for the plasmid used is to include an origin of replication, allowing replication independent from the bacterial chromosome. In addition the presence of an antibiotic resistance gene in the plasmid allows the selection of colonies that have been successfully transformed. If the bacterial host is sensitive to a particular antibiotic, only the colonies formed from transformed cells that have acquired the resistance marker will grow. The use of *Escherichia coli* competent cells such as XL-10 from Stratagene, DH5-α from New England Biolabs® Inc. or One shot® TOP10 from Invitrogen™ that possess high transformation efficiency, are recombination and endonuclease deficient and lack restriction endonucleases and T7 RNA polymerase ensures high yield of plasmid DNA.

The typical transformation experiment was carried out with 1 µL of cloning/expression vector DNA and 10 µL of competent cells. Once the DNA was added to the cells, the reaction was incubated on ice for 30 min. To promote the uptake of DNA by the cells, a heat shock step of 30 to 45 s at 42 ºC was performed followed by
incubation on ice for 2 min. Finally 250 µL of LB media was added to the reaction which was then incubated at 37 °C for one hour. The bacterial cells were then plated onto a semi solid growth medium made of LB medium, agar and 0.1 mM of the desired antibiotic for selection purposes. The plates were incubated at 37 °C overnight. When bacterial colonies appeared, these were used to inoculate liquid cultures prepared with LB medium and antibiotic. Those cultures were incubated at 37 °C for 24 h. The plasmid DNA was then isolated by standard mini-prep procedure using the QIAprep® Spin Miniprep kit from QIAGEN.

2.5. DNA Sequencing

The sequence of the gene of interest in the cloning/expression vector was confirmed by DNA sequencing. The samples were sent to the University of Michigan DNA Sequencing Core, Ann Arbor, MI. The method used is an automated version of the chain termination method. In this method, *E. coli* DNA Polymerase I is used to make complementary copies of the DNA being sequenced. Since the enzyme adds the nucleotides in the 5’ to 3’ direction, primers are needed to give the initial 3’ end. In the cases of the pET vectors, the T7 promoter and T7 terminator primers are commonly used. For the sequencing reaction, the DNA molecule to be sequenced is mixed with the four deoxynucleotides substrates, the primer, the DNA polymerase I and a small amount of fluorescent 2’,3’-dideoxyribonucleotides. Since those dideoxynucleotides are lacking the 3’-hydroxyl group required to continue DNA elongation, they terminate the polynucleotide chain, hence the name of the method. By using four different fluorescent dyes representing the four bases in the same experiment combined with an electrophoresis step, the exact sequence of the DNA sample is determined. The U. of M.
DNA sequencing core uses capillary electrophoresis and a UV laser to determine the sequences. The results are given as a fluorescence-based chromatogram representing the nucleotide sequence. Analysis of those results is carried out using the ExPASy website.\textsuperscript{31} The nucleotide sequence is converted into an amino acid sequence using the Translate tool.\textsuperscript{32} In addition to the primary sequence of the protein, the presence of typical sequences such as a poly histidine tag or protease-sensitive sites help the selection of the reading frame. Once the appropriate sequence is identified, the amino acid sequence is subjected to a BLAST search,\textsuperscript{33} an algorithm comparing the given sequence with a protein sequence database. The Blast results are used to align the translated sequence from the DNA sequencing reaction with the desired protein sequence. Any difference can then be visually determined, thereby confirming if the desired protein is encoded by the sequenced gene.

3. Protein purification

After the successful cloning of the gene of interest, the overexpression of the chimeric plasmid in a bacterial host produces large quantities of the encoded protein. The protein is then purified by a fractionation process based on its physicochemical properties such as solubility, ionic charge and binding specificity. This section describes the different steps required to produce the recombinant protein and purify it.

3.1. Protein expression

The first step in the production of the protein of interest is the overexpression of the protein. This requires the transcription of the plasmid DNA into RNA and then the translation of the messenger RNA into a polypeptide chain.
3.1.1. Transformation

The transformation step presented here is very similar to the transformation step presented in the molecular biology section of this chapter. The procedure is exactly the same; the only change is the type of bacterial competent cells used. The cloning/expression vector possesses a promoter site and encodes a ribosome binding site both required for transcription and translation respectively, therefore as long as the host cell contains the machinery required for protein synthesis the protein can be synthesized. Typically bacterial strains such as T7 express from New England Biolabs® Inc. or BL21 (DE3) from EMD Biosciences are preferred expression hosts since they are endonuclease and protease deficient and posses a bacteriophage T7 RNA polymerase gene on their bacterial chromosome that binds specifically to the T7 promoter just upstream of the gene of interest.

The typical transformation experiment was carried out with 1 µL of cloning/expression vector and 10 µL of competent cells. Once the DNA was added to the cells, the reaction was incubated on ice for 30 min. To promote the uptake of DNA by the cells, a heat shock step of 30 to 45 s at 42 ºC was performed followed by incubation on ice for 2 min. Finally 250 µL of LB media was added to the reaction which was then incubated at 37 ºC for one hour. The bacterial cells were then plated onto a semi solid growth medium made of LB medium, agar and 0.1 mM of the desired antibiotic for selection purposes. The plates were incubated at 37 ºC overnight. When bacterial colonies appeared, these were used to inoculate liquid cultures prepared with LB medium and antibiotic. The liquid cultures were grown to an optical density of 0.6 at 600 nm and
glycerol stocks of the cells were prepared by adding 20 % v/v of glycerol to the culture
tube. The stocks were flash frozen in liquid nitrogen and stored at – 80 ºC.

3.1.2. Inducing protein expression through the lac operon

In the expression vector, the gene of interest is located in the multiple cloning
sites region between the T7 promoter and the T7 terminator sequences. The bacterial
strains used to synthesize protein posses the gene encoding T7 RNA polymerase in their
bacterial chromosome. T7 RNA polymerase is faster than E. coli RNA polymerase
allowing overexpression of the gene of interest more rapidly. Also by supplying another
RNA polymerase; the synthesis of the protein of interest does not take away from the
maintenance of the E. coli proteome. Moreover the gene of interest is placed after the lac
operon and its controlling elements. The expression of the enzyme is therefore inducible.
The E. coli lac operon is made of three consecutive genes Z, Y and A that encode
proteins required to metabolize lactose. In the absence of lactose, a protein called lac
repressor binds to a specific DNA sequence located near the promoter of the lac operon,
the lac operator. This inhibits the transcription of any genes following the lac operon.
When lactose is present in the growth medium, allolactose, a lactose metabolite,
specifically binds the lac repressor causing it to change shape. The repressor can no
longer bind to the operator sequence, allowing transcription of the lac operon genes. In
the case of recombinant protein expression, isopropyl-β-D-thiogalactoside (IPTG), a non-
metabolizable analog of allolactose, is used as an inducer.

Typically, transformed cells were grown in culture medium at 37 ºC to a high
concentration, estimated by optical density (OD) measured on a benchtop
spectrophotometer. For LB medium, the optimal OD is 0.6 at 600 nm; for the richer
Terrific broth (TB) medium, the optimal OD is 1.2 at 600 nm. Once dense enough, synthesis of the protein of interest was induced by addition of IPTG to a final concentration of 1 mM. The culture was then incubated at 16 ºC for 16 to 20 h.

3.2. Harvesting the cells

After the induction at 16 ºC, the cells were harvested. There are two alternatives: centrifugation or tangential flow filtration (TFF). Centrifugation uses the centrifugal force to separate components of a mixture whereas TFF uses a membrane to separate components in a liquid suspension based on size or molecular weight. Centrifugation was carried out in a 5810-R Eppendorf centrifuge with swing-bucket rotors. The harvest of the bacterial cells by centrifugation was performed at a speed of 4,000 rpm at 4 ºC. TFF was performed on a Millipore Pellicon system, whose molecular weight cut-off is 30,000 kDa. The concentrated suspension of cells was then submitted to a centrifugation step at 4,000 rpm at 4 ºC. The pelleted cells were then resuspended in the desired buffer and stored at –80 ºC until purification was initiated.

3.3. Protein purification

The purification protocol is a stepwise fractionation process. The protein of interest is synthesized inside the bacterial cell. It is therefore necessary to start the preparation by lysis of the cells. The crude lysate is then submitted to a series of chromatography steps using the different physicochemical properties of the protein.

3.3.1. Cell Lysis

The frozen aliquot of resuspended cells was thawed at room temperature and then placed on ice. The catalytic hydrolysis of undesired components was initiated by addition
of 1 mM of lysozyme and 0.1 mM of DNase I, both purchased from Sigma-Aldrich, Inc. Lysozyme catalyzes the hydrolysis of the bacterial cell wall. The enzyme specifically degrades the beta (1-4) linkages between N-acetylmuramic and N-acetyl-D-glucosamine in the peptidoglycan and between N-acetylglucosamine in chitodextrins. The catalysis step was performed for 20 minutes on ice prior to lysis of the cells using the Misonix Sonicator® 3000. The crude lysate was centrifuged at 11,000 rpm at 4 ºC to remove the cellular debris using a fixed-angle rotor. The pellet component was discarded and the supernatant liquid was decanted and filtered through an immobilized 0.22 µm Millipore Millex®-GP membrane. The filtrate was then submitted to chromatography steps.

3.3.2. Chromatography

Chromatography techniques are used to separate mixtures; they involve the interaction of a mobile and a stationary phase. Typically the mixture of substances to be separated is dissolved into a liquid, the mobile phase, and is percolated through a column containing a porous matrix, the stationary phase. When substances interact with the stationary phase, their progression is retarded. The type of interactions between the components of the mixture and the stationary phase is dictated by the type of chromatography used.

3.3.2.1. Metal affinity chromatography

The protein of interest was first crudely purified by immobilized metal affinity chromatography. The expression constructs used in this work were designed to harbor a metal-binding tag such as a poly histidine tag. The poly histidine tag is made of six consecutive histidine residues placed at either the N- or the C-terminus of the
recombinant protein. The tag has a strong affinity for metal cations such as Co$^{2+}$ and Ni$^{2+}$. The metal used in the laboratory is Ni$^{2+}$. The metal binds to the carbohydrate-based matrix of the chromatography column by simple metal-ligand affinity. When the sample to be purified is loaded onto the metal-affinity column, only the protein bearing the metal-chelating group is retained. A buffer containing imidazole is then used to elute the protein off the column. Bearing a very similar structure, imidazole competes for metal binding with histidine. At a given concentration of imidazole, typically near 125 mM, the 6xHis tag no longer displays the higher affinity for the nickel and the protein elutes from the column. The histidine tag can be removed by the action of a specific protease. The recognition sequence for such an enzyme is typically engineered between the sequence coding for the poly histidine tag and the gene of interest. The engineered sequence in the plasmids used in this study is LEVLFGP, which is specifically recognized by PreScission™. After action of the protease the solution is run over the metal-affinity column once more. This time only the cut tag and the tagged protease stick to the column while the protein of interest, no longer bearing a tag, goes through the column.

The typical experiment was run on a HisTrap™ chelating HP column purchased from GE Healthcare. The column was loaded with Ni$^{2+}$ (50 mM NiSO$_4$) and used on an Äkta®FPLC™ fast protein liquid chromatography system commercialized by Amersham Biosciences. The filtered cell lysate was loaded onto the column, which was then washed with binding buffer. A gradient of increasing concentration of elution buffer, a modified binding buffer in which 0.5 M of imidazole has been added, was then run through the column. The purity of the fractions containing the protein of interest was
visualized using gel electrophoresis *vide infra*. The fractions containing the protein of interest were pooled and were dialyzed, if needed, in the presence of PreScission™ protease at 4 °C overnight to remove the poly histidine tag. Another nickel column was run. The flow-through fractions containing the protein of interest were pooled and the solution was dialyzed at 4 °C against the buffer required for the next purification step.

3.3.2.2. Ion exchange chromatography

For some of the proteins purified in the course of this work, it was necessary to add an ion exchange chromatography step. This technique uses the ionic charge of the protein as a separation criterion. Depending on the ionic strength of the buffer and its pH value, the protein displays a given charge. For the ion exchange chromatography experiment, the charged protein molecules are attracted to oppositely charged groups that are chemically bound to a matrix. For instance, anions bind to cationic group of an anion exchanger. The binding affinity of a particular protein depends on the presence of other ions that compete with the protein for binding to the ion exchange material. A buffer with high ionic strength is typically used to elute the protein from the column.

Typically, a solution of protein dialyzed into crystallization buffer was loaded on a HiTrap™ Q FF column, purchased from GE Healthcare. The column was washed with crystallization buffer. A gradient of elution buffer, a modified crystallization buffer with 1 M NaCl, was subsequently run through the column. The fractions containing the protein of interest were identified using gel electrophoresis, then pooled and dialyzed against crystallization buffer. In preparation for the last purification step, the protein solution was concentrated to a smaller volume by either a salting out procedure or the use of Millipore centrifugal filter devices.
3.3.2.3. Size exclusion chromatography also called gel filtration

This technique separates molecules according to their size and shape. The stationary phase is made of gel beads containing pores. Because of their size, large molecules flow through the column more rapidly. The typical gel filtration experiment was run on a Superdex™ 200 column mounted on the Amersham Biosciences AktaPrime™ system. The concentrated protein solution was injected onto the system and the crystallization buffer was run isocratically. The fractions containing the purified protein of interest were identified by gel electrophoresis and pooled. The protein solution was then concentrated to 5-10 mg/mL and stored at – 80 °C.

3.3.3. Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Gel electrophoresis separates molecules according to the overall charge density, the size and the shape of the molecule. In presence of the detergent sodium dodecyl sulfate (SDS) the proteins assume a rod-like shape and the large negative charge that the SDS imparts masks the protein intrinsic charge. Therefore SDS-PAGE separates molecules according to their mass only. The relative motilities of protein vary approximately linearly with the log of their molecular weight. Visualization of the proteins is achieved by soaking the gel in a solution of stain that binds the protein such as Coomassie blue.

3.3.4. Salting out as a concentration method

Because a protein contains multiple charged groups, its solubility depends on the concentration of dissolved salts in solution, the polarity of the solvent, the pH and the
temperature. The salting out procedure uses the competition of added salt ions and the
the protein for the molecules of solvent. At very high salt concentration, so many of the
added ions are solvated that there is less bulk solvent available to dissolve the protein
which precipitates. This is an efficient and rapid concentration technique. A typical
ammonium sulfate precipitation experiment used 2.4 M of the salt at 4 °C. Once all the
ammonium sulfate was in solution, the obtained suspension was centrifuged at 10,000
rpm at 4 °C for 30 min. The supernatant was discarded and the pelleted, precipitated
protein was dissolved in a minimal amount of crystallization buffer. The solution was
then dialyzed against crystallization buffer at 4 °C to remove any excess (NH₄)₂SO₄.

3.3.5. Dialysis

Dialysis is mainly used to rapidly change the components of a protein solution.
This process uses the difference in rates of diffusion of molecules through a semi-
permeable membrane. Small salt molecules are able to go through the membrane
whereas the protein molecules cannot fit through the pores. Therefore equilibrium is
established between the protein solution placed in the sealed dialysis tubing and the
buffer used for the dialysis. In the laboratory, the typical molecular weight cut off is
10,000 Da.

4. Spectrophotometric applications

Spectrophotometry has different uses in the protein purification process. It is
used for evaluation of cell culture density, detection of protein during the
chromatography steps and calculation of the purified protein concentration. It is also an
important component of the assay described in Chapter 3. All those applications follow
the Beer-Lambert law. All the protein purification related experiments were run on a GENESYS™ 6 UV-Vis spectrophotometer by Thermo Electron Corporation. The assay experiments were run on an absorbance plate reader Spectramax® 340PC by Molecular Devices.

4.1. Introduction to the Beer-Lambert law

The Beer-Lambert Law is an empirical relationship that relates the absorption of light to the properties of the medium through which the light is travelling. It states that there is a logarithmic dependence between the transmission of light through a substance ($T$), the product of the absorption coefficient of the substance ($\alpha$) and the distance the light travels through the material ($l$), also called path length. The equation is given in Equation 1, where $I_0$ is the intensity of the incident light and $I_1$ is the intensity of the light after it went through the material.

$$T = \frac{I_1}{I_0} = 10^{-\alpha l}$$  

Equation 1

The absorption coefficient can also be described as the product of the molar absorptivity of a substance ($\varepsilon$) and its concentration as shown in Equation 2.

$$\alpha = \varepsilon C$$  

Equation 2

Instead of transmission, the Beer-Lambert law is often express in absorbance ($A$) giving Equation 3.

$$A = -\log \left( \frac{I_1}{I_0} \right) = \varepsilon l C$$  

Equation 3
4.2. Monitoring the protein purification process using spectrophotometry

The chromatography steps of the protein purification process are carried out on a mostly automated platform. As part of the process, the FPLC systems measure the absorbance of the solution running through the system at 280 nm and display the corresponding chromatogram. A wavelength of 280 nm was chosen because aromatic moieties in tyrosine and tryptophan residues strongly absorb light at that wavelength. The chromatogram gives an indication of where proteins can be found in the different fractions obtained from the FPLC run.

4.3. Calculation of protein concentration

The Beer-Lambert law states that the relationship between absorbance and concentration is linear. Since the molar absorptivity coefficient of a protein can be easily calculated from its primary sequence,\textsuperscript{31} an absorbance reading performed at 280 nm allows the calculation of protein concentration.\textsuperscript{34}

4.4. Purity determination of protein samples

In addition to the 280 nm absorbance reading, a reading at 260 nm can be used to determine the purity of the protein.\textsuperscript{35} Since nucleic acids strongly absorb light at 260 nm, while protein strongly absorbs at 280 nm, the calculation of the 280/260 ratio indicates the purity of the sample. Typically, a 280/260 ratio above 1.8 corresponds to a protein sample free of major nucleic acid contamination.
5. Macromolecular crystallography

5.1. Introduction to macromolecular crystallography

The determination of the three dimensional structure of a protein is usually performed by single crystal crystallography using X-ray diffraction. This method uses the ability of matter to interact with electromagnetic radiation. In order for an object to diffract electromagnetic radiation, the wavelength of the radiation must be no larger than the object. In the case of atoms and bonds, the wavelength of the radiation must be in the X-ray range. Diffraction is only produced by highly ordered objects; therefore the molecules of protein need to form an orderly three dimensional array to diffract X-rays. This type of array of molecules is called the crystalline state. Crystals of proteins grow by slow and controlled precipitation from aqueous solution of pure protein.

If a high quality crystal is obtained and a diffraction pattern is observed upon exposure of the crystal to X-rays, a data set is collected. The information collected includes the directions and the intensities of the X-ray beams diffracted by the crystal. Computer software is then used to get an image of the content of the crystal. This image is interpreted by computer graphics and a molecular model is built.

5.2. Crystallization methods

Several types of crystallization experiments are available, but only vapor diffusion experiments were carried out.

5.2.1. Hanging drop vapor diffusion crystallization method

In this method, a solution of purified protein is mixed with a reservoir solution whose precipitant concentration is optimal for producing crystals. The mixture is then
suspended as a droplet underneath a cover slip which is sealed onto the top of the reservoir. Since the system is closed, the difference of precipitant concentration between the droplet and the reservoir requires equilibration through the transfer of water molecules. During this process, the precipitant concentration that causes crystallization can be reached slowly, allowing the slow formation of crystals.

5.2.2. Microbatch crystallization method

The microbatch method is very similar to the hanging drop method. In this method, the droplet of protein solution and reservoir solution is placed under a drop of oil. The equilibration process with the atmosphere occurs slowly by evaporation through the layer of oil.

5.3. Experimental method

After purification, the proteins were dialyzed onto crystallization buffer. Crystallization trays were then set up using commercial screening kits from Nextal and Hampton research. The crystallization method was chosen according to the type of screen and the amount of protein available. The crystal trays were placed in an incubator at 18 ºC and were observed periodically to check for crystal growth. When hits were identified, the crystallization conditions were expanded to achieve high quality crystals. Once diffraction quality crystals were obtained, they were screened for diffraction on the home diffractometer. It is a Rigaku FRE High Brilliance X-Ray Generator with R-AXIS IV and Saturn 92 CCD detectors. It uses a monochromatic source that produces single-wavelength X-rays at 0.154 nm.
Chapter Three

Development of a High-Throughput Assay to Test the Activity of Antigen 85

1. Introduction to the drug discovery process

The resurgence of TB as a fatal disease emphasizes the urgent need of new antitubercular treatments; however the discovery of new drugs is a lengthy process requiring numerous steps. From the initial idea to the final drug, the process can take up to fifteen years. One of the possible routes for drug discovery is the screening of libraries of natural and synthetic chemicals for activity against the etiological agent of a disease. A simplified version of the process is presented in Figure 7. Typically, the drug discovery process begins with the identification of a new target which can be an oligonucleotide or a polypeptide. The subsequent step is target validation where initial experiments confirm the target involvement in the disease and that interference with its function causes a therapeutic effect. Once the target has been validated, an assay is developed to allow screening of molecular libraries of molecules in order to characterize their effect on the target. If during the primary screening any hits appear, they are confirmed and the molecules showing promising results are promoted to lead status. The lead compounds are then optimized and developed into a drug candidate or new chemical
entity (NCE). After pre-clinical development the NCE is submitted to clinical trials and if approved is finally released to the market.

![Drug Discovery Process Diagram]

**Figure 7** Drug Discovery Process

It is often difficult to identify good drug targets, but it is sometimes even more difficult to find an appropriate method to assay the activity of the target. Indeed, in order to be successful, this type of approach requires the use of methods that allow screening in a high-throughput format.
With the emergence of drug-resistant strains of *M. tb*, there is a growing interest in the study of the mechanisms of resistance against the available drugs. In addition to molecular mechanisms, *M. tb* antibiotic resistance also arises from the reduced permeability of its thick hydrophobic cell envelope. Therefore molecules involved in the maintenance of the bacterial cell wall are of high importance to the mycobacterium. The development of antitubercular drugs such as INH, EMB, ETH and thiacetazone (TAC) targeting components of the cell wall illustrates this theory. INH, ETH and TAC inhibit the synthesis of mycolic acids while EMB inhibits the synthesis of the arabinogalactan layer of the cell wall. Recently, the characterization of more components of the cell envelope and the identification of enzymes implicated in their synthesis provided more possible drug targets. Among the known enzymes involved in the maintenance of the mycobacterial cell wall, the antigen 85 complex (Ag85) stands out as an attractive drug target. The existing assay testing Ag85 activity does not meet the high-throughput requirements of practicality, cost and adaptability to automation; therefore a new assay is needed. After a brief introduction of Ag85, this chapter presents the development of a high-throughput assay to test the activity of a well-characterized enzyme from *M. tb*, antigen 85C.

2. The antigen 85 Complex

2.1. Introduction to the antigen 85 complex

The different components of the antigen 85 complex (Ag85) have been given different names over the years such as α-antigen or Fibronectin-binding proteins (Fbps). This family of proteins was first isolated as the major secreted component of *M. tb* culture filtrates. Three proteins compose Ag85: antigen 85A (Ag85A), antigen 85B
(Ag85B) and antigen 85C (Ag85C). The genes encoding Ag85A, Ag85B and Ag85C are called fbpA, fbpB, and fbpC2, respectively. After their isolation from culture filtrates, further characterization demonstrated the presence of these proteins in the cell wall and their ability to bind fibronectin. In addition the proteins are not only secreted in laboratory culture but also in their natural host, the macrophage. These observations prompted the hypothesis of the involvement of Ag85 in pathogenesis of the disease and the focus of Ag85 studies on its immunological properties. Since M. tb is a slow growing bacterium, isolation of the proteins from culture media is not a convenient method to obtain large quantities of Ag85. Therefore cloning, sequence determination and recombinant expression of Ag85A, B and C were investigated. The amino acid sequences revealed the presence of a 40 amino-acid signal peptide, involved in the secretion process. The proteins bear a high sequence identity (68-79%) in their mature secreted forms but, surprisingly, the genes encoding Ag85 are in separate loci of the M. tb genome. It is believed that the gene of the original antigen 85 protein was duplicated in the mycobacterial genome. Consequently, the three proteins are considered paralogs. In addition to being the immunodominant antigen of M. tb, the proteins display enzymatic activity. Their roles as mycolyltransferases was revealed in 1997 by Belisle et al. In the in vitro assay, described in detail in a later section, the three paralogs purified from M. tb H37Ra culture filtrate transferred mycolic acid moieties from trehalose monomycolate (TMM) onto another TMM molecule, thereby forming trehalose dimycolate (TDM) as shown in Figure 9. Ag85C displayed the highest activity while Ag85A showed a level of activity comparable to Ag85C. Ag85B had a reduced activity, only 20% of that of Ag85C. The mycolate transfer reaction suggested that the enzymes
possess a carboxylesterase domain with a characteristic Gly-X-Ser-X-Gly consensus sequence and a Ser, Asp/Glu, His catalytic triad in the active site. Sequence alignments with known proteins identified Ag85A, B and C as members of the serine hydrolase family. The unraveling of Ag85C crystal structure identified its catalytic triad as Ser 124, His 260 and Glu 228, and a reaction mechanism was proposed. Ag85 enzymes operate through a ping pong mechanism, or double displacement mechanism presented in Figure 8. In general, the enzyme (E) binds the substrate (A) to form the enzyme-substrate complex (EA) that reacts to yield the activated enzyme intermediate (E*P) (step 1, Figure 8). The first product (P) is then released freeing the activated enzyme (E*) that is therefore available to bind the second substrate (B) and form the activated enzyme-substrate complex (E*B). In the final step (Step 2, Figure 8), the activated enzyme-substrate complex reacts to form the enzyme-product intermediate (EQ). Finally, the product (Q) is released and the enzyme is regenerated (E), ready for the next enzymatic reaction.

![Figure 8 General ping pong mechanism](image)

The mechanism using TMM as a substrate is detailed in Figure 9. The nucleophilic Ser124 of Ag85C attacks the carbonyl ester of the TMM substrate (A) to
form a mycolated enzyme \((E^*P)\) (Step 1, Figure 9), also called the acyl-enzyme intermediate. Another molecule of TMM \((B)\) binds and promotes the release of the enzyme intermediate \((E^*)\) by acting as a nucleophile to attack the ester linkage of the mycolic acid. The final product TDM \((Q)\) is released and Ag85C is regenerated. This mechanism was devised for Ag85C but the high sequence homology and the conservation of the active site catalytic residues in the three paralogs suggested that the mechanism of action would be the same for all enzymes. Crystal structures of Ag85B and Ag85A strengthen the proposed mechanism.\(^{48,49}\)

These enzymes are monomeric proteins displaying an \(\alpha/\beta\) hydrolase fold with a central \(\beta\)-sheet bordered by \(\alpha\)-helices.\(^{47}\) This fold is characterized by an \(\alpha/\beta\)-sheet core of five to eight \(\beta\)-strands connected by \(\alpha\)-helices with the catalytic triad residues located on the loops between the helices and the strands.\(^{50}\) This fold is widespread with forty-six families of enzymes including acetylcholinesterases (AChE), lipases, cutinases, thioesterases (TE) and carboxyesterases. In fact, this fold is so extensive, that a database was created to gather information on this superfamily of enzymes.\(^{51}\)

The active site of Ag85 is located in a cleft at the C-terminal ends of the \(\beta\)-strands composing the \(\beta\)-sheet. The substrate binding pocket includes two chambers: a carbohydrate binding pocket and a fatty acid binding pocket that can easily accommodate TMM. The crystal structures also revealed the presence of a secondary carbohydrate binding pocket in addition to the active site.
Figure 9  Mechanism of Ag85C
2.2. The antigen 85 complex is a validated drug target

Evidence regarding the importance of the mycolic acid component of the mycobacterial cell wall has been presented in the first chapter. When disclosing the mycolyltransferase assay, Belisle et al. also reported that the addition of 6-azido-6-deoxy-\(\alpha,\alpha'\)-trehalose (ADT) to the reaction inhibited the activity of Ag85C by 60\%\(^4\). The addition of ADT to culture media of *Mycobacterium aurum*, a screening surrogate for *M. tb*, caused a complete inhibition of mycobacterial growth. Moreover at lesser ADT concentrations, the composition of the cell wall changed in treated cultures, showing 44\%, 87\% and 62\% inhibition of TMM, TDM and cell wall-bound mycolic acids synthesis, respectively. The inactivation of the *fbpC2* gene as reported by Jackson et al. resulted in 40\% less cell wall bound-mycolates but no change in the quantity of non-covalently bound lipids like TMM and TDM.\(^5\) The decrease in bound mycolates greatly increased the permeability of the cell wall for hydrophobic compounds such as chenodeoxycholate and small hydrophilic compounds such as glycerol, however the ability of INH to go through the cell envelope was not affected. Another observation was the lack of effect on the growth of the mycobacteria. Overall, this study demonstrated that the *in vitro* mycolyltransferase activity could also be observed *in vivo* and that, since Ag85A and B did not compensate for the inactivation of *fbpC2*, the three enzymes of Ag85 were not redundant. A related important conclusion was that, for more efficiency, a drug would need to target at least two enzymes of the complex.

Similar studies performed on Ag85A and B surprisingly gave different conclusions. The *fbpA* and *fbpB* inactivated mutants did not show any difference in cell envelope mycolate composition when compared to the wild type strain.\(^5\) However, in
another study disruption of fbpA resulted in poor growth in macrophage-like cell line, ultimately leading to the death of the mycobacteria while disruption of fbpB did not have any effect.\textsuperscript{54} This particular observation suggested a partial redundancy of the enzymes for their substrates and confirmed the need of targeting more than one enzyme. Besides the gene disruption experiments, RNA interference work by Harth and co-workers showed a significant inhibition of \textit{M. tb} growth in culture when the three gene transcripts are targeted by antisense modified oligodeoxyribonucleotides.\textsuperscript{55} Even if no clear evidence of the bactericidal effect of inactivation of Ag85 was reported, the simple hypothesis of a bacteriostatic activity represented a great achievement. Combination of Ag85-targetting compounds and existing anti-bacterial drug could produce a bactericidal effect that improved upon that seen for INH alone.

In addition to the evidence presented so far, the extracellular localization of Ag85 makes it a pertinent drug target. The usual bacterial artillery against antibiotic, modifying enzymes, efflux pumps and target modification, involves intracellular molecules and processes. By targeting secreted proteins, the development of resistance becomes less likely. Moreover, if the drug targets all three enzymes from the complex, the odds of mutations of all three genes occurring at the same time are infinitesimal. Finally, since the mycolyltransferases are only found in mycobacteria and they act upon substrate not found in humans, a drug targeting those enzymes should not exhibit significant side effects in humans.

In summary, Ag85 possesses numerous characteristics that make it an excellent drug target for tuberculosis therapies. With this in mind, we proceeded to develop new
ideas and tools to promote the drug discovery process and a way to test potential inhibitors.

2.3. Existing assay

The discovery of Ag85 inhibitors necessitates a rapid and easy assessment of the activity of the enzymes in the presence of inhibitors. In 1997, Belisle et al. reported a mycolyltransferase assay. The assay involves the transfer of non-radioactive mycolic acids from a lipid-soluble trehalose monomycolate molecule to a radioactive water-soluble trehalose molecule as presented in Figure 10. Subsequent extraction of the reaction mixture allows the assessment of lipid-soluble radioactive trehalose monomycolate and trehalose dimycolate formation. This protocol assesses the activity of the enzyme in a qualitative fashion.

Figure 10  Mycolyltransferase assay developed by Belisle and coworkers
Although the radiometric assay works well, it requires the use of radioactively-labeled molecules and is relatively slow since it necessitates an extraction step to visualize the results. High-throughput screenings of large numbers of potential inhibitors is not feasible using this assay, therefore another assay was needed.

3. Development of the high-throughput assay

To be considered an improvement from the existing one, the novel assay required ease of manipulation and immediate visualization of the results. A spectrophotometric method seemed the best candidate to meet those requirements. The chosen spectrophotometric format is versatile enough to allow colorimetric or fluorescence-based experiments. Previous results from Belisle and coworkers showed that Ag85C is the most active of the enzymes in vitro, thus the development phase of the assay was carried out with Ag85C.

3.1. Design

Structural and mechanistic considerations guided the design of the assay. Since Ag85 catalyzes the transfer of mycolate to trehalose, the assay called for a carbohydrate-based substrate. TMM, the postulated acyl donor of the mycolyl transfer reaction, was chosen as the template for the substrate design. The active site of the enzyme accommodates disaccharide molecules such as trehalose, as well as glycoconjugates containing alkyl moieties such as octylthioglucoside. Thus a glucose-derivative with an acyl chain should bear sufficient structural similarities with the natural substrate to be recognized by the enzyme. At this point, the spectrophotometric detection still required the addition of a chromophoric moiety in the substrate. Because of the esterase activity
of Ag85C, it did not seem possible to obtain the release of the chromophore by the enzyme. In fact, coupling the activity of Ag85C with another enzyme seemed the best possible approach.\textsuperscript{57} In the carbohydrate field, glycosidases and \(p\)-nitrophenylated compounds are extensively used,\textsuperscript{58,59} therefore a glucose-based derivative harboring a \(p\)-nitrophenyl moiety was chosen. Measuring the \(p\)-nitrophenolate release is ideal for a spectrophotometric experiment since the anion absorbs light at a wavelength of 405 nm.\textsuperscript{60} In summary, the substrate for the assay, shown in Figure 11 is a glucose derivative esterified at the six position with an eight-carbon acyl chain and a \(p\)-nitrophenyl moiety attached by a \(\beta\)-linkage onto C1, allowing the use of \(\beta\)-glucosidase.

![Figure 11 Substrate used in the assay](image)

The assay proceeds in two steps presented in Figure 12. Nucleophilic attack of the octanoyl chain of the substrate by Ag85C releases \(p\)-nitrophenylglucoside, which is then hydrolyzed by \(\beta\)-glucosidase into glucose and \(p\)-nitrophenolate. The rate of \(p\)-nitrophenolate release is observed by direct measurement of the absorbance at 405 nm over time. The initial velocity of the reaction is directly calculated from the raw data. To promote release of the octanoyl moiety from the acyl-enzyme intermediate, an excess of D-glucose was added to the reaction to function as an acyl acceptor.
3.2. Synthesis of the substrate

The compound was synthesized by our collaborators in two steps with an overall yield of 39% as shown in Figure 13. Reaction of octanoyl chloride with thiazolidine-2-thione yielded the activated N-octanoylthiazolidine-2-thione (88%). Esterification of p-nitrophenylglucoside with the activated thione gave the desired substrate in a modest 44% yield.\(^\text{61, 62}\)
3.3. Cloning of \( fbpC2 \), expression and purification of Ag85C

The cloning of \( fbpC2 \), the expression and the purification of Ag85C were the subject of several publications however a new protocol was devised following difficulties with the published procedures. An additional chromatography step was added to get higher protein purity.

3.3.1. Cloning of \( fbpC2 \)

The portion of the \( M.\text{tb.} \ fbpC2 \) gene encoding the mature secreted form of Antigen 85C lacking a stop codon was amplified by polymerase chain reaction using the following primers 5’-CTC CTC CTC CTC GAG GGC GGC CGG CGC AGC-3’ and 5’-GGA GAT ATA CAT ATG GCA TTC TCT CGG CCG-3’. It was then placed in pET-29 using NdeI and XhoI restriction enzymes (NEB) in a traditional digestion/ligation procedure. The resulting open reading frame produces a protein with a C-terminal 6X-histidine tag.

3.3.2. Expression and purification of Ag85C

The expression plasmid pET29-Ag85C was transformed into T7 express cells. The bacterial cells were grown at 37 °C in Terrific Broth, induced with IPTG at 16 °C for 16-20 h and harvested by centrifugation. The pelleted cells were resuspended in Ag85 binding buffer (20 mM sodium phosphate pH 7.5, 5 mM \( \beta \)-mercaptoethanol) and stored at -80 °C until protein purification was initiated. The frozen aliquot of resuspended cells was thawed at room temperature and then placed on ice. An enzymatic hydrolysis of undesired components was performed for 20 minutes on ice prior to lysis of the cells using a sonicator. The obtained suspension was clarified by high-speed centrifugation.
The pellet component was discarded and the supernatant was filtered. The protein was first crudely purified by immobilized metal affinity chromatography. A HisTrap FF 5mL column was used on a fast protein liquid chromatography system. The chromatogram is presented in Figure 14. The fractions containing Ag85C were pooled and diluted 5 fold into Ag85 crystallization buffer (10 mM Tris pH 7.5, 2 mM ethylenediaminetetraacetic acid (EDTA) and 1 mM dithiothreitol (DTT)) and subsequently run on an anion exchange chromatography column. The protein was eluted using a modified Ag85 crystallization buffer (10 mM Tris pH 7.5, 1 M sodium chloride, 2 mM EDTA and 1 mM DTT). The chromatogram is shown in Figure 15. The fractions containing Ag85C were pooled and solid ammonium sulfate was added to a concentration of 2.4 M. Only ammonium sulfate precipitation permitted concentration of the protein. The use of centrifugal filtering devices did not work for Ag85C. The precipitated protein was pelleted by centrifugation and resuspended in crystallization buffer. The final purification step uses size exclusion chromatography on a Superdex™ 200 prep grade column equilibrated with crystallization buffer. The chromatogram is presented in Figure 16. The fractions containing Ag85C were pooled and stored at -80 °C until needed for the assay.

The described preparation yielded about 10 mg of pure protein per liter of culture. The concentration of Ag85C was calculated from absorption measurement at 280 nm and the predicted extinction coefficient of the recombinant Ag85C with a value of 84,340 M⁻¹ cm⁻¹. The purity of the protein was assessed by SDS-PAGE and by calculation of the 280/260 absorption ratio. On average, the ratio values ranged from 1.8 to 1.95 indicating pure protein solution containing minimal amounts of nucleic acids.
Figure 14    Metal affinity chromatogram and the corresponding SDS-PAGE
Figure 15  Anion exchange chromatogram and the corresponding SDS-PAGE
Figure 16  
Size exclusion chromatogram and the corresponding SDS-PAGE gel
3.4. Analysis of the data using enzyme kinetics

The chemical kinetics analysis developed to characterize the reaction rate of uncatalyzed reaction can be expanded to enzymatic catalysis. Enzyme kinetics allows the determination of the properties of an enzymatic reaction. In order to study enzymes, a simplified mathematical model was developed. The Michaelis-Menten model requires the fulfillment of certain conditions such as an irreversible reaction and a large excess of the substrate compared to the enzyme (usually a hundred to a thousand fold). The conditions of the assay were designed to meet these requirements. The second reaction ensures the irreversibility of the Ag85C reaction by driving the reaction forward using the product of the first reaction. Also for this analysis, the acyl donor will be considered as the sole substrate of the reaction. The acceptor, D-glucose, is used in large excess so it is not a factor influencing the rate of the reaction and the concentration is fixed.

This section describes the Michaelis-Menten model in a very general manner that is nevertheless applicable to the assay. The simple reaction substrate (S) yields product (P) shown in Equation 4 if under enzymatic catalysis can be rearranged as Equation 5.

\[
S \rightarrow P \quad \text{Equation 4}
\]

\[
E + S \rightarrow P + E \quad \text{Equation 5}
\]

When the requirements of the Michaelis-Menten model are satisfied, the reaction is conceptualized as a first order bimolecular reaction between the enzyme (E) and the substrate to form the enzyme-substrate complex (ES) followed by a unimolecular reaction that yields the product and regenerates the enzyme, as shown in Equation 6.
At low substrate concentration, the bimolecular reaction is considered to be a rapid equilibrium defined by the rate constants $k_1$ and $k_{-1}$. When the substrate concentration is increased to the point where the enzyme has entirely been converted to $ES$, the unimolecular reaction is the rate-determining step of the reaction and is defined as a single catalytic step with the rate constant $k_2$ also called $k_{cat}$. The rate of the reaction is defined as the rate of formation of product as shown in Equation 7.

$$v = \frac{d[P]}{dt} = k_2[ES]$$

Equation 7

Since the concentration of $ES$ is unknown, more assumptions need to be taken into account. For example, in the time interval studied, the concentration of $ES$ is considered constant. This is the steady-state assumption where the overall reaction rate is independent of the substrate concentration and can be described by Equation 8.

$$\frac{d[ES]}{dt} = 0 = k_1[E][S] - (k_{-1} + k_2) [ES]$$

Equation 8

Also the enzyme only exists in two forms: the free enzyme $E$ and the enzyme-substrate complex $ES$, leading to the definition of the total concentration of enzyme presented in Equation 9.

$$[E]_{total} = [E] + [ES]$$

Equation 9

Rearrangement and combination of Equations 8 and 9 yield Equation 10.

$$[ES] = \frac{k_1([E]_{total} - [E])[S]}{k_{-1} + k_2}$$

Equation 10
To simplify this expression, the Michaelis constant ($K_M$) was defined. It is described in Equation 11.

$$K_M = \frac{k_{-1} + k_2}{k_1}$$  \hspace{1cm} \text{Equation 11}

Further rearrangement and introduction of $K_M$ yield Equation 12.

$$[ES] = \frac{[E]_{\text{total}}}{[S] + K_M}$$  \hspace{1cm} \text{Equation 12}


$$v = k_2 \frac{[E]_{\text{total}}}{[S] + K_M}$$  \hspace{1cm} \text{Equation 13}

Finally, the maximal rate of reaction ($V_{\text{max}}$), defined in Equation 14, is achieved when all the molecules of enzyme have substrate bound. At this point it is assumed that total enzyme concentration is approximately equal to the concentration of the enzyme-substrate complex. All this combined yields the Michaelis-Menten equation reported in Equation 15.

$$V_{\text{max}} = k_2 [ES]$$  \hspace{1cm} \text{Equation 14}

$$V_i = \frac{V_{\text{max}} [S]}{K_M + [S]}$$  \hspace{1cm} \text{Equation 15}

The measurement of the initial velocity values at different substrate concentrations allows the calculation of the $V_{\text{max}}$ and $K_M$ values.
3.5. Preliminary experiments

Before running the assay, a few preliminary experiments were necessary to determine the assay parameters. The optimal wavelength of the chromophore as well as its molar extinction coefficient under the assay conditions needed to be resolved first.

3.5.1. Wavelength determination

To determine the optimal wavelength to monitor the release of \( p \)-nitrophenolate during the assay, a wavelength scan was performed with a stock solution of \( p \)-nitrophenol. A 0.2-mM solution of \( p \)-nitrophenolate was prepared using the buffer conditions of the assay (50 mM sodium phosphate [pH 7.4], 3 mM D-glucose, and 2% DMSO). A wavelength scan experiment (340–800 nm) was then performed on the absorbance plate reader (SpectraMax340PC, Molecular Devices). The data were exported from Molecular Devices SoftMaxPro5 software and analyzed in GraphPad Prism 5® software. The complete scan (340–800 nm) is presented in Figure 17 and an expansion of the region of interest is in the inset. The curve reaches a maximum at 395 nm. However to be consistent with current literature, an absorbance of 405 nm was chosen for the assay.
3.5.2. Determination of the molar extinction coefficient of \( p \)-nitrophenolate

Since the value of the molar extinction coefficient varies depending on the assay conditions and the wavelength, it was determined from calibration curve studies with standard \( p \)-nitrophenolate solutions. A range of solutions (0.5–300 \( \mu \)M) was prepared by serial dilution of a stock of 0.1 M \( p \)-nitrophenolate. The plotted values of absorbance over path length versus \( p \)-nitrophenolate concentration yielded a linear plot with a correlation coefficient \((r^2)\) of 0.9979 as shown in Figure 18. The extinction coefficient was calculated from these data to be 15,300 M\(^{-1}\)cm\(^{-1}\).
3.6. Experimental methods

The assay was run under atmospheric pressure at room temperature in a 96-well plate format. A master mix containing 2 units (U) of almond β-glucosidase purchased from Sigma-Aldrich®, 3 mM of D-glucose purchased from Fisher Scientific Inc., and 0.05 M sodium phosphate pH 7.4 was prepared for each well. Typically 2 μL of the substrate in DMSO solution was added to each well to get to the desired substrate concentration followed by addition of Ag85C to a final concentration of 100 nM. The release of the signaling molecule, p-nitrophenolate (extinction coefficient \( \varepsilon = 15,300 \text{ M}^{-1} \text{ cm}^{-1} \) at pH 7.4), was monitored by measurement of the absorbance at 405 nm every 30s for 1 hour in a Molecular Devices absorbance plate reader Spectramax® 340PC. The
data were exported from the Molecular Devices SoftMax® Pro5 software and imported into Microsoft Excel® where the initial velocity numbers (Vi) were calculated. These values were plotted versus the substrate concentration using GraphPad Prism 5®.

4. Enzymology Results

4.1. Positive and negative control experiments for Antigen 85C

4.1.1. Proof-of-principle

The initial experiment confirming that the assay worked is shown in Figure 19. Three reactions were carried out in parallel with 175 µM of substrate and 3 mM of D-glucose. In the first reaction, both enzymes were added (200 nM of Ag85C and 5 units of β-glucosidase). In the second reaction, only β-glucosidase was added and the third one had both. The slope of the absorbance versus time reveals the initial velocity of the reaction.
In the absence of either Ag85C or β-glucosidase, no signal was observed. The addition of both β-glucosidase and 100 nM Ag85C caused a significant absorbance increase over time. The control reaction with only β-glucosidase revealed this enzyme displays a low activity with the synthetic substrate to produce \( p \)-nitrophenolate at a very low background rate. Since the hydrolytic rate of the substrate by β-glucosidase can be measured by the assay, it can also be easily subtracted and thus will not interfere with rigorous analysis of the data.

4.1.2. Turbidity issues

The presence of the octanoyl chain confers to the substrate limited solubility in aqueous buffer. Even at low concentration (1–5 µM), the reaction mixture was turbid. Addition of dimethyl sulfoxide (DMSO) to the reaction mixture improved the solubility of the substrate. A DMSO concentration course experiment, where only the concentration of DMSO is varied (2–15% v/v), shown in Figure 20, revealed that at high concentration, DMSO affected the enzymatic activity of Ag85C. Therefore 2% (v/v) DMSO was chosen as the optimal DMSO concentration, compromising between the solubility of the substrate and the optimal activity of the enzyme.
4.1.3. Rate-determining step

The next parameter adjusted was the ratio of β-glucosidase and Ag85C. In order for the assay to directly reflect the activity of Ag85C it is necessary that the step catalyzed by Ag85C is rate-determining. This was confirmed by an Ag85C concentration course experiment as presented in Figure 21. Several reactions were carried out with a fixed concentration of β-glucosidase (2 U) and increasing concentrations of Ag85C (10–1500 nM).

The linear regression analysis of the initial velocity values obtained from this experiment, presented in Figure 22, gave a first-degree order equation with a $r^2$ value of 0.9979 and revealed that there is a linear relationship between the reaction rate and the concentration of Ag85C.
This result confirmed that β-glucosidase was in excess and that Ag85C catalyzed the slowest step of the coupled assay. In addition, a similar concentration course experiment was performed with varying concentration of β-glucosidase (0.1–5 U per reaction) and at a fixed Ag85C concentration of 200 nM. The data shown in Figure 23 illustrate that the initial velocity for each reaction increases with respect to β-glucosidase concentration of 0.1–1 U. For β-glucosidase concentrations of 1 unit/reaction and higher, the initial velocity peaks near 0.02 µM/s, indicating that the Ag85C activity is now limiting the velocity of the reaction.
Based on these data, a concentration of 2 units/reaction of β-glucoSIDase was used for all subsequent assays.
4.1.4. Optimal assay parameters

The control experiments helped determine the optimal assay conditions presented in Table 2. These conditions were used for most of the subsequent assays performed.

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity (per reaction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag85C</td>
<td>100 nM</td>
</tr>
<tr>
<td>β-glucosidase</td>
<td>2 U</td>
</tr>
<tr>
<td>D-glucose</td>
<td>3 mM</td>
</tr>
<tr>
<td>Substrate:</td>
<td>175 µM</td>
</tr>
</tbody>
</table>

Table 2 Optimal reaction conditions

An Ag85C concentration of 100 nM appeared to be the minimal concentration of Ag85C needed to reliably detect an absorbance change as shown in Figure 21. At this concentration, 2 units of β-glucosidase guaranteed that Ag85C was responsible for the rate-determining step. The substrate concentration was chosen because it is at the limit of solubility of the compound at 2% DMSO.

4.2. Kinetic parameters of Ag85C

4.2.1. Introduction

The developed assay allows the determination of kinetic parameters of Ag85C. To determine those parameters using the synthetic substrate, multiple assays were performed using substrate concentrations ranging from 10–200 µM. The limits were chosen for both theoretical and practical reasons. A concentration of 10 µM was defined as the lower limit because of theoretical requirements. The Michaelis-Menten model, applied to analyze the results, is only valid when the substrate of the reaction is present in
great excess compared to the enzyme. Therefore if a Ag85C concentration of 100 nM is used, a substrate concentration of 10 μM represents a 100-fold excess. The upper limit, 200 μM, was chosen because of technical limitations. At this concentration, the reaction mixture started to present traces of turbidity meaning that the maximum solubility of the substrate in 2% DMSO had been reached.

4.2.2. Results

The initial velocity at the different substrate concentrations was analyzed to obtain values for Michaelis-Menten parameters. The obtained curves are presented in Figure 24 and Figure 25. Under the current assay regime, the Vmax and the K_M using the synthetic substrate are 0.0062 ± 0.0003 μM/s and 0.047 ± 0.008 mM, respectively. The turnover number of the enzyme is 0.062 s⁻¹. The K_M value reflects a moderate affinity of the enzyme for the substrate and the calculated turn-over number shows a low efficiency of the enzyme.

The K_M value determined for the Ag85C assay using the synthetic substrate is similar to that seen for other members of the α/β hydrolase superfamily of enzymes. Kinetic studies with AChEs from the electric eel and PKS–TE from *Bacillus brevis* have shown these enzymes to possess K_M values of 100 and 3 μM, respectively.⁶³, ⁶⁴ These K_M values are quite similar to the value observed for Ag85C, indicating that the synthetic Ag85 substrate performs well in mimicking the natural substrate. In contrast, the k_cat value for Ag85C is much lower than that observed for other α/β hydrolases such as AChEs⁶³ and other carboxylesterases.⁶⁵ These enzymes have very high turnover numbers on the orders of 10⁴ or 10⁵ s⁻¹.
Figure 24  Absorbance versus time graphs for different substrate concentrations

Figure 25  Michaelis-Menten Fit of Vi values versus substrate concentration
However, the $k_{\text{cat}}$ for Ag85C is more reflective of the 1 s$^{-1}$ value measured for the PKS–TE. The reasons for these differences are not clear, but one can rationalize that the $k_{\text{cat}}$ values may reflect the identity of the nucleophile in the second step of the enzymatic reaction. AChEs and carboxylesterases use water as the nucleophile to disrupt the acyl-enzyme intermediate, whereas Ag85 and PKS–TE do not. Rather, PKS–TE relies on an intramolecular nucleophilic attack at the thioester linkage that produces the cyclic product and allows release from the active site. The relatively low turnover number possibly reflects the requirement of a structural change in the substrate to promote this attack. The Ag85 mechanism goes a step further and requires an intermolecular nucleophilic attack, which further depresses the turnover rate with respect to PKS–TE. However, although the described assay uses D-glucose as the acyl acceptor at a 3 mM concentration, this might not reflect the concentration encountered by the enzyme in vivo.

5. Conclusion on the enzymology

The Ag85 complex of M. tb, involved in the maintenance of the cell envelope and in the phagocytosis of the bacterium by macrophages through their fibronectin-binding ability, was identified as a possible antitubercular drug target. To speed the drug discovery process, a new assay was needed because the existing radiometric assay is cumbersome and time-consuming. A rapid colorimetric assay was developed to test the activity of Ag85. The initial development phase successfully produced a reliable experimental protocol for Ag85C. The design, synthesis and purification of the main components of the assay as well as proof-of-principle, control experiments and preliminary results were presented in this chapter. The kinetic parameters of Ag85C were established using the synthetic substrate designed for the assay. The affinity of the
enzyme for the synthetic compound was not great but it should allow the use of the assay for the screening of compound libraries.

6. Future work

Since the substrate showed limited solubility in the assay conditions, new synthetic substrates will be designed to improve the solubility. The sensitivity of the method could be improved by the use of a fluorescence-based assay instead of a colorimetric one.

The development phase of the project focused on Ag85C because it is more active than the other two paralogs. The assay protocol will be validated for Ag85A and Ag85B. Different applications of the assay are under development: the kinetic characterization of the three paralogs and the screening of small libraries of synthetic and natural compounds.
Introduction to screening applications

1.1. Validation of the assay for high throughput screening application

The high-throughput coupled colorimetric assay presented in Chapter 3 was developed to assess the enzymatic activity of Ag85C. It uses the acyltransferase activity of Ag85C to remove an octyl chain from the substrate and the glycosidase activity of β-glucosidase to release the chromophore. Under the optimal conditions, the assay runs perfectly for kinetic parameters measurement. Another possible application for the assay is the screening of libraries of compounds. Indeed by adding potential inhibitor to the assay and comparing the obtained results to a reference assay without that compound, it is possible to evaluate the inhibitory activity of the compound on the studied enzyme. In this case, the initial velocity values \( V_i \) obtained from the absorbance measurements over time are compared. An increase in \( V_i \) translates an enhancement of the enzyme activity whereas a decrease in \( V_i \) means inhibition of the enzyme activity. The latter is of greater interest in this work. Since the ultimate goal is to use the assay for high throughput screening (HTS), it is necessary to confirm that the assay is compatible with HTS requirements of practicality, cost and capability of automation. Moreover, the signal-to-
noise ratio and the reproducibility need to be excellent. Zhang et al. developed a statistical analysis of the assay to assess its potential as HTS assay.\textsuperscript{66}

The statistical parameter $Z'$ was developed for the evaluation of the inherent signal-to-noise ratio and to quantitatively represent assay reproducibility. The $Z'$ factor is defined in Equation 16, where $\sigma_p$ is the standard deviation of positive controls, $\sigma_n$ the standard deviation for negative controls, $\mu_p$ the mean value of positive controls and $\mu_n$ the mean value of negative controls.

$$Z' = 1 - \frac{3 \times (\sigma_p + \sigma_n)}{|\mu_p - \mu_n|}$$ \hspace{1cm} \text{Equation 16}

A $Z'$ factor value of 1 would be ideal and when approached it reflects a huge dynamic range and small standard deviation. Typically, values between 0.5 and 1 describe an excellent assay; values between 0 and 0.5 describe a marginal assay and values below 0 describe an assay useless for screening purposes.

To determine this parameter for the described assay, the $V_\text{i}$ values for a series of negative and positive controls were measured. The positive control reactions were performed using assay conditions mentioned previously (100 nM of Ag85C, 2 U of $\beta$-glucosidase, 3 mM of D-glucose and 175 $\mu$M of substrate). For the negative control assay, reactions were performed in the presence of $\beta$-glucosidase but lacking Ag85C. Two separate sets of 16 control experiments were used to evaluate $Z$. From these data, the calculated $Z'$ value was determined to be 0.81 ± 0.06, thereby indicating an excellent signal-to-noise ratio for the assay and that it is suitable for HTS applications.
1.2. Secondary control assay

The main application envisioned for the assay is the screening of compound libraries to identify potential anti-tubercular drugs. This type of application requires the development of a secondary control assay to ensure that the inhibition occurs on the enzyme of interest (Ag85C) and not the coupling enzyme (β-glucosidase). To create conditions similar to the main assay, the secondary assay, presented in Figure 26, uses the substrate encountered by β-glucosidase. The experimental set-up is the same. The comparison of the release rate of p-nitrophenolate in the absence and presence of the potential inhibitors will reveal what the target of the compounds actually is.

![Figure 26 Control reaction for β-glucosidase](image)

1.3. Type of libraries screened

After validation of the assay for HTS, it was used to screen libraries of synthetic and natural compounds. Some of the synthetic compounds were obtained through collaboration with Dr. Sucheck. Another source of synthetic compounds was the National Institute of Health (NIH) Clinical Collection (NCC). The natural compounds were acquired through collaboration with Dr. Hacker.
2. Synthetic compounds

2.1. Introduction

In recent years, Ag85 was the target of different classes of synthetic inhibitors. After the report by Belisle et al. that 6-azido-6-deoxytrehalose (ADT) caused inhibition of Ag85 activity in their assay, other inhibitors were designed and synthesized. To date, only four libraries of inhibitors were reported. The different libraries were based upon phosphonate compounds, sulfonate compounds and two types of trehalose analogs. The phosphonate and sulfonate compounds were designed from structural and mechanistic studies as transition state analogs. Once synthesized, the compounds were assayed using the radiometric assay described by Belisle and coworkers producing concentration values of compound needed to achieve half of the maximal activity (IC$_{50}$). The first trehalose-based library was designed before the crystal structure of Ag85C was reported, based on the mycolyl transferase activity of the enzymes. The second trehalose-based library was derived from mechanistic studies and included substrate analogs. The two libraries of compounds were assayed by a disk diffusion assay yielding a diameter of zone of inhibition (DZI) for a given drug concentration or by a macrodilution broth assay giving a minimum inhibitory concentration value (MIC). The structures of the compounds as well as their inhibitory activities are presented in Figure 27. Only the most potent compounds are reported here.
2.2. Screening of a library of methyl 5-$S$-alkyl-5-thio-D-arabinofuranosides

As part of the collaboration with Dr. Sucheck’s laboratory, a small library of compounds was screened for inhibitory activity against Ag85C. The synthetic molecules were designed based on mechanistic and structural features of the enzymes. For example a modified ester linkage on a trehalose derivative is likely to promote the nucleophilic attack of the serine residue but it would trap the acylated intermediate. The compounds possess an alkyl arabinosyl backbone similar to the mycolated arabinogalactan and harbor a thioether linkage instead of an ester, as shown in Figure 28. This particular feature should allow the potential trapping of the intermediate formed.
after nucleophilic attack of the enzyme. The library was made of two classes of compounds with an alpha or beta methyl group and variable carbon chain length.

![Diagram of methyl 5-S-alkyl-5-thio-D-arabinofuranosides](image)

**Figure 28** Library of methyl 5-S-alkyl-5-thio-D-arabinofuranosides

### 2.2.1. Experimental method

A master mix was prepared, for each reaction, containing 0.05M sodium phosphate buffer pH 7.5, 2 units of β-glucosidase and 3 mM of D-glucose. A volume of 85.84 µL of the master mix was added to each well of the assay, followed by 10 µL of a 10% v/v DMSO stock solution of substrate (0.1-2.5 mM), and 1 µL of a DMSO solution containing the compound to be tested. To initiate the reaction 4.16 µL of a 45 µg/mL solution of purified Ag85C in 10mM Tris-HCl buffer pH 7.5, 2 mM EDTA and 1 mM DTT was added. The control reaction was prepared with 1 µL of DMSO instead of the tested compound.

### 2.2.2. Results

The arabinofuranoside library of compounds was screened using the assay described in Chapter 3. Unfortunately, the compounds showed low solubility under the assay conditions, especially the C6 and C8 compounds. The C4 compounds were
screened but didn’t display any activity as shown in Figure 29. Although C8 was not assayed, the compound displayed antibacterial activity in *M. smegmatis* cultures. Efforts to improve the solubility of the compounds are under way.

![Absorbance versus time curve for C4α at different concentrations](image)

Figure 29  
Absorbance versus time curve for C4α at different concentrations

2.3. Screening of arabinofuranoside and trehalose esters library

Another small library of substrate analogs was designed as potential mechanism-based inhibitors. Similarly to the previous library, the compounds harbor a modified ester moiety on the trehalose or arabinofuranoside backbone. This type of modified ester compound is likely to promote the nucleophilic attack of the serine residue but it would slow the reaction of the acylated intermediate. The compounds were synthesized in Dr. Suchec’s laboratory.

The library is composed of methyl β-D-arabino-heptafuranosiduronic acid methyl ester, referred to as methyl ester from now on, methyl 7-keto-β-D-arabino-
octafuranosiduronic acid methyl ester, referred to as keto ester, methyl N-butyl-7-keto-β-D-arabino-octafuranosiduronamide, referred to as keto amide and bis(methyl-α-D-glucopyranosyluronate)ether referred to as trehalose methyl ester. The structures of the compounds are presented in Figure 30.

Figure 30 Structure of the mechanism-based inhibitors

Only the methyl ester and the keto ester have been tested so far.
2.3.1. Experimental methods

The different ester derivatives displayed different solubility patterns; therefore different experimental protocols were required. The methyl ester was readily soluble in the assay buffer and the keto ester was only soluble in a buffer-DMSO mixture (50% v/v).

For the methyl ester, a master mix was prepared, for each reaction, containing 0.05 M sodium phosphate buffer pH 7.4, 2 U of β-glucosidase and 3 mM of D-glucose. A volume of 89.84 µL of the master mix was added to each well of the assay, followed by 2 µL of an 8.75 mM stock solution of substrate in DMSO, and 0.1-2.5 µL of a 0.1 M solution of methyl ester for final concentrations of 1-25 mM. To initiate the reaction 4.16 µL of a 45 µg/mL solution of purified Ag85C in 10mM Tris-HCl buffer pH 7.5, 2 mM EDTA and 1 mM DTT was added. The control reaction was prepared with 1 µL of buffer instead of the tested compound. After several preliminary experiments, better results were noticed when the methyl ester compound was incubated with the enzyme on ice for 30 min prior to the assay.

For the keto ester, a master mix was prepared with, for each reaction, 0.05 M sodium phosphate buffer pH 7.4, 2 U of β-glucosidase and 3 mM of D-glucose. A volume of 77.29 µL of the master mix was added to each well of the assay, followed by 2 µL of an 8.75 mM stock solution of substrate in DMSO, and 11.55 µL of a 0.0866 M solution of keto ester in buffer-DMSO (50%) for a final concentration of 10 mM. To initiate the reaction 4.16 µL of a 45 µg/mL solution of purified Ag85C in 10mM Tris-HCl buffer pH 7.5, 2 mM EDTA and 1 mM DTT was added. The control reaction was
prepared with 9.25 µL of buffer instead of the tested compound and 2.3 µL of DMSO since the keto ester reaction contained 4.3 % DMSO overall.

2.3.2. Results

From the slope of the absorbance versus time curves, the initial velocity values of the reactions were calculated. Comparison of the $V_i$ of the reaction with methyl ester to the control reaction without compounds gave the percent activity of Ag85C in presence of methyl ester as shown in Figure 31. For example at 25 mM of methyl ester, Ag85C is only 38% active after incubation, whereas in the assay run without pre-incubation, Ag85C is 78% active. A high concentration of the compound is required to see inhibition; therefore it does not qualify as an interesting compound for drug development.

![Figure 31](image)

Figure 31 Percent activity of Ag85C in presence of methyl ester
More experiments were run at different substrate concentrations to try to determine the type of inhibition achieved by the methyl ester; however, the obtained results were not conclusive.

Similar experiments were carried out with the keto ester. The incubation did not matter as much. At a 10 mM concentration of keto ester, a 20% inhibition was achieved. Because of the high concentration of compound required to accomplish this low inhibition, the keto ester was not considered as a lead for drug development.

2.4. National Institute of Health Clinical Collection

The NIH clinical collection (NCC) is a plated array of 360 small molecules that have been tested as potential drugs in human clinical trials. It is distributed by the Molecular Libraries Small Molecule Repository (MLSMR) operated by BioFocus DPI©, a Galapagos company.

2.4.1. Experimental methods

The compounds are at a concentration of approximately 10 mM in 100% DMSO. Therefore the assay parameters had to be adjusted. A master mix was prepared with, for each reaction, 0.05M sodium phosphate buffer pH 7.5, 2 units of \( \beta \)-glucosidase and 3 mM of D-glucose. A volume of 87.84 µL of the master mix was added to each well of the assay, followed by 2 µL of a 8.75 mM DMSO stock solution of substrate (for a final concentration of 175 µM), and 1 µL of a DMSO solution of tested compound. To initiate the reaction 4.16 µL of a 45 µg/mL solution of purified Ag85C in 10mM Tris-HCl buffer pH 7.5, 2 mM EDTA and 1 mM DTT was added. The control reactions were prepared with 1 µL of DMSO instead of the tested compound.
2.4.2. Results

Ten compounds resulted in a significant inhibition of Ag85C, yielding a hit rate of 3%. The most potent compound gave a 90% inhibition the first time the assay was run. The results were confirmed by running the assay multiple times. The seventeen compounds that gave consistently the best inhibition were then subjected to the secondary assay. Among those, six compounds produced a significant inhibition on the reaction rate of $\beta$-glucosidase, leaving only eleven compounds for further testing.

Besides their inhibitory effect on Ag85C, the ability of the compounds to cause growth inhibition of Mycobacteria was assayed by disk diffusion antibiotic sensitivity testing or Kirby Bauer testing. This assay uses antibiotic-impregnated wafers to test if bacteria are susceptible to a given antibiotic. The compound to be tested is impregnated onto a wafer that is then placed on a media plate. Bacteria cells are then spread onto the plate and cell growth is monitored. If the bacteria are not capable of growing because of the presence of the tested compound, a zone of inhibition can be observed on the plate. Dr. Sucheck preformed the Kirby Bauer disk assay. Out of the eleven compounds tested, five inhibited the growth of Mycobacterium smegmatis, a known non-pathogenic model for M. tb studies. The results for two of the compounds, Clofazimine (C1) and Ebselen (C2), are presented in Figure 32. The structure of those two compounds is presented in Figure 33. The compounds produced a zone of inhibition when assayed alone and gave a synergistic effect when combine with the known TB drug INH.

This result suggests that the growth of M. smegmatis was partly inhibited through a mechanism involving inhibition of Ag85C, thus confirming the enzyme is a good drug
target. In order to learn more about the mode of action of Clofazimine and Ebselen, crystallographic studies of the enzyme-inhibitor complex are under way.

Figure 32 Pictures of the disk diffusion assay plates (courtesy of Dr. Sucheck)

Figure 33 Structure of Clofazimine and Ebselen
3. Natural compounds

3.1. Introduction

An alternative to the use of synthetic compounds is to screen natural products. The natural samples can either be pure characterized molecules or extracts. In the latter case, the extracts may originate from organisms or microbial cultures. The screening methodology is described in Figure 34. The discovery of inhibitors using this particular route goes through an iterative process. The initial culture extracts are first screened for identification of hits. Any hit is then confirmed by repeating the assay. The sample is then recultured and screened again for activity. If the inhibitory activity is confirmed, the initial sample is then submitted to an organic/aqueous fractionation. Both fractions are then tested and the one displaying inhibitory activity is then further purified.

3.2. Screening of natural extracts

As part of the collaboration with Dr. Hacker, screening of micro-organism culture extracts was undertaken using the enzymatic assay. The tested samples are culture extracts from organisms collected in unique biological niches such as underground lakes and highly polluted Superfund sites. The organisms are uncharacterized and it is expected that they produce novel substances that may possess interesting pharmacologic activities.
Figure 34  Screening procedure for the natural extracts.
3.2.1. Experimental method

For each well, a master mix containing 0.05M sodium phosphate buffer pH 7.5, 2 units of β-glucosidase, 3 mM of D-glucose was prepared. A volume of 1 µL of the sample to be tested was added to the well, followed by 1 µL of a 17.5 mM DMSO solution of the substrate and the master mix. To initiate the reaction Ag85C was added to a final concentration of 100 nM. Control reactions were prepared by omitting the sample to be tested and replacing it with DMSO. Comparison of the initial velocity values obtained for each sample and the control showed no effect on Ag85C activity.

3.2.2. Results

The results are scored using Equation 17. The initial velocity of the reaction containing the sample is compared to the initial velocity obtained for the control reaction run with DMSO.

\[
Score = \frac{V_i(sample)}{V_i(control)} \times 100
\]  

Equation 17

Three different trends have been identified: an increase with a score higher than 110%, no change with a score between 90 and 110% or a decrease in the initial velocity with a score lower than 90%. The latter is the only one of interest in the present study. An increase in initial velocity might result from the presence of enzymes catalyzing the production of the p-nitrophenolate or the reaction of the substrates with molecules present in the sample. No change in initial velocity indicated that there was nothing interfering with the assay. For the observed decrease in activity, the samples with a score of less than 50% were identified as hits. So far five hundred samples have been screened in the laboratory and after the first round of screening, forty hits were identified. The confirmation of those hits is currently under way.
3.3. Conclusion

In the context of the screening of compounds, the assay fulfilled the expectations. It allows a robust testing of different substances, synthetic or natural, pure or mixtures at a fast pace. The high-throughput assay allowed testing of hundreds of compounds in a matter of days. This represents significant improvement over the radiometric assay. The absence of extraction/purification step makes the assay easier; also the assay does not necessitate any unusual equipment or hazardous substance handling; it only requires a spectrophotometer and uses aqueous solutions. The format can easily be changed from a single experiment to a 96-, 384- or 1586-well plate. In conclusion, the newly developed Ag85 assay is a powerful tool that permits the rapid screening and identification of possible anti-tubercular drugs.

3.4. Future work

Presently, the rate limiting step is the data analysis. The screening of libraries of compounds generates a tremendous amount of data that needs to be manually processed. Improvements on the data processing such as the use of application macros in Microsoft Office Excel are being developed. Also the format of the assay could be changed form a kinetic analysis to an endpoint analysis, therefore minimizing the amount of data.

After identification and validation of the inhibitor, it will be kinetically characterized. The inhibitory constant (K_i) of the enzyme-inhibitor complex as well as the concentration of compounds that provides 50% inhibition (IC_{50}) can be easily calculated using the assay.

The compound screening is one step of the drug development process. Once an inhibitor is identified and confirmed to inhibit Ag85C only, it will be promoted to lead
compound status. Crystallographic study of the lead-enzyme complex will help a structure-based drug development approach. By modifying the lead to achieve better binding to the enzyme, better inhibitors will be discovered.
Chapter Five

Study of the acyl acceptor specificity of the antigen 85 complex
through kinetic and structural characterization

1. Introduction to the antigen 85 complex substrate paradox

The kinetic parameters of Ag85C were determined using the assay presented in Chapter 3. The same procedure was used for the kinetic characterization of the other members of *M. tb* antigen 85 complex. By varying the concentration of one of the reaction substrates, either the acyl donor or the acyl acceptor, the affinity of the enzyme for that substrate can be determined. This application of the assay will be used for the kinetic characterization of the antigen 85 complex and the study of its substrate specificity.

The radiometric mycolyltransferase assay developed by Belisle *et al.* revealed the *in vitro* mycolyltransferase activity of Ag85A, B and C. As part of the observations, Ag85B was less active than Ag85A and C. Gene disruption experiments confirmed the *in vivo* involvement of Ag85 with the transfer of mycolates to the mycobacterial cell wall. To date, there is no definitive explanation for the existence of three enzymes encoded by three independent genes that display the same enzymatic function in *M. tb*. Different rationales have been proposed for this phenomenon such as the differential expression of the genes depending on the mycobacterial environment and growth phase,
and a different affinity for the proposed substrates. Early studies demonstrated that Ag85A, B and C are expressed in different ratios, 2:3:1 respectively.\textsuperscript{45} Gene-expression studies using a reverse-transcriptase polymerase chain reaction (RT-PCR) showed a differential expression of the Ag85 complex depending on the growth conditions.\textsuperscript{75} In synthetic Sauton medium, both Ag85B and C expression was detected, whereas in resting macrophages, only Ag85A was expressed. Finally, in activated macrophages, only Ag85C was actively expressed. A similar study carried out in a \textit{M. tb} clinical isolate and a laboratory strain demonstrated slightly different results.\textsuperscript{76} In the virulent laboratory strain H37Rv, the expression of Ag85B and C was detected in resting macrophages, while in the clinical isolate, these enzymes were expressed in activated macrophages. The expression of Ag85A was negligible in both cases. These results confirmed the differential expression of the Ag85-encoding genes in \textit{M. tb} depending on environmental conditions.

In addition to the changes in expression of the enzymes, their different affinity for the substrates remains unclear. The disruption of Ag85C resulted in a 40\% decrease of the cell wall-bound mycolates suggesting a preference of the enzyme to use arabinogalactan as a substrate.\textsuperscript{52} Similar experiments conducted on Ag85A and B did not yield any change in the mycolate content of the cell envelope.\textsuperscript{53} However, the overexpression of Ag85A and B in the Ag85C-deficient strain experiments restored the wild-type cell envelope composition. The experimental evidence was that large amounts of Ag85A and B compensated for the absence of the very active Ag85C. This study concluded that the enzymes posses partial redundancy in their biological roles.\textsuperscript{54}
Moreover the absence of mycolyl transfer to trehalose molecules in vivo was in contradiction with the results from the in vitro assay developed by Belisle.\textsuperscript{4} This was interpreted as an experimental artifact; because of the in vitro nature of the assay, the enzymes displayed wider substrate specificity.\textsuperscript{54} Recently, a study carried out in \textit{M. smegmatis} confirmed the possibility of different substrate affinities. The inactivation of the \textit{fbpA} gene resulted in a 45\% decrease in TDM production.\textsuperscript{77} Because of the high sequence homology between Ag85A from \textit{M. tb} and \textit{M. Smegmatis}, this result suggested that Ag85A prefers trehalose as a substrate while previous reports proposed that Ag85C prefers arabinogalactan. By showing that the apo-enzymes possess identical active sites,\textsuperscript{48} the structural studies of the three paralogs did not improve the understanding of that matter.

To better understand this apparent substrate preference, many methodologies can be used. Because of its easily tunable design, our assay will help clarify the affinity of the three enzymes for different substrates. Trehalose was originally suggested as the mycolyl acceptor of the in vitro mycolyltransferase reaction.\textsuperscript{4} But in vivo experiments determined that the arabinogalactan moiety of the cell envelope was the substrate in the mycobacterial cell.\textsuperscript{52} In order to study the substrate specificity of Ag85, two model substrates were chosen for this preliminary study: trehalose and methyl 2-\textit{O}-\beta-\textit{D}-arabinosylfuranosyl-\alpha-\textit{D}-arabinofuranoside, hereafter called arabinofuranosyl arabinoside (AA), a synthetic substrate used as a model for the arabinogalactan.

During the kinetic experiments, the acyl donor, \textit{p}-nitrophenyl containing substrate, was conserved and the acyl acceptor was varied. Our initial studies used \textit{D}-glucose, trehalose and AA shown in Figure 35.
1.1. Experimental methods

This section describes the purification protocols of Ag85A and Ag85B as well as the experimental procedure for the assay.

1.1.1. Cloning of $fbpA$ and $fbpB$

Antigen 85A and 85B were cloned, expressed and purified following a protocol very similar to the one developed for Antigen 85C. The portion of the $M. tb. fbpA$ and $fbpB$ genes encoding the mature secreted forms of Ag85A and Ag85B, respectively, were amplified by PCR using the following primers 5'-CAC CAT CAT ATG TTT TCC CGG CCG GGC TTG CCG GT-3' and 5'-TTG GAT CCC TAG GCG CCC TGG GGC GCG G-3' for Ag85A and 5'-CAC CAT CAT ATG TTC TCC CGG CCG GGG CTG CC-3' for Ag85B. The PCR products were then placed into pET-28 using NdeI and BamHI restriction enzymes in a traditional digestion/ligation procedure. In both cases, the resulting open reading frame produces proteins with an N-terminal 6XHis tag.
1.1.2. Expression and purification of Ag85A and B

The plasmids pET28-Ag85A and pET28-Ag85B were used to transform T7 Rosetta cells. Those competent cells were prepared in the laboratory by transformation of T7 Express cells (New England BioLabs) with the pRos plasmid (Stratagene). The pRos plasmid confers the ability to express genes encoding tRNA for codons rarely used by E. coli. The bacterial cells were grown in LB Broth at 37 °C, induced with IPTG and harvested by centrifugation after 16-20 hours of incubation at 16 °C. The pelleted cells were resuspended in Ag85 binding buffer and stored at -80 °C until protein purification was initiated. The frozen aliquot of resuspended cells was thawed at room temperature and then placed on ice. An enzymatic hydrolysis step was performed for 20 minutes on ice prior to lysis of the cells using a sonicator. The obtained lysate was clarified by high-speed centrifugation. The pellet component was discarded and the supernatant was filtered. The protein of interest was first crudely purified by immobilized metal affinity chromatography, as shown in Figure 36 for Ag85A and in Figure 37 for Ag85B. The fractions containing Ag85A or B were pooled, dialyzed against Ag85 crystallization buffer at 4 °C for 15-20 h. The protein solution was subsequently subjected to anion exchange chromatography. The protein was eluted using the modified Ag85 crystallization buffer. The chromatograms are presented in Figure 38 for Ag85A and in Figure 39 for Ag85B.

The fractions containing the protein of interest were pooled and solid ammonium sulfate was added to a concentration of 2.4 M. Similarly to Ag85C, only ammonium sulfate precipitation permitted the concentration the proteins. The precipitated protein was pelleted by centrifugation and resuspended in crystallization buffer. The final
purification step uses size exclusion chromatography. The chromatograms are presented in Figure 40 for Ag85A and in Figure 41 for Ag85B. The fractions containing Ag85A or Ag85B were pooled and stored at -80 °C until needed for the assay.

The N-terminal tag was not cleaved off the recombinant protein because it should not interfere with the enzymatic assay. If any crystallization experiment were to be conducted, the tag would be removed using the tag removal protocol described in the experimental methods chapter (Chapter 2).

Ag85A was obtained in high purity according to the final SDS-PAGE gel shown in Figure 40. From the SDS-PAGE gel presented in Figure 41, Ag85B was considered pure enough to carry out enzymatic experiments despite the presence of a faint band at a lower molecular weight of 25 kDa. The concentrations of the proteins were calculated from absorbance readings at 280 nm and the calculated molar extinction coefficient of Ag85A and Ag85B, also an additional absorbance reading at 260 nm confirmed the absence of nucleic acids in the sample.
Figure 36  Ag85A Nickel affinity chromatogram and the corresponding SDS-PAGE gel
Figure 37  Ag85B Nickel affinity chromatogram and the corresponding SDS-PAGE gel
Figure 38  Ag85A anion exchange chromatogram and the corresponding SDS-PAGE gel
Figure 39   Ag85B anion exchange chromatogram and the corresponding SDS-PAGE gel
Figure 40  Ag85A size exclusion chromatogram and the corresponding SDS-PAGE gel
Figure 41 Ag85B size exclusion chromatogram and the corresponding SDS-PAGE gel
1.1.3. Assay

The same optimal assay conditions determined in Chapter 3 were used for the kinetic experiments with the exception of the enzyme concentration. In the previous *in vitro* assay, the three paralogs displayed different levels of activity. While Ag85A displayed a comparable activity, Ag85B displayed only 20% of the activity of Ag85C. As a result if 100 nM was the minimal amount of Ag85C to be used in order to get a detectable absorbance change in this assay, it was necessary to increase the enzyme concentration for Ag85A and especially Ag85B. A test experiment was run with 100 nm of each enzyme, 2 U of β-glucosidase and the optimized assay conditions described earlier. The results are presented in Figure 42.

![Graph showing absorbance over time for different enzyme concentrations](image)

Figure 42    Test reaction with 100 nM of each enzyme
Ag85A and Ag85B are known to be less active than Ag85C. As expected, the change in absorbance caused by Ag85A or B was not significant compared to the reaction without antigen 85. The initial velocity values were 0.722 nM/s, 0.674 nM/s and 0.472 nM/s for Ag85A, Ag85B and β-glucosidase only respectively. There is no significant difference between the reaction rate of β-glucosidase and the reaction rates of Ag85A and Ag85B. By comparison, at the same concentration, Ag85C displayed a $V_i$ of 2 nM/s that is approximately 5-fold higher, compare to β-glucosidase. The concentration of Ag85A and B was consequently increased to 500 nM for subsequent assays. If Ag85C was to be used at this concentration, the reaction would be so fast, that it would not produce any consistent results.

A typical reaction was carried out under atmospheric pressure at room temperature. The experiments were carried out in 96-well plates with a final volume of 100 µL. A master mix specific to each enzyme was prepared with, for each well, the 0.05 M sodium phosphate pH 7.4 buffer, the desired enzyme at a final concentration of 500 or 100 nM and 2 units of β-glucosidase. Typically 2 µL of an 8.75 mM solution of the substrate in DMSO was added to each well followed by addition of the 0.1 M acyl acceptor solution in buffer to get to the desired concentration (1.25-5 mM). The reaction was initiated by addition of the master mix containing the enzymes to the well.

1.2. Kinetic characterization of Ag85A and B using D-glucose as the acyl acceptor

Figure 43 presents the results obtained when D-glucose was used as the acyl acceptor. The calculated kinetic parameters ($V_{max}$, $K_M$ and $k_{cat}$) for Ag85 A, B and C are shown in Table 3.
Table 3  
Kinetic parameters with D-glucose as the acyl acceptor

The $K_M$ and $k_{cat}$ values of Ag85A and B for D-glucose are essentially the same. The $K_M$ values for all three enzymes are comparable; however, because of the standard deviation values obtained for Ag85B and Ag85C, these experiments need to be improved. The turnover number shows that Ag85C is 5-fold more efficient than Ag85A or B.
1.3. Acyl acceptor specificity

1.3.1. Kinetic parameters determination for arabinofuranosyl arabinoside as the acyl acceptor

The assay was run using arabinofuranosyl arabinoside as the acyl acceptor; the results are presented in Figure 44 and in Table 4.

Ag85C was still the most active and Ag85A and B displayed comparable low levels of activity. Ag85A and B are essentially indistinguishable from one another on the graph and all the kinetic parameters of Ag85A and B are the same if the standard deviation values are taken into account. If only this substrate was considered, the redundancy of Ag85A and B would be confirmed; however it is not the case (vide infra).
Table 4  Kinetic parameters with arabinofuranosyl arabinoside as the acyl acceptor

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$V_{\text{max}}$ (nM/s)</th>
<th>$K_M$ (mM)</th>
<th>$k_{\text{cat}}$ (s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag85A</td>
<td>4.1 ± 0.10</td>
<td>0.70 ± 0.080</td>
<td>0.0082</td>
</tr>
<tr>
<td>Ag85B</td>
<td>3.8 ± 0.30</td>
<td>0.61 ± 0.25</td>
<td>0.0076</td>
</tr>
<tr>
<td>Ag85C</td>
<td>6.0 ± 0.13</td>
<td>0.071 ± 0.050</td>
<td>0.06</td>
</tr>
</tbody>
</table>

The turnover value of Ag85C was about 7.3-fold and 7.9-fold higher than that of Ag85A and B, respectively. As for the affinity for the substrate, there is a 10-fold difference in $K_M$ values between Ag85A or B and Ag85C. This result proved that Ag85C displays a higher affinity for AA than the other two enzymes.

1.3.2. Kinetic parameters determination for trehalose as the acyl acceptor

The results obtained when trehalose was used as an acyl acceptor are presented in Figure 45 and in Table 5. This experiment confirms the observation of Belisle et al. about the difference in activity. Ag85C is the most active, followed by Ag85A whereas Ag85B displayed a much slower rate of reaction. When trehalose is used as the acyl acceptor, Ag85A displays a 2-fold higher turnover number than Ag85B even though their $K_M$ values are comparable. Ag85C shows a 2-fold increase of $k_{\text{cat}}$ compared to Ag85 A. Ag85C catalyzes the reaction so quickly that more experiments are needed to accurately determine its $K_M$ value. In these assay conditions, Ag85C reaches $V_{\text{max}}$ at concentrations lower than 1.5 mM of trehalose, the minimal concentration used here. More experiments will be carried out at lower concentrations of acceptor.
Figure 45  Michaelis-Menten fit when trehalose is used as an acceptor

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$V_{\text{max}}$ (nM/s)</th>
<th>$K_M$ (mM)</th>
<th>$k_{\text{cat}}$ (s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag85A</td>
<td>$14.1 \pm 0.11$</td>
<td>$0.83 \pm 0.27$</td>
<td>$0.028$</td>
</tr>
<tr>
<td>Ag85B</td>
<td>$7.0 \pm 0.57$</td>
<td>$0.67 \pm 0.27$</td>
<td>$0.014$</td>
</tr>
<tr>
<td>Ag85C</td>
<td>$22.4 \pm 0.81$</td>
<td>N.A.</td>
<td>$0.045$</td>
</tr>
</tbody>
</table>

Table 5  Kinetic parameters when trehalose is used as an acceptor
1.4. Acyl acceptor specificity

The assay allowed the kinetic characterization of the three enzymes for the different substrates. Since it has been reported that the use of $k_{\text{cat}}/K_M$ is erroneous when comparing different enzymes,\textsuperscript{78} this parameter will not be reported when the enzymes are being compared for a given substrate.

For an overall understanding of the results presented so far, the comparison between the different substrates is shown in Table 6.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Acceptor</th>
<th>$V_{\text{max}}$ (µM/s)</th>
<th>$K_M$ (mM)</th>
<th>$k_{\text{cat}}$ (s$^{-1}$)</th>
<th>$k_{\text{cat}}/K_M$ (M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag85A</td>
<td>D-glucose</td>
<td>3.8 ± 0.13</td>
<td>0.36 ± 0.09</td>
<td>0.0075</td>
<td>20.8</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>4.1 ± 0.10</td>
<td>0.70 ± 0.08</td>
<td>0.0082</td>
<td>11.7</td>
</tr>
<tr>
<td></td>
<td>Trehalose</td>
<td>14.1 ± 0.11</td>
<td>0.83 ± 0.27</td>
<td>0.028</td>
<td>33.7</td>
</tr>
<tr>
<td>Ag85B</td>
<td>D-glucose</td>
<td>3.2 ± 0.22</td>
<td>0.35 ± 0.20</td>
<td>0.0063</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>3.8 ± 0.30</td>
<td>0.61 ± 0.25</td>
<td>0.0076</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td>Trehalose</td>
<td>7.0 ± 0.57</td>
<td>0.67 ± 0.27</td>
<td>0.014</td>
<td>20.9</td>
</tr>
<tr>
<td>Ag85C</td>
<td>D-glucose</td>
<td>18.5 ± 1.8</td>
<td>0.22 ± 0.24</td>
<td>0.037</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>6.0 ± 0.13</td>
<td>0.071 ± 0.050</td>
<td>0.06</td>
<td>845</td>
</tr>
<tr>
<td></td>
<td>Trehalose</td>
<td>22.4 ± 0.81</td>
<td>N.A.</td>
<td>0.045</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

Table 6  
Kinetic Parameters of Ag85 with different acceptors

The most striking feature is that, in this assay, Ag85C did not display a substrate preference. All the kinetic parameters are in the same range and almost equal if the standard errors are taken into account. This result tends to confirm the hypothesis that Ag85C has a broader substrate range \textit{in vitro}, however this particular trend does not
apply to the other two enzymes. Ag85B displayed similar activity in all cases, with slightly better results for trehalose. It either means that another carbohydrate is the preferred acceptor for Ag85B or that the relative abundance of this paralog compensates for its weak activity. Based on the turnover number, Ag85A favored trehalose as a substrate with a 3.5-fold difference in favor of trehalose. This result is confirmed by the catalytic efficiency values, $k_{\text{cat}}/K_M$, with a 2.8-fold difference over AA.

Those surprising results need to be further investigated with an emphasis on a more accurate determination of the $K_M$ value for each enzyme/acyl-acceptor system. This can be achieved by running a very similar experiment with lower concentration of acyl acceptor. Data from a previous set of experiments using different 96-well plates is shown in Figure 46 as an example.

![Figure 46](image.png)

**Figure 46** Comparison of the Michaelis Menten fit obtained for Ag85A (red), Ag85B (green) and Ag85C (blue). The dashed lines represent data from reactions carried out with AA whereas the solid lines correspond to trehalose data.
1.5. Conclusion

In conclusion, more experiments need to be carried out to solve the debate over the substrate preference of Ag85. While Ag85C does not display any substrate preference, it seems that trehalose is a preferred substrate for Ag85A. The case of Ag85B remains enigmatic. A screening of different acyl acceptors needs to be carried out for Ag85B. Indeed there are many mycolated components in the bacterial cell wall that could result from the mycolyltransferase activity of Ag85B.

An *in vitro* approach might not be sufficient to unravel all the aspects of Ag85 activity in *M. tb*. From the literature review and the present work, it seems that many parameters are involved. First, a differential expression of the genes depending on the *M. tb* strain and its environment has been demonstrated. Secondly the experiments showing inactivation of the *M. tb* Ag85 genes or homologous genes in parent species did not give consistent results. Consequently rigorous gene disruption experiments of all three *M. tb* paralogs under controlled environmental conditions such as medium and growth phase, need to be undertaken to achieve a clear understanding of Ag85 activity *in vivo*.

The difference in catalytic behavior of the Ag85 paralogs remains surprising. The high sequence homology and the conservation of the catalytic residues between Ag85A, B and C suggest that the difference in acyl donor affinity and catalytic activity might reside in structural details that change during the various steps of catalysis.
2. Structural study of the Antigen 85 complex in relation to its substrate specificity

The structure of the apo-form of Ag85C and the structure of the enzyme modified with diethyl phosphate (DEP) were solved.\(^\text{47}\) Ag85C was reacted with diethyl \(p\)-nitrophenyl phosphate, a known mechanism-based inhibitor of acetylcholinesterase, to yield the modified Ag85C-DEP that mimics the transition state of the enzyme. The phosphorus atom undergoes a nucleophilic attack by the hydroxyl oxygen of the catalytic serine; as a result \(p\)-nitrophenolate is released as shown in Figure 47.

![Figure 47: DEP modification of Ag85C](image)

Comparison of the apo and the DEP-modified structures revealed that upon formation of the transition state intermediate one loop near the active site moves as shown in Figure 48. The importance of this conformational change is not completely understood but it is clear that it has an effect on the acyl acceptor binding in Ag85C.
In order to study the different catalytic behavior of Ag85 A, B and C, the conformational change observed in Ag85C was investigated through structural studies of modified Ag85A and Ag85B as well as kinetic characterization of mutants of those two enzymes.

2.1. Structural study of modified Ag85A and Ag85B

Ag85A and Ag85B were reacted with diethyl p-nitrophenyl phosphate to produce Ag85A-DEP and Ag85B-DEP according to the procedure used for Ag85C. The partial modification of the enzyme was confirmed by comparison of the enzymatic assay results between the apo enzymes and the DEP-modified samples. In the case of Ag85A, a 50% inhibition of the modified enzyme was observed. For Ag85B, the observed inhibition was 77%. These results indicate that the enzymes were partially modified. Repeated attempts to confirm the modification by electrospray ionization mass spectrometry failed.
The DEP-modified protein samples were used to set up crystallization trays. The Classics Lite Suite (Qiagen) was used for the initial screening. Only a few conditions gave crystals and unfortunately, the crystals formed were not diffraction quality. Ag85B-DEP forms clusters of plates in 0.1 M ammonium sulfate, 0.1 M sodium acetate and 1% polyethylene glycol (PEG) 2000 MME. Ag85A-DEP tends to form very fine needle-like crystals. Pictures of the crystals are shown in Figure 49 and Figure 50. Recent efforts to improve these crystals have been unsuccessful.

Figure 49 Picture of Ag85B-DEP crystals
After the unsuccessful attempts to solve the structure of Ag85A-DEP and Ag85B-DEP, another approach was chosen.

2.2. Kinetic study of Ag85A and Ag85B mutants

The formation of the transition intermediate during the first step of the acyltransferase Ag85C reaction causes the movement of the C210-E229 region of the protein. This characteristic has yet to be confirmed for Ag85A and Ag85B. Since no structural data is yet in hand, a study of that particular region of the enzymes was initiated. Sequence alignment of this region of the polypeptide chain reveals a highly conserved sequence as shown in Figure 51. Compared to Ag85C, the observed differences are K213T, N220D and L222I for Ag85A and N215S, E216D, A220D and E225K for Ag85B, where the first letter indicates the amino acid present in Ag85A or B, the number is the residue position in the polypeptide chain and the second letter is the corresponding amino acid present in Ag85C.
DNA primers for mutagenesis experiments were designed for Ag85A K213T, N220D, L222I and Ag85B N215S/E216D, A220D and E225K. The mutagenesis experiments were carried out using standard site-directed mutagenesis procedures. All the Ag85A mutations were completed, however only A220D and E225K have been successful for Ag85B. The mutations were confirmed by DNA sequencing and T7 Rosetta cells were transformed with the mutated plasmids. The mutated proteins were purified following the procedure described for the wild-type proteins.

The different mutants were then assayed to compare their activity with the apo enzymes. The results are presented in Table 7.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Percent activity compared to the apo enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag85 A K213T</td>
<td>175%</td>
</tr>
<tr>
<td>Ag85A N220D</td>
<td>368%</td>
</tr>
<tr>
<td>Ag85A L222I</td>
<td>132%</td>
</tr>
<tr>
<td>Ag85B A220D</td>
<td>88%</td>
</tr>
<tr>
<td>Ag85B E225K</td>
<td>60%</td>
</tr>
</tbody>
</table>

Table 7 Enzymatic activity of the Ag85A and Ag85B mutants

The mutations performed on Ag85A led to an increase of activity. The K213T and L222I mutations did not significantly improve the enzymatic activity whereas the N220D is almost four times more active than the wild-type enzyme. The Ag85B mutants
did not show any enhancement of the enzymatic activity; on the contrary, the wild-type enzyme was more active than the mutants. Since position 220 was mutated for both enzymes, it is quite surprising that the effects are so different. The explanation for the slow activity of Ag85B might not reside in this portion of the enzyme. This preliminary study produced interesting results that needs to be investigated further.

3. Conclusion

The assay developed to assess the enzymatic activity of Ag85C was successfully used to test the activity of Ag85A and B. The previously published results concerning the difference in enzymatic activity were confirmed. Ag85C is more active than the other two enzymes. Moreover, this assay allowed the initial study of the substrate specificity of Ag85. Even if the three paralogs are very similar, Ag85A prefers trehalose as an acyl acceptor while it is not clear if Ag85C or Ag85B have any substrate preferences.

The comparison of the structures of unmodified Ag85C and DEP-modified Ag85C revealed a conformational change in the enzyme prompted by the acylation of the catalytic serine. The conformational change occurs between the positions 210 and 229 of the polypeptide chain of Ag85C. Preliminary crystallographic studies of the modified Ag85A and B have been successful but have not yet yielded structural information. Mutagenesis experiments to change the 210-229 region of Ag85A and Ag85B into Ag85C have given interesting results. The single mutation of N220D in Ag85A caused a 3.7-fold increase in activity. A similar change in Ag85B, A220D, did not produce such a striking result. These preliminary studies will be continued to explore the substrate and activity paradox of Ag85.
4. Future work

The substrate specificity study of Ag85 needs to be performed on more substrates since no preferred substrate was identified for either Ag85C or Ag85B. The presence of mycolated glycerol, mycolated glucose and other mycolic esters in the bacterial cell wall suggests that there are many molecules to be investigated for the substrate specificity of the mycolyltransferase.

The difference of activity of Ag85A, B and C needs to be studied in detail. The mutagenesis experiments that have been done so far revealed interesting trends for Ag85A. Since three separate mutations produced an increase in enzymatic activity, the study of combined mutations or even the triple mutant might help understand the intrinsic differences of the apo enzymes.

Also it would be interesting to determine if the conformational change observed in Ag85C-DEP also occurs in Ag85A-DEP and Ag85B-DEP. Improvement of the crystallization conditions are under way to get diffraction-quality crystals.
Chapter Six

Crystallization Studies of the Enhanced Intracellular Survival Protein

1. Introduction

Chapter 1 emphasized the need to discover new drug targets to fight mycobacterial infection. The complete sequencing of *M. tb* in 1998\(^5,6\) opened the path to more anti-tubercular research. The discovery of 600 unknown open-reading frames suggested that there was a lot more to be learned. The identification and study of virulence factors of *M. tb* became a major goal in mycobacterial research. Indeed any gene product linked with pathogenicity seems to represent an attractive drug target. One such gene product is the enhanced intracellular survival protein or Eis that was identified by Friedman *et al.* in 2000.\(^5\)

One of the main differences between pathogenic and non-pathogenic mycobacteria is that non-pathogenic strains do not survive inside macrophages. As it was noted in Chapter 1, the resistance of *M. tb* to macrophage killing arises from different mechanisms such as inhibition of the phagosome-lysosome fusion, inhibition of the acidification of the phagosome, resistance to killing by reactive intermediates or modification of the composition of the cell membrane. In order to identify gene products involved in the resistance to macrophage killing, a plasmid library based on the genome
of the virulent *M. tb* strain H37Rv was introduced in the avirulent *Mycobacterium smegmatis*. The mycobacteria were then used to infect human macrophage-like cells. Selection of transformants yielded a clone with enhanced intracellular survival phenotype. The analysis of the clone revealed three open reading frames but only one with a promoter region and a ribosome binding site; the associated gene (Rv2416c) was named *enhanced intracellular survival (eis)*. The putative gene product, Eis, was characterized as a 42-kDa protein. To continue the characterization of *eis* and its product Eis, Friedman *et al.* investigated the subcellular localization of Eis in the mycobacterial cell.\(^7\) Mainly hydrophilic, Eis presents a small hydrophobic region, a possible transmembrane domain, and is not glycosylated. The protein was determined to be primarily present in the cytoplasm but to also appear in smaller concentrations in the cell envelope and the culture media. To further investigate the localization issue, pulmonary sera of infected patients were analyzed and showed that in forty percent of the cases Eis antibodies were present, confirming that the protein can also be released across the mycobacterial cell wall. The molecular characterization of the *eis* promoter revealed that the expression of the gene can be regulated.\(^8\) Work by Mariani *et al.* demonstrated that expression of Eis is increased following the activation of human macrophages infected with a clinical isolate strain of *M. tb*.\(^7\) The authors identified this phenomenon as a “counteraction” response to macrophage activation. In a similar fashion, induced stringent response in *M. smegmatis*, used as a model for *M. tb*, showed negative regulation of Eis expression.\(^8\) More recently, Friedman *et al.* characterized the production of Eis and its effects from an immunological standpoint.\(^8\) They confirmed that Eis belongs to the mycobacterial effectors of virulence since it is produced as early as four hours after
infection and it is released into the cytoplasm of the macrophage. Their work also demonstrated that Eis modulates the secretion of tumor necrosis factor α (TNF-α) and interleukin-10 (IL-10), changing the immunological state of the host. By increasing IL-10 and decreasing TNF-α levels, Eis might prevent cell apoptosis. Further immunological studies by Sharma and coworkers established that Eis inhibits T-cell proliferation in vitro. It also affects the production of cytokines such as TNF-α, interleukins 4 and 10 and interferon γ (IFN-γ).\textsuperscript{83} They confirmed Friedman’s conclusions: increased levels of IFN-γ and IL-10 and decreased levels of TNF-α and IL-4. The main consequence of modified levels of those particular cytokines is suggested to impair T cell regulation.

Finally, bioinformatic analysis of Eis revealed that it belongs to the GCN5 superfamily of N-acetyltransferases, which contains histone and aminoglycoside acetyltransferases.\textsuperscript{83} To date there is no report of a three-dimensional structure and there are only speculations about the exact function of Eis and its substrates.

2. Cloning, protein expression and purification of Eis.

The discovery of Eis is fairly recent. So far the study of the protein was limited to the immunological properties of the protein. From the nucleotide sequence, the different parameters of the gene and the protein have been computed as seen in Table 8.\textsuperscript{31}

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleotides</td>
<td>1209bp</td>
</tr>
<tr>
<td>Amino acid</td>
<td>402 amino acids</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>43.8kDa</td>
</tr>
<tr>
<td>pI</td>
<td>6.08</td>
</tr>
<tr>
<td>Extinction coefficient</td>
<td>52160 M\textsuperscript{-1}.cm\textsuperscript{-1}</td>
</tr>
</tbody>
</table>

Table 8      Predicted parameters of Eis using ProtParam tool on the ExPASy website.
The biochemical and biophysical characterizations of Eis is expected to help understand its exact function during pathogenesis. Once solved, the crystal structure should confirm the acetyltransferase domain and give insight on what the possible substrates are. The focus of this study was to obtain crystals and solve the crystal structure of Eis.

2.1. Cloning of the eis gene

When this project was initiated, there was no report in the literature of recombinant Eis. The initial research plan was to insert the gene into a pET vector and produce recombinant protein from *E. coli*. The initial molecular biology was carried out by a previous member of the laboratory. The *eis* gene was amplified by PCR using the following primers 5’- AGG AGG ATC CTC AGT GGT GGT GGT GGT GGT GGT GGA ACT CGA ACG CGG TCT GG -3’ and 5’- CAC CAT CAT CAT ATG CCA CAG TCG GAT TCT GTG A -3’. The gene was then introduced into the pET-28 plasmid vector using BamHI and NdeI restriction enzymes in a traditional digestion/ligation procedure. The resulting open reading frame produces a protein with an N-terminal 6X-Histidine tag. Unfortunately after sequencing of the pFLEIS plasmid, it turned out that there was a mutation in the plasmid leading to the R39L mutation in the protein sequence as shown in Figure 52.

In 2007 both Friedman and Sharma’s research group reported cloning of Eis from *M. tb* H37Rv and its overexpression in *E. coli* for immunological properties studies.\(^8^3\) Since attempts to replicate the published results were unsuccessful, a new strategy had to be devised. Attempts to amplify the *eis* gene using the initial PCR program and reaction
conditions were unsuccessful. New sets of primers were designed but unfortunately those reactions did not yield the desired PCR product. In addition to traditional PCR/ligation methods, Topo and Gateway technology were tried without success.

![Sequence alignment of the eis gene plasmid with the R69L mutation using BLAST software on the Expasy website](image)

At this point, a new strategy was adopted. Mutagenesis experiments were used to fix the mutation on the original plasmid. The primers used were 5'-CGC CTG GCG GAC CCT GGT GCC CAC CG-3' and 5'-CGG TGG GCA CCA GGG TCC GCC AGG CG-3'.

Figure 52  Sequence alignment of the eis gene plasmid with the R69L mutation using BLAST software on the Expasy website
The reaction mixture was composed of:

- 2 µL of template DNA pFLEIS,
- 39 µL of sterile water,
- 5 µL of Pfx50\textsuperscript{TM} 10x PCR buffer,
- 1 µL of each primer,
- 1 µL of Pfx50\textsuperscript{TM} DNA Polymerase,
- 1 µL of dNTPs mix.

The PCR experiment was run at 51 °C for 16 cycles. After DpnI digestion, the reaction mixture was transformed into XL10 competent cells onto a kanamycin agar plate. The cells were cultured overnight at 37 °C. The culture yielded two colonies. Small cultures were started from the two colonies. The cells were grown for 24 h at 37 °C and submitted to a plasmid preparation. The samples were sent to sequencing. Once the sequence was confirmed as seen in Figure 53 and Figure 54, Eis was purified.

2.2. Protein expression.

The plasmid pET28-Eis was transformed into T7 express cells. The bacterial cells were cultured in LB broth at 37 °C, induced with IPTG and harvested by centrifugation after 16 hours of incubation at 16 °C. The pelleted cells were resuspended in binding buffer (20 mM sodium phosphate pH 7.5, 300mM NaCl, 5 mM β-mercaptoethanol) and were stored at -80 °C until protein purification was initiated.
**Figure 53**  
N-terminus sequence of pET28-Eis. (The mutation is shown by an arrow)
Figure 54  C-terminus sequence of pET28-Eis

2.3. Protein purification

The frozen aliquot of resuspended cells was thawed at room temperature and then placed on ice. The catalytic hydrolysis of undesired components was performed by addition of 1 mM of lysozyme and 0.1 mM of DNase1 to the suspension. The catalysis step was performed for 20 min on ice prior to lysis of the cells using a sonicator. The obtained suspension was clarified by high-speed centrifugation. The pellet component was discarded and the supernatant was filtered. The protein was first crudely purified by immobilized metal affinity chromatography as shown on Figure 55. The fractions
containing Eis were pooled, transferred into dialysis membrane with addition of PreScission™ protease and dialyzed for 15-20 h at 4 °C against binding buffer. A Tag removal method was run to further purify the protein as presented in Figure 56. The flow through fractions containing Eis were pooled and dialyzed into crystallization buffer (10 mM Tris pH 8.5, 2 mM EDTA and 1 mM DTT). The protein was then concentrated. The final purification step uses size exclusion chromatography. The concentrated protein from the previous step was injected in the AktaPrime™ system. From the chromatogram shown in Figure 57, Eis is probably a tetramer in solution since it is eluted fairly early during the size exclusion chromatography. The fractions containing Eis were pooled and concentrated by centrifugation. The protein was then stored at 4 °C until crystallization set up. The yield of the protein preparation was low with about 1 mg/L of culture. The use of different buffer conditions did not improve the overall yield.
Figure 55  Nickel affinity chromatogram and corresponding SDS-PAGE
Figure 56  Tag Removal Chromatogram and corresponding SDS-PAGE
Figure 57  Size Exclusion Chromatogram and corresponding SDS-PAGE
3. Crystallization Studies

The general strategy adopted in the laboratory is to first screen numerous conditions using commercial screens. Nextal™ Classics, JCSG+, PEGs and AmSO₄ screens were set up using microbatch under mineral oil and Eis protein at a concentration of 1.5 to 2.2 mg/mL depending on the protein batches. Once set up, the crystallization trays were placed in an incubator at 18 °C. A few hits were identified with microcrystal formation. Two conditions from the PEGs screen gave single crystals shown in Figure 58: 0.2 M di-ammonium tartrate, 20% (w/v) PEG 3350 and 0.2 M di-ammonium phosphate, 20% (w/v) PEG 3350.

![Figure 58](image)

Figure 58 Initial hits visualized by optical microscopy (Nikon SMZ1500, magnification 10x)

Only the more promising hits with single crystals were pursued. A two-dimensional expansion experiment was prepared with varying concentrations of diammonium phosphate and PEG3350. The expansion experiments produced bigger crystals as seen in Figure 59. However those single bigger crystals were only observed
for a couple of days. They seemed to exist only transiently. Unfortunately the crystals did not diffract when screened with the macromolecular X-Ray diffractometer.

![Initial crystal hit visualized by optical microscopy (courtesy of Dr. T. Bigioni)](image)

Since Eis is believed to be an acetyltransferase, it might be useful to add a possible substrate such as the acetyl donor. Therefore, acetyl-CoA was added to the crystallization set-up. At a concentration of 1.25 mM, acetyl-CoA helped to stabilize the crystals, no longer transient. In some drops the crystals seemed to be slightly bigger as shown in Figure 60. When screened for X-Ray diffraction, the addition of acetyl-CoA did not help: no diffraction was observed. More additives are currently under investigation.
4. Conclusion

After difficulties were encountered to reproduce the published purification procedure, a new protocol was devised and Eis was successfully purified. The protocol was low yielding but enough protein was obtained to set-up crystallization trays. Crystals of about 50 µm were obtained. For the protein that was crystallized, however, no diffraction was observed on the home source.

5. Future work

The crystallization studies of Eis should be continued. The obtained hits are reproducible and the addition of additives seems to help stabilize the crystals. The screening of additives should be pursued in order to get bigger crystals that diffract.

It would be interesting to use biophysical experiments to further prove that Eis is an acetyltransferase. Isothermal calorimetry (ITC) experiment with acetyl-coA would reveal the binding affinity of Eis for acetyl-CoA revealing if it is indeed the acetyl donor. ITC could also be used to screen peptides as well as oligonucleotides to attempt to identify the possible substrates.
References


71. Rose, J. D.; Maddry, J. A.; Comber, R. N.; Suling, W. J.; Wilson, L. N.; Reynolds, R. C., Synthesis and biological evaluation of trehalose analogs as potential inhibitors of mycobacterial cell wall biosynthesis. *Carbohydrate Research* 2002, 337, (2), 105-120.


76. Cappelli, G.; Volpe, P.; Sanduzzi, A.; Sacchi, A.; Colizzi, V.; Mariani, F., Human Macrophage Gamma Interferon Decreases Gene Expression but Not Replication of Mycobacterium tuberculosis: Analysis of the Host-Pathogen Reciprocal Influence on


