INVESTIGATING THE STRUCTURE AND FUNCTION OF METALLO-β-LACTAMASES WITH DIRECTED EVOLUTION

by

Raymond Daniel Socha

B.Sc., The University of Guelph, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Biochemistry and Molecular Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2018

© Raymond Socha, 2018
The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Investigating the structure and function of metallo-β-lactamases with directed evolution

submitted by Raymond Daniel Socha in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Biochemistry and Molecular Biology

Examining Committee:

Dr. Nobuhiko Tokuriki
Supervisor

Dr. Harry Brumer
Supervisory Committee Member

Dr. Natalie Strynadka
Supervisory Committee Member

Dr. Steve Withers
University Examiner

Dr. Michael Murphy
University Examiner
Abstract

Metallo-β-lactamases (MBLs) are powerful enzymes capable of conferring pathogenic bacteria with effective resistance against all major classes of β-lactam antibiotics. Their continuing global dissemination, paired with a lack of therapeutic inhibitors, has combined to pose a significant threat to human health. This thesis aims to use an evolutionary perspective to better understand the structure, function, and behaviour of the MBLs. The comprehensive characterization of eight MBLs in three different host organisms, *Escherichia coli*, *Klebsiella pneumoniae*, and *Pseudomonas aeruginosa*, revealed that host specific constraints can limit the effective periplasmic expression of the enzymes, and as a result, might restrict the dissemination of MBLs to certain bacteria. The directed evolution of NDM-1, VIM-2, and IMP-1 for the provision of greater ampicillin resistance in *Escherichia coli* exposed the mechanisms by which MBLs may adapt to overcome these expression barriers, while revealing the critical role that the signal peptide plays in host adaptation. The subsequent directed evolution of the same three MBLs with two other β-lactam antibiotics, cefotaxime and meropenem, demonstrated the relative robustness of the family’s broad substrate specificity, as only two of seven complete trajectories featured a narrowing of specificity and changing the selection pressure on one of these trajectories swiftly restored broad specificity. The long-term genetic drift of VIM-2 under purifying selection at different thresholds revealed the plasticity of the MBL’s sequence and structure, but also the robustness of its activity and function. Overall, the results presented in this thesis contribute to our understanding of the MBL family and will help to develop better treatment strategies in the future.
Lay Summary

Metallo-β-lactamases are tiny molecular machines that destroy antibiotics. Due to the heavy medical use of antibiotics in recent years, more and more bacteria have learned how to build these machines and employ them for protection against the bacterial-killing action of the antibiotics. If this trend continues, it could lead to a substantial worldwide increase in the number of deaths from infectious diseases. Developing a solution for this problem has proven difficult due to the fact that there are many different kinds of these machines and we cannot reliably predict their behaviour. Here, I sought to determine how bacteria learn to effectively use these machines while also probing their structure and function, and investigating how they currently behave so that we may be able to predict how they will do so in the future. This work will aid in the development of better strategies to stop their progressive advance.
Preface

All work presented in this thesis was performed in the laboratory of my graduate supervisor, Dr. Nobuhiko Tokuriki, at the University of British Columbia in Vancouver, British Columbia, between 2012 and 2018 (the publication of this thesis).

Chapter 1 and 6 were written solely by me, and edited by my supervisor, Dr. Tokuriki. The experiments in Chapter 2 were conceived and designed by Dr. Tokuriki and me. I performed all experiments and wrote the chapter in collaboration with Dr. Tokuriki. The work presented in Chapter 3 was conceived and designed by Dr. Tokuriki and me. I performed all experiments and wrote the chapter entirely by me. Dr. Tokuriki subsequently edited the chapter.

The experiments in Chapter 4 were conceived and designed by Dr. Tokuriki and me. Parts of Chapter 4 were performed by John Chen, a PhD student in the same laboratory, and Jason Stephany, a technician in the laboratory of Doug Fowler at the University of Washington. Together, we created a comprehensive codon mutagenized library for the enzyme VIM-2. John Chen then performed the deep mutational scanning screen and data analysis. Jason Stephany performed the sequencing. I performed all other experiments and data analysis, interpreted the results, and wrote the chapter in full. Dr. Tokuriki edited the chapter.

The work presented in Chapter 5 was conceived and designed by Dr. Tokuriki and me. Parts of Chapter 5 were performed by three undergraduate students in the same laboratory: Ben Life, Linda Kiritchkov, and Ayşe Nisan Erdoğan. Together with Ben Life, we established the experimental mutagenesis and screening methods. Ben Life performed the first 25 rounds of genetic drift; I performed the second 25 rounds; Ayşe Nisan Erdoğan performed the remaining rounds of the experiment. Together with Linda Kiritchkov, we developed the fitness determination protocol, which Linda Kiritchkov used to measure the fitness of the first 50 rounds
of genetic drift. Ayşe Nisan Erdoğan subsequently used the method to characterize the remaining rounds of the experiment and measured the fitness of selected rounds under different antibiotic selection. I performed all other experiments and data analysis, interpreted the results, and wrote the chapter in full. Dr. Tokuriki edited the chapter.
**Table of Contents**

Abstract............................................................................................................................ iii

Lay Summary ................................................................................................................... iv

Preface............................................................................................................................. v

Table of Contents ........................................................................................................... vii

List of Tables .................................................................................................................. xiii

List of Figures ................................................................................................................ xvi

List of Symbols .............................................................................................................. xx

List of Abbreviations ................................................................................................... xxi

Acknowledgements ........................................................................................................ xxiv

Dedication ....................................................................................................................... xxvi

**Chapter 1: Introduction** ............................................................................................. 1

1.1 On the Importance of Antibiotics............................................................................... 1

1.2 β-lactam antibiotics .................................................................................................. 3
   1.2.1 Evolutionary history ......................................................................................... 3
   1.2.2 Initial discovery to mass production ............................................................... 4
   1.2.3 β-Lactam mechanism of action........................................................................ 5
   1.2.4 Initial resistance and expansion of β-lactam antibiotics .................................... 7

1.3 β-lactam resistance mechanisms ............................................................................... 9
   1.3.1 Modification of target enzymes ....................................................................... 9
   1.3.2 Restricted cellular entry and increased efflux ................................................ 10
1.3.3 Enzymatic degradation ................................................................. 11
1.4 Serine-β-lactamases ...................................................................... 12
  1.4.1 Origins, emergence, evolution, and spread ........................................ 12
  1.4.2 Structure and mechanism ............................................................... 13
1.5 Metallo-β-lactamases .................................................................... 15
  1.5.1 Evolutionary origins ................................................................. 15
  1.5.2 Clinical emergence ................................................................ 18
  1.5.3 Sequence and structural diversity ............................................... 19
  1.5.4 Catalytic mechanism ............................................................... 22
1.6 Current perspective on β-lactam antibiotics ................................... 24
1.7 Experimental evolution .................................................................. 25
  1.7.1 Directed evolution ................................................................ 25
  1.7.2 Genetic drift ........................................................................ 29
1.8 Thesis objectives and outline ......................................................... 32

Chapter 2: Hidden phenotypic variation limits the dissemination of metallo-β-lactamases to new host organisms ............................................................................................................. 34

  2.1 Summary .................................................................................. 34
  2.2 Introduction ............................................................................. 34
  2.3 Methods .................................................................................. 37
    2.3.1 Cloning and expression of MBLs in E. coli, P. aeruginosa, and K. pneumoniae ............ 37
    2.3.2 Determination of minimum inhibitory concentration (MIC) values ....................... 37
    2.3.3 Purification of Strep-tagged MBLs ........................................................................ 38
    2.3.4 Enzyme assays to determine kinetic parameters ...................................................... 39
    2.3.5 Cellular fractionation and the β-lactamase activity measurement in the cell lysate .... 39
Chapter 3: Host specific expression constraints may be overcome through adaptive evolution

3.1 Summary ........................................................................................................................................... 59
3.2 Introduction ......................................................................................................................................... 60
3.3 Methods ............................................................................................................................................... 62
  3.3.1 Creation of mutagenized libraries ................................................................................................. 62
  3.3.2 Application of selection .................................................................................................................. 64
  3.3.3 Measuring the minimum inhibitory concentration of selected variants ........................................ 64
  3.3.4 Sanger sequencing of selected variants ........................................................................................ 65
  3.3.5 Purification of Strep-tagged MBLs .................................................................................................. 65
  3.3.6 Enzyme assays to determine kinetic parameters .......................................................................... 65
  3.3.7 Cellular fractionation and β-lactamase activity measurement in the cell lysate .......................... 66
3.3.8 Calculation of the number of functional MBLs per cellular fraction ........................................ 66
3.3.9 Cloning, expression, and fitness testing of MBLs in the three host organisms ..................... 66
3.3.10 Replacement of the MBL signal peptides with the PelB leader sequence ........................... 66

3.4 Results ........................................................................................................................................ 67
3.4.1 Directed evolution of the MBLs resulted in increased resistance to ampicillin ...................... 67
3.4.2 Mutations predominantly accumulated in the signal peptide .............................................. 70
3.4.3 The catalytic efficiency of the enzymes increased only slightly during the evolution .......... 75
3.4.4 Increasing periplasmic expression drives the increase in resistance ................................. 78
3.4.5 Periplasmic expression is improved through increased cytoplasmic expression and a higher
efficiency of translocation .............................................................................................................. 80
3.4.6 Host specific constraints on expression can be overcome through adaptation ................. 85
3.4.7 Mutations in the signal peptide largely drive the adaptation for increased resistance ....... 86
3.4.8 Laboratory evolution predicts natural mutations ................................................................. 90

3.5 Discussion .................................................................................................................................. 93

Chapter 4: The broad substrate specificity of the metallo-β-lactamases is an ingrained trait,
but can be disrupted .......................................................................................................................... 96

4.1 Summary .................................................................................................................................... 96
4.2 Introduction ................................................................................................................................ 97

4.3 Methods ..................................................................................................................................... 98
4.3.1 Creation of mutagenized libraries ......................................................................................... 98
4.3.2 Application of selection .......................................................................................................... 98
4.3.3 Measuring the minimum inhibitory concentration for selected variants ............................ 99
4.3.4 Sanger sequencing of selected variants ................................................................................ 99
4.3.5 Purification of Strep-tagged MBLs ....................................................................................... 99
4.3.6 Enzyme assays to determine kinetic parameters .......................................................... 100

4.3.7 Deep mutational scanning of a VIM-2 comprehensive codon mutagenized library ........ 100

4.4 Results ................................................................................................................................. 101

4.4.1 Directed evolution resulted in increased resistance in all trajectories ......................... 101

4.4.2 The shift in substrate specificity is reflected in the changing catalytic ability of VIM-2 ... 104

4.4.3 Mutations accumulated predominantly in the signal peptide, but also at specific positions throughout the catalytic domain of each MBL ........................................................................ 107

4.4.4 The substrate specificity change is the sum of the effects of multiple mutations .......... 115

4.4.5 Broad substrate specificity was swiftly restored upon changing selection pressure ........ 120

4.5 Discussion .......................................................................................................................... 127

Chapter 5: The experimental genetic drift of VIM-2 reveals long term evolutionary trends
................................................................................................................................................. 131

5.1 Summary ............................................................................................................................. 131

5.2 Introduction ......................................................................................................................... 131

5.3 Methods ............................................................................................................................. 133

5.3.1 Creation of mutagenized libraries ............................................................................... 133

5.3.2 Application of purifying selection ............................................................................... 134

5.3.3 Determining the fitness of the drifted populations ....................................................... 134

5.3.4 Sanger sequencing selected variants and the development of a python program for sequence data analysis ........................................................................................................ 135

5.4 Results ................................................................................................................................ 135

5.4.1 Establishment of a comprehensive, long-term genetic drift experiment ..................... 135

5.4.2 The fitness of each trajectory reflected the conditions of the purifying selection .......... 138

xi
5.4.3 The mutational load steadily increased in each trajectory over the course of the genetic drift
.................................................................................................................................................... 143
5.4.4 The distribution of mutations reveals the adaptability of VIM-2 ................................................. 146
5.4.5 Broad β-lactamase activity is maintained by selection with a single antibiotic ....................... 156
5.5 Discussion .................................................................................................................................... 158

Chapter 6: Conclusion and Future Outlook.................................................................................... 166

6.1 General summary and conclusion ................................................................................................. 166
6.2 Future outlook ............................................................................................................................... 168
   6.2.1 Investigating phenotypic variation in regulatory elements...................................................... 168
   6.2.2 Examining the key features of evolutionary optimization processes .................................... 169
   6.2.3 Thoroughly investigating the broad substrate specificity of the MBLs under different
       selection conditions .................................................................................................................. 172
   6.2.4 Utilizing experimental genetic drift as a platform for long-term antibiotic resistance
       evolution ....................................................................................................................................... 173
   6.2.5 Creating and screening indel libraries ..................................................................................... 174
6.3 Conclusion .................................................................................................................................... 175

Bibliography .................................................................................................................................... 176

Appendices ....................................................................................................................................... 192

Appendix A Supplementary Figures and Tables ............................................................................ 192
Appendix B Sequencing Data ........................................................................................................... 208
List of Tables

Table 3.1: Mutations identified in the sequence round 18 variants from the NDM-1, VIM-2, and IMP-1 ampicillin trajectories ............................................................................................................ 71
Table 3.2: Measured kinetic parameters of the wild type and their round 18 evolved counterpart for NDM-1, VIM-2, and IMP-1, for seven β-lactam substrates .............................................................. 77
Table 3.3: Ampicillin MIC values (µg/mL) for wild type and PelB MBL constructs in the three tested organisms................................................................................................................................. 89
Table 4.1: Measured kinetic parameters for VIM-type natural and evolved variants .......... 106
Table 4.2: Catalytic domain mutations identified in the sequenced variants from the VIM-2 meropenem trajectory .................................................................................................................. 111
Table 5.1: Mutation location statistics produced from analysis of eight sequence variants from the 40th round of the T10, TOsc, and T1000 trajectories .................................................................................. 150
Table A.1: The sequence and structural similarity data for all MBL variants examined in Chapter 2............................................................................................................................................. 193
Table A.2: MIC measurements of all MBL variants for all tested antibiotics in three host organisms from Chapter 2.................................................................................................................. 194
Table A.3: Kinetic parameters for all eight MBL variants for CENTA, cefotaxime, cefazidime, meropenem, imipenem, ampicillin, and penicillin.................................................................................. 198
Table A.4: Measured copies of functional enzyme in the periplasm and cytoplasmic fractions for each variant in each organism ........................................................................................................... 200
Table A.5: Measured amounts of DNA, RNA, and protein in both the cytoplasm and periplasm for the eight MBL variants investigated in Chapter 2.......................................................................... 203
Table A.6: MIC measurements of all MBL variants with the PelB leader sequence for all tested antibiotics in three different host organisms from Chapter 2. .......................................................... 204
List of Figures

Figure 1.1: The molecular structures of the core of the β-lactam antibiotic classes and example β-lactam antibiotics from each class. ................................................................. 9

Figure 1.2: The general catalytic mechanism of the serine-β-lactamases. ......................... 14

Figure 1.3: Sequence similarity networks for the metallo-β-lactamase superfamily ............... 17

Figure 1.4: Multiple sequence alignment of selected B1-MBL amino acid sequences .......... 20

Figure 1.5: Representative structures of each metallo-β-lactamase family subclass .......... 21

Figure 1.6: The general B1 metallo-β-lactamase catalytic mechanism proposed in by Lisa et. al. in 2017. ........................................................................................................................................ 23

Figure 1.7: Schematic representation of directed evolution as applied to improving antibiotic resistance ........................................................................................................... 27

Figure 2.1: Measured minimum inhibitory concentrations (MICs) for the metallo-β-lactamases with representative β-lactam antibiotics in E. coli, P. aeruginosa, and K. pneumoniae .......... 44

Figure 2.2: Relationship between kcat/KM and ampicillin MICs for the metallo-β-lactamases in the three organisms. ...................................................................................................................... 46

Figure 2.3: Relationship between MIC, kcat/KM, and [Ep] for each MBL in the three organisms. ............................................................................................................................................. 49

Figure 2.4: The variation between MBLs at each level of expression compared to the variation in ampicillin resistance when expressed in E. coli ...................................................................................................................... 52

Figure 2.5: The effects of the replacement of the native signal peptide with the PelB leader sequence on MBL-conferred resistance .............................................................................................................. 54

Figure 3.1: The average absolute improvement in ampicillin resistance exhibited by the NDM-1, VIM-2, and IMP-1 ampicillin trajectories over 18 rounds of directed evolution .......... 69

xv
Figure 3.2: The average relative improvement in resistance exhibited by the NDM-1, VIM-2, and IMP-1 ampicillin trajectories over 18 rounds of directed evolution.................................................. 69
Figure 3.3: The distribution and frequency of mutations accumulated in the 18-round NDM-1, VIM-2, and IMP-1 ampicillin trajectories................................................................. 72
Figure 3.4: The frequency of nonsynonymous mutations per residue, per round, in the signal peptide and catalytic domain regions of the NDM-1, VIM-2, and IMP-1 ampicillin trajectories. .................................................................................................................................................. 74
Figure 3.5: The location of high frequency and fixed mutations from the 18th round of the NDM-1, VIM-2 and IMP-1 ampicillin trajectories identified on their respective structures.................. 75
Figure 3.6: A comparison between the improvement in catalytic efficiency and ampicillin resistance over the NDM-1, VIM-2, and IMP-1 ampicillin trajectories............................................ 78
Figure 3.7: The ampicillin MIC of variants from rounds 3, 6, 9, 12, 15, and 18 of VIM-2 and IMP-1 compared to their level periplasmic CENTA activity................................................................. 79
Figure 3.8: The molecular changes in the fractional expression of NDM-1, VIM-2, and IMP-1 throughout the 18 rounds of directed evolution............................................................................ 84
Figure 3.9: The fold-improvement between the evolved round 18 variants and wild type enzymes for each trajectory in three organisms................................................................. 88
Figure 3.10: The distribution of mutations in natural NDM, VIM, and IMP variants. ................... 92
Figure 4.1: The improvement in the average MIC in each trajectory over 18 rounds of selection for increased resistance to cefotaxime, meropenem, and ampicillin............................................. 102
Figure 4.2: The distribution of accumulated mutations in randomly selected variants from VIM-2 cefotaxime, meropenem, and ampicillin trajectories......................................................... 108
Figure 4.3: Mutations acquired by the 18th round of directed evolution in a purified VIM-2 variant, V2-MEM, which was selected for increased meropenem resistance. ................... 114

Figure 4.4: The relative enrichment under selection for cefotaxime, meropenem, and ampicillin, for each amino acid mutation at each of the 10 VIM-2 catalytic domain positions that had acquired a mutation in the V2-MEM round 18 variant.................................................. 117

Figure 4.5: The predicted and measured substrate specificity ratios for the V2-MEM variant from the 18th round of the VIM-2 meropenem trajectory................................................................. 119

Figure 4.6: The changing substrate specificity throughout the VIM-2 meropenem trajectory. . 121

Figure 4.7: The location of mutations in the round 21 variants from the VIM-2 meropenem trajectory after three rounds of selection for ampicillin resistance............................................... 123

Figure 4.8: The relative enrichment under selection for cefotaxime, meropenem, and ampicillin, for each amino acid mutation at each of the six VIM-2 catalytic domain positions that were mutated at high frequency in the three rounds of directed evolution towards increased ampicillin resistance........................................................................................................ 125

Figure 4.9: The predicted and measured substrate specificity ratios for the round 21 variants of the VIM-2 meropenem trajectory. .......................................................... 126

Figure 5.1: The selection threshold applied to each round for the T0, T1, T10, T1000, and TOsc trajectories.................................................................................................................. 137

Figure 5.2: The fitness of the drifted populations in the T1000 high functional threshold T10 low functional threshold, T1 sub-inhibitory threshold, and T0 control trajectories. ...................... 139

Figure 5.3: The fitness of the drifted populations in the TOsc oscillating trajectory............ 140

Figure 5.4: The fitness of every second round in each of the five trajectories as measured by EC50 value........................................................................................................... 142
Figure 5.5: The nucleotide and amino acid mutational rates for the five trajectories. ............ 144
Figure 5.6: The ratio of transitions to transversions for the five trajectories....................... 145
Figure 5.7: The location of non-synonymous mutations acquired by the 40th round of T10, T1000, and TOsc mapped on VIM-2’s primary sequence and compared to secondary structure. ................................................................................................................................. 148
Figure 5.8: The location of non-synonymous mutations acquired by the 40th round of T10, T1000, and TOsc mapped onto a 3D model of VIM-2’s crystal structure separately and combined........................................................................................................................................ 154
Figure 5.9: The location of unique non-synonymous mutations acquired by the 40th round of the TOsc trajectory, not present in the T10 and T1000 trajectories, mapped onto a 3D model of VIM-2’s crystal structure........................................................................................................................................ 155
Figure 5.10: The MICs of cefotaxime, meropenem and ampicillin for variants from the 60th round of the T10 and T1000 genetic drift trajectories............................................................................................................. 157
Figure A.1: Measured minimum inhibitory concentrations (MICs) for the metallo-β-lactamases with representative β-lactam antibiotics in E. coli, P. aeruginosa, and K. pneumoniae.......... 195
Figure A.2: The initial rates of reaction for each enzyme-substrate pair over the testable substrate range for BcII, IND-1, IMP-1, CcrA, VIM-1, VIM-2, SPM-1, and NDM-1 ................. 197
Figure A.3: Relationship between $k_{\text{cat}}/K_M$ and selected β-lactam antibiotic MICs for the metallo-β-lactamases in the three organisms .............................................................................................................. 199
Figure A.4: The variation within MIC, $k_{\text{cat}}/K_M$, and $[E_p]$ for each antibiotic, organism, and MBL. .................................................................................................................................................................................. 201
Figure A.5: Relationship between MIC and the product of $k_{\text{cat}}/K_M$ and $[E_p]$ for each MBL in the three organisms.................................................................................................................. 202
Figure A.6: Relationship between $k_{cat}/K_M$ and selected β-lactam antibiotic MICs for the three organisms harbouring the PelB-metallo-β-lactamase constructs .................................................. 205

Figure A.7: The initial rates of reaction for each enzyme-substrate pair over the testable substrate range for NDM-1, VIM-2, IMP-1, and a selected variant from the 18th round of each ampicillin trajectory, NA, VA, and IA .................................................................................................................. 206

Figure A.8: The initial rates of reaction for each MBL-substrate pair ........................................ 207

Figure B.1: A map showing the location of each mutation acquired in the NDM-1 directed evolution trajectories ........................................................................................................ 209

Figure B.2: A map showing the location of each mutation acquired in the VIM-2 directed evolution trajectories ........................................................................................................ 210

Figure B.3: A map showing the location of each mutation acquired in the IMP-1 directed evolution trajectories ........................................................................................................ 211

Figure B.4: A heat map displaying the relative enrichment of each possible amino acid mutation at each position of VIM-2 under selection from three different β-lactam antibiotics .......... 212

Figure B.5: A map showing the location of mutations identified in the sequenced variants of selected rounds from each genetic drift trajectory ........................................................................ 213
### List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>alpha</td>
</tr>
<tr>
<td>$\beta$</td>
<td>beta</td>
</tr>
<tr>
<td>$\mu$</td>
<td>micro</td>
</tr>
<tr>
<td>$%$</td>
<td>percent</td>
</tr>
<tr>
<td>$W$</td>
<td>fitness</td>
</tr>
<tr>
<td>$f$</td>
<td>function</td>
</tr>
<tr>
<td>$[E_p]$</td>
<td>periplasmic expression</td>
</tr>
<tr>
<td>$k_{\text{cat}}$</td>
<td>turnover number</td>
</tr>
<tr>
<td>$K_M$</td>
<td>Michaelis constant</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$</td>
<td>specificity constant</td>
</tr>
<tr>
<td>$&gt;$</td>
<td>greater than</td>
</tr>
<tr>
<td>$&lt;$</td>
<td>less than</td>
</tr>
<tr>
<td>EC$_{50}$</td>
<td>half maximal effective concentration</td>
</tr>
</tbody>
</table>
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMP</td>
<td>ampicillin</td>
</tr>
<tr>
<td>6-APA</td>
<td>6-aminopenicillanic acid</td>
</tr>
<tr>
<td>CAZ</td>
<td>ceftazidime</td>
</tr>
<tr>
<td>CCM</td>
<td>comprehensive codon mutagenesis</td>
</tr>
<tr>
<td>CTX</td>
<td>cefotaxime</td>
</tr>
<tr>
<td>DMS</td>
<td>deep mutational scanning</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxynucleic acid</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>ESBL</td>
<td>extended-spectrum β-lactamase</td>
</tr>
<tr>
<td>EXAFS</td>
<td>extended X-ray absorption fine structure</td>
</tr>
<tr>
<td>g</td>
<td>gravity</td>
</tr>
<tr>
<td>HGT</td>
<td>horizontal gene transfer</td>
</tr>
<tr>
<td>HMM-PBP</td>
<td>high molecular mass penicillin-binding protein</td>
</tr>
<tr>
<td>ID</td>
<td>identification</td>
</tr>
<tr>
<td>IMI</td>
<td>imipenem</td>
</tr>
<tr>
<td>K. pneumoniae</td>
<td><em>Klebsiella pneumoniae</em></td>
</tr>
<tr>
<td>LMM-PBP</td>
<td>low molecular mass penicillin-binding protein</td>
</tr>
<tr>
<td>MEM</td>
<td>meropenem</td>
</tr>
<tr>
<td>MBL</td>
<td>metallo-β-lactamase</td>
</tr>
<tr>
<td>MIC</td>
<td>minimum inhibitory concentration</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleic acid</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
</tbody>
</table>
OD\textsubscript{XXX} \hspace{5em} \text{optical density at XXX nm}

\textit{P. aeruginosa} \hspace{5em} \textit{Pseudomonas aeruginosa}

PDB \hspace{5em} \text{Protein Data Bank}

PCR \hspace{5em} \text{polymerase chain reaction}

PEN \hspace{5em} \text{penicillin}

PBP \hspace{5em} \text{penicillin-binding protein}

qPCR \hspace{5em} \text{quantitative polymerase chain reaction}

RMSD \hspace{5em} \text{root mean standard deviation}

RNA \hspace{5em} \text{ribonucleic acid}

SBL \hspace{5em} \text{serine-β-lactamase}

UBC \hspace{5em} \text{University of British Columbia}

Vs. \hspace{5em} \text{versus}

Amino acid abbreviations:

A (or Ala) \hspace{5em} \text{alanine}

C (or Cys) \hspace{5em} \text{cysteine}

D (or Asp) \hspace{5em} \text{aspartate}

E (or Glu) \hspace{5em} \text{glutamate}

F (or Phe) \hspace{5em} \text{phenylalanine}

G (or Gly) \hspace{5em} \text{glycine}

H (or Hist) \hspace{5em} \text{histidine}

I (or Ile) \hspace{5em} \text{isoleucine}
<table>
<thead>
<tr>
<th>Letter (or Abbreviation)</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>K (or Lys)</td>
<td>lysine</td>
</tr>
<tr>
<td>L (or Leu)</td>
<td>leucine</td>
</tr>
<tr>
<td>M (or Met)</td>
<td>methionine</td>
</tr>
<tr>
<td>N (or Asn)</td>
<td>asparagine</td>
</tr>
<tr>
<td>P (or Pro)</td>
<td>proline</td>
</tr>
<tr>
<td>Q (or Qln)</td>
<td>glutamine</td>
</tr>
<tr>
<td>R (or Arg)</td>
<td>arginine</td>
</tr>
<tr>
<td>S (or Ser)</td>
<td>serine</td>
</tr>
<tr>
<td>T (or Thr)</td>
<td>threonine</td>
</tr>
<tr>
<td>V (or Val)</td>
<td>valine</td>
</tr>
<tr>
<td>W (or Trp)</td>
<td>tryptophan</td>
</tr>
<tr>
<td>Y (or Tyr)</td>
<td>tyrosine</td>
</tr>
</tbody>
</table>
Acknowledgements

A PhD can often feel like a solitary experience, but reflection quickly reveals an army of supporters who have made it possible. First, I would like to thank my supervisor, Dr. Nobuhiko Tokuriki, for creating a stimulating and always challenging research environment that encouraged me to grow and become a strong problem solver and capable scientist. Further, his flexibility and open-mindedness allowed me to take control of my professional development, explore different career paths, develop new skills, and pursue unique opportunities not normally afforded to graduate students. As a result, I am leaving UBC a much better person than when I arrived.

My committee, Dr. Harry Brumer and Dr. Natalie Strynadka, have also provided guidance, valuable research advice, and sober second thought throughout these past few years and this thesis would not be what it is without them.

While the contribution of undergraduate students in research is often overlooked, the substantial impact that a team of co-op students has had on my PhD cannot go unacknowledged. Over the past few years, I have had the honour to work with Ben Life, Linda Kiritchkov, and Ayşe Nisan Erdoğan. They have laboured tirelessly to help obtain a treasure trove of data and create an experimental platform that should continue to serve the lab for many years to come. What they have achieved is a testament to their growing ability and commitment to hard work, and I wait in eager anticipation to see all of the amazing things that they will accomplish in the future.

I have also been most fortunate to have two wonderful mentors, Dr. Rob Maurus and Dr. Warren Williams, who have helped me to develop my teaching ability through their constant encouragement and helpful feedback during my tenure as a teaching assistant in ten of their
biochemistry courses. This helped me to mature from a nervous graduate student into a confident and effective teacher ready to take on new challenges. Dr. Katharina Rothe also imparted much teaching wisdom as we worked together to re-launch a comprehensive business course for non-business academics, and I will be forever grateful for her assistance in tackling such a challenging endeavour.

I would also like to thank the large and diverse group of graduate students and postdoctoral fellows in the Michael Smith Laboratories who have provided support, guidance, and mentorship during my time at UBC. Dr. Jennifer Rowland, Dr. Stefanie Vogt, and Dr. Charisse Petersen proved particularly helpful with providing a big-picture perspective on graduate school. Dr. Eric Brown and Dr. Chris Roach encouraged me to pursue a unique opportunity that has since developed into a new career path. Many others have also assisted or provided support throughout my time at UBC, including Dr. Dave Anderson, Kylynda Bauer, Dr. Alina Chan, Vivienne Chan, Linda Horianopoulos, Steve Hur, Kelsey Huus, Katelyn Knuff, Scott Lambie, David Morales, Stefanie Novakowski, Jorge Peña, Nina Radisavljevic, Dr. Lisa Reynolds, Lisa Thorson, Dr. Joris van der Heijden, Sarah Woodward, and Hansel Wu. Their contributions are often intangible, but their support has proved invaluable. Graduate school would certainly have been a lonelier and darker place without them.

Lastly, I would like to thank my family for their patience and support as I pursued my PhD on the other side of the country. My parents have always looked out for my best interest, even when I haven’t, and their support as proved critical. And Lauren, my friend who became my wife, I couldn’t have done this without you. Thank you for being eternally patient and supportive during this journey.
For Lauren and those who came before us.
Chapter 1: Introduction

1.1 On the Importance of Antibiotics

On the morning of 2 September 1918, Private John William Smith (#3106176) of the 14th Battalion advanced with the rest of the Canadian 1st Division under heavy fire to seize the Drocourt-Quéant line. In the intense fighting that followed, he received a gunshot wound to his right ankle, requiring immediate evacuation from the front. He was transported to No. 54 General Hospital, a temporary medical facility near Boulogne, where the bullet was extracted 48 hours later. On 7 September, Pvt. Smith was expatriated from France and admitted to the Royal Herbert Hospital in London, where he was to be monitored for the next three months. His temperature remained elevated at approximately 102°F until a minor operation on 3 October appeared to stabilize his condition. However, on 17 October, his fever spiked to 105°F and amputation of the leg became medically necessary. Six days after the operation, his temperature returned to normal (Library and Archives Canada 1917-1919; Cook 2008).

Pvt. Smith was one of 5,600 casualties taken by the Canadians on that day, but he was one of the fortunate ones: he had survived not just the battle, but bacterial infection in an era where 90% of battlefield wounds became infected with up to 10 different bacterial species, and the only treatments available were weak antiseptics, such as carbolic acid, hypochlorite solution, potassium permanganate, and salt water (Cameron 1907; Dakin 1915; Fraser et al. 1918). Despite the seemingly rudimentary nature of these treatments, the First World War saw some of the most advanced medical techniques at that time developed and implemented on a wide scale in an effort to reduce the exceptionally high mortality rates from bacterial infections (Runcie 2015). It spurred the interest of leading scientists in the search for compounds that could be used to treat infections, yet it took a serendipitous discovery 10 years later, and the urgency of the
Second World War, for effective antibiotics to be isolated and mass-produced. The discovery of antibiotics was hailed as one of the greatest medical triumphs and is commonly credited with the substantial decrease in mortality from infectious diseases in the 20\textsuperscript{th} century.

The total societal impact of antibiotics is incalculable. Penicillin alone is estimated to have saved tens of millions of lives since it was first introduced in 1943, while also driving the development of a number of beneficial technologies like microbial strain improvement and the different chemical engineering methods used to mass produce fermentation products (Kardos & Demain 2011). Yet some studies have challenged the role that antibiotics have played in the reduction of deaths caused by infectious diseases (Hemminki & Paakkulainen 1976; Guyer et al. 2000). For example, while infectious disease-related deaths dropped 99.7\% in the United States between 1900 and 2000, a 90\% decrease had already occurred by 1940, before the widespread use of antibiotics (Guyer et al. 2000). However, these studies likely fail to capture the full impact that antibiotics have had: they are focused upon developed countries where rapid increases in hygiene during the first half of the century may have muted the antibiotics’ potential impact, which may not be the case globally. They also do not account for the myriad of more advanced medical procedures that are now reliant upon the established use of antibiotics (\textit{i.e.}, invasive surgeries) (Runcie 2015).

Regardless, there is no debate that the loss of our modern arsenal of antibiotics would be catastrophic. In the intervening years between their first widespread use and today, significant bacterial resistance has emerged for virtually every clinically approved antibiotic. As a result, the Centers for Disease Control and Prevention (CDC) estimated in 2013 that each year, at least two million people are infected with antibiotic resistant bacteria in the United States alone, which results in at least 23,000 direct deaths (CDC 2013). As with the total impact of antibiotics in the
20th century, the impact of growing resistance is difficult to quantify, and current estimates vary drastically (Naylor et al. 2018). For example, a recent study attempted to estimate the total economic cost of antimicrobial resistance and determined that it could reduce the global population by 11 to 444 million by 2050, while costing the economy somewhere between $2.1 to $124.5 trillion (Taylor 2014).

As the number of new, effective antibiotics reaching the clinic slows, the necessity for understanding the nature of antibiotic resistance grows. To ensure the security of our treatment options, we need to understand (1) where antibiotic resistance comes from, (2) how antibiotic resistance behaves, and (3) how it might change in the future. This section will detail the history of antibiotic use and resistance, and outline how this thesis aims to understand the behaviour of current resistance determinants and predict future trends.

1.2 β-lactam antibiotics

1.2.1 Evolutionary history

Humans did not invent antibiotics. Most of the antibiotics used today, including β-lactam antibiotics, are semi-synthetic compounds that have been co-opted from microbial organisms and further modified to increase their efficacy (Abraham 1983). Where then did antibiotics originally come from and what was their biological function? The origin and role of antibiotics in nature has recently become controversial. The classical explanation is that antibiotics are weapons used in bacterial warfare. Bacteria are engaged in a constant battle over resources within complex microbial communities and have evolved a number of tools for attack and defence, including antibiotics (Abrudan et al. 2015). However, an alternative theory has emerged and gained popularity in recent years: antibiotics may actually be signaling molecules that coordinate
interactions between coexisting bacteria (Linares et al. 2006). This is based on the simple observation that bacteria are unlikely to produce antibiotics at clinically relevant concentrations and the efficacy of lower concentrations as antibacterial agents is questionable. While no clear consensus has emerged, it is likely that both occur to some degree in different situations, thereby complicating the matter.

### 1.2.2 Initial discovery to mass production

Of course, Alexander Fleming did not know of the evolutionary history of antibiotics in 1928 when a confluence of random events led to their accidental discovery. A mould, *Penicillium notatum* (now known as *Penicillium chrysogenum*), contaminated a plate that Fleming had seeded with staphylococci before his vacation. A ‘cold streak’ allowed the mould to become fully established on the plate before a return to seasonal weather promoted bacterial growth (Abraham 1983). Upon Fleming’s return, he noted a halo around the mould of lysed staphylococci and began a series of tests to investigate: He determined that broth in which the mould had been grown had gained bacteriolytic properties, and that when filtered, could be used to inhibit the growth of many common pathogenic bacteria (Fleming 1929). He named this new agent penicillin and further demonstrated that it was non-toxic for mice, rabbits, and humans. Unfortunately, penicillin proved difficult to isolate and the antibacterial activity was easily lost, masking the full medical potential from Fleming.

In 1938, Howard Florey and Ernst Boris Chain began a project to survey all known antibacterial agents produced by microorganisms and came across penicillin. By 1940, they had crudely, yet successfully isolated the compound and demonstrated that subcutaneous injections could save mice from fatal bacterial infections (Chain et al. 1940). A significant effort was made
over the next year to produce enough of the drug for a small clinical trial, which was carried out at Oxford with five patients suffering from severe streptococcal or staphylococcal infections (Abraham 1941). Despite extreme limitations with the low purity and available quantity of the antibiotic, this trial demonstrated the validity of treating bacterial infections with penicillin.

With the Second World War raging above the United Kingdom, Florey left for the United States in 1941 in the hopes that large-scale production would be possible there. With a substantial investment from the government, the Rockefeller Foundation, and several major pharmaceutical companies, in addition to input from many American collaborators, more penicillin was produced and several large-scale clinical trials were performed with favourable results (Demain & Elander 1999). The yield of penicillin was quickly increased 100-fold during this time through strain improvement and medium modifications, such as the use of deep fermentation instead of surface cultures, and within three years of Florey’s trip to the United States, Allied forces had enough penicillin to supply every casualty during the Normandy landings on D-Day (Abraham 1983).

1.2.3 β-Lactam mechanism of action

All antibiotics work to derail normal function and induce bacterial cell death or limit cell growth by disrupting critical biological processes while minimizing adverse effects on the patient. β-lactam antibiotics achieve this by inhibiting synthesis of the bacterial cell wall, a peptide cross-linked glycan network (peptidoglycan) that provides structural strength to bacteria and resists osmotic pressure from the cytoplasm (Strominger et al. 1959). Peptidoglycan consists of a disaccharide (N-acetylglucosamine-N-acetylmuramic acid) that is polymerized into long strands, which are intermittently connected together by cross-linked pentapeptides (Anderson et al. 1966;
Zhao et al. 2017). It forms a thin layer between the inner and outer membranes in Gram-negative bacteria and a thick outer layer in Gram-positive bacteria; however, the enzymes and pathways involved in the peptidoglycan biosynthesis are conserved in both. As it is critical for the survival of most bacteria, but absent in humans and animals, disruption of the cell wall is an ideal drug target for the treatment of bacterial infections (Bugg et al. 2011).

β-lactam antibiotics disrupt peptidoglycan synthesis by inhibiting the transpeptidase activity of the penicillin-binding proteins (PBPs), directly resulting in the loss of cell wall integrity and subsequent lysis (Park & Strominger 1957; Tipper & Strominger 1965). PBPs cross-link the glycan strands in peptidoglycan by catalyzing the connection of bridging pentapeptides. Specifically, a PBP will bind the donor glycan strand, allowing the active site serine to perform a nucleophilic attack on the terminal peptide bond of the pentapeptide (D-Ala-D-Ala), forming an acyl-enzyme intermediate and releasing the terminal D-Ala. An amino group from the pentapeptide of the acceptor glycan strand can then attack the acyl-intermediate to complete the transpeptidation, releasing the enzyme (Lovering et al. 2012). PBPs have been traditionally divided into two groups: high molecular mass (HMM) PBPs, which are responsible for the bulk of peptidoglycan synthesis, and low molecular mass (LMM) PBPs, which perform maintenance, remodeling, and recycling roles (Sauvage & Terrak 2016). HMM-PBPs are further divided into two subgroups based upon their sequence, structure, and function. Class A HMM-PBPs are bifunctional and possess two domains: an N-terminal transglycosylase domain and a standard transpeptidase domain. By contrast, class B HMM-PBPs are monofunctional and only exhibit transpeptidase activity. However, the general transpeptidase mechanism between the class A and B is conserved.
This permits β-lactam antibiotics to be effective against a wide-range of PBPs, although the efficacy varies between each specific antibiotic and enzyme pair (Nangia et al. 1996; Kocaoglu & Carlson 2015). β-lactam antibiotics mimic the PBP substrate (D-Ala-D-Ala), leading to the amide bond in the β-lactam ring being attacked in the same manner as of that in the pentapeptide. However, this instead results in irreversible formation of an acyl-enzyme product and complete inactivation of the enzyme. The loss of the PBPs not only leads to a significant disruption of the cell wall, resulting in the bacteria becoming susceptible to cell lysis and death, but may also incite a futile cycle of peptidoglycan synthesis and degradation, putting an additional burden on the cell and promoting enzyme-mediated lysis (Cho et al. 2014; Yang et al. 2017).

1.2.4 Initial resistance and expansion of β-lactam antibiotics

The first sign of resistance against β-lactam antibiotics was actually detected from an E. coli strain in 1940, before the wide spread clinical use of penicillin began (Abraham & Chain 1940). Increasing numbers of resistant staphylococci strains in hospitals soon followed its clinical introduction (Abraham 1983). However, a series of three major discoveries in the 1950s spurred the development of new β-lactam antibiotics that could overcome this resistance. First, the natural variation of the penicillin side chains was determined, and second, the core nucleus of the penicillin molecule, a five-membered thiazolidine ring fused to the 2’ and 3’ positions of the β-lactam ring and called 6-aminopenicillanic acid (6-APA), was efficiently obtained in high yields from Penicillin chrysogenum (Batchelor et al. 1959). This allowed the semi-synthetic production of a wide-range of penicillin-type (penam) β-lactam antibiotics, including methicillin, which proved to not only be effective against previously resistant staphylococci, but also exhibited
higher activity against a number of Gram-negative strains that were unaffected by the original penicillin. The third discovery was the isolation of a new type of β-lactam antibiotic, called Cephalosporin C, which was obtained from the fungi Cephalosporium acremonium (Abraham et al. 1953). Applying the same principles learned with penicillin, the core nucleus of this molecule was identified to be a dihydrothiazine ring attached to the β-lactam ring and a range of new first, and later second-generation cephalosporins were semi-synthetically produced. Like the new penicillins, these cephalosporins showed an extended spectrum of activity and are useful against many Gram-negative and Gram-positive bacteria.

However, as before, bacteria developed resistance to many of these new penams and cephalosporins, prompting the development of a third and fourth generation of cephalosporins, two entirely new classes of β-lactam antibiotics, and a range of inhibitors to prevent the enzymatic degradation of the antibiotics by bacterial enzymes. Clavulanic acid, sulbactam, tazobactam, avibactam were discovered as weak β-lactam antibiotics that function better as inhibitors by irreversibly binding to the enzymes that were degrading the other β-lactam antibiotics (Demain & Elander 1999). They are often prescribed in combination with other penams and cephalosporins in order to sustain their utility. Two new classes of β-lactam antibiotics were also developed at this time: the carbapenems and the monobactams (Papp-Wallace et al. 2011; Sykes & Bonner 1985). The carbapenem core features a 2,3 unsaturated ring fused to the β-lactam ring, whereas monobactams consist of a monocyclic β-lactam ring with a sulfonic acid group attached to the nitrogen. Unlike the penicillins and cephalosporins, which were derived from fungi, the carbapenems and monobactams were appropriated from bacteria (Sykes et al. 1981). And yet, despite the diverse array of β-lactam antibiotics assembled, resistance has emerged for every single compound in each of the four subclasses, as well as the
inhibitors (Figure 1.1). Even the carbapenems, ‘antibiotics of last resort’, have elicited resistance, putting the future utility of the β-lactam antibiotics is in jeopardy.

![Molecular structures of β-lactam antibiotic classes and example β-lactam antibiotics from each class.](image)

**Figure 1.1:** The molecular structures of the core of the β-lactam antibiotic classes and example β-lactam antibiotics from each class.

### 1.3 β-lactam resistance mechanisms

#### 1.3.1 Modification of target enzymes

β-lactam antibiotics irreversibly acetylate the transpeptidase domain of the PBPs. However, as previously mentioned, the efficiency at which they inactivate the enzymes varies, and is dictated by the specific antibiotic-enzyme combination (Kocaoglu & Carlson 2015). Some PBPs escape
β-lactam inhibition entirely: PBP2a from *Staphylococcus aureus* is one such enzyme that has been acquired through horizontal gene transfer and allows for broad-spectrum β-lactam resistance (Lim & Strynadka 2002). This enzyme, with low affinity for the majority of β-lactam antibiotics, has been a key driver behind a worldwide surge of methicillin-resistance *S. aureus* (MRSA) infections (Diekema et al. 2001). Despite the development of new antibiotics that were effective against this PBP2a, resistance is already beginning to emerge. Several mutations that decrease binding affinity and increase resistance have already been identified (Katayama et al. 2004; Mendes et al. 2012). Thus, PBP2a specifically highlights the two strategies for PBP-based resistance in bacteria: The recruitment of new PBPs that are unaffected, or, the evolution of endogenous PBPs to have lower affinity for the antibiotics (Zapun et al. 2008).

1.3.2 **Restricted cellular entry and increased efflux**

In addition to the modification of the target PBPs, bacteria may escape the action of antibiotics through the restriction of their entry into the periplasm or, through their increased export out of the periplasm via efflux pumps (Fernández & Hancock 2012). Antibiotics can enter the periplasm of Gram-negative bacteria through porins, which are large β-barrel membrane proteins that allow passive diffusion of molecules, discerned only by gross physiochemical properties, across the membrane (Kojima & Nikaido 2013). Bacteria can have several different porins (*e.g.*, *E. coli* has OmpF, PhoE, and OmpC), with each exhibiting slight differences in their selectivity for which molecules are allowed to pass through. Similarly, the β-lactam antibiotics themselves diffuse through a given porin at different rates (Yoshimura & Nikaido 1985). Bacterial regulation of the permeability of the outer membrane is very complex, but allows for several effective ways in which antibiotic resistance can arise. Often, resistance emerges through the reduction in
expression of the certain porins that are permitting the transport of the β-lactam antibiotics (Jaffe et al. 1982). The loss of these porins can be compensated by the increased expression of other endogenous porins. Several other porin-mediated resistance mechanisms also exist, including the expression of mutated porins with altered specificity and the use of channel blockers to prevent the antibiotics from entering through the porin (Pagès et al. 2008).

Efflux pumps present another route by which Gram-negative bacteria may effectively resist antibiotics. Efflux pumps allow the bacteria to export drugs, compounds, and even toxic proteins out of the cell. They consist of three components: an integral membrane protein, which is inserted in the inner membrane; a large channel protein, which is anchored in the outer membrane and acts as an ‘exit duct’; and a soluble periplasmic adaptor protein, which connects the two other components together (Eswaran et al. 2004). Compounds can enter the pump from both the cytoplasm and periplasm, with the periplasmic entry facilitating the exit of the β-lactam antibiotics. Bacteria can also carry multiple versions of the efflux pumps, with some being acquirable through horizontal gene transfer. Some pumps are specific to a single drug class, but others can actively transport a wide-range of drugs, giving rise to multi-drug resistance efflux pumps (Marquez 2005).

1.3.3 **Enzymatic degradation**

The last major pathway to resistance is through the enzymatic degradation of the antibiotics. Enzymes that break down β-lactam antibiotics are referred to as β-lactamases, of which there are two major types: the metallo-β-lactamases (MBLs), which rely upon one or two zinc ions for catalysis, and the serine-β-lactamases (SBLs), which employ an active-site serine for activity (Ambler 1980; Bebrone 2007). Both MBLs and SBLs inactivate β-lactam antibiotics through the
cleavage of the amide bond within the antibiotic’s β-lactam ring, and have widely disseminated to many different pathogenic bacteria through horizontal gene transfer. They are discussed in greater detail below.

1.4 Serine-β-lactamases

1.4.1 Origins, emergence, evolution, and spread

The SBLs are estimated to have arisen over two billion years ago, when they diverged from DD-peptidases into three distinct groups defined by their sequence similarity: A, C, and D (Hall & Barlow 2004). Despite their long evolutionary history and growing observations of β-lactam resistance throughout the 1950s, SBLs were not considered a threat until 1965, when a class A SBL, TEM-1, was identified in an Salmonella typhi isolate from an Athenian patient called Temoniera (Ruiz 2018). The identification of TEM-1 as a plasmid-encoded gene caused the most concern as it could be easily spread to different bacteria through horizontal gene transfer, and as predicted, it was subsequently found in Neisseria gonorrhoeae isolates worldwide (Datta & Kontomichalou 1965; Ashford et al. 1976). The arrival of TEM-1 was followed by the emergence of several other clinically-relevant SBLs over the next 20 years, including class A SBLs SHV-1 and CTX-M, class C SBL AmpC, and the class D SBL OXA-1 (Roupas & Pitton 1974; Bauernfeind et al. 1990; Hedges et al. 1974).

To combat these new threats, third-generation cephalosporins were heavily employed. While they proved to be effective for a short time, it was not long before these SBLs adapted to and could provide effective resistance against these new antibiotics (Turner 2005). The discovery of SHV-2 in 1985 heralded a new age of antibiotic resistance in which antibiotic resistance genes were shown to be rapidly evolving to adapt to the deployment of new antibiotics. These enzyme
variants were deemed ‘extended-spectrum β-lactamases’ (ESBLs) due to their expanded activity profile (Medeiros 1997). Currently, there are over 170 TEM-type, over 50 SHV-type, and over 40 CTX-M-type β-lactamases with expanded activity that includes all β-lactam antibiotics. Many of these variants are also unaffected by several β-lactamase inhibitors that had been developed to prolong the duration of the β-lactam antibiotics usefulness.

1.4.2 Structure and mechanism

The three classes of SBLs are so genetically diverged that all traces of homology at the sequence level have been lost, but nonetheless, their structural similarity supports a homologous origin (Hall & Barlow 2004). The overall structure consists of a β and an α/β domain, with the active site situated between the two. Four motifs are critical for substrate binding and catalysis and are found in each of the three classes: (1) an active-site serine, (2) a leaving group proton donor (a serine in class A and D, tyrosine in class C), (3) a conserved lysine residue, and (4), the omega loop.

All three classes of the SBLs share the same general mechanism for the hydrolysis of β-lactam antibiotics (Figure 1.2). First, a conserved active-site serine is deprotonated and performs a nucleophilic attack upon the amide bond, briefly forming a tetrahedral intermediate before the ring is cleaved and the tetrahedral intermediate breaks down into a transient acyl-enzyme intermediate. Second, an activated water molecule attacks the acyl-enzyme intermediate, releasing the product from the active site. The differences between the classes emerge in the specific details surrounding this mechanism, such as which residue donates a proton to the nitrogen leaving group and how the serine is initially deprotonated. For example, in TEM-1, Ser70 performs a nucleophilic attack on the antibiotic after Lys73, acting as a general base,
abstracts a proton from Ser70. Lys73 later transfers the proton to the thiazolidine ring nitrogen via Ser130. Glu166 acts as a general base and facilitates the activation of a water molecule to perform the deacylation (Strynadka et al. 1992).

![Diagram of catalytic mechanism]

**Figure 1.2: The general catalytic mechanism of the serine-β-lactamases.**

The SBLs rapidly evolved into ESBLs upon the widespread usage of third and fourth generation cephalosporins, inhibitors, and carbapenems in the 1980s. Many of the mutations found driving this transition are in the omega loop and active site of the enzymes. In TEM-type ESBLs, mutations have resulted in an increase in conformational heterogeneity in order to accommodate the side chains of the new, larger antibiotics. In CTX-M-type ESBLs, the mutations expanding their spectrum instead appear to affect the flexibility and conformation of the active site (Palzkill 2018).
1.5 Metallo-β-lactamases

1.5.1 Evolutionary origins

The metallo-β-lactamases, or class B β-lactamases, emerged from the MBL superfamily, an ancient group of enzymes that can be found throughout the genomes of bacterial, archaeal, and eukaryotic organisms. Members of the superfamily were found to perform 16 different biological functions in an early survey of the group, which included in addition to the β-lactamases, glyoxylases, oxidoreductases, arylsulfatases, polyketide synthases, enzymes that process the 3’-ends of mRNA transcripts, DNA repair enzymes, DNA uptake proteins, choline-binding proteins, alkylphosphonate uptake proteins, CMP-N-acetylneuraminic acid hydroxylases, other antibiotic resistance enzymes, alkylsulfatases, carbofuran hydrolases, methyl parathion hydrolases, and phosphodiesterases (Daiyasu et al. 2001). Eight additional distinct functional groups have since been identified, including dehalogenases, phosphorylcholine esterases, quorum quenching enzymes, nitric oxide reductases, dioxygenases, and quinone biosynthesis enzymes (Baier & Tokuriki 2014). Greater sequencing data will likely continue to increase the known number of functions that this superfamily can perform.

The large variation in function is reflected by the genetic diversity within the superfamily, with some members sharing no greater than 5% sequence identity (Baier & Tokuriki 2014). However, despite this wide genetic and phenotypic diversity, the structural similarity within the superfamily is strikingly conserved: the majority of variants feature a shared αβ/βα-fold and a mono- or binuclear metal-binding active site (Bebrone 2007).

Where did this diverse superfamily come from and how does it perform such a variety of biological functions? New enzymatic activities are thought to develop through the optimization of promiscuous functions that enzymes often exhibit at low levels (Jensen 1976). This principle
has been demonstrated artificially in the laboratory: directed evolution was used to turn a phosphodiesterase, that primarily catalyzes the hydrolysis of paraoxon, but also weakly catalyzes the hydrolysis of 2-naphthyl hexanoate, into a full arylesterase that efficiently catalyzes 2-naphthyl hexanoate, while only weakly hydrolyzing the paraoxon (Tokuriki et al. 2012). Indeed, this is likely how the MBL superfamily formed, with repeating cycles of duplication and optimization of promiscuous functions leading to gradual divergence. This hypothesis is also supported by a recent survey of promiscuous functions in the MBL superfamily, which showed that there was a high level of connectivity between the catalytic functions of its members (Baier & Tokuriki 2014).

While this explains the general mechanism by which the MBLs may have arisen from their superfamily, it does not provide a specific explanation. How did \( \beta \)-lactamase activity emerge from the MBL superfamily? Three subclasses of MBLs have been defined based upon sequence similarity and structural alignments: B1, B2, and B3 (Garau et al. 2004). Interestingly, phylogeny and analysis of a sequence similarity network for the MBL superfamily suggests that the B1 and B2 groups are relatively closely related, whereas the B3 MBLs likely developed \( \beta \)-lactamase activity independently (Hall et al. 2004; Baier & Tokuriki 2014) (Figure 1.3). The nearest neighbour to the B1 and B2 MBLs in the superfamily with a confirmed activity is a dehalogenase, whereas the B3 MBLs are closest to the dioxygenases and glyoxylases. Interestingly, in both cases, the nearest neighbours to the MBLs do exhibit weak \( \beta \)-lactamase activity. While it is possible that the MBLs evolved from ancestors with their respective neighbours’ function, the true ancestral activity for the MBLs remains unknown.
Figure 1.3: Sequence similarity networks for the metallo-β-lactamase superfamily. (Top) The representative sequence similarity network of the entire MBL superfamily, displaying 3,811 nodes that represent 71,489 (sequences with >40% identity merged into one representative node), showing wide divergence of the sequence relationships within the superfamily and the location of the MBL B1 (blue), B2 (light green), and B3 (red) families when generated by the Enzyme Function Initiative and visualized with an alignment score of 15. (Bottom) The sequence similarity network within the MBL B1 family with 1224 sequences visualized with a BLAST e-value cutoff of $10^{-55}$. The sequence of selected B1-MBLs is shown as large circles and highlighted with colours: BcII (yellow), IND-1 (purple), CcrA (black), NDM-1 (blue), VIM-1 (pink), VIM-2 (red), IMP-1 (green), SPM-1 (orange), and the B2 family (light green).
1.5.2 Clinical emergence

The first known MBL, BcII, was identified from a *Bacillus cereus* lab strain in 1966 (Sabath & Abraham 1966). For the next few decades, it remained a laboratory curiosity with no clinical relevance. However, following the sudden increase in the number of available β-lactam antibiotics during the 1980s, other B1 MBLs began to appear in clinical settings with increasing frequency. IMP-1, identified in a *Serratia marcescens* clinical isolate from Japan in 1994, was the first in a series of clinically relevant MBLs to be discovered in rapid succession (Osano et al. 1994). It was followed by VIM-1 in 1999 and SPM-1 in 2002, both from *P. aeruginosa*, and NDM-1 in 2008 from *K. pneumoniae* (Lauretti et al. 1999; Toleman et al. 2002; Yong et al. 2009).

Interestingly, it appears that these B1 MBLs are specifically driving nosocomial infections. This may be due to their mobilization onto plasmids, whereas B2 and B3 MBLs, like CphA or L1, are largely relegated to the genomes of their host organisms. Indeed, IMP- and VIM-type MBLs are encoded on gene cassettes, mobile genetic elements that consist of a gene and a recombination site. These cassettes are often contained in integrons with other antibiotic resistance genes, facilitating their transfer between different bacteria (Bennett 2008). By contrast, NDM-type MBLs have not been found in integrons, but are instead present on many different plasmids (Nordmann et al. 2011). Today, the rapid worldwide spread of the acquirable B1 MBLs has resulted in their dissemination to a number of pathogenic Enterobacteriaceae and other bacteria, including *Pseudomonas aeruginosa, Escherichia coli, Klebsiella pneumoniae, Serratia marcescens, and Acinetobacter baumannii* (Palzkill 2012). As with the SBLs, the MBLs have continued to rapidly evolve: over 40 variants of IMP-type and VIM-type MBLs with sequence
identity varying by up to 10% have been identified since their initial discovery (Mojica, Bonomo, et al. 2015).

1.5.3 Sequence and structural diversity

The MBLs are a strikingly divergent family. Variants within each subclass often only share between 20-40% amino acid sequence identity. For example, a multiple sequence alignment of eight selected B1 MBLs reveals that only 16 residues are conserved between them, a paltry ~7% of their entire sequence (Figure 1.4). A full pairwise comparison between these selected variants can be found in Appendix A, Table A.1. Between subclasses, the sequence identity is even less, which highlights the great evolutionary distance between the groups.

Despite this high sequence diversity, the MBLs exhibit remarkably high structural similarity overall. Enzymes from all three subclasses display the superfamily’s characteristic $\alpha\beta/\beta\alpha$ sandwich fold, with two zinc-binding sites nestled in between several active site loops (L3 and L10) (Palzkill 2012). The sequence diversity begins to emerge in the details of their structure. For example, the zinc binding sites vary slightly between each subclass, but are generally conserved within each group. The first zinc-binding site in the B1 MBLs is composed of three histidine residues, while the second contains an aspartate, cysteine, and histidine (Figure 1.5). Most B3 MBLs only differ in the replacement of the cysteine in the second zinc site with another histidine residue. B2 MBLs, by contrast, exhibit the greatest divergence: while the remnants of the first zinc site remain to some degree, these enzymes demonstrate a maximum level of activity with only a single zinc occupying the canonical second zinc site, which is composed of an aspartate, cysteine, and histidine, as in the B1 group (Bebrone 2007).
Figure 1.4: Multiple sequence alignment of selected B1-MBL amino acid sequences. Red highlighted residues indicate a fully conserved position between VIM-1, VIM-2, BclI, IND-1, CcrA, NDM-1, IMP-1, and SPM-1. Red text residues indicate similarity. Periods denote insertions and deletions.
Figure 1.5: Representative structures of each metallo-β-lactamase family subclass. Cartoon presentation of the crystal structures of B1 MBL VIM-2 (PDB ID 1K03), B2 MBL CphA (PDB ID 1X8G), and B3 MBL FEZ-1 (PDB ID 1K07) with close-up views of their respective active sites. Zinc ions are shown as spheres in light orange, with metal binding residues displayed as sticks.
1.5.4 Catalytic mechanism

The diversity present within the MBL active site structures has complicated the elucidation of their catalytic mechanism and resulted in several questions surrounding which zinc sites are essential for function, the identity of the nucleophile, proton donor, and intermediates, and how the substrate binds the enzymes (Meini, Llarrull, et al. 2015). The recent use of rapid-mixing techniques coupled to several spectroscopic techniques (EXAFS, NMR) has allowed for a comparison of β-lactam catalysis across the MBL subclasses and the proposal of the most comprehensive mechanism to date for all three subclasses (Lisa et al. 2017). The proposed mechanism for the B1 MBLs is shown in Figure 1.6.
Figure 1.6: The general B1 metallo-β-lactamase catalytic mechanism. The general reaction mechanism proposed in Lisa et. al. in 2017 features a branched pathway. (Lisa et al. 2017). The carbapenem core structure is shown as an example β-lactam.
1.6 Current perspective on β-lactam antibiotics

Currently, β-lactam antibiotics are one of the most commonly prescribed antibiotics in the world. For example, over 65% of injectable antibiotic prescriptions in the US between 2004 and 2014 belonged to this class of antibiotic (Bush & Bradford 2016). However, the usage of each β-lactam subclass has changed since their initial deployment. The initial resistance observed in the 1960s against the penams has become widespread, limiting their therapeutic use as a monotherapy. Early generation cephalosporins shared the same fate. While the development of several β-lactam inhibitors has allowed a few to continue to be used, some ESBLs have even evolved to evade inhibition and maintain effective resistance (Demain & Elander 1999). Third and fourth generation cephalosporins, as well as the monobactam aztreonam, still continue to serve an important role in the treatment of serious Gram-negative pathogens, yet growing resistance from the SBLs and MBLs also continue to threaten their effective use. Even carbapenems, the ‘antibiotics of last resort’, have elicited growing resistance in recent years (Bush 2013).

With no single β-lactam antibiotic free from resistance, their future utility is seriously threatened. Past strategies of dealing with resistance are unlikely to be effective, as existing resistance mechanisms have repeatedly shown an ability to adapt to new antibiotics. The decreasing attractiveness in the development of new antibiotics from existing scaffolds is likely reflected in the reduced investment from pharmaceutical companies (Cole 2014; Butler et al. 2017). Although the development of new inhibitors shows promise (King et al. 2016), there is no guarantee on how effective and for how long these compounds would be able to stem the tide. A better understanding of the current resistance mechanisms is required in order to design strategies that ensure the long lasting security of our healthcare system. However, one of the
largest challenges of understanding resistance is that it is not static: it evolves and adapts to our societal actions. This makes experimental evolution a vital component to understand how resistance determinants behave.

1.7 Experimental evolution

Experimental evolution, a method of artificially recreating the natural process of evolution in the laboratory, is a field that has been intricately linked with the development of antibiotics and the study of antibiotic resistance. In the early years of penicillin development, experimental evolution was used to isolate strains with improved antibiotic yields (Backus & Stauffer 1955; Vournakis & Elander 1983). Beyond this, antibiotic resistance genes, particularly SBLs, have often been utilized as model enzymes in modern experimental evolution systems due to the ease with which antibiotic resistance can be selected (i.e., survival screening). This has generated a wealth of data that not only validated the various experimental systems and provided deep insights into both molecular evolution and protein function, but has also revealed information regarding the behaviour of the antibiotic resistance determinants. Here, I detail two applications of directed evolution that have been used to investigate antibiotic resistance genes: directed evolution and genetic drift.

1.7.1 Directed evolution

Directed evolution consists of repeated rounds of mutagenesis and selection (Packer & Liu 2015). This simple yet powerful process is frequently employed to replicate natural selection in order to engineer proteins with desired features (Turner 2009; Bornscheuer et al. 2012). For example, a gene would be subjected to mutagenesis and the resulting mutated proteins screened
for a desired trait. Those variants that best display the trait are then selected and used as the template for the next cycle of directed evolution, while the remainder is discarded. In this way, the selected trait gradually improves among the population until it reaches the desired level. The two most commonly employed mutagenesis techniques are error-prone PCR, in which genes are duplicated by error-prone polymerases or in the presence of base analogues that induce mismatch mutations, and DNA shuffling, in which diverse pools of mutant genes are recombined to create new combinations of mutations. The subsequent selection of the mutant libraries is dependent upon what trait is desired. Past examples include using directed evolution to improve catalytic activity, binding, stability, expression, and function in non-natural environments (Arnold 1993; Cirino & Arnold 2002; Collins et al. 2003; Socha & Tokuriki 2013).

Aside from the obvious biotechnological applications, directed evolution can also be used to investigate the natural evolution of proteins, such as β-lactamases, while probing their structure and function. As demonstrated by the transition of SBLs into ESBLs in the 1980s, antibiotic resistance determinants are continuing to evolve under selective pressure from increasing antibiotic use (Medeiros 1997). Directed evolution can mimic this process, allowing for the identification and characterization of future potential mutations and evolutionary pathways (Hall 2004b) (Figure 1.7). This information could allow researchers and clinicians to act in advance of clinical resistance trends. For example, it could inform more robust treatment plans and provide direction on research efforts.
Figure 1.7: Schematic representation of directed evolution as applied to improving antibiotic resistance. The initial gene is mutagenized, creating a diverse pool of mutants that are expressed in bacteria. The fitness of the variants is determined by plating their host bacteria on a series of agar plates with an increasing concentration of antibiotics. The variants that are the most fit (i.e., those conferring the highest level of resistance to their host and facilitating their growth on the highest concentration) are selected and used as a template for the next round. As the level of resistance improves, the range of antibiotic concentrations screened is gradually shifted upwards to match.

This strategy was validated in 1998 when a clinical isolate of *K. pneumoniae* was found to harbour TEM-52, a new TEM-type SBL variant with higher β-lactam resistance (Poyart et al. 1998). TEM-52 featured a combination of three mutations that had been previously characterized four years earlier in a laboratory after the triple combination was obtained through the random DNA shuffling of other TEM-type variants (Stemmer 1994). This triple combination mutant was later independently obtained by selecting TEM-1 for increased cefotaxime resistance using error-prone PCR (Orenicia et al. 2001). Also in 1998, directed evolution with error-prone PCR was able to recapitulate all four of the natural single mutations that conferred resistance to a combination ampicillin-clavulanate treatment (Vakulenko et al. 1998). While these experiments
were limited in scope (i.e., one target enzyme, one selected activity, few rounds of directed evolution), a systematic experiment in 2002 thoroughly demonstrated that directed evolution was able to mimic natural evolutionary processes when the selection of mutated TEM-1 libraries with third and fourth-generation cephalosporins and aztreonam resulted in the recovery of seven of nine mutations that have arisen independently multiple times in nature, with six recovered in multiple independent trajectories (Barlow & Hall 2002).

Many more directed evolution experiments on the β-lactamases have followed in the years since (Barlow & Hall 2003b; Barlow & Hall 2003a; Novais et al. 2008; Guthrie et al. 2011; Salverda et al. 2011). Unfortunately, most have focused solely upon TEM-1: by 2010, 25 separate TEM-1 laboratory evolution studies had been performed alone (Salverda et al. 2010). For those experiments that did look beyond TEM-1 to other β-lactamases, they were not particularly thorough. For example, the selection of IMP-1 for increased resistance to imipenem in a single round of directed evolution failed to return any mutations that improved imipenem resistance (Hall 2004a). While it was concluded under the circumstances that IMP-1 could not evolve to provide increased resistance, the experiment neglected the examination of double or triple mutations that might combine to provide higher resistance. In other situations, the result is often hard to translate to an effective strategy to combat resistance. A recent study in which the MBL AIM-1 was subjected to five rounds of directed evolution with selection for increased cefoxitin resistance was able to identify a single mutation that could greatly improve resistance, but was unable to determine a mechanism for how it did so (Hou et al. 2017).

Moving beyond the simple prediction of mutations, several studies have used adaptive directed evolution of the β-lactamases to deeply probe their structure and function. Tomatis et al. identified four point mutations that occurred with a high frequency in four rounds of
evolution attempting to increase the cephalexin resistance provided by BcII (Tomatis et al. 2005). A thorough characterization of the mutations was able to reveal the mechanism by which a second-shell ligand mutation (G262S) could alter the substrate specificity of the enzyme and facilitate the increased cephalexin resistance. They were further able to demonstrate that the combination of G262S and another one of the mutations, N70S, created an additive effect that further improved resistance and produced a broader substrate spectrum (Tomatis et al. 2008).

While this experiment provided deep insight into the structure, function, and substrate specificity of BcII, whether this result is transferable to the remainder of the B1-MBL family is unknown, as the high level of diversity between MBLs complicates direct comparisons. In order to provide pertinent information that may be widely applied to the family, a more systematic approach is required.

Overall, while these experiments were found to generally mirror the natural evolution of their target enzymes, they provide little actionable information for the remaining β-lactamases. If accurate predictions that can influence treatment strategies and research efforts are to be made, adaptive directed evolution will have to be applied in a systematic way to more β-lactamases under more conditions.

1.7.2 Genetic drift

The second mechanism by which genetic variation can become fixed in nature is through genetic drift, in which neutral (or sometimes deleterious) mutations can spread through a population by chance. Genetic drift can also be easily replicated with experimental evolution by adapting the directed evolution method. Instead of gradually increasing selective pressure in order to obtain
beneficial mutations, the pressure is kept constant at a specific level, which allows neutral mutations to accumulate.

Experimental evolution has only been used sparingly over the past two decades to replicate genetic drift or neutral evolutionary processes. The objectives of these experiments are often concerned with probing the structure, function, and evolutionary potential of proteins. In 2006, Bershtein et al. performed 10 rounds of random mutational drift with a purifying selection at three levels of selection (high, low, and none) on the SBL TEM-1, and determined that contrary to the many results from mutational studies at the time, most mutations do in fact have a deleterious effect on fitness (Bershtein et al. 2006). The deleterious effects of most mutations in single-mutant studies are usually masked by the threshold robustness (i.e., margin of excess stability and function). However, genetic drift led to the accumulation of multiple deleterious mutations, reducing this margin so that the majority of the deleterious effects could be seen. This implied that threshold robustness is inherently linked to the epistatic interactions between residues.

This was followed by another experiment from the same lab in which the DNA methyltransferase M.HaeIII was subjected to 17 rounds of genetic drift (Rockah-Shmuel et al. 2015). Deep sequencing of the libraries revealed that deleterious mutations arose at a high frequency in the early rounds of the experiment, but were substantially reduced by the 17th generation. Essentially, genetic drift allowed for the cancellation of the ‘noise’ that comes from the margin of excess stability and function, and for the subsequent elucidation of the relative fitness effects of all possible nucleotide mutations. 67% of mutations were shown to be deleterious, with a further 16% likely to be deleterious. This showed that genetic drift can be a
good method for the measurement of the fitness effects of protein mutations that better replicates natural evolutionary forces.

The next year, two independent experiments demonstrated that genetic drift could facilitate the accumulation of mutations that alter the adaptive potential of an enzyme for another promiscuous function, without compromising the primary function of the enzyme. Amitae et. al. mutagenized the serum paraoxonase, PON1, and selected neutral variants exhibiting wild-type lipo-lactonase activity and expression using in vitro compartmentalization in double emulsions and FACS (Gupta & Tawfik 2008). Three rounds of genetic drift, and the subsequent characterization of 311 neutral variants, showed that mutations that are neutral with respect to PON1’s native lactonase function could alter its promiscuous acyl esterase, phosphotriesterase, and aromatic lactonase promiscuous functions, thereby modulating the evolutionary potential for adaptation to these new functions. This was demonstrated with cytochrome P450 the same year, and beta-glucoronidase shortly after (Bloom, Romero, et al. 2007; Smith et al. 2011).

Another intriguing observation made from a genetic drift experiment also occurred in 2007, when Bloom et. al., continuing their drift experiment with cytochrome P450, determined that genetic drift of sufficiently large populations favours mutational robustness of proteins (Bloom, Lu, et al. 2007). This was again confirmed shortly later by Bershtein et. al. with their continuing TEM-1 genetic drift experiment, which provided a mechanism by which this occurs: genetic drift results in the accumulation of ‘global suppressor’ mutations that increase the stability of the enzymes and allow for increased adaptive potential (Bershtein et al. 2008).

While the previous experiments seemed to support the notion that genetic drift plays an extensive role in the expansion of genetic diversity, Petrie et. al. used a continuous evolution system to drift two different RNA ligase ribozymes and demonstrated that the distribution of
genotypes throughout the experiment remained clustered around the initial sequences. This suggests that neutral drift may not be the primary driver of genetic change – though it does provide an initial increase in genetic variation that may be selected for during adaptation (Petrie & Joyce 2014). Whether this phenomenon is common or restricted to this specific platform with this specific level of selection is unknown. Our understanding of the dynamics of genetic drift is still incomplete due to the limited number of genetic drift experiments that have been performed and the short time-scales they are often restricted to.

1.8 Thesis objectives and outline

The rapid emergence and global dissemination of the metallo-β-lactamases, enzymes capable of granting pathogenic bacteria with resistance against almost all β-lactam antibiotics, is threatening healthcare systems worldwide. With reduced numbers of new antibiotics making it from bench to bedside, the necessity of better understanding the nature of antibiotic resistance in order to reduce its significant societal cost has never been greater. The aim of this thesis is to employ directed evolution and related techniques to investigate the B1-MBLs, leading to a better understanding of their origin, current structure and function, and potential future.

Chapter 2 describes the functional characterization of eight B1-MBLs (BcII, IND-1, IMP-1, CcrA, VIM-1, VIM-2, NDM-1, and SPM-1) in three commonly associated host organisms (E. coli, K. pneumoniae, and P. aeruginosa). I identify how genetic diversity in the family manifests as hidden phenotypic diversity, and determine how cellular expression systems may present as a barrier to the successful dissemination of family members to new bacteria.

Chapter 3 describes the directed evolution of three B1-MBLs (NDM-1, VIM-2, and IMP-1) towards improved ampicillin resistance in E. coli and the subsequent characterization of the
evolved variants. I identify how MBLs may respond to continued exposure to β-lactam antibiotics and determine that improved resistance is largely driven by adaptation to the host organism.

Chapter 4 describes the directed evolution of the same three B1-MBLs (NDM-1, VIM-2, and IMP-1) towards improved resistance for different β-lactam antibiotics (cefotaxime and meropenem) in *E. coli* and the subsequent characterization of a VIM-2 evolved variant that exhibited an altered substrate specificity pattern. I identify the mutations that drive this substrate specificity change and determine that it is a product of epistatic interactions.

Chapter 5 describes the set-up and performance of a long-term genetic drift experiment, in which an evolved VIM-2 population (from Chapter 3) was propagated at five different selective thresholds for over 60 rounds of directed evolution. I determined that even a low antibiotic selection could maintain a population of MBLs at a clinically relevant level of fitness over long evolutionary time periods.

Chapter 6 provides an overview of the major conclusions drawn from this work, and presents several paths forward for the further investigation of the MBLs.
Chapter 2: Hidden phenotypic variation limits the dissemination of metallo-β-lactamases to new host organisms

2.1 Summary
Understanding how phenotypic variation can shape the distribution of antibiotic resistance genes throughout the microbial world is critical for the development of strategies to prevent further dissemination. I sought to determine the extent of hidden phenotypic variation in the B1 Metallo-β-Lactamase (MBL) family and how it may affect successful horizontal gene transfer by systematically characterizing eight MBL orthologs in three organisms representative of common pathogenic hosts (E. coli, P. aeruginosa, and K. pneumoniae). MBLs provide diverse levels of resistance in each organism. Catalytic ability does not sufficiently explain this diversity alone; rather, the amount of functional periplasmic enzyme is needed in conjunction to predict the variation in resistance. Functional periplasmic expression was found to vary dramatically between both MBLs and hosts. This diversity emerged in the quantity of mRNA, the amount of MBL expressed, and the amount of functional enzyme translocated to the periplasm. Thus, the interaction between each MBL and each host’s underlying cellular processes (transcription, translation, and translocation) determines the amount of functional periplasmic enzyme. These host-specific processes may constrain the effective spread and deployment of MBLs to certain hosts and this may explain their current distribution bias in nature.

2.2 Introduction
Orthologs are genetically diverged genes that perform the same functional role in different host organisms. The genetic variation among orthologs is thought to be largely neutral and acquired
through genetic drift, a process in which random mutations accumulate by chance (Kimura 1968). However, some genetic changes may still be driven by adaptation. For example, genetic variation may result from adaptation to transient historical changes in the level of selection pressure or in the cellular environment (e.g., pH, temperature, the concentration of nutrients and antibiotics). Moreover, some genetic changes during genetic drift may be driven by co-evolution with other proteins and molecules within each organism (Lovell & Robertson 2010; Ghadie et al. 2017). While these sequence changes can be neutral within each organism, they may prove deleterious or advantageous when the proteins are expressed in other host organisms, revealing their “hidden phenotypic variation”. This is most frequently demonstrated during heterologous protein expression in conventional laboratory host organisms such as *Escherichia coli*, which often results in low or even no protein expression due to incompatibility with the host because of reasons such as low codon optimality, the lack of endogenous chaperones, existence of heterologous proteases, or other molecular constraints (Lambertz et al. 2014; Bershtein et al. 2013; Bershtein et al. 2012). Such hidden phenotypic variation may also play an important role for genes that are transferred between organisms through horizontal gene transfer (HGT), such as antibiotic resistance genes, by restricting the ability of the gene to transfer. These sequence changes appear neutral within the organism, but may prove deleterious upon expression in other host organisms (Porse et al. 2018; Bershtein et al. 2015). However, our understanding of how and to what extent hidden phenotypic variation affects the dissemination of antibiotic resistance determinants to different organisms is still limited.

The B1 Metallo-β-Lactamases (MBL) family is one such group of antibiotic resistance genes that has been subjected to extensive dissemination through HGT to a wide variety of organisms (Maltezou 2009). Despite high levels of genetic variation (pairwise amino acid
identities as low as 20%), these orthologs feature a shared αβ/βα-fold, an identical active-site architecture (two zinc binding sites coordinated by H-H-H and D-C-H), and the ability to confer resistance to most β-lactam antibiotics through the hydrolysis of the β-lactam ring (Bebrone 2007). First isolated from the genome of *Bacillus cereus* in 1966, MBLs have since been identified in a growing number of bacteria, including: *Bacillus fragilis, Chryseobacterium indologenes, Empedobacter brevis, Serratia marcescens Shewanella frigidimarina, and Shewanella livingstonensis* (Cuchural et al. 1986; Bellais et al. 1999; Osano et al. 1994; Poirel et al. 2005; Bellais et al. 2002). However, in the last two decades, some MBLs, such as NDM-type, VIM-type, IMP-type, and SPM-type have become acquirable, and have subsequently disseminated through HGT to diverse Gram-negative pathogens, such as, *Pseudomonas aeruginosa, Acinetobacter baumannii, Klebsiella pneumoniae, Enterobacter cloacae*, and *Escherichia coli*, (Bebrone 2007; Walsh et al. 2005). This has hastened the spread of multi-drug resistance pathogens that threaten healthcare systems worldwide.

Interestingly, these acquired genes show a bias with respect to the bacterial hosts that harbour them. For example, SPM-type MBLs have been noted for their strong association with *P. aeruginosa* (González et al. 2014), whereas NDM-type MBLs are predominantly found in *K. pneumoniae* and *E. coli* (Berrazeg et al. 2014). Are there molecular differences that account for this bias? Moreover, what molecular features separate the acquirable MBLs from those that remain chromosomally encoded? Understanding the genetic and molecular causes underlying the features of highly transferable genes is important in developing our ability to prevent and control dissemination of multi-drug resistance genes.

Here, I conducted a comprehensive characterization of eight MBL enzymes to determine their resistance level in three bacterial hosts, and unveil to what extent genetic variation causes
hidden phenotypic variation among these MBLs. I further performed diverse biochemical and biophysical characterizations of the enzymes to reveal the underlying molecular basis for such hidden phenotypic variation.

2.3 Methods

2.3.1 Cloning and expression of MBLs in E. coli, P. aeruginosa, and K. pneumoniae

The eight MBL genes were synthesized (BioBasic) and subcloned in a broad-host-range vector, pBBR1MCS-2, along with the P_{BAD} promoter, which is inducible in the presence of arabinose (5.9-vector). The plasmids were transformed with the three bacterial strains: E. coli 10G (chemical transformation), P. aeruginosa PA01 (electroporation) (Choi et al. 2006), and K. pneumoniae ATCC13883 (electroporation) (Fournet-Fayard 1995).

2.3.2 Determination of minimum inhibitory concentration (MIC) values

To determine the MIC for each β-lactam antibiotic, single colonies of E. coli 10G, P. aeruginosa PA01, and K. pneumoniae ATCC 13883 that were transformed with the MBL plasmids were picked, grown in quadruplicate overnight at 30°C in a 96-well plate in 500 µL of LB medium with 1% glucose to suppress expression and 40 µg/mL of kanamycin for selective resistance. The overnight culture was used at a 1:40 dilution to start a new 200 µL culture of LB medium supplemented with 40 µg/mL of kanamycin and 0.02% arabinose to induce expression. The cells were grown for 6 hours at 37°C before being transferred with replicator pins onto a series of LB agar plates containing two-fold increases in the concentration of antibiotic and 0.02% arabinose and grown overnight at 37°C (the range of concentrations screened for each antibiotic was as follows: cephalosporins, 0.032 µg/mL to 4096 µg/mL; carbapenams, 0.016 µg/mL to 64 µg/mL;
penams, 2 µg/mL to 32768 µg/mL). The MICs were determined at the concentration of antibiotics by which no growth was observed at least three of the four replicates.

2.3.3 Purification of Strep-tagged MBLs

All MBL variants were cloned into a pET-26(b) vector without their signal peptide and with a C-terminal Strep-tag (GNSGSAWSHPQFEK). Each enzyme was expressed in *E. coli* BL21 (DE3) cells in TB auto-induction medium (EMD Millipore) supplemented with 1% (w/v) glycerol, 200 µM ZnCl₂, and 40 µg/mL kanamycin. 200 mL cultures were inoculated with 5 mL of overnight culture (LB medium, 40 µg/mL of kanamycin) and incubated at 30°C for 6 hours before further incubation at 18°C for 10 hours. Cells were harvested by centrifugation at 3200 × g and pellets were frozen at -80°C overnight. Cell pellets were resuspended in the lysis buffer containing 50% B-PER protein extraction reagent (Thermo Scientific) in Buffer A (50 mM Tris–HCl (pH 7.5), 100 mM NaCl and 200 µM ZnCl₂) and 100 µg/mL of lysozyme, and incubated on ice for 1 hour. The cell lysates were centrifuged at 25,000 × g for 30 minutes at 4°C and the Strep-tag fusion proteins were purified from the clarified lysate according to the manufacturer’s instruction with Strep-tactin resin (IBA Lifesciences). The purified protein solution was desalted using Econo-Pac 10DG Column (Bio-Rad) and eluted in 4 mL of Buffer H (20 mM HEPES pH 7.5, 100 mM NaCl₂, 200 µM ZnCl₂). The concentration of each protein was determined by spectrophotometer. The A₂₈₀ was measured for each sample with the following extinction coefficients, which were calculated with ExPASy ProtParam (Gasteiger et al. 2005): BcII, 34,950 M⁻¹cm⁻¹; IND-1, 40,910 M⁻¹cm⁻¹; IMP-1, 50,420 M⁻¹cm⁻¹; CcrA, 46,410 M⁻¹cm⁻¹; VIM-1, 33,920 M⁻¹cm⁻¹; VIM-2, 35,410 M⁻¹cm⁻¹; NDM-1, 33,460; SPM-1, 36,440 M⁻¹cm⁻¹.
2.3.4 Enzyme assays to determine kinetic parameters

The catalytic ability of each MBL enzyme was measured for seven β-lactam substrates (CENTA, cefotaxime, ceftazidime, meropenem, imipenem, ampicillin, and benzylpenicillin) in Buffer H supplemented with 0.2% Triton X-100. The rate of hydrolysis of CENTA was determined by measuring changes in the absorbance at 405 nm (the extinction coefficient of CENTA at 405 nm is 6400 M\(^{-1}\)cm\(^{-1}\)). The rate of hydrolysis of the β-lactam ring in the six antibiotics was determined by measuring changes in absorbance at the following wavelengths with these reported extinction coefficients: cefotaxime, 260 nm, 7500 M\(^{-1}\)cm\(^{-1}\); ceftazidime, 260 nm, 9000 M\(^{-1}\)cm\(^{-1}\); meropenem, 300 nm, 6500 M\(^{-1}\)cm\(^{-1}\); imipenem, 300 nm, 9000 M\(^{-1}\)cm\(^{-1}\); ampicillin 235 nm, 820 M\(^{-1}\)cm\(^{-1}\); and benzylpenicillin, 235 nm, 775 M\(^{-1}\)cm\(^{-1}\). The initial rates of reaction were measured in triplicate over the range of substrate concentrations (1 µM to 400 µM for CENTA, cefotaxime, ceftazidime, meropenem and imipenem, and 25 µM to 2000 µM for ampicillin and benzylpenicillin). The rates were used to determine the kinetic constants for each enzyme-substrate pair by fitting the data with the Michaelis-Menten equation using KaleidaGraph (Synergy). For those pairs where substrate saturation of the enzyme was not possible, the linear portion of the Michaelis-Menten plot was used to determine the \(k_{cat}/K_M\) values.

2.3.5 Cellular fractionation and the β-lactamase activity measurement in the cell lysate

To quantify the expression of the MBLs in the periplasm, cytoplasm, and whole cell fractions, single colonies from *E. coli* 10G, *P. aeruginosa* PA01, and *K. pneumoniae* ATCC 13883 transformed with the MBL plasmids were picked, grown in quadruplicate overnight at 30°C in a 96-well plate in 200 µL of LB medium with 1% glucose to suppress expression and 40 µg/mL of kanamycin for selective resistance. A 1:40 dilution was used to start a 500 µL LB medium
expression culture that was grown for 6 hours at 37°C with 0.02% arabinose for induction and 40 µg/mL of kanamycin. The cells were collected by centrifugation at 3200 × g for 10 minutes. The periplasmic fractions were isolated using the osmotic shock protocol (Imperi et al. 2009). The pellets were suspended in OS1 buffer (30 mM Tris-HCl, pH 7.1, 20% sucrose, and 1mM phenylmethylsulfonyl fluoride) and incubated for 30 minutes at room temperature before centrifugation (3200 × g for 10 minutes). The pellets were then resuspended in OS2A buffer (ice-cold 0.5 mM MgCl₂) and incubated on ice for 5 minutes before centrifugation (3200 × g for 10 minutes). 100 µL of the supernatant was collected and mixed with 100 µL OS2B (40 mM HEPES, pH 7.5, 200 mM NaCl, 400 µM ZnCl₂, and 0.4% Triton). The initial rate of β-lactamase activity was measured by mixing the isolated periplasm fraction and 50 µM of CENTA at a 1:10 ratio in Buffer H supplemented with 0.2% Triton X-100. The remaining pellets were then frozen overnight at -20°C. To obtain the cytoplasmic fraction, the frozen pellets were resuspended in 200 µL of lysis buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 200 µM ZnCl₂, 0.2% Triton, 200 ug/mL Lysozyme, 1 U Benzonase (Millipore), and 0.25 mM MgCl₂), and incubated at room temperature for 1 hour. After centrifugation (3200 × g for 10 minutes), the supernatant was removed and β-lactamase activity was measured to determine the level of enzymes in the cytoplasm. The β-lactamase activity of the whole cell was prepared in the same way as the cytoplasmic fraction without the initial isolation of the periplasmic fraction.

**2.3.6 Calculation of the number of functional MBLs per cellular fraction**

The initial rate of reaction with CENTA for the fractional or whole cell lysates was used in conjunction with the initial rate measured with purified enzyme to determine the relative amount of enzyme in each fraction: 

\[
[E] = \left(\frac{A_{\text{lysate}}}{A_{\text{purified}}} \times [E_{\text{purified}}] \times V \times N_A\right) \div (OD_{600} \times 1 \times 10^9)
\]
CFU/OD\textsubscript{600}), where $A\text{lysate}$ denotes the rate of CENTA hydrolysis in the lysate, $A\text{purified}$ denotes the rate of CENTA hydrolysis of purified enzyme, [$E\text{purified}$] denotes the concentration of purified enzyme, $V$ denotes the culture volume, $N_A$ denotes the Avogadro constant, and OD\textsubscript{600} denotes the absorbance of the cultures at 600 nm.

2.3.7 Determination of MBL gene expression

After induction of the *E. coli* harbouring each MBL as described above, plasmid DNA was isolated from 500 µL of culture with a QIAprep Miniprep Kit (Qiagen). For isolation of RNA, 500 µL of the induced culture was mixed with 500 µL of RNAProtect Bacteria Reagent (Qiagen) and incubated at room temperature for 5 minutes before RNA extraction as per the manufacturer’s instructions with the RNeasy Mini Kit (Qiagen). To ensure removal of contaminating DNA, the samples were processed with a DNA-Free kit (Ambion). cDNA was then prepared using the QuantiTect Reverse Transcription Kit (Qiagen).

Quantitative PCR was performed on a 7500 Fast Real-Time PCR System (Applied Biosystems) with the following cycling conditions: 20 seconds at 95°C, followed by 40 cycles of 95°C for 3 seconds and 58°C for 30 seconds. The primers used for the DNA quantification were designed to detect the pBBR1-pBAD vector: F-CCAACACGCGATTCGTCTGG, R-AGCCAGAAGACACTTTTCAAGC. The forward primers used for the RNA quantification were designed for each MBL: BcII, F-GATTTAGGAAACGTGGCGATGC; IND-1, F-CAATGTATTGGATGGCTGTC; CcrA, F-GCATGGCCGAAAAACTCTCG; NDM-1, F-CTCGGCAATCTCGGTGATGC; VIM-1, F-GGAAGCAGAGGTCGTCATTCC; VIM-2, F-GGAAGCAGAGGTCGTCATTCC; IMP-1, F-TAGAAGCTTGCGGCAAGTCC; and SPM-1, F-AACTTGGTTATCTGGGAGATGCC. The reverse primer was the same for all eight MBLs:
R-GCAACGCAATTAATGTGAGTTAGC. The primers for the reference gene, histidyl-RNA synthetase (hisS), were: F-GCTCCGGCATTAGGTGATTA and R-TCAAGCAGTTTGCACAGACC. Normalized expression units for each MBL gene were calculated using the ΔΔCt method relative to hisS, whereas the absolute value of the number of DNA copy number was determined with known standards of the vector, pBBR1-pBAD, and normalized by the OD$_{600}$ of the original culture.

2.3.8 Replacement of MBL signal peptides with the PelB leader sequence

The native signal peptides for each MBL were replaced in the pBBR1-pBAD vector with the PelB leader sequence (MGKYLLPTAAAGLLLLAAQPAMAMDSG) using Golden Gate assembly cloning with the Type IIS restriction enzyme, BsaI.

2.4 Results

2.4.1 Metallo-β-lactamases provide diverse levels of antibiotic resistance

I chose a diverse set of eight MBLs to investigate the phenotypic diversity among enzymes in the B1-MBL family: Three enzymes represent chromosomally-encoded MBLs (BcII, IND-1, and CcrA), and five represent those acquired on mobile genetic elements through HGT (NDM-1, VIM-1, VIM-2, IMP-1, SPM-1). These enzymes also exemplify the large sequence diversity within the B1-MBL family (amino acid identities range from 23% to 36%, except between VIM-1 and VIM-2, which is 90%) (Appendix A, Table A.1). The MBL genes were subcloned into a modified broad-host-range pBBR1MCS-2 vector along with the inducible P$_{BAD}$ promoter (pBBR1-pBAD) and subsequently transformed with the three bacterial strains: *Escherichia coli* E. cloni® 10G, *Pseudomonas aeruginosa* PA01, and *Klebsiella pneumoniae* ATCC13883. These
three organisms were chosen because they represent major opportunistic pathogens that have recently acquired MBLs through HGT. The minimum inhibitory concentration (MIC) for the strains harbouring the MBL genes was determined for six different compounds representing the three major classes of β-lactam antibiotics: cephalosporins, cefotaxime (CTX) and ceftazidime (CAZ); carbapenems, meropenem (MEM) and imipenem (IMI); and penams, ampicillin (AMP) and penicillin (PEN). The growth was measured on agar plates supplemented with each antibiotic (Appendix A, Table A.2). It should be noted that the background MICs vary considerably between the bacterial strains due to their intrinsic resistance.

The MBLs confer diverse levels of resistance (Figure 2.1; Appendix A, Figure A.1). For example, there is an over 4000-fold range in the ceftazidime MIC between MBLs in P. aeruginosa, from 2 µg/mL for CcrA to 8192 µg/mL for NDM-1. Nonetheless, within each organism, the relative order of the MBLs by resistance was similar for the six different antibiotics. For example, IND-1 and VIM-1 in E. coli confer the highest MICs, followed by BcII, VIM-2, NDM-1, IMP-1, SPM-1, and lastly, CcrA. This indicates that the all MBLs exhibit similar and broad substrate specificity, but other factors influence the specific level of resistance.
Figure 2.1: Measured minimum inhibitory concentrations (MICs) for the metallo-β-lactamases with representative β-lactam antibiotics in *E. coli*, *P. aeruginosa*, and *K. pneumoniae*. MICs for cefotaxime (blue), meropenem (green), and ampicillin (red) measured for each MBL (BcII, IND-1, IMP-1, CcrA, VIM-1, VIM-2, NDM-1, and SPM-1) in *E. coli* 10G, *P. aeruginosa* PA01, and *K. pneumoniae* ATCC 13883. MICs were determined with the concentration at which at least three of four replicates did not grow. The grey represents the background resistance of the organisms without MBL expression. The chemical structures of cefotaxime, meropenem, and ampicillin as well as the concentrations screened to determine the MICs are shown below. MICs for ceftazidime, imipenem, and penicillin are presented in Appendix A, Figure A.1.

By contrast, a substantially different trend emerges when comparing the resistance levels across the three different organisms. For example, SPM-1 confers one of the lowest MICs among the eight MBLs in *E. coli* (e.g., 64-fold lower ampicillin MIC than BcII) but is equivalent to BcII.
and NDM-1 in *P. aeruginosa*. CcrA, whose expression appears to be lethal in *E. coli*, is tolerated better in the other two organisms and even confers substantial resistance to *K. pneumoniae* (*e.g.*, meropenem MIC 64-fold over background). The results suggest that the relationship between the MBL genes and host organisms plays a strong role in determining the relative resistance provided by each MBL.

### 2.4.2 The catalytic efficiency of the MBLs cannot explain the level of antibiotic resistance

The level of antibiotic resistance is generally considered to be strongly associated with the catalytic efficiency of the MBL enzymes (*k*_cat/_K*_M) and the determination of the kinetic parameters is usually one of the first experiments performed on newly discovered MBL variants (Liu et al. 2018; Papagiannitsis et al. 2015; Jeannot et al. 2012). I examined to what extent the _k*_cat/_K*_M of the MBLs can explain the variation observed in the MICs. The mature MBL genes (with their signal peptides removed) were fused with a C-terminal Strep-tag by subcloning into the pET-26(b) vector, and overexpressed in *E. coli* BL21 (DE3). The MBLs were subsequently purified and their kinetic parameters (*k*_cat, _K*_M, and _k*_cat/_K*_M) were determined for the six β-lactam antibiotic substrates, in addition to a generic β-lactamase substrate, CENTA (Appendix A, Figure A.2 and Table A.3). Overall, the MBLs are moderately to highly efficient enzymes for all seven substrates, with _k*_cat/_K*_M values ranging from 10^4 to 10^7 M^{-1}s^{-1}. For each substrate, variation in _k*_cat/_K*_M between the MBLs is relatively small, with the difference between the highest and lowest enzymes being within two orders of magnitude, and often much closer (*e.g.*, the widest range observed was with ceftazidime where BeII’s _k*_cat/_K*_M value was 1.3×10^4 compared to NDM-1’s value of 1.9×10^6). Interestingly, the catalytic efficiency of the enzymes rarely accounted for the level of resistance that the enzyme confers to the bacteria (Figure 2.2, Appendix A, Figure A.3).
Of the six antibiotics assessed with the three different organisms (18 combinations), the relationship between $k_{\text{cat}}/K_M$ and MIC was inversely correlated in 10 of the 18 cases. The 8 remaining cases showed positive, yet weak correlations. This is highlighted in *E. coli* in particular, where SPM-1 exhibits a 10-fold higher $k_{\text{cat}}/K_M$ with ampicillin compared to VIM-2 ($2.7 \times 10^6$ vs. $2.7 \times 10^5 \text{M}^{-1}\text{s}^{-1}$), yet SPM-1 confers 256-fold lower resistance against ampicillin than VIM-2 (64 vs. 16,384 µg/mL). Thus, whereas the existence of $\beta$-lactamase activity is essential to confer resistance to bacteria, the level of resistance that the bacteria obtains cannot be explained by only the kinetic parameters ($k_{\text{cat}}/K_M$), but must be strongly affected by other factors that are associated with the relationship between the host and the enzyme.

![Figure 2.2: Relationship between $k_{\text{cat}}/K_M$ and ampicillin MICs for the metallo-$\beta$-lactamases in the three organisms.](image)

The measured ampicillin MIC values for each MBL in relation to their $k_{\text{cat}}/K_M$ for each organism are shown with the background resistance for each organism denoted by the grey box. The relationship for the other 5 antibiotics is shown in Appendix A, Supplementary Figure A.3.

### 2.4.3 The combination of catalytic efficiency and periplasmic expression of MBLs determines the level of antibiotic resistance they provide

The level of antibiotic resistance, or the fitness of the host strain ($W$), that is conferred by the expression of MBLs, can be associated not only with their catalytic ability ($f$, or $k_{\text{cat}}/K_M$), but also
with the concentration of functional enzyme in the periplasmic fraction of the cell ([E\textsubscript{p}]). This relationship can be defined as \( W = f \times [E\textsubscript{p}] \) (Tokuriki & Tawfik 2009). To determine if this relationship can sufficiently explain the patterns of observed resistance levels, I determined the [E\textsubscript{p}] of each enzyme in each organism. Briefly, the MBL enzymes were expressed as with the MIC determination in each host. The cells were harvested, the periplasmic fractions were isolated using the osmotic shock method, and the level of β-lactamase activity in the periplasmic fraction was determined using 50 µM of CENTA. The number of functional MBL enzymes per cell in the periplasm fraction [E\textsubscript{p}] was then calculated from the kinetic parameters obtained in this study and the cell density of the cultures (Appendix A, Table A.4). It should be noted that, as a previous study demonstrated, NDM-1 is not a soluble periplasmic enzyme but rather localizes to outer membrane vesicles (King & Strynadka 2011; González et al. 2016). This unique localization explains why NDM-1 exhibits low concentrations in the periplasm despite consistently providing high levels of resistance. IND-1 exhibited unexpectedly low level of activities in the lysate enzymatic assay, and I speculate that IND-1 may be unstable in our assay buffer. Thus, I omitted these enzymes from the following analyses.

Of the six remaining MBLs, there was significant variation in the calculated [E\textsubscript{p}]. The overall range in the variation of [E\textsubscript{p}] was generally much greater than the range of variation of \( k_{\text{cat}}/K_M \) (Figure 2.3A, Appendix A, Figure A.4). For example, the [E\textsubscript{p}] of MBLs differs by more than >600-fold in \textit{E. coli}, with SPM-1 expressing <20 molecules per cell and VIM-1 producing >8000 molecules per cell, whereas the largest difference between catalytic efficiencies was only ~100-fold between BcII and NDM-1 for ceftazidime. In general, the average variation between the [E\textsubscript{p}] values for the six MBLs is 130-fold in \textit{E. coli}, 85-fold in \textit{P. aeruginosa}, and 60-fold in \textit{K. pneumoniae}. The relative order of the MBLs in terms of [E\textsubscript{p}] is similar across the three
organisms (Appendix A, Table A.4); VIM-1 is consistently the most highly expressed enzyme, and out-produced the next highest MBL by approximately 10-fold across all three organisms. By contrast, SPM-1 and CcrA are the lowest expressed enzymes in the three organisms. Nonetheless, the specific level of $[E_p]$ of each enzyme varies depending on the host organism. For example, VIM-2’s periplasmic expression in *P. aeruginosa* is tripled compared to *E. coli*, whereas BcII’s expression is halved. Thus, the expression of MBL enzymes is highly dependent on their sequences, but it can also be dependent on the host organism.

Despite substantial variation in $k_{cat}/K_M$ and $[E_p]$ for each enzyme in each organism, variation in the antibiotic resistance levels conferred by MBLs is well explained using the equation, $W = f \times [E_p]$ (Figure 2.3B, Appendix A, Figure A.5). However, our observations suggest that genetic variation among the MBLs exerts greater influence on the overall fitness ($W$) through expression ([$E_p$]) rather than function ($f$).
Figure 2.3: Relationship between MIC, $k_{cat}/K_M$, and $[E_p]$ for each MBL in the three organisms. (a) The variation within MIC and the $k_{cat}/K_M$ for ampicillin, and the $[E_p]$ within each organism for the MBLs. Each data point represents the fold-difference between the values for each MBL divided by the lowest value present for the group. The relationship for the other 5 antibiotics is shown in Appendix A, Figure A.4. (b) The relationship between the product of $k_{cat}/K_M$, and $[E_p]$ with MIC. NDM-1 and IND-1 were not included in the fit as NDM-1 is a known to be bound to the outer membrane and could therefore not be accurately measured with the assay, while IND-1 activity was not detectable in the assay. The relationship for the other 5 antibiotics is shown in Appendix A, Figure A.5.
2.4.4 The variation in periplasmic expression emerges during the transcription, translation, and translocation of the MBLs

Next, I sought to identify what underlying molecular mechanisms influence MBL periplasmic expression. To do so, I analyzed the variation in the relative quantity of MBL precursors at every major stage in the MBL production process (i.e., DNA, RNA, protein, and finally, periplasmic protein). The variation at each stage is cumulative and a product of several properties, for example, the variation in mRNA would reflect not just the different transcription efficiencies of the MBLs, but also mRNA stability and turnover as well. However, overall, the variation at each stage is specific enough to identify what cellular processes or properties may be contributing to the observed variation in periplasmic expression. The absolute amount of plasmid DNA and relative amount of mRNA was quantified with qPCR, and compared to the values previously calculated for the whole cell and periplasmic expression of each MBL (Appendix A, Table A.5). The variation in each category was determined by normalizing each measurement to the lowest measurement obtained in the category. This variation was then compared across categories to determine at what stage is it reflective of the variation in the periplasmic expression (Figure 2.4).

The variation in the amount of plasmid DNA present for each MBL was small, implying that there are few differences between the MBLs at this stage and that variation must emerge through downstream processes. The range of variation between the eight MBLs increased at each subsequent level of expression, from 16-fold at the mRNA level, to 210-fold at the total cellular protein level, to 646-fold at the periplasmic protein level, suggesting that variation steadily increases at each stage. More specifically, the initial emergence of variation at the mRNA level suggests that varying transcription rates and/or mRNA stability is the first way in which hidden phenotypic variation can emerge from the sequence of the MBLs. Subsequently, the >10-fold
increase in variation among the cellular protein measurements suggests that differences in translation efficiency and/or protein stability contribute the most to phenotypic variation. Lastly, protein translocation to the periplasm appears to add another layer of variation, as the total variation in the group increases by >3-fold. Interestingly, the order of the MBLs also changes between the mRNA, cellular protein, and periplasmic protein levels. For example, CcrA has the highest expression of mRNA, yet one of the lowest levels of cellular protein expression, suggesting that it is poorly translated or that its protein product is largely inactive. BcII, which appears to be better expressed at the cellular level than VIM-2, has lower periplasmic expression, suggesting that it is translocated less efficiently. This indicates that not only is variation increasing throughout the production of the MBLs, but that the relative expression of each MBL can be modified.
2.4.5 A universal signal peptide does not fully abrogate phenotypic diversity

Lastly, I sought to determine the effect of the signal peptide sequence on the MIC variation observed among MBLs. It has been demonstrated that the beginning of the coding region of a gene is strongly associated with the level of protein expression through its influence on translation initiation (Kudla et al. 2009). In the case of the MBLs, this region also corresponds to the N-terminal signal peptide of the translated protein, which would exert significant influence on the efficacy of the MBL’s translocation to the periplasm (or outer membrane for NDM-type MBLs) through its interactions with the signal recognition particle, trigger factor, or SecA, in
addition to the SecYEG translocase (Tsirigotaki et al. 2017). This region also represents the highest genetic variation between the eight MBL genes chosen for this study, with no fully conserved residues and diverse lengths between 27 and 43 residues (as defined by where the first secondary structure element begins). Thus, I hypothesized that substitution of the native signal peptide of each MBL with a universal sequence may reduce the observed hidden phenotypic variation.

The PelB leader sequence is a 22 amino acid signal peptide sequence that was originally identified from the pectate lyase B gene from *Erwinia carotovor* and is extensively used for recombinant periplasmic protein expression in *E. coli* (Power et al. 1992). I replaced the native signal peptide of each MBL with the PelB sequence, generating PelB-MBL fusion genes in the pBBR1-pBAD vector, and determined the MICs of the three organisms harbouring the constructs for the six β-lactam antibiotics (Appendix A, Table A.6). As predicted, the replacement of the signal peptide altered the MIC values; in some cases, the PelB peptide caused increase in MIC, such as the 2 to 8-fold increases observed for NDM-1 and SPM-1 in *E. coli*. However, the substitution of the native signal peptides with PelB generally either led to no change or a decrease in MIC (up to 128-fold, as for VIM-1 in *E. coli*) (Figure 2.5A). Consequently, the correlation between the PelB-MBL MICs and the measured catalytic efficiency became stronger compared to the MICs of the native MBLs: the correlation is positive for 12 of the 18 organism-antibiotic combinations with the universal PelB signal peptide while only 8 were positive with the native sequence (Figure 2.5B, Appendix A, Figure A.6). However, these correlations are still not as strong as those that take into account both catalytic efficiency and functional periplasmic expression. This suggests that while the sequences in the signal peptide region substantially contribute to the expression of the MBLs, and thus the antibiotic resistance of the host
organisms, the hidden phenotypic variation observed amongst the family cannot by explained solely by the signal peptide sequences alone. Genetic variation throughout the entire gene plays a role in determining the level of functional periplasmic expression.

Figure 2.5: The effects of the replacement of the native signal peptide with the PelB leader sequence on MBL-conferred resistance. (a) Relationship between the MICs for each metallo-β-lactamase with their native signal peptide and with the PelB leader sequence for ampicillin, cefotaxime, and meropenem in *E. coli*, *P. aeruginosa*, and *K. pneumoniae*. (b) The relationship between the catalytic efficiency of each PelB-MBL (*k*_{cat}/*K*_{M}) and their ampicillin MICs. The relationship for the other 5 antibiotics is shown in Appendix A, Figure A.6. MICs were determined in quadruplicate, with the concentration at which three or more replicates did not grow used as the value.
2.5 Discussion

The systematic characterization of eight MBL enzymes revealed that the genetic variation among the B1-MBL family can cause substantial hidden phenotypic variation, which is revealed upon expression in different host organisms. Moreover, I show that stronger variation emerges in the functional periplasmic expression of these enzymes than in their catalytic efficiency. This suggests that some of the genetic variation observed in the family is the result of co-evolution with the cellular expression systems in their host organisms; each MBL sequence may have adapted to changes in the host’s transcription, translation and translocation systems.

Our study has important implications for the genetic compatibility of genes that are horizontally transferred to new host organisms and the relationship between genetic sequence and gene expression. When a heterologous gene is introduced by HGT, its compatibility with the new host’s systems would be largely dependent on how pre-disposed the gene and resulting protein are to its new host (i.e., the further diverged a gene is from the host, the less likely it would be compatible with that host). But what molecular properties determine this genetic predisposition? Heterologous protein expression is in essence a ‘wicked problem’ (i.e., a problem that is difficult to solve due to contradictory requirements). Much work has been done to identify the causes of genetic incompatibility in order to effectively express proteins in conventional laboratory host organisms (Ahmad et al. 2018; Rosano & Ceccarelli 2014), and many molecular properties have been proposed to influence gene compatibility. A prevailing theory, supported by several recent genome-wide investigations, is that optimal codon usage can regulate expression through the modification of transcription efficiency, mRNA stability, translation efficiency, and co-translational folding (Stergachis et al. 2013; Kudla et al. 2009; Z. Zhou et al. 2016; Presnyak et al. 2015; Pechmann & Frydman 2013). In contrast, a recent study
that assessed 200 diverse antibiotic resistance genes in *E. coli* demonstrated that many proposed properties such as codon optimality, GC content, and mRNA-folding energy could not sufficiently explain the compatibility of the heterologous genes. Instead, only phylogenetic distance (*i.e.*, evolutionary distance between the original and new hosts) and mechanistic compatibility with the host (*i.e.*, dependency upon specific components of the host’s physiology and metabolism) provided a generally accurate estimation of protein fitness (Porse et al. 2018). It is possible that phylogenetic distance is a better predictor for heterologous expression despite its unsophisticated nature because it roughly reflects the probability that the nuanced and overlapping hidden codes for good expression, which have been identified at the genome-level but cannot yet be reliably applied to specific genes, will translate between the original and new host organisms. Regardless, such contradictions demonstrate the complexity of the underlying molecular mechanisms that determine genetic compatibility.

Our results suggest, in addition to the genetic compatibility of the MBLs with the host expression systems, that compatibility between the MBL’s signal peptide sequences with the host’s translocation system is one of the important mechanisms that can influence the success of an HGT event. With proteins expressed to the periplasm, interactions between the nascent protein and the translocation system levy further genetic constraints that may not be captured by codon optimality. Whereas a general trend in the signal peptide sequences has been identified for each translocation system (*e.g.*, the Sec protein translocation pathway), we have little understanding of the relationship between genetics and the efficacy of translocation. It is essential to develop such knowledge to further understand the universal rules for the relationship between sequence and expression.
HGT is the primary mechanism underlying the dissemination of antibiotic resistance genes to different bacterial hosts. Needless to say, associated genetic elements such as the transcription and translation initiation sequences, and the origin of replication (in the case of the plasmids) would significantly affect the outcome of the HGT of genetic material to a particular host. However, our results suggest that genetic variation of the coding region itself plays an important role in the gene’s dissemination. Essentially, the compatibility between the antibiotic resistance genes and host expression systems (including transcription, translation, and translocation) is one of the overriding factors for successful HGT to diverse hosts. MBLs that are pre-disposed to the new host organism will express well with little to no fitness cost and provide effective resistance. On the other hand, MBLs that are less pre-disposed to the host may not provide sufficient levels of resistance and likely be eliminated. Interestingly, no specific constraints on the effective deployment of SPM-1 were found in this study, despite its noted association with only *P. aeruginosa*. Yet CcrA exemplified how a poorly pre-disposed protein may incur a significant fitness cost for the host and prove toxic: while conferring substantial resistance in *K. pneumoniae*, its expression in *E. coli* appears lethal. This toxicity is likely the result of a specific incompatibility between CcrA gene and the translocation process in *E. coli* as it can be cytoplasmically overexpressed as a mature protein without its signal peptide in *E. coli* without any toxic effect.

Beyond some examples of host specificity described above, our study did not reveal any particular trend that differentiates acquirable MBLs from non-mobile MBLs. Indeed, the chromosomal MBL BcII, from *Bacillus cereus*, consistently provided one of the strongest resistance levels in all three organisms, but has yet to spread to pathogenic bacteria and has never been isolated in clinics. Again, other genetic elements such as the ribosome-binding site,
promoter, and the association with mobile genetic elements may explain this (Jiang et al. 2018).

Moreover, other specific characteristics of MBLs may be associated with their ability to confer high antibiotic resistance to their hosts in the human body. For example, Bahr et al. and Stewart et al. both recently demonstrated that NDM-1 variants are evolving to maintain activity in a low zinc concentration environment (Stewart et al. 2017; Bahr et al. 2018). Nonetheless, it is likely that other chromosomal MBLs will eventually transfer and confer multi-drug resistance to pathogenic bacteria in the future.

Our results emphasize the importance of studying the hidden phenotypic variation of antibiotic resistance genes in a more holistic, universal, and consistent experimental setup. I found that while the catalytic ability of the MBLs may be a necessary component for providing effective resistance (i.e., the presence of an enzyme with β-lactamase activity is a basic requirement), the expression of MBLs, particularly periplasmic expression, substantially affects the level of resistance provided by that enzyme and this can be specific to each host organism. However, in general, more emphasis has been placed on determining the enzymatic efficiencies of newly identified MBL isolates, and the direct comparison of antibiotic resistance level between variants is less valued: each study uses different expression vectors and expression protocols for the MIC assay. More effort should be directed to develop a universal procedure and protocols, e.g., a universal expression vector for multiple host expression ability and representative clinically related model bacterial hosts to assessing MBL variants in situ to compare new clinical threats and identify hidden molecular properties causing higher antibiotic resistance.
Chapter 3: Host specific expression constraints may be overcome through adaptive evolution

3.1 Summary

How will MBLs respond to continued selective pressure from antibiotics? I address this question by performing a directed evolution experiment in which the three MBLs, NDM-1, VIM-2, and IMP-1, were subjected to a repeated cycle of mutagenesis and selection for the conferral of increased ampicillin resistance in *E. coli*. The resistance provided by each MBL increased 48 to 80-fold over the course of 18 rounds of directed evolution. The majority of this improvement came from an increase in functional periplasmic expression, largely driven through an increase in overall expression, but influenced by the translocation efficiency as well. These adaptations were predominantly caused by mutations in the signal peptide, as substitution with the PelB leader sequence negated all beneficial effects from the improvement in periplasmic expression. Further, these changes demonstrated host specific tendencies, as the improvement in resistance was diminished in *K. pneumoniae* and *P. aeruginosa*. This suggests that MBLs may continue to evolve largely through adaptation to their host organisms and that this co-evolution may drive the accumulation of further mutations in the signal peptide regions of clinical variants. This additionally demonstrates the rapid rate at which heterologous genes may overcome host specific constraints through evolutionary optimization, and serves to highlight the danger presented by antibiotic resistance genes that disseminate via horizontal gene transfer. Future work should focus on understanding the hidden genetic codes that influence the efficiency of transcription, translation, and translocation, which are important for understanding the functional constraints
that restrict the dissemination of antibiotic resistance determinants, and may allow the development of strategies that capitalize on these restraints to limit their spread.

3.2 Introduction

The emergence of the serine-β-lactamases (SBLs) in response to the widespread deployment of β-lactam antibiotics threatened their continued clinical utility in the 1960’s. Only the production of semi-synthetic penicillins, which were modified to escape destruction, as well as the development of new unaffected classes of β-lactam antibiotics, such as the cephalosporins, sustained the impact of these newfound treatments (Abraham 1983). However, this proved to be only a temporary solution as the SBLs continue to evolve and adapt to provide resistance to new antibiotics. Indeed, over 170 variants of the SBL, TEM-1, have been isolated since it was first characterized in 1963 (Datta & Richmond 1966; Salverda et al. 2010). Originally, TEM-1 was found to provide resistance against penicillin, but its descendants have gradually adapted to confer resistance against later generations of antibiotics and inhibitors, including cephalosporins and monobactams (Salverda et al. 2010). This has spurred the continuous development of new antibiotics and inhibitors to stem the tide of resistance (Matagne et al. 1999). Currently, a number of β-lactam antibiotics and SBL inhibitors are still used in combination to effectively treat infections worldwide, although the continued utility of these treatments is once again threatened by growing resistance (Llarrull et al. 2010).

In contrast to the SBLs, the recent emergence of the MBL B1 family has been rapid and severe: the MBLs have hastened the global spread of pathogens that are resistant to the latest generation of β-lactam antibiotics, the carbapenems. It is believed that this will place a severe strain on healthcare systems worldwide, while inflicting a staggering economic cost on society
(Carlet et al. 2012; Codjoe & Donkor 2017). No clinical inhibitors currently exist to mitigate this threat as the depth of the diversity present in the family has frustrated the search for broad-spectrum therapeutic inhibitors (Rotondo & Wright 2017). At the same time, these enzymes have also demonstrated an ability to rapidly evolve, further complicating inhibitor development. Indeed, 16 variants of the MBL NDM-1 have been found since it was first reported in 2009 (Yong et al. 2009; Bahr et al. 2018).

How will the MBLs evolve in the future? If the SBLs are any indication of the evolvability of antibiotic resistance genes, the MBLs may have great potential to adapt to and overcome any new antibiotics or inhibitors deployed against them. The MBLs already confer resistance against all β-lactam antibiotics save the monobactam, aztreonam (Bebrone 2007). How then, will they respond to selective pressure from increased antibiotic use? Will their specificity narrow as their catalytic ability for a smaller subset of antibiotics increases or will they maintain their broad substrate specificity? An ability to predict how MBLs will adapt to antibiotics would be useful for developing new treatments and treatment strategies. Detailed comparisons between mutations identified in natural variants and those obtained in the laboratory evolution of TEM-1 suggests that directed evolution can be used to mimic the natural course of the evolution of antibiotic resistance determinants in the clinic and provide for a relatively accurate prediction (Salverda et al. 2010). Further, these techniques have been previously applied in small-scale experiments to individual MBLs (Hall 2004a; Tomatis et al. 2005; Tomatis et al. 2008; Meini et al. 2014; Meini, Tomatis, et al. 2015). However, these studies have been limited in scope: they have focused solely upon the catalytic domain of a single enzyme, neglecting the signal peptide region, and have restricted the extent of the directed evolution to only a few rounds. This disregards family-wide and longer-term trends that might emerge. A more
comprehensive and holistic approach is required to better predict the evolutionary future of the MBL B1 family.

Here, I conducted a systematic directed evolution experiment in which NDM-1, VIM-2 and IMP-1 were repeatedly mutagenized and selected for the conferral of increased ampicillin resistance in *E. coli* for 18 rounds, in order to investigate the evolutionary potential of the MBLs and determine possible future evolutionary trajectories in the clinic. These MBL genes were chosen for this study due to their increasing clinical relevance: these particular MBLs are readily acquirable and have swiftly disseminated to multi-drug resistant pathogens throughout the world. I further performed diverse biochemical and biophysical characterizations of the evolved enzymes to reveal the mechanism by which MBLs are likely to provide increased resistance.

3.3 Methods

3.3.1 Creation of mutagenized libraries

Randomly mutagenized libraries of NDM-1, VIM-2, and IMP-1 were created by error-prone PCR with the nucleotide analogues, 8-oxo-2’-deoxyguanosine-5’-triphosphate (8-oxo-dGTP) and 2’-deoxy-P-nucleoside-5’-triphosphate (dPTP) (TriLink). For each library, two independent PCRs were used to ensure a balanced mutation rate with only one of the analogues featured in each reaction. Each 25 µL PCR consisted of 1 x GoTaq Buffer (Promega), 3 µM MgCl₂, 0.1 µM of each primer, 0.2 mM of dNTPs, 1.25 U of GoTaq DNA polymerase (Promega), 1 ng of template plasmid, and either 100 µM of 8-oxo-dGTP or 1 µM of dPTP. The ‘mutagenesis’ PCR was set-up and ran as follows: an initial denaturation (95°C for 2 minutes), followed by 20 cycles of denaturation (95°C for 30 seconds), annealing (58°C for 60 seconds), and extension (72°C for 60 seconds), before a final extension step (72°C for 3 minutes). The PCR products were
subsequently purified with the Cycle Pure PCR Purification Kit (ENZA), quantified, and used as template in a combined second ‘amplification’ PCR. Each 50 µL ‘amplification’ PCR consisted of 1 x GoTaq Buffer (Promega), 3 µM MgCl₂, 0.1 µM of each primer, 0.25 mM of dNTPs, 1.25 U of GoTaq DNA polymerase (Promega), and 5 ng of each PCR product from the two previous separate reactions. The reaction was run as before with the exception of an increase in the number of cycles from 20 to 35. Each product was purified again with the Cycle Pure PCR Purification Kit (ENZA), treated with DpnI (Fermentas) to remove the original supercoiled plasmid template, NcoI (Fermentas), and either and HindIII, XhoI, or EcoRI for NDM-1, VIM-2, and IMP-1 respectively, in a 1 hour incubation at 37°C to prepare the PCR products for cloning.

Three versions of the pIDR-5.1 plasmid were prepared for use as the vector in this experiment. The selective resistance marker in the plasmid was replaced with kanamycin resistance for use with NDM-1, chloramphenicol resistance for VIM-2, and tetracycline resistance for IMP-1. Each plasmid was then prepared with the restriction enzymes flanking each of their corresponding genes (NcoI and HindIII for NDM-1, NcoI and XhoI for VIM-2, and NcoI and EcoRI for IMP-1). Each pIDR-5.1 plasmid was concurrently prepared by incubation with their set of restriction enzymes for 3 hours at 37°C for use as the vector in this experiment. While the digested plasmid was subsequently isolated via gel electrophoresis, the digested PCR products were again purified with the Cycle Pure PCR Purification Kit (ENZA) before ligation with the prepared vector. Each 10 µL ligation consisted of 1 × T4 DNA ligase buffer (ThermoFisher), 5 U of T4 DNA Ligase (ThermoFisher), 10 ng of prepared vector, and 30-40 ng of prepared mutagenized insert, before incubation at room temperature for 1 hour. The ligations were then purified with a Microelute Kit (ENZA) and eluted in 20 µL of water.
3.3.2 Application of selection

The mutagenized libraries were electroporated with E. cloni® 10G E. coli cells (Lucigen), suspended in 1 mL of LB medium, and allowed to recover for 1 hour at 37°C. They were then grown overnight at 30°C in 10 mL of LB medium supplemented with each trajectory’s selective antibiotic (40 µg/mL of kanamycin for NDM-1, 34 µg/mL of chloramphenicol for VIM-2, and 15 µg/mL of tetracycline for IMP-1). The next day, a 1:100 dilution was made of each overnight culture in LB medium and 100 µL was plated on each selection plate. Each trajectory was plated onto a series of LB agar plates containing 2-fold increases in the concentration of ampicillin from 2 to 8192 µg/mL, in addition to the constant selective antibiotic. The plate with the highest concentration of ampicillin that had between 100 and 1000 colonies was collected. The plasmids were then extracted and used as the template for the next round.

3.3.3 Measuring the minimum inhibitory concentration of selected variants

To assess the fitness of individual variants within the trajectories, single colonies were obtained from selected rounds and grown in a deep-96-well plate overnight at 30°C in 500 µL of LB medium with 40 µg/mL of kanamycin for the NDM-1 trajectory, 34 µg/mL of chloramphenicol for the VIM-2 trajectory, and 15 µg/mL of tetracycline for IMP-1 trajectory. The next day, 5 µL of overnight culture was used to inoculate 195 µL of LB medium with each trajectory’s corresponding selective resistance in quadruplicate in a 96-well plate and grown for 3 hours at 37°C. The cultures were then plated with 96-well replicator pins on a series of LB agar plates with increasing levels of antibiotics (two-fold increases in cefotaxime, meropenem and ampicillin from 0.032 to 4096 µg/mL, 0.016 to 64 µg/mL, and 2 to 32768 µg/mL respectively). The agar plates were subsequently incubated overnight at 37°C. The next day, the minimum
inhibitory concentration (MIC) was determined by identifying the concentration of antibiotics by which no growth was observed in at least three of the four replicates for each variant.

3.3.4 Sanger sequencing of selected variants

Between 5-10 variants from selected rounds of each trajectory were isolated and sent for Sanger sequencing (Genewiz). The sequence results were visually inspected in Geneious® (8.1.9) and the mutations were identified.

3.3.5 Purification of Strep-tagged MBLs

NDM-1, VIM-2, and IMP-1, in addition to a single variant obtained from the 18th round of each trajectory (NA, VA, and IA) were cloned into a pET-26(b) vector without their signal peptide and with a C-terminal Strep-tag (GNSGSAWSHPQFEK). Each enzyme was expressed and purified as described in Chapter 2. The A$_{280}$ was measured for each sample with the following extinction coefficients, which were calculated with ExPASy ProtParam (Gasteiger et al. 2005):

NDM-1, 33,460 M$^{-1}$cm$^{-1}$; VIM-2, 35,410 M$^{-1}$cm$^{-1}$; IMP-1, 50,420 M$^{-1}$cm$^{-1}$; NA, 33,460 M$^{-1}$cm$^{-1}$; VA, 33,920 M$^{-1}$cm$^{-1}$; and IA, 50,420 M$^{-1}$cm$^{-1}$.

3.3.6 Enzyme assays to determine kinetic parameters

The catalytic ability of each MBL enzyme was measured for seven β-lactam substrates (CENTA, cefotaxime, ceftazidime, meropenem, imipenem, ampicillin, and benzylpenicillin) as described in Chapter 2.
3.3.7 Cellular fractionation and β-lactamase activity measurement in the cell lysate

To quantify the expression of the MBLs in the periplasm and cytoplasm, each fraction was obtained and the CENTA activity measured as described in Chapter 2.

3.3.8 Calculation of the number of functional MBLs per cellular fraction

The initial rate of reaction with CENTA for each fraction was used to calculate the total number of functional MBLs per cell as described in Chapter 2.

3.3.9 Cloning, expression, and fitness testing of MBLs in the three host organisms

A single variant from the 18th round of the three directed evolution trajectories were subcloned in a broad-host-range vector, pBBR1MCS-2, along with the P_BAD promoter, which is inducible in the presence of arabinose (pBBR1-pBAD). The three plasmids were then transformed with the three bacterial strains: *E. coli* 10G (chemical transformation), *P. aeruginosa* PA01 (electroporation) (Choi et al. 2006), and *K. pneumoniae* ATCC13883 (electroporation) (Fournet-Fayard 1995). The ampicillin MICs were determined as previously described with the pIDR-5.1 expression plasmid.

3.3.10 Replacement of the MBL signal peptides with the PelB leader sequence

The native signal peptides for each MBL were replaced in the pBBR1-pBAD vector with the PelB leader sequence (MGKYLLPTAAAGLLLLAAQPAMAMDSG) using Golden Gate assembly cloning with the Type IIS restriction enzyme, *BsaI*.
3.4 Results

3.4.1 Directed evolution of the MBLs resulted in increased resistance to ampicillin

NDM-1, VIM-2, and IMP-1 were subjected to 18 rounds of directed evolution to investigate how MBLs adapt and confer greater resistance to *Escherichia coli* while under a strong selective pressure from higher concentrations of ampicillin. The three MBLs were subcloned into a low-copy number plasmid (pIDR5.1) and placed under the expression of a weak, constitutive promoter (blaTEM). The genes were mutagenized, re-cloned into the vector, and transformed with *E. coli* E. cloni® 10G cells, which were then selected for increased ampicillin resistance by plating the cells on a series of agar plates supplemented with increasing concentrations of ampicillin. The plate with the highest concentration of ampicillin and with growth of between 100 to 1000 colonies was collected, and the plasmids were isolated for use as the template in the next round of mutagenesis and selection. This directed evolution cycle was repeated while the range of ampicillin concentrations screened was gradually increased over 18 rounds from approximately 256 µg/mL to 8,192 µg/mL for NDM-1 and VIM-2, and 4 µg/mL to 256 µg/mL for IMP-1, by which point, the improvement in resistance had plateaued.

To examine and compare how the level of resistance increased over the course of the evolution, 6 to 10 variants were isolated from rounds 1, 2, 3, 6, 9, and 18 and the minimum inhibitory concentration (MIC) for the *E. coli* harbouring each variant was determined for three different compounds representing the three major classes of β-lactam antibiotics: the cephalosporin, cefotaxime (CTX); the carbapenem, meropenem (MEM); and the penam, ampicillin (AMP). The growth was measured on agar plates supplemented with each antibiotic (the range of concentrations screened for each antibiotic was: cefotaxime, 0.032 µg/mL to 4,096 µg/mL; meropenem, 0.016 µg/mL to 64 µg/mL; ampicillin, 2 µg/mL to 32,768 µg/mL).
The ampicillin MICs for the three wild type MBLs differ significantly: IMP-1’s MIC was only 16 µg/mL, compared to 512 and 256 µg/mL for NDM-1 and VIM-2 respectively (Figure 3.1). Over the 18 rounds of directed evolution, the three trajectories experienced a similar increase in the level of ampicillin resistance, with the average ampicillin MIC increasing 80, 60, and 48-fold over the wild type NDM-1, VIM-2, and IMP-1 respectively (Figure 3.2). Consequently, the final ampicillin MIC value for the three MBLs also differs significantly: isolated NDM-1 variants from the 18th round have an average MIC of approximately 41,000 µg/mL, whereas the average for the VIM-2 and IMP-1 trajectories are 15,000 µg/mL and 800 µg/mL respectively. Interestingly, the evolved IMP-1 variants provide a similar level of resistance as the initial NDM-1 and VIM-2 enzymes. Similar increases in MIC were observed with the two other antibiotics, cefotaxime and meropenem, which were not used for selection. The increase in resistance for cefotaxime and meropenem in the NDM-1 trajectory was 24 and 25-fold respectively, 32 and 26-fold in the VIM-2 trajectory, and 88 and 143-fold in the IMP-1 trajectory. This suggests that the mutations that were acquired in the trajectories did not alter the substrate specificity of the MBLs, and that the three MBLs have maintained their broad substrate specificity during the directed evolution. Taken together, while each MBL possesses an evolutionary potential to further increase their ability to provide resistance, the level at which the fitness plateaus can differ significantly among the three enzymes.
Figure 3.1: The average absolute improvement in ampicillin resistance exhibited by the NDM-1, VIM-2, and IMP-1 ampicillin trajectories over 18 rounds of directed evolution. Each data point represents the average increase in ampicillin MIC over the wild type MBL for 6 to 10 variants randomly chosen from selected rounds. The size of the data points incorporates the error for each average. The blue, red, and green lines represent the NDM-1, VIM-2, and IMP-1 trajectories respectively.

Figure 3.2: The average relative improvement in resistance exhibited by the NDM-1, VIM-2, and IMP-1 ampicillin trajectories over 18 rounds of directed evolution. Each data point represents the average increase in MIC over the wild type MBL for 6 to 10 variants randomly chosen from selected rounds. The solid red line represents the increase in resistance for ampicillin, whereas the blue and green dotted lines signify the increase in resistance for the unselected antibiotics, cefotaxime and meropenem, respectively.
3.4.2 Mutations predominantly accumulated in the signal peptide

To determine what genetic changes drive the observed increase in resistance, the variants that were previously isolated from rounds 3, 6, 9, and 18 of each trajectory were sequenced. The overall mutation rate varied between the three trajectories: on average, by the 18th round, 32, 41, and 52 nucleotide mutations were acquired by each gene in the NDM-1, VIM-2, and IMP-1 trajectories, respectively. The cumulative result of this was that 104 of the 266 mutable codon positions (39%) were mutated at least once in one of the sequenced variants from the NDM-1 trajectory, 139 of 267 mutable positions in the VIM-2 trajectory (52%), and 135 of 265 mutable positions in the IMP-1 trajectory (51%). However, only approximately half of these positions acquired non-synonymous mutations (25% of NDM-1, 27% of VIM-2, and 27% of IMP-1). Each trajectory enriched a number of non-synonymous mutations during the evolution, with 18 positions (7%) of NDM-1, 19 positions (7%) of VIM-2, and 33 positions (12%) of IMP-1 mutated at greater than 50% frequency (Table 3.1). While synonymous mutations are equally distributed across the entire gene, non-synonymous mutations were clustered in certain regions, which illustrates that while the entirety of each gene was generally subjected to mutagenesis at equal rates, beneficial and neutral mutations non-synonymous are localized in certain regions under the strong selection pressure to increase the level of antibiotic resistance (Figure 3.3).
Table 3.1 Mutations identified in the sequenced round 18 variants from the NDM-1, VIM-2, and IMP-1 ampicillin trajectories.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Round 18 Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NDM-1 Ampicillin Trajectory</strong></td>
<td></td>
</tr>
<tr>
<td>Fixed Mutations (100%)</td>
<td>N5K, M7I, H8L, A11T, G29P, I35M, M154L, A172E, S213G, A263V</td>
</tr>
<tr>
<td>High Frequency</td>
<td>M39L, T41P, G42S, F70L, V118A, D130G, N142S, E170D</td>
</tr>
<tr>
<td><strong>VIM-2 Ampicillin Trajectory</strong></td>
<td></td>
</tr>
<tr>
<td><strong>IMP-1 Ampicillin Trajectory</strong></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.3: The distribution and frequency of mutations accumulated in the 18-round NDM-1, VIM-2, and IMP-1 ampicillin trajectories. Data was extracted from the sequencing of 6 to 10 variants from rounds 3, 6, 9, and 18 for each trajectory and mapped onto the primary
sequence of each MBL. The secondary structure of VIM-2 is shown aligned above the sequences in black. The line represents loops and turns in the crystal structure (PDB ID 1KO3), whereas the arrow represents a β-strand and the cylinder represents α-helices. Non-synonymous mutations are identified by frequency by shade of blue, while synonymous mutations not already supplanted by non-synonymous mutations are identified in green. Beige indicates insertions and deletions between the three genes.

The strongest trend observed among all three trajectories is that non-synonymous mutations are most enriched in the signal peptide sequence. The non-synonymous mutation frequency per residue per round was determined separately for the signal peptide and the catalytic domain for rounds 3, 6, 9, and 18 of each MBL trajectory (Figure 3.4). The frequency was approximately constant in the catalytic domain in all four investigated rounds for all three enzymes (approximately 0.002 to 0.008 mutations per residue per round). By contrast, the mutation frequency in the signal peptide region was much higher (3 to 7-fold) than in the catalytic domain in round 3 (0.018, 0.020, and 0.065 for NDM-1, VIM-2, and IMP-1 respectively). This disparity decreased substantially by round 18, with the frequency of non-synonymous mutations in the signal peptide steadily decreasing by 2 to 5-fold to be similar to the mutation frequency in the catalytic domain. These results suggest that some of the mutations accumulated in the signal peptide in the early rounds of evolution are strongly associated with the improvement in antibiotic resistance and likely provide the highest ‘return on investment’ during the evolution.
Despite the catalytic domain being less frequently mutated compared to the signal peptide, some mutations were enriched during the evolution. Highly enriched non-synonymous mutations (present in greater than 50% of sequenced variants in round 18) in the catalytic domain of each MBL were mapped onto the tertiary structure of the enzymes (Figure 3.5, Table 3.1). Several mutations were found to occur within the active site of the enzymes, particularly on the L3 and L10 loops, which define the binding pocket of the MBLs. Specifically, this included: F70L and S213G in NDM-1, Q59R, A65T, I223V, and Y224S in VIM-2, and N62D, N237D, and I241L in IMP-1. These mutations may affect the substrate and transition state binding in the active site of the enzyme, and thereby alter kinetic parameters of the MBLs. Additional mutations are present in other areas throughout the enzymes, but most mutations are largely found in the loops and turns throughout the structure. While these mutations may exert no influence over catalysis, it has been previously shown that remote mutations can affect the
function of MBLs through the alteration of conformational dynamics and other long-range network effects. For example, a camelid nanobody was recently shown to allosterically inhibit VIM-4 despite binding far from the active site (Sohier et al. 2013). Thus these distant mutations may also affect the catalytic function of the MBLs, however, I speculate many, if not most, of these mutations could be neutral ‘hitch hiker’ mutations due to high mutation rate in the evolutionary experiment.

![Figure 3.5](image)

**Figure 3.5:** The location of high frequency and fixed mutations from the 18th round of the NDM-1, VIM-2 and IMP-1 ampicillin trajectories identified on their respective structures. Mutations that were found to be mutated in 50% or greater of the six sequenced variants of the round 18 libraries are highlighted as spheres on the crystal structure of each enzyme (NDM-1 PDB ID 3SPU; VIM-2 PDB ID 1KO3; IMP-1 PDB ID 1DDK).

3.4.3 **The catalytic efficiency of the enzymes increased only slightly during the evolution**

To elucidate if changes in the catalytic ability of the MBLs resulted in the observed increase in resistance, I determined the kinetic parameters for an isolated variant from the 18th round of each trajectory and compared them to the values of the wild type MBLs (Table 3.2, Appendix A, Figure A.7). The mature MBL genes (with their signal peptides removed) were fused with a C-terminal Strep-tag by subcloning into the pET-26(b) vector. The MBL proteins were then
overexpressed in *E. coli* BL21 (DE3), subsequently purified, and their kinetic parameters (*k*<sub>cat</sub>, *K*<sub>M</sub>, and *k*<sub>cat/KM</sub>) were determined for the six β-lactam antibiotic substrates, in addition to the β-lactamase substrate, CENTA. Only slight increases were observed in the *k*<sub>cat/KM</sub> values for each of the enzymes, with NDM-1 improving 1.3-fold from 8.4×10<sup>6</sup> to 1.1×10<sup>7</sup> M<sup>-1</sup>s<sup>-1</sup>, VIM-2 improving 6.3-fold from 2.7×10<sup>5</sup> to 1.7×10<sup>6</sup>, and IMP-1 improving 2.6-fold from 1.5×10<sup>5</sup> to 3.9×10<sup>5</sup>. With NDM-1 and VIM-2, this change was brought on by an increase in *k*<sub>cat</sub>, offset by a slight increase in *K*<sub>M</sub>. This may be a reflection of the selection conditions: the ampicillin concentration used for screening the VIM-2 and NDM-1 trajectories surpassed 2048 µg/mL (6 mM) for a large portion of the directed evolution for these two enzymes (the last 12 and 15 rounds respectively). With a concentration of substrate far above the *K*<sub>M</sub> of the enzymes (86 µM for NDM-1 and 83 µM for VIM-2), the rate at which the enzymes degrade the ampicillin would be largely dependent on both the *k*<sub>cat</sub> and the amount of enzyme ([*E*]). Therefore, in such a situation, *K*<sub>M</sub> could be compromised for higher *k*<sub>cat</sub> in order to better increase resistance. In contrast, the *k*<sub>cat</sub> of IMP-1 decreased 4.5-fold, while the *K*<sub>M</sub> decreased over 10-fold. IMP-1 began with a lower resistance than the other two MBLs and was therefore less exposed to higher ampicillin concentrations, only reaching screening concentrations of 64 µg/mL (180 µM) in the last 12 rounds. Additionally, the *K*<sub>M</sub> of the wild type IMP-1 is 610 µM, the highest initial *K*<sub>M</sub> of the three MBLs, and thus, decreasing *K*<sub>M</sub> likely caused a substantial increase in resistance in the IMP-1 trajectory. Nonetheless, the improvement in the catalytic parameters (1.3 to 6.3-fold) was not sufficient in explaining the observed improvement in resistance (48 to 80-fold) (Figure 3.6).
### Table 3.2: Measured kinetic parameters of the wild type and their round 18 evolved counterpart for NDM-1, VIM-2, and IMP-1, for seven β-lactam substrates

<table>
<thead>
<tr>
<th>Kinetics</th>
<th>NDM-1</th>
<th>VIM-2</th>
<th>IMP-1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WT</td>
<td>R18</td>
<td>WT</td>
</tr>
<tr>
<td>$k_{\text{cat}}$ (s$^{-1}$)</td>
<td>200 ± 10</td>
<td>430 ± 20</td>
<td>33 ± 1</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>8.0 ± 1.4</td>
<td>11 ± 2</td>
<td>8.4 ± 0.9</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>2.6 × 10$^7$</td>
<td>4.0 × 10$^7$</td>
<td>3.9 × 10$^8$</td>
</tr>
<tr>
<td>$k_{\text{cat}}$ (s$^{-1}$)</td>
<td>68 ± 2</td>
<td>240 ± 10</td>
<td>38 ± 2</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>7.2 ± 0.8</td>
<td>25 ± 2</td>
<td>15 ± 3</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>9.5 × 10$^6$</td>
<td>9.7 × 10$^6$</td>
<td>2.5 × 10$^8$</td>
</tr>
<tr>
<td>$k_{\text{cat}}$ (s$^{-1}$)</td>
<td>97 ± 2</td>
<td>n.d.</td>
<td>1.1 ± 0.1</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>51 ± 2</td>
<td>n.d.</td>
<td>66 ± 8</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>1.9 × 10$^6$</td>
<td>1.0 × 10$^6$</td>
<td>1.6 × 10$^5$</td>
</tr>
<tr>
<td>$k_{\text{cat}}$ (s$^{-1}$)</td>
<td>26 ± 1</td>
<td>230 ± 20</td>
<td>0.90 ± 0.03</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>21 ± 2</td>
<td>250 ± 30</td>
<td>7.4 ± 1.0</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>1.2 × 10$^6$</td>
<td>9.2 × 10$^5$</td>
<td>1.2 × 10$^5$</td>
</tr>
<tr>
<td>$k_{\text{cat}}$ (s$^{-1}$)</td>
<td>123 ± 3</td>
<td>n.d.</td>
<td>8.7 ± 0.3</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>56 ± 4</td>
<td>n.d.</td>
<td>8.4 ± 1.1</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>2.1 × 10$^6$</td>
<td>7.5 × 10$^5$</td>
<td>1.0 × 10$^6$</td>
</tr>
<tr>
<td>$k_{\text{cat}}$ (s$^{-1}$)</td>
<td>720 ± 30</td>
<td>2400 ± 100</td>
<td>23 ± 1</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>86 ± 15</td>
<td>220 ± 20</td>
<td>83 ± 15</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>8.4 × 10$^6$</td>
<td>1.1 × 10$^7$</td>
<td>2.7 × 10$^5$</td>
</tr>
<tr>
<td>$k_{\text{cat}}$ (s$^{-1}$)</td>
<td>320 ± 20</td>
<td>950 ± 40</td>
<td>76 ± 3</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>61 ± 11</td>
<td>130 ± 20</td>
<td>51 ± 10</td>
</tr>
<tr>
<td>$k_{\text{cat}}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>5.3 × 10$^6$</td>
<td>7.3 × 10$^6$</td>
<td>1.5 × 10$^6$</td>
</tr>
</tbody>
</table>

n.d. not determined
3.4.4 Increasing periplasmic expression drives the increase in resistance

Previously, in Chapter 2, I determined that the antibiotic resistance provided by an MBL can be generally described by the equation $W = f \times [E_p]$, wherein $W$ represents the resistance, $f$ represents the catalytic efficiency, and $[E_p]$ represents the amount of functional periplasmic enzyme. If $f$ could not fully account for the striking increase in resistance, I sought to determine if the change in $[E_p]$ could sufficiently explain this improvement. To determine the relative amount of functional periplasmic enzyme, 28 variants were selected from rounds 3, 6, 9, 12, 15, and 18 from the VIM-2 and IMP-1 trajectories. NDM-1 was not included in this analysis because NDM-1 has been shown to be incorporated into outer membrane vesicles, and subsequently, its enzymatic activity cannot be measured in the isolated periplasmic fraction of the cell (González et al. 2016). The cultures were grown for 3 hours before the cells were harvested, the periplasmic fractions were isolated via osmotic shock, and the level of $\beta$-lactamase activity in the periplasmic
fraction was determined with 50 µM of CENTA. At the same time, the ampicillin MIC values were determined for the same variants for a direct comparison to the detected β-lactamase activity in the periplasm (Figure 3.7). A strong, positive correlation is observed for the relationship between the ampicillin MIC and CENTA activity ($R^2 > 0.9$), which verified that the β-lactamase activity level in the periplasmic fraction is a strong indicator of resistance.

Figure 3.7: The ampicillin MIC of variants from rounds 3, 6, 9, 12, 15, and 18 of VIM-2 and IMP-1 compared to their level periplasmic CENTA activity. Each data point represents one of 28 variants isolated from each round in the VIM-2 and IMP-1 ampicillin trajectories. CENTA activity was measured in duplicate and the MICs were determined by identifying the lowest concentration at which at least three of four replicates did not grow.
The periplasmic activity of IMP-1 increased 56-fold over the course of the evolution, whereas its catalytic parameters for CENTA do not dramatically shift, suggesting that \([E_p]\) is likely to be the dominant factor in explaining the 48-fold increase in ampicillin resistance. The level of periplasmic activity of VIM-2 increased 9-fold from the wild type to the variants in round 18, while the \(k_{\text{cat}}\) for CENTA increased about 3-fold. Again, in this situation, the \(K_M\) of the VIM-2 variants is 5 to 10-fold lower than the concentration used in screening (50 µM), thus \(k_{\text{cat}}\) would likely better reflect the rate of reaction. Taking the change in catalytic ability into account with the rising periplasmic activity, the change in \([E_p]\) in the VIM-2 trajectory would be approximately 3-fold. Thus, the 64-fold increase in MIC can be explained as a result of a balanced contribution from both \(f\) and \([E_p]\). While I was unable to measure the activity of NDM-1 due to its localization, I deduce that the greatest contribution to its increase in resistance came from increased expression, as the \(k_{\text{cat}}/K_M\) of NDM-1 was virtually unchanged in round 18.

Taken together, the results suggest that the directed evolution resulted in a large increase in ampicillin resistance through the combination of the catalytic ability of each enzyme and increasing functional periplasmic expression. However, the level of contribution for each factor differs in each trajectory: the adaptation of IMP-1 and NDM-1 was predominantly driven through an increase in \([E_p]\), whereas the adaptation of VIM-2 was a result of an increase in both \([E_p]\) and catalytic ability.

### 3.4.5 Periplasmic expression is improved through increased cytoplasmic expression and a higher efficiency of translocation

I sought to identify any molecular traits that are associated with the observed increase in functional periplasmic expression in the VIM-2 and IMP-1 trajectories. First, I assessed the
change in the total expression of the MBLs; I determined the level of cytoplasmic activity by isolating the cytoplasmic fractions of the same 28 variants that were previously assessed for their ampicillin MIC and periplasmic activity from the VIM-2 and IMP-1 trajectories. The cytoplasmic activity was compared to the periplasmic activity, which was found to strongly positively correlate ($R^2 > 0.9$); suggesting that overall expression of the MBLs is a large determining factor in the functional periplasmic expression (Figure 3.8A). The percent of the total activity (i.e., sum of cytoplasmic and periplasmic activity) that was detected in the periplasm was relatively consistent across the VIM-2 trajectory, with values varying around 45% (Figure 3.8B). IMP-1, by contrast, had a much wider range. It started with a lower percentage at 23%, before decreasing to 11% in the third round. Afterwards, the percent of activity in the periplasm steadily increased to 42% by the eighteenth round. This suggests that the overall translocation efficiency of the MBLs can vary dramatically, affecting the total amount of functional periplasmic expression, and subsequently, the level of resistance they can provide. It is clear that for the VIM-2 trajectory, an efficient level of translocation was maintained throughout the directed evolution, while the overall expression of the MBL increased. IMP-1 by contrast, required optimization of both translation and translocation efficiency. This suggests that VIM-2 might be more predisposed to expression in *E. coli* than IMP-1, which appears to be less compatible with the host’s expression systems (i.e., transcription, translation, translocation). Upon HGT, genetic incompatibility with these processes may therefore present as significant barriers to the effective deployment; however, in some situations, these barriers may be readily overcome through evolutionary adaptation.

To determine if any molecular properties of the enzymes can explain the change in protein expression, the thermostability for the wild-type and R18 variant of each MBL measuring
denaturation temperature was assessed using SYPRO Orange® dye (Figure 3.8D). No trend was apparent with respect to MBL thermostability and the resistance that they provide: The $T_M$ of NDM-1 increased slightly from 57.6 to 60.7 °C while the $T_M$’s of VIM-2 and IMP-1 decreased substantially from 52.6 to 45.4 °C and 68.6 to 50.8 °C respectively. Such a drop would be expected because many functional mutations would be expected to compromise stability until a functional threshold is reached (Socha & Tokuriki 2013). It is possible that NDM-1 is already at its threshold and thereby maintained its $T_M$ around this level, whereas VIM-2 and IMP-1 were further from their threshold and could accommodate several destabilizing mutations.

Next, I calculated the pI values for the wild type NDM-1, VIM-2, IMP-1, and their round 18 counterparts (Figure 3.8C). In contrast to thermodynamic stability, the convergence of the isoelectric point values (pI) in each trajectory demonstrated a clear trend. While the pI of NDM-1 only decreased slightly from 5.9 to 5.6 and the pI of VIM-2 increased slightly from 5.1 to 5.5, the pI of IMP-1 decreased substantially from 8.4 to 5.3, bringing all three enzymes to within a 0.3 range. Interestingly, this also brought all three MBLs into alignment with the pI of one of the largest segments of proteins within the *E. coli* proteome. Analysis of the entire *E. coli* genome had revealed a bimodal distribution in pI, with the majority of proteins clustering around either a pI of 5 to 6 or 9 to 10 (VanBogelen et al. 1999). It was believed that the bimodal pI distribution was caused by the need for different pI values in different sub-cellular localizations, with cytoplasmic proteins constituting the acidic modality and integral membrane proteins constituting the basic modality (Schwartz et al. 2001). However, several subsequent studies called this theory into question and proposed that the bimodality came about as a result of a ‘mathematical whim’, based on the inherent properties of the proteins themselves (*i.e.*, amino acid $pK_A$ values, amino acid composition, and molecular weight) (Weiller et al. 2004; Wu et al.
2006). Regardless of the origin of the distribution, its establishment has very real biological consequences: A recent report demonstrated that positively charged proteins diffuse up to 100-fold slower in the cytoplasm than negative or neutral proteins (Mullineaux 2017; Schavemaker et al. 2017). The dramatic decrease in the pI of IMP-1 may therefore have contributed to the increase in its functional periplasmic expression. Its original pI of 8.4 would indicate that IMP-1 is positively charged at physiological pH, and it would diffuse much more slowly through the cytoplasm. This might negatively affect the rate at which it is transported to the periplasm. The decrease in pI and subsequent shift to being negatively charged at physiological pH may reverse this and increase periplasmic expression through quicker transport. Alternatively, the change in pI may facilitate more productive interactions with binding partners, such as the Sec machinery, which may also expedite translocation. A third potential avenue by which pI may influence expression is through changes in solubility and functionality within the cytoplasm and periplasm. With several possible theories, further experiments are required to determine the mechanism by which a decreased pI results in higher IMP-1 expression. However, regardless of the specific mechanism by which pI may affect IMP-1’s expression, a question remains as to why the pI of IMP-1 is not already optimized for bacterial expression. Indeed, the majority of most cytoplasmic proteins have a net negative charge at physiological pH in most organisms, including archaea, eukaryotes, and prokaryotes (Schwartz et al. 2001).
Figure 3.8: The molecular changes in the fractional expression of NDM-1, VIM-2, and IMP-1 throughout the 18 rounds of directed evolution. (A) The relationship between periplasmic and cytoplasmic β-lactamase activity. Each data point represents the measured rate of CENTA hydrolysis for the periplasmic and cytoplasmic fractions of E. coli expressing selected MBL variants from rounds 3, 6, 9, 12, 15, and 18 of the VIM-2 and IMP-1 ampicillin trajectories. Variants from the VIM-2 trajectory are denoted as red squares, whereas variants from the IMP-1 trajectory are denoted by green circles. The colour progressively darkens from round 3 variants to round 18 variants. The wild type MBLs are coloured in grey. (B) The percent of β-lactamase activity detected in the periplasm for each round. A black line for each round denotes the average percent. Each variant is represented by a single data point that follows the same scheme as above. (C) The relationship between pI and the ampicillin MIC for the wild type MBLs and their round 18 counterparts. VIM-2 and IMP-1 are represented as before, while NDM-1 is represented by a blue triangle. The wild type MBLs are connected to their round 18 counterparts with a black line. (D) The relationship between thermostability and the ampicillin MIC for the wild type MBLs and their round 18 counterparts. The MBLs are represented as before.

3.4.6 Host specific constraints on expression can be overcome through adaptation

In Chapter 2, I demonstrated that the expression level of MBLs could vary substantially depending on their bacterial host. The directed evolution here resulted in improved expression in E. coli, which suggests that the MBLs adapted to the E. coli expression and translocation systems. This result poses a question as to how the adaptation in E. coli would affect the expression of the MBLs in other host organisms. To investigate this, the wild type MBLs and their round 18 counterparts were subcloned into the broad-host range vector (pBBR1-pBAD) and expressed in E. coli, K. pneumoniae ATCC 13883, and P. aeruginosa PA01. The ampicillin MIC was then determined for each variant in each of the three organisms (Table 3.3, Figure 3.9A). Interestingly, the increase in MIC between the wild type and round 18 variants in the other two organisms was generally lower than in E. coli. For example, the improvement in resistance demonstrated by the NDM-1 trajectory in E. coli (32-fold) decreased to exhibit only a 4-fold
improvement between the wild type and round 18 variant in *P. aeruginosa*, while the improvement of the VIM-2 trajectory also decreased 2-fold between the two organisms. Moreover, the IMP-1 R18 variant provided no benefit over its wild type counterpart in *P. aeruginosa*, despite a 4-fold improvement observed in *E. coli*. The overall effect is lessened in *K. pneumoniae*, where the improvement in the NDM-1 and VIM-2 trajectories remains the same, while the 4-fold improvement of the IMP-1 trajectory once again disappears entirely. This suggests that the changes in the adaptive trajectories are at least partially host specific, and that the MBLs have started to become ‘domesticated’ in their new *E. coli* host. As the effects become more pronounced the further the phylogenetic distance is from *E. coli*, it is likely that given enough evolutionary time and adaptation, there would be a substantial host organism trade off between the evolved variants and the original wild types. Indeed, this may explain why a recent study that assessed 200 diverse antibiotic resistance genes in *E. coli* found that phylogenetic origin is one of the overriding factors in determining fitness, over such sequence-level traits as GC content, codon optimality, and mRNA-folding energy (Porse et al. 2018).

### 3.4.7 Mutations in the signal peptide largely drive the adaptation for increased resistance

As the signal peptide was heavily mutated in the three evolutionary trajectories, I sought to determine to what extent these mutations contributed to the increase in resistance, while also quantifying their contribution to the previously observed change in host specificity, by replacing the native signal peptide of each MBL with this PelB sequence and measuring their resistance in *E. coli, K. pneumoniae*, and *P. aeruginosa*. PelB-MBL fusion constructs were generated in the pBBR1-pBAD vector, and the ampicillin MIC was determined in the three organisms harbouring the constructs. In *E. coli*, substitution of the native signal peptide resulted in a 2-fold decline of
the improvement in resistance for the VIM-2 trajectory (Table 3.3, Figure 3.9B). The resistance provided by the round 18 VIM-2 variant is still 8-fold above the wild type, which likely reflects the 6.3-fold improvement in the $k_{\text{cat}}/K_M$ of the round 18 variant. By contrast, the improvement in resistance was entirely abolished for the NDM-1 and IMP-1 trajectories upon signal peptide substitution. As the change in $k_{\text{cat}}/K_M$ was minimal for both the NDM-1 and IMP-1 trajectories (<2-fold), this suggests that not only was the improvement in resistance driven by an increase in functional periplasmic expression, but also that this increase was largely, if not completely, caused by mutations in the signal peptide.

In *K. pneumoniae* and *P. aeruginosa*, no round 18 variant exhibited any resistance greater than that of their wild type. Indeed, the difference between IMP-1 and its round 18 counterpart in *P. aeruginosa* had actually reversed, with the wild type providing 2-fold higher resistance than the evolved counterpart. Similarly, expression of the evolved VIM-2 variant proved lethal in *K. pneumoniae*. This suggests that mutations in the catalytic domain for these two MBLs may also be affecting the host specificity, in addition to changes in the signal peptide.
Figure 3.9: The fold-improvement between the evolved round 18 variants and wild type enzymes for each trajectory in three organisms. (A) The improvement in ampicillin MIC for the three MBL ampicillin trajectories was determined in *E. coli* 10G, *K. pneumoniae* ATCC 13383, and *P. aeruginosa* PA01. An arrow displays the order in phylogenetic distance from *E. coli*. (B) The improvement in ampicillin MIC for the three trajectories with their native signal peptides replaced by the PelB leader sequence was determined in the three organisms.
Table 3.3: Ampicillin MIC values (µg/mL) for wild type and PelB MBL constructs in the three tested organisms.

<table>
<thead>
<tr>
<th>Signal Peptide</th>
<th>MBL Variant</th>
<th>E. coli 10G</th>
<th>K. pneumoniae ATCC13883</th>
<th>P. aeruginosa PA01</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDM-1 Ampicillin Trajectory</td>
<td>WT</td>
<td>1024</td>
<td>512</td>
<td>4096</td>
</tr>
<tr>
<td></td>
<td>R18</td>
<td>32768</td>
<td>16384</td>
<td>16384</td>
</tr>
<tr>
<td></td>
<td>Fold-Change</td>
<td>32</td>
<td>32</td>
<td>4</td>
</tr>
<tr>
<td>PelB</td>
<td>WT</td>
<td>4096</td>
<td>512</td>
<td>4096</td>
</tr>
<tr>
<td></td>
<td>R18</td>
<td>4096</td>
<td>512</td>
<td>4096</td>
</tr>
<tr>
<td></td>
<td>Fold-Change</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>VIM-2 Ampicillin Trajectory</td>
<td>WT</td>
<td>1024</td>
<td>512</td>
<td>2048</td>
</tr>
<tr>
<td></td>
<td>R18</td>
<td>16384</td>
<td>8192</td>
<td>16384</td>
</tr>
<tr>
<td></td>
<td>Fold-Change</td>
<td>16</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>PelB</td>
<td>WT</td>
<td>32</td>
<td>512</td>
<td>2048</td>
</tr>
<tr>
<td></td>
<td>R18</td>
<td>256</td>
<td>0</td>
<td>2048</td>
</tr>
<tr>
<td></td>
<td>Fold-Change</td>
<td>8</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>IMP-1 Ampicillin Trajectory</td>
<td>WT</td>
<td>256</td>
<td>512</td>
<td>2048</td>
</tr>
<tr>
<td></td>
<td>R18</td>
<td>1024</td>
<td>512</td>
<td>2048</td>
</tr>
<tr>
<td></td>
<td>Fold-Change</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PelB</td>
<td>WT</td>
<td>128</td>
<td>512</td>
<td>2048</td>
</tr>
<tr>
<td></td>
<td>R18</td>
<td>128</td>
<td>512</td>
<td>1024</td>
</tr>
<tr>
<td></td>
<td>Fold-Change</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>
3.4.8 Laboratory evolution predicts natural mutations

IMP-1 was first reported in 1994, followed by VIM-2 in 1999, and NDM-1 in 2009 (Osano et al. 1994; Poirel et al. 2000; Yong et al. 2009). Since their first detection, new variants have emerged at a steady pace, suggesting that these enzymes continue to evolve and adapt in response to antibiotic treatments in the clinic (Meini et al. 2014). I sought to determine if our directed evolution experiment accurately reflects this natural evolution by comparing the mutations obtained in nature to the high frequency mutations acquired by round 18 in the three trajectories. The frequency of natural non-synonymous mutations was mapped onto the primary sequences of the three MBLs (Figure 3.10). Synonymous mutations are unfortunately not highlighted in any of the reported variants, despite the important role that they may have with respect to host specific expression. The majority of the positions that are mutated at a high frequency by round 18 in the VIM-2 and IMP-1 trajectories are also mutated in natural variants (68% and 76% respectively). Interestingly, NDM-1 does not display the same trend – only 2 of 19 laboratory-mutated positions are found mutated in nature (11%). It is unknown whether this is a result of reduced sampling (i.e., NDM-type variants have only one to three mutated positions, whereas VIM-2 and IMP-1 variants are further diverged, with upwards of 50 positions mutated in certain isolates) or if the selection used in this experiment does not adequately mirror the situation in nature. Indeed, NDM-1 localizes to outer membrane vesicles, whereas VIM-2 and IMP-1 are soluble periplasmic enzymes (King & Strynadka 2011; González et al. 2016). It is possible that selection on agar plates better reflects the selection pressure applied to VIM-2 and IMP-1 in nature, rather than NDM-1. Further, it is possible that the mutagenesis strategy does not fully capture all possible mutations at the same probability that could be expected in nature. The libraries were constructed with base analogs and the mutation rates could therefore be skewed from what is
observed in nature. It is possible that the artificial library construction may constrain the trajectories to different degrees, and be more apparent in the NDM-1 trajectory.

While few of the high frequency mutations are identical to those found in nature (only 11% of the mutations in NDM-1, 26% in VIM-2, and 18% in IMP-1), the majority of the mutations are similar (i.e., F28V in the IMP-1 trajectory, compared to F28L in IMP-2, IMP-8, IMP-11, IMP-12, IMP-13, IMP-14, IMP-16, IMP-19, IMP-20, IMP-21, IMP-22, IMP-24, IMP-27, and IMP-33). Further, several of the identical mutations have been shown to play critical roles in the function of the MBLs. For example, M154L in NDM-1 has been associated with optimizing the binding of zinc in low zinc conditions (Bahr et al. 2018). This suggests that the directed evolution system is fairly accurately capturing the selective conditions experienced by the MBLs in nature. As such, these experiments may be useful in illuminating potential evolutionary trajectories of the MBLs in the clinic.
Figure 3.10: The distribution of mutations in natural NDM, VIM, and IMP variants. The frequency at each position was determined from 16 NDM, 25 VIM, and 27 IMP natural variants.
(Bahr et al. 2018; Widmann et al. 2012) mapped onto the primary sequence of NDM-1, VIM-2, and IMP-1, with VIM-2’s secondary structure in black for comparison. Dark blue represents positions that are mutated in greater than 50% of the variants, whereas medium and light blue represent positions that are mutated in between 25 to 49% and 10 to 24% of variants, respectively. White indicates that no mutations are present in these positions. Beige indicates insertions and deletions between the three genes.

3.5 Discussion

The systematic directed evolution of three MBL genes towards improved resistance against the β-lactam ampicillin demonstrated the high evolvability of enzymes within the MBL family. Interestingly, while there was some modification to the kinetic parameters of the round 18 variants, the MBLs appear to provide higher resistance primarily through an increase in functional periplasmic expression. This suggests that these moderately efficient enzymes do not have an easily accessible path towards developing higher catalytic efficiency for the hydrolysis of ampicillin. It is theorized that the highest efficiency an enzyme can achieve is limited by the substrate diffusion rate, and as such, a ‘perfect enzyme’ could only reach a maximum efficiency of approximately $10^9$ to $10^{10}$ M$^{-1}$s$^{-1}$ (Zhou & Zhong 1982; Bar-Even et al. 2011). Why then are the MBLs limited to catalytic efficiency values between $10^4$ and $10^7$ M$^{-1}$s$^{-1}$ for the various β-lactam antibiotics? A possible explanation for this is Pareto optimality: no design can be optimal for the performance of multiple tasks at the same time, and thus, functional trade-offs will arise in evolutionary optimization processes (Shoval et al. 2012; Tokuriki et al. 2012; Tendler et al. 2015). Substantially increasing catalytic efficiency for one particular antibiotic would likely come at the cost of decreasing efficiency for others. This could result in a significant disadvantage – specificity in bacterial warfare would leave an organism vulnerable to other threats. Perhaps broad substrate specificity has been selected for in the MBLs simply because
having a moderately efficient, generalist enzyme can better protect the organism. Over time, selection for broad substrate specificity may have become so ingrained in the sequence and structure of the MBLs, that they now occupy an inescapable fitness peak, from which no additional selection pressure from β-lactam antibiotics can produce improved catalytic efficiency. Indeed, the active site of the MBLs is large and wide, resulting in few specific contacts with the β-lactam antibiotics (King & Strynadka 2011). Thus, extensive remodeling of the active site would likely be required to promote the strong, specific binding needed for higher catalytic efficiency (Bar-Even et al. 2011). Overall, the inability of the catalytic efficiency to greatly improve highlights our previous result detailed in Chapter 2: catalytic efficiency is necessary for effective resistance, but expression is just as if not more important.

Our investigation into the increase in functional periplasmic expression revealed that it arose largely through improved overall expression. However, the process of translocation should not be overlooked. While the VIM-2 trajectory maintained a relatively consistent ratio between the number of enzymes in the periplasm and those in the cytoplasm, the variation in the IMP-1 trajectory highlighted how the efficiency of translocation can exert strong control over periplasmic expression, and thus, resistance. The dramatic increase in functional periplasmic expression therefore raises the question: Why were the MBLs not already optimized for expression and translocation? Interestingly, the improvement in the directed evolution appeared to be partially or fully host specific, which suggests that the MBLs were adapting to the host E. coli’s expression systems at the cost of decreased compatibility with other organisms. I previously demonstrated in Chapter 2 that the genotypic variation among MBLs emerges largely as variation in expression among different bacterial hosts and posited that this reflected varying levels of compatibility with each host’s expression systems. Genetic incompatibility may present
as a significant barrier to the deployment of each MBL and subsequent conferral of effective resistance. However, this experiment demonstrates that these barriers may be swiftly overcome through adaptation.

This has grave implications for the dissemination of MBLs and other antibiotic resistance determinants through horizontal gene transfer. If these constraints can be readily overcome through evolutionary optimization, then there is likely little to contain the spread of antibiotic resistance to pathogenic bacteria. However, this is not the situation that has been observed in the clinic: it has been noted that antibiotic resistance is not total or universal (Waglechner & Wright 2017). Further understanding the hidden codes that control functional protein expression and thus, shape protein evolution, is important for understanding the functional constraints that restrict the dissemination of antibiotic resistance determinants and explaining this discrepancy. It may also lead to the development of strategies that capitalize on these restraints to limit the spread of antibiotic resistance genes in the clinic.
Chapter 4: The broad substrate specificity of the metallo-β-lactamases is an ingrained trait, but can be disrupted

4.1 Summary

The broad substrate specificity of the metallo-β-lactamases is a unique feature that distinguishes them from other β-lactamases. But how did this feature emerge? I probe this question by performing a long term directed evolution experiment in which NDM-1, VIM-2, and IMP-1 were selected for three different β-lactam antibiotics to determine if functional trade-offs result in a narrowing of their substrate spectrum. Of seven completed trajectories, only two exhibited noticeable substrate specificity changes: IMP-1, evolved under selection for cefotaxime, demonstrated increased cefotaxime and meropenem resistance, while VIM-2, evolved under selection for meropenem, demonstrated only increased meropenem resistance. Further analysis of the VIM-2 trajectory revealed the mechanism by which its substrate spectrum narrowed: three active site mutations and one mutation in the enzyme’s core contributed the majority of the change but did not dramatically alter VIM-2’s ability to catalyze the hydrolysis of meropenem. In fact, these mutations were only slightly beneficial to neutral with respect to meropenem activity, but proved deleterious for ampicillin and cefotaxime. To determine if VIM-2’s broad substrate specificity could be restored, three additional rounds of directed evolution were performed in which the diminished ampicillin activity was selected for. The broad specificity phenotype of VIM-2 was readily returned, while the genotype diverged farther from wild type sequence. Interestingly, despite the increased divergence, the reversion of two mutations acquired in the initial evolution appeared to be critical to the restoration of broad specificity. Overall, this suggests that while selection has likely played a significant role in the development
and long-term maintenance of the MBLs’ broad substrate specificity, it is now a relatively ingrained trait that cannot be easily altered, and can be readily restored when lost.

### 4.2 Introduction

One of the unique features that differentiate the MBLs from other β-lactamases is their broad substrate specificity. In contrast to the SBLs, which emerged in the 1960s and have slowly expanded their ability to degrade antibiotics, from penicillins to cephalosporins to carbapenems, the recent clinical emergence of the B1-MBLs has been defined by a remarkable ability to provide effective resistance against nearly all clinical β-lactam antibiotics, save the monobactam, aztreonam (Bebrone 2007). This broad substrate specificity appears to be a common feature within the B1-MBL family. Indeed, despite sharing less than 35% amino acid sequence identity between each other, acquired MBLs like NDM-1, VIM-2 and IMP-1, have all been shown to efficiently hydrolyze penicillins, cephalosporins, and carbapenems with $k_{cat}/K_M$ values between $1 \times 10^5$ to $5 \times 10^6$ M$^{-1}$s$^{-1}$ (Chapter 2). Even ElBla2, a B1-MBL that was isolated far from any clinical environment in the North Sea, confers resistance to a wide range of β-lactam antibiotics when expressed in *E. coli* (Jiang et al. 2018). As it is unlikely that this environmental variant has been exposed to any of the clinical antibiotics in use, its broad substrate spectrum suggests that this feature in the B1-MBL family is not the result of selection from clinical antibiotics. Two possibilities then remain for the origin of the family’s broad substrate specificity: first, that this trait has evolved over long time periods under selection from natural antibiotics. Second, that the broad specificity is not a selected trait, but rather an emergent property that is a byproduct of their unique active-site structure.
How did the MBLs’ broad substrate specificity arise? Can this broad substrate specificity be altered and will there be functional trade-offs? What molecular features allow MBLs to provide such broad substrate specificity? I sought to answer these questions by using directed evolution to narrow the substrate spectrum of three MBLs by selecting for one of three β-lactam antibiotics over 18 rounds of evolution. If broad substrate specificity is an inherent property of the MBLs, then it should prove difficult to disrupt. However, if functional trade-offs result in the narrowing of their broad spectra, it would suggest that the broad specificity is the result of continual selection from a broad range of antibiotics. Any isolated variants with altered substrate specificity will be biochemically characterized to determine the mechanisms by which MBLs may modify their substrate specificity to reveal what molecular features grant it in the first place.

4.3 Methods

4.3.1 Creation of mutagenized libraries

Randomly mutagenized libraries of NDM-1, VIM-2, and IMP-1 were created as described in Chapter 3.

4.3.2 Application of selection

The mutagenized libraries were screened as in Chapter 3, but with the cefotaxime and meropenem. Each trajectory was plated onto a series of LB agar plates containing 2-fold increases in the concentration of each β-lactam antibiotic: meropenem, from 0.008 to 8 µg/mL, and cefotaxime, from 0.063 to 512 µg/mL. As before, the plate with the highest concentration of each antibiotic that had between 100 and 1000 colonies was collected. The plasmids were then extracted and used as the template for the next round.
To restore broad substrate specificity, the VIM-2 meropenem trajectory was selected with ampicillin for three additional rounds (19 to 21). The same protocol was employed, except each round was plated onto a series of LB agar plates containing 2-fold increases in the concentration of ampicillin, from 2 to 8192 μg/mL.

4.3.3 Measuring the minimum inhibitory concentration for selected variants
To assess the fitness of individual variants within the trajectories, the minimum inhibitory concentration (MIC) was determined by identifying the concentration of antibiotics by which no growth was observed in at least three of the four replicates for each variant, as described in Chapter 3.

4.3.4 Sanger sequencing of selected variants
Between 5-10 variants from selected rounds of each trajectory were isolated and sent for Sanger sequencing (Genewiz). The sequence results were visually inspected in Geneious® (8.1.9) and the mutations were identified.

4.3.5 Purification of Strep-tagged MBLs
VIM-1, VIM-2, and a round 18 variant from the VIM-2 meropenem and VIM-2 ampicillin trajectories were cloned into a pET-26(b) vector without their signal peptide and with a C-terminal Strep-tag (GNSGSAWSHPQFEK). Each enzyme was expressed and purified, as described in Chapter 2. The concentration of each protein was determined by spectrophotometer. The $A_{280}$ was measured for each sample with the following extinction coefficients, which were

99
calculated with ExPASy ProtParam (Gasteiger et al. 2005): VIM-1, 33,920 M⁻¹cm⁻¹; VIM-2, 35,410 M⁻¹cm⁻¹; V2-MEM, 35,410 M⁻¹cm⁻¹; and V2-AMP, 33,920 M⁻¹cm⁻¹.

4.3.6 Enzyme assays to determine kinetic parameters

The catalytic ability of each MBL enzyme was measured for seven β-lactam substrates (CENTA, cefotaxime, ceftazidime, meropenem, imipenem, ampicillin, and benzylpenicillin) as described in Chapter 2.

4.3.7 Deep mutational scanning of a VIM-2 comprehensive codon mutagenized library

Saturation mutagenesis was used to create a comprehensive codon mutagenized library for VIM-2, in which every codon position was mutated to every possible codon (64 in total). The library for each position was combined together and transformed with *E. coli* 10G, which was then grown for 6 hours in the presence of either cefotaxime, meropenem, or ampicillin before the variants were collected and the frequency of each codon was determined with next-generation sequencing. The enrichment of each variant was obtained by dividing the frequency of each variant in the selected condition (*i.e.*, grown with antibiotic) over the frequency of the same variant in the non-selected condition (*i.e.*, grown without antibiotic). The wild type enrichment was calculated in the same manner. The fitness score is calculated by taking the base 2 logarithm of the variant enrichment divided by the enrichment of the wild type.
4.4 Results

4.4.1 Directed evolution resulted in increased resistance in all trajectories

NDM-1, VIM-2, and IMP-1 were subjected to 18 rounds of directed evolution as described in Chapter 3, but here, these enzymes were selected for an improved ability to provide resistance against cefotaxime and meropenem. The three MBLs were mutagenized and transformed with *E. coli*, which were then separately selected for increased resistance towards each of the two antibiotics. In combination with the ampicillin trajectories from Chapter 3, this created nine separate evolutionary trajectories (three enzymes × three antibiotics). After completion of the trajectories, the MICs for six to twelve randomly selected variants from rounds 1, 2, 3, 6, 9, and 18 were determined for each antibiotic. The average fold-improvements in MIC over the wild-type enzyme were then calculated and plotted (Figure 4.1). These variants were sequenced in order to identify the mutations that are responsible for the MIC changes. Unfortunately, it was determined that the meropenem trajectories for NDM-1 and IMP-1 were contaminated with genes from the cefotaxime trajectories after the 9th and 6th rounds respectively, as sequence patterns from earlier rounds were completely supplanted by those from the cefotaxime trajectories. Thus these trajectories were omitted from further analysis.
Figure 4.1: The improvement in the average MIC in each trajectory over 18 rounds of selection for increased resistance to cefotaxime, meropenem, and ampicillin. The fold-improvement in MIC for cefotaxime is shown in blue, meropenem in green, and ampicillin in red. Each data point represents the average MIC for 6-12 variants from that round. The solid line and larger data point represents the antibiotic that was used in the selection of that trajectory, whereas the dotted line and smaller data points reflect the change in MIC for the two other antibiotics not under selection. The NDM-1 and IMP-1 meropenem trajectories were terminated at round 9 and 6 respectively, due to contamination from the cefotaxime trajectories.

The improvement in MIC values in the remaining cefotaxime and meropenem trajectories exhibited a similar pattern as to what was observed with the three previous ampicillin trajectories. By round 18, the average cefotaxime MIC for the NDM-1, VIM-2, and IMP-1 cefotaxime trajectories increased 32, 48, and 320-fold respectively, while the average meropenem MIC for the VIM-2 meropenem trajectory increased 116-fold. Not only was the magnitude of the improvement in these trajectories similar to the average improvement observed in the three ampicillin trajectories, which were 80, 60, and 48-fold for NDM-1, VIM-2, and IMP-
1 respectively, but all of the trajectories also demonstrated diminishing returns, as the largest improvements in MIC occurred early in the trajectories.

Five out of the seven completed trajectories (all three trajectories for ampicillin, and the NDM and VIM trajectories for cefotaxime) showed the same universal increase for all antibiotics (i.e., as resistance increased for the antibiotic under selection, it also increased for the other two antibiotics in a relatively similar manner). Interestingly, however, two trajectories exhibited a dramatic substrate specificity change. The IMP-1 cefotaxime trajectory produced an equivalent rise in resistance for both cefotaxime and meropenem (approximately 730-fold), while ampicillin resistance did not improve significantly (<1.5-fold), resulting in a dramatic shift in IMP-1’s substrate specificity. This can be measured in terms of the substrate specificity ratio, which was calculated by dividing the MICs of a given variant by the MICs of the wild type enzyme and determining the ratio between the products. This gives the wild type enzyme a balanced ratio of 1:1:1 for the three antibiotics, cefotaxime, meropenem, and ampicillin, respectively, which reflects its broad substrate specificity. If the substrate specificity shifts (e.g., cefotaxime resistance increases 2-fold relative to the other antibiotics, the new ratio would be 2:1:1, reflecting this change. The average substrate specificity ratio of R18 variants from the IMP-1 cefotaxime trajectory was 233:531:1, reflecting the drastic shift in substrate specificity. Similarly, the VIM-2 meropenem trajectory exhibited a large increase in meropenem resistance (approximately 120-fold), while resistance to the other two antibiotics only slightly improved (<4-fold), resulting in substrate specificity shifting greater than 33-fold, and a new substrate specificity ratio of 1.4:33:1. By contrast, the ratios for the remaining trajectories suggest the broad substrate specificity of the wild type enzymes is largely intact. The ratios for the NDM-1 and VIM-2 cefotaxime trajectories were 6:8:1 and 3.6:2.7:1, respectively, while the ratios for the
NDM-1, VIM-2, and IMP-1 ampicillin trajectories were 1:1:3.3, 1.2:1:2.3, and 1.8:3:1, respectively. This suggests that, while it can be altered, broad substrate specificity is a relatively robust trait of the MBLs while they are under positive selection for β-lactamase activity.

4.4.2 The shift in substrate specificity is reflected in the changing catalytic ability of VIM-2

In order to determine the molecular mechanisms by which such substrate specificity changes could occur, the VIM-2 meropenem trajectory was further investigated. I first sought to determine what effect the directed evolution had on the biochemical and biophysical properties of the population by purifying a single representative variant from the 18th round of the trajectory and measuring its kinetic parameters. The kinetic parameters for this isolated variant, V2-MEM, was determined and compared to those of two wild type VIM-2 and V2-AMP (a R18 variant from the VIM-2 ampicillin trajectory). As described in Chapter 2, the mature VIM genes were fused with a C-terminal Strep-tag by subcloning into the pET-26(b) vector, and overexpressed in *E. coli* BL21 (DE3). The VIM enzymes were subsequently purified and their kinetic parameters (\(k_{\text{cat}}\), \(K_M\), and \(k_{\text{cat}}/K_M\)) were determined for the six β-lactam antibiotic substrates, in addition to the β-lactamase substrate, CENTA (Table 4.1, Appendix A, Figure A.8). Interestingly, for V2-MEM, both \(k_{\text{cat}}\) and \(K_M\) increased for meropenem by approximately 18-fold, resulting in only a slight increase in the \(k_{\text{cat}}/K_M\) from \(1.2 \times 10^5\) to \(1.6 \times 10^5\) M\(^{-1}\)s\(^{-1}\). Unlike the previously discussed ampicillin trajectories, the concentration of meropenem used in selection (0.2 to 10 µM) was not high enough to fully saturate VIM-2 (with a \(K_M\) of 7.4 µM), which suggests that the \(K_M\) of the enzyme would still influence the rate of reaction and would be biologically relevant in the trajectory. It is unclear then, as to what advantage the increased \(K_M\) would provide in such a
situation, but perhaps it was offset by the slightly higher increase in $k_{\text{cat}}$, which provided a small benefit. The kinetic parameters for the other tested carbapenem, imipenem, were not significantly altered.

By contrast, the kinetic parameters for the other antibiotics of the cephalosporin and penam $\beta$-lactam subclasses that were not under selection demonstrated significant changes: the $K_M$ values for the V2-MEM enzyme could not be obtained as they had increased beyond the experimentally-permissible concentration range for cefotaxime, ceftazidime, and ampicillin in their respective kinetic assays, suggesting that there was a substantial increase in $K_M$ for all three antibiotics. This is also reflected by penicillin, which, while it could be determined, exhibited a 7-fold increase in $K_M$, with no change to the $k_{\text{cat}}$. The biological consequences of this are shown in the enzymes diminished ability to provide strong resistance against these antibiotics.

Overall, this suggests that the ~120-fold increase in meropenem resistance was driven not by an increase in catalytic ability, but likely by increased functional periplasmic expression, as with the ampicillin trajectories. The stagnant cefotaxime and ampicillin resistance (<5-fold change) was therefore a reflection of decreasing catalytic efficiency with these substrates offset by increasing expression. Overall, this data suggests that the observed substrate specificity change is largely a product of changes in the catalytic ability of the enzyme, specifically, a decreased ability to hydrolyze other $\beta$-lactam antibiotic subclasses not under selection.
Table 4.1: Measured kinetic parameters for VIM-type natural and evolved variants.

<table>
<thead>
<tr>
<th>Kinetics</th>
<th>VIM-1</th>
<th>VIM-2</th>
<th>V2 MEM</th>
<th>V2 AMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>72 ± 2</td>
<td>33 ± 1</td>
<td>39 ± 3</td>
<td>110 ± 10</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>54 ± 5</td>
<td>8.4 ± 0.9</td>
<td>24 ± 5</td>
<td>4.1 ± 0.7</td>
</tr>
<tr>
<td>$k_{cat}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>$1.3 \times 10^6$</td>
<td>$3.9 \times 10^6$</td>
<td>$1.7 \times 10^6$</td>
<td>$2.6 \times 10^7$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cefotaxime</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>190 ± 10</td>
<td>38 ± 2</td>
<td>n.d.</td>
<td>80 ± 4</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>73 ± 6</td>
<td>15 ± 3</td>
<td>n.d.</td>
<td>11 ± 2</td>
</tr>
<tr>
<td>$k_{cat}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>$2.7 \times 10^6$</td>
<td>$2.5 \times 10^6$</td>
<td>$1.2 \times 10^6$</td>
<td>$7.1 \times 10^6$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cefazidime</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>46 ± 4</td>
<td>1.1 ± 0.1</td>
<td>n.d.</td>
<td>12 ± 1</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>270 ± 40</td>
<td>66 ± 8</td>
<td>n.d.</td>
<td>170 ± 20</td>
</tr>
<tr>
<td>$k_{cat}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>$1.7 \times 10^5$</td>
<td>$1.6 \times 10^4$</td>
<td>$5.6 \times 10^4$</td>
<td>$7.0 \times 10^4$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Meropenem</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>19 ± 1</td>
<td>0.90 ± 0.03</td>
<td>18 ± 1</td>
<td>2.6 ± 0.1</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>140 ± 10</td>
<td>7.4 ± 1.0</td>
<td>120 ± 10</td>
<td>4.1 ± 0.5</td>
</tr>
<tr>
<td>$k_{cat}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>$1.4 \times 10^5$</td>
<td>$1.2 \times 10^5$</td>
<td>$1.6 \times 10^5$</td>
<td>$6.2 \times 10^5$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Imipenem</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>6.8 ± 0.3</td>
<td>8.7 ± 0.3</td>
<td>13 ± 1</td>
<td>72 ± 2</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>24 ± 4</td>
<td>8.4 ± 1.1</td>
<td>24 ± 3</td>
<td>30 ± 2</td>
</tr>
<tr>
<td>$k_{cat}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>$2.8 \times 10^5$</td>
<td>$1.0 \times 10^6$</td>
<td>$5.3 \times 10^5$</td>
<td>$2.4 \times 10^6$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ampicillin</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>130 ± 10</td>
<td>23 ± 1</td>
<td>n.d.</td>
<td>180 ± 5</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>140 ± 30</td>
<td>83 ± 15</td>
<td>n.d.</td>
<td>110 ± 10</td>
</tr>
<tr>
<td>$k_{cat}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>$9.2 \times 10^5$</td>
<td>$2.7 \times 10^5$</td>
<td>$2.7 \times 10^5$</td>
<td>$1.7 \times 10^6$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Penicillin</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>580 ± 30</td>
<td>76 ± 3</td>
<td>69 ± 5</td>
<td>185 ± 5</td>
</tr>
<tr>
<td>$K_M$ (µM)</td>
<td>170 ± 40</td>
<td>51 ± 10</td>
<td>360 ± 70</td>
<td>69 ± 6</td>
</tr>
<tr>
<td>$k_{cat}/K_M$ (M$^{-1}$•s$^{-1}$)</td>
<td>$3.5 \times 10^6$</td>
<td>$1.5 \times 10^6$</td>
<td>$1.9 \times 10^6$</td>
<td>$2.7 \times 10^6$</td>
</tr>
</tbody>
</table>

n.d. not determined
4.4.3 Mutations accumulated predominantly in the signal peptide, but also at specific positions throughout the catalytic domain of each MBL

To determine what mutations might be altering the kinetic parameters, and subsequently, the substrate specificity of the variants in the VIM-2 meropenem trajectory, the six to twelve variants from each selected round (3, 6, 9, and 18), whose MICs were previously measured, were sequenced and their mutations were compared to those obtained in the VIM-2 cefotaxime and ampicillin trajectories (Figure 4.2, Appendix B, Figure B.2). Mutations primarily accumulated in the signal peptide region in all three VIM-2 trajectories, a trend that was previously described for the ampicillin trajectories of NDM-1, VIM-2, and IMP-1 in Chapter 3. As before, this is likely increasing the functional periplasmic expression of each enzyme by increasing the level of expression and the efficiency of translocation and would explain why the variants could provide substantially higher resistance towards meropenem without a corresponding change in the underlying catalytic ability of the enzymes.
Figure 4.2: The distribution of accumulated mutations in randomly selected variants from VIM-2 cefotaxime, meropenem, and ampicillin trajectories. The frequency at each position was determined from 6-12 variants isolated and sequenced from the 3rd, 6th, 9th, and 18th rounds of each trajectory and mapped onto the primary sequence of each enzyme with VIM-2’s secondary structure displayed in black for comparison. The frequency at each position is indicated by blue shading.
In addition to the signal peptide, there were also many positions that were heavily mutated within the catalytic domain. Interestingly, of the 32 mutations identified in the catalytic domain in the 18th round of the VIM-2 meropenem trajectory, 9 mutations were also found in the cefotaxime and ampicillin trajectories (V66A, A97V, I99V, S215R, I223V, H254R, L265R, L297P, and K301R), while 11 mutations were unique, but located at positions that were also mutated in the other two trajectories (E37V, Y46N, I56T, D78S, A93T, T198A, I202V, R228H, Q253R, K301T, and T304A). The remaining 12 mutations were entirely unique to the VIM-2 meropenem trajectory (S54T, F61Y, F61H, L107H, D176N, F180L, S197T, S207F, T229A, S230T, I248L, and P261S) and a subset of these mutations is likely driving the substrate specificity change. However, the mutations in the former group cannot be entirely discounted, as they may interact with the other mutations to produce an effect that cannot be easily predicted in a phenomenon termed epistasis.

To further isolate which mutations are likely contributing the most to the substrate specificity change, the mutations in each round were categorized by frequency and related to the relative improvement of each antibiotic (Table 4.2). As the substrate specificity change is a global trend within the trajectory (i.e., all variants exhibit the change), it is likely that the mutations that arise frequently within the trajectory are the primary drivers of the specificity switch. Five mutations, E37V, I56T, S197T, S215R, and R228H, were entirely fixed in the 18th round of the trajectory (i.e., mutations present in 100% of the sequenced variants), while eight more, F61Y, S207F, I223V, S230T, I248L, H254R, L265R, and K301T, were present at a high frequency (i.e., mutations present in greater than 50% of the sequenced variants). The five fixed mutations were also fixed by round 6, however, when the change in specificity was less drastic (substrate specificity ratio of 2.5:8.5:1 in round 6, compared to the round 18 ratio of 1.4:33:1),
suggesting that the lower frequency mutations acquired in the 9\textsuperscript{th} and 18\textsuperscript{th} round are also contributing the change.
Table 4.2: Catalytic domain mutations identified in the sequenced variants from the VIM-2 meropenem trajectory

<table>
<thead>
<tr>
<th>Round</th>
<th>Frequency</th>
<th>Mutations(^1)</th>
<th>Substrate Specificity Ratio(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Fixed Mutations (100%)</td>
<td>AMP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R228H</td>
<td>1 2 4 8 16 32 64</td>
</tr>
<tr>
<td></td>
<td>Other Mutations</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fixed Mutations (100%)</td>
<td>AMP</td>
</tr>
<tr>
<td>6</td>
<td>High Frequency Mutations (&gt;50%)</td>
<td>K301T</td>
<td>1 2 4 8 16 32 64</td>
</tr>
<tr>
<td>9</td>
<td>High Frequency Mutations (&gt;50%)</td>
<td>K301T, K301R</td>
<td>1 2 4 8 16 32 64</td>
</tr>
<tr>
<td>18</td>
<td>High Frequency Mutations (&gt;50%)</td>
<td>F61Y, S207F, I223V, S230T, I248L, H254R, L265R, K301T</td>
<td>1 2 4 8 16 32 64</td>
</tr>
</tbody>
</table>

\(^1\) Mutations that are unique to the VIM-2 meropenem trajectory are bolded, with those at positions not mutated in other trajectories underlined.

\(^2\) The substrate specificity ratio was calculated by dividing the average MIC for each antibiotic for each round by the MICs of VIM-2, and determining the ratio between the values.
The mutations present in the V2-MEM purified enzyme were mapped onto the tertiary structure of VIM-2 (Figure 4.3). This included the five fixed mutations, E37V, I56T, S197T, S215R, and R228H, as well as five other mutations that had appeared at lower frequencies in the 18th round of the trajectory, F61Y, L107H, Q253R, H254R, and K301T. With exception to L107H and Q253R, all of these mutations were found at a >50% frequency in the other round 18 variants. Of the ten catalytic domain mutations present in V2-MEM, three mutations, F61Y, S197T, and R228H, are located within the active site, and one, I56T, is located on a β-strand in the interior of the enzyme, while the other six mutations (E37V, S197T, L107H, Q253R, H254R, and K301T) are located on the enzyme’s surface and distant from the active site. While these surface mutations could potentially affect the substrate specificity of the enzyme, it is far more likely that mutations present within the active site (and to a lesser extent, I56T in the interior) result in the modification of catalysis and/or substrate binding due to their proximity to the reaction. S197T is located beside one of the metal binding residues for the first zinc site, H196, and directly below Loop 10. Thus, S197T may alter substrate specificity directly through influencing the metal binding coordination and/or the shape of the active site. R228H is located on the N-terminal side of Loop 10 near the second metal binding site (coordinated by D120, C221, H263). Interestingly, R228 is frequently mutated in clinical isolates (Garcia-Saez et al. 2008; King & Strynadka 2011; Lassaux et al. 2011), and the effect of R228 in VIM-type variants has been extensively investigated. Indeed, it was recently shown that the substitution of Arg with Leu in VIM-24 produces a 6-fold increase in the $k_{cat}/K_M$ for ceftazidime over VIM-2 (Mojica, Mahler, et al. 2015). While the specific mutation, R228H, has not yet been identified in natural variants, it is likely that any mutation at this position could alter the catalytic activity and substrate specificity of the enzyme. Likewise, F61Y, located on Loop 3, has not been identified
in a clinical isolate (only F61L in VIM-7), but F61 is believed to play a role in aromatic-aromatic interactions with the substrate and any change may influence this interaction, catalysis, and subsequently, the substrate specificity (Widmann et al. 2012; Docquier 2003).
Figure 4.3: Mutations acquired by the 18th round of directed evolution in a purified VIM-2 variant, V2-MEM, which was selected for increased meropenem resistance. Fixed mutations, defined as positions mutated in all 6 sequenced variants from the 18th round, are represented as red spheres on the structure of VIM-2 (PDB ID 1KO3), whereas other mutations identified at lower frequencies are represented as orange spheres.
4.4.4 The substrate specificity change is the sum of the effects of multiple mutations

In order to positively ascertain the identity of the mutations that are driving the observed substrate specificity switch, the fitness effect of each mutation in isolation was determined. A deep mutational scanning experiment was performed to determine the general effect that each of the ten identified mutations in the V2-MEM variant had on VIM-2’s ability to confer resistance against cefotaxime, meropenem, and ampicillin (Appendix B, Figure B.4). Saturation mutagenesis was used to create a comprehensive codon mutagenized library for VIM-2, which was transformed with *E. coli*. The library was then grown for six hours in the presence of either cefotaxime, meropenem, or ampicillin before collection. The frequency of each mutation was present within the screened libraries was determined with next-generation sequencing. The enrichment of each variant was then obtained by dividing the frequency of each variant in the selected condition (i.e., grown with antibiotic) over the frequency of the same variant in the non-selected condition (i.e., grown without antibiotic). The wild type enrichment was calculated in the same manner. The fitness score is calculated by taking the base 2 logarithm of the variant enrichment divided by the enrichment of the wild type.

The fitness scores for each of the 20 possible mutations at each of the 10 positions that were mutated in the V2-MEM variant were extracted (Figure 4.4). Overall, 59% of the possible mutations at these positions affected the fitness of the enzyme negatively, with an average 2.7-fold decrease in relative enrichment among the deleterious mutations. By comparison, the average increase in relative enrichment for the remaining 41% was only 1.6-fold. This is consistent with previous observations that the majority of mutations are deleterious. However, mutations were not consistently deleterious between all ten positions: The fitness scores for
different amino acids at each position varied greatly overall, from S197, where most amino acid substitutions did prove deleterious, to R228, where many substitutions provided a fitness advantage. Substitutions at E37, S215, Q253, H254, and K301 largely had no effect on fitness scores compared to the wild type (i.e., mutations are neutral). This may reflect the structural and functional constraints placed upon each position. S197 is located in the active site, adjacent to the first zinc-binding site and immediately next to one of the metal-binding histidine residues. Variation at this position would likely have a substantial disruptive effect on the ability of the enzyme to bind the zinc ion, negatively affecting catalysis and the overall fitness of the enzyme. By contrast, K301 is located on a distal, C-terminal α-helix, with the side chain oriented into the solvent. Variation at this position would likely have little effect on the structure and function of VIM-2.
Figure 4.4: The relative enrichment under selection for cefotaxime, meropenem, and ampicillin, for each amino acid mutation at each of the 10 VIM-2 catalytic domain positions that had acquired a mutation in the V2-MEM round 18 variant. The wild type residue is coloured black. The residue that the position was mutated to in the trajectory is also highlighted by a black box. Mutations that were absent in the data set are coloured grey. Relative enrichment is shown by shading and colour: dark teal indicates that the mutation is beneficial compared to wild type, whereas dark brown indicates that the mutation is deleterious compared to the wild type.
Based on the fitness scores, the 10 mutations obtained in the V2-MEM variant could be divided into two categories: mutations with global effects (i.e., uniform effect on fitness, regardless of selection condition) and mutations with specific effects (i.e., varied effect on fitness, dependent upon the selection condition). E37V, S215R, Q253R, H254R, and K301T belong to the former category. While the particular effect of each of these mutations ranges from no change in relative enrichment (i.e., neutral) to a slight increase (1.8-fold increase), their relative enrichment differs by no more than 30% between selection conditions. The second group includes I56T, F61Y, S197T, and R228H, which have different effects depending on the selection condition. For all four mutations, the relative enrichment for meropenem is greater than their enrichment for cefotaxime and ampicillin, sometimes by up to 5-fold higher. This suggests that these four mutations are driving the substrate specificity change. Unfortunately, no fitness score could be calculated for L107H, as this specific mutant was not present at a high enough frequency in the initial comprehensive codon mutagenized library.

To determine if the overall substrate specificity change could be accurately predicted from the individual effects of each of the nine mutations for which data exists, the theoretical combined relative enrichment for each selection condition was calculated and the predicted substrate specificity ratio from these values was compared to the measured ratio (Figure 4.5). For example, the relative enrichment for F61Y and S197T under meropenem selection was 4.4 and 3.5 respectively, so the theoretical relative enrichment of this combination for that selection condition would be the product of 4.4 × 3.5, or 15.4. The theoretical values for the different conditions could then be used to predict the substrate specificity ratio for a given set of mutations. The predicted ratio of the improvement each antibiotic was 3.2:59:1 for cefotaxime, meropenem, and ampicillin respectively, whereas the measured ratio for the V2-MEM variant was 1:64:1.
Therefore, the predicted substrate specificity ratio aligns closely with the measured values. This suggests that the effects that the mutations acquired in the trajectory have on substrate specificity are additive and can be predicted. The majority of the predicted ratio can be recapitulated by four mutations that were identified to have specific effects on fitness. Alone, they produce a substrate specificity ratio of 3:47:1, suggesting that by and large, the observed substrate specificity change is a product of these four mutations, I56T, F61Y, S197T, and R228H.

Figure 4.5: The predicted and measured substrate specificity ratios for the V2-MEM variant from the 18th round of the VIM-2 meropenem trajectory.
4.4.5 Broad substrate specificity was swiftly restored upon changing selection pressure

The VIM-2 meropenem trajectory demonstrated that in some circumstances, the broad substrate spectrum of the MBLs could narrow to preclude several β-lactam antibiotics that they are not under active selection for. But how permanent is this change? To determine if the broad substrate specificity could be reestablished in the VIM-2 meropenem trajectory, three additional rounds of directed evolution were performed (rounds 19 to 21) in which the previously diminished ampicillin resistance was selected for, instead of meropenem. Eight variants from each of the three restoration rounds were randomly selected and the cefotaxime, meropenem, and ampicillin MICs were determined and compared to the values from the previous 18 rounds (Figure 4.6). The broad substrate specificity was readily restored in the trajectory, but not to the same extent as the wild type enzyme: Round 21 variants exhibited an average substrate specificity ratio of 1:2.5:1 for cefotaxime, meropenem, and ampicillin respectively, down from the 1:64:1 ratio exhibited by the round 18 V2-MEM variant and the 1.4:33:1 average ratio for the 18th round of the trajectory, but still not reaching the 1:1:1 ratio that would indicate a native level of broad specificity.
Figure 4.6: The changing substrate specificity throughout the VIM-2 meropenem trajectory. Rounds 1 to 18 were subjected to selection for cefotaxime (green shading). Rounds 19 through 21 were subjected to selection for ampicillin (pink shading). The fold-improvement in MIC for cefotaxime is shown in blue, meropenem in green, and ampicillin in red. Each data point represents the average MIC for 6-12 variants from that round. The solid line and larger data point represents the antibiotic that was used in the selection of that trajectory, whereas the dotted line and smaller data points reflect the change in MIC for the two other antibiotics not under selection.

To determine what mutations might be restoring the broad substrate specificity in the population, four variants from round 21 were sequenced. Interestingly, all four variants had the same nucleotide sequence, suggesting that the population in round 21 has little diversity, and could potentially be monoclonal. Comparison to the V2-MEM round 18 variant revealed that there were seven new mutations in the four round 21 variants (V40A, P68S, I103V, I223V, I248L, E249G, and E255D) that were not previously identified in the trajectory, while five mutations that were in V2-MEM were no longer present (F61Y, L107H, R228H, Q253R, and H254R). The large distance between the round 21 sequences and V2-MEM suggests that the round 21 variants arose from another lineage within the population. Examination of the other
sequenced round 18 variants shows several that are more similar to the round 21 sequences. Indeed, the closest round 18 variant differs by only the reversion of F61Y and R228H, as well as the additions of V40A, P68S, I103V, and E249G.

The mutations in the round 21 variants were mapped onto the tertiary structure of VIM-2 (Figure 4.7). The only new mutation found in the active site was P68S, which appears at the base of loop 3 and may play a role in altering the flexibility of the loop. The other six new mutations were located either on the surface (V40A, I103V, E249G, and E255D) or in the interior of the enzyme below loop 10 (I223V and I248L). As previously discussed, F61Y and R228H are both located in the active site on loops 3 and 10 respectively, and their loss or reversion would likely have the opposite effect that was shown before, and result in a decrease in meropenem resistance relative to cefotaxime and ampicillin. Q253R and H254R were located on the exterior of the enzyme and had no effect on the substrate specificity before, so their absence likely has no effect on the substrate specificity of the R21 variants. L107H was located on the exterior of the enzyme as well, but its role in substrate specificity could not be directly identified due to the lack of DMS data for that particular mutation. While it is unlikely that a mutation so remote from the active site would alter substrate specificity, and other mutations at this position appear to have no effect on substrate specificity, it cannot be entirely discounted as there is precedence for long-distance interactions affecting the catalytic properties of the MBLs (Sohier et al. 2013).
Figure 4.7: The location of mutations in the round 21 variants from the VIM-2 meropenem trajectory after three rounds of selection for ampicillin resistance. High frequency mutations, defined as positions mutated in at least 50% of four sequenced variants from each round (19, 20 and 21), are represented as spheres on the structure of VIM-2 (1KO3). Round 19 mutations are shown in red, while mutations from earlier in the trajectory are shown in blue.
In order to positively ascertain the identity of the mutations that are driving the restoration of broad substrate specificity, the individual effects of each mutation were examined through deep mutational scanning. The fitness scores for each of the 20 possible mutations at each of the seven newly mutated positions in the round 21 variants were extracted (Figure 4.8). In general, substitution of I103, I223, or I248 with any other residue aside from leucine or valine was largely deleterious, whereas V40, P68, E249, and E255 were far more amenable to substitution with most of the possible residues. Again, the specific mutations fell into two categories: the fitness effects of P68S, I223V, E249G, and E255D were global and did not vary with between selection conditions, whereas V40A, I103V, and I248L affected fitness differently, depending on the selection. Interestingly, only V40A had greater relative enrichment when selected with cefotaxime and ampicillin than with meropenem. I248L demonstrated the exact opposite trend, with the relative enrichment under meropenem selection being 2 to 4-fold higher than that of cefotaxime and ampicillin selection respectively. However, as I248L was present in some of the round 18 variants, it is likely that it is a holdover from the meropenem selection and did not arise while under selection for ampicillin in rounds 19 through 21.
Figure 4.8: The relative enrichment under selection for cefotaxime, meropenem, and ampicillin, for each amino acid mutation at each of the six VIM-2 catalytic domain positions that were mutated at high frequency in the three rounds of directed evolution towards increased ampicillin resistance. The wild type residue is coloured black. The residue that the position was mutated to in the trajectory is also highlighted by a black box. Mutations that were absent in the data set are coloured grey. Relative enrichment is shown by shading and colour: dark teal indicates that the mutation is beneficial compared to wild type, whereas dark brown indicates that the mutation is deleterious compared to the wild type.

To determine if the overall substrate specificity change observed in the three restoration rounds could be recapitulated from the individual effects of the mutations, the theoretical
combined relative enrichment for each condition was calculated for the new round 21 variants, and the substrate specificity ratio was subsequently determined and compared to the measured ratio (Figure 4.9). While the predicted ratio was relatively close to the measured value, it was evident that I248L largely contributed the difference between the predicted and measured ratios. The discrepancy indicates that epistatic interactions are affecting the substrate specificity of the enzyme, and these interactions most likely involve I248L. Indeed, unlike V2-MEM where the mutations were largely isolated from each other, I248L is in direct contact with I223V. This suggests that the interactions between them may be serving to alter the overall effect that both mutations have on the substrate specificity of the enzyme, although more work is needed to conclusively demonstrate this.

![Figure 4.9: The predicted and measured substrate specificity ratios for the round 21 variants of the VIM-2 meropenem trajectory.](image)

Figure 4.9: The predicted and measured substrate specificity ratios for the round 21 variants of the VIM-2 meropenem trajectory.
4.5 Discussion

The completion of seven separate directed evolution trajectories in which three MBL genes were selected for an increased ability to confer resistance against different β-lactam antibiotics demonstrated the relative robustness of the MBLs’ broad substrate specificity. Of the seven completed trajectories, only two exhibited a noticeable substrate specificity shift, while the other five remained broadly promiscuous, suggesting that strong, positive selection for a single β-lactam antibiotic can often maintain effective resistance for the other β-lactam antibiotics. However, further examination of the VIM-2 meropenem trajectory, which exhibited an altered average substrate specificity ratio of 1.4:33:1 for cefotaxime, meropenem, and ampicillin respectively, demonstrated that this is not always the case and the MBLs’ broad substrate specificity can be disrupted. The pliability of active site residues allowed for enzymes in this trajectory to incorporate mutations that were beneficial or neutral with respect to meropenem selection, but deleterious to the degradation of the other antibiotics. Interestingly, as the catalytic ability of the enzymes to hydrolyze meropenem did not drastically improve, it is likely that these mutations are either ‘hitchhikers’ that accrued by chance or were selected for an effect on the expression of the enzyme. Taken together, this suggests that while the broad substrate specificity of the MBLs is relatively resistant to modification, it can still be disrupted by chance.

The restoration of broad substrate specificity was quickly achieved with three additional rounds of directed evolution in which ampicillin was used for selection instead of meropenem. The ease with which this reversion was achieved stands in contrast to several recent experiments: In 2009, Bridgham et. al. examined the evolutionary history of a glucocorticoid receptor with ancestral sequence reconstruction and determined that the accumulation of several ‘restrictive’ mutations makes it unlikely that the receptor would be able to directly revert to its ancestral
function (Bridgham et al. 2009). Kaltenbach et. al. then demonstrated in 2015 the phenotypic reversion of previously evolved arylesterase back towards its native phosphotriesterase activity. They concluded that while phenotypic reversion was possible, epistatic interactions preclude the complete reversion of the genotype (Kaltenbach et al. 2015). The restoration of broad substrate specificity demonstrated similar trends: the genotype of VIM-2 was not reestablished, but rather, further diverged due to the accumulation of new mutations. However, in contrast to the arylesterase’s trajectory, it is evident that the restoration of VIM-2’s native substrate specificity appeared to be largely driven by at least two critical reversions (Y61F and H228R). Deep mutational scanning allowed for a full account of the fitness effects of each mutation and revealed that the reversion of these two mutations contributed substantially to the reversal of phenotype. Why do reversions feature so prominently in the restoration of VIM-2’s broad substrate specificity but not in the reversion of an arylesterase back into a phosphotriesterase?

There are likely two reasons for this. First, the level of mutagenesis is different between the two experiments: the arylesterase accumulated 25 mutations in its catalytic domain whereas the V2-MEM had only accumulated approximately 10. A greater number of mutations increases the probability that ‘restrictive’ mutations are acquired and that epistasis will dominate. Second, the evolution of VIM-2 was to enhance its native function, whereas the phosphotriesterase was evolved for its promiscuous arylesterase activity. This necessitated the complete remodeling of its active site for the new activity, whereas VIM-2 required only minor adjustments. As a result, the mutations acquired in the VIM-2 trajectory were further apart (e.g., the active site mutations were located on separate loops). This isolation may have facilitated the ease with which they were reverted, as their sudden appearance in the trajectory was not ‘locked in’ by any other restrictive mutation in proximity. Indeed, the relative accuracy with which the substrate
specificity ratio was initially calculated contrasts with the discrepancy between the prediction and measured ratios for the restored variants. The primary difference between these two enzymes was that the restoration variant contained mutations in positions near each other. Regardless of the specific mechanism by which the restoration was achieved, the prompt reversion suggests that even when the broad substrate specificity of the MBLs is lost, it can be swiftly restored upon changing selective pressure.

How then did the MBLs’ broad substrate specificity develop? Is it an emergent property, arising as a byproduct of its structure, or has it been specifically selected for? Our results suggest that both explanations are likely true to some degree. The broad substrate specificity of the enzymes is relatively robust and often maintained while under selection with a single β-lactam antibiotic. When narrowed, their broad spectrum can be readily restored upon a change in selection. This supports the proposition that broad specificity is an inherent structural trait. Indeed, the active site architecture of the MBLs certainly lends itself to broad specificity: the separated loops act as isolated islands, which are flexible and can accommodate many different substrates. However, this isolation and flexibility facilitates the acquisition of active site mutations that can alter the substrate specificity. The fact that substrate specificity can be disrupted suggests that there must be some selection acting on the MBLs to maintain their broad substrate specificity over long evolutionary time periods. As previously discussed, bacteria may be engaged in constant bacterial warfare and are likely exposed to a plethora of natural antibiotics (Abrudan et al. 2015). Exposure to this wide range of natural β-lactam antibiotics has likely maintained the broad substrate specificity of MBLs since their divergence. However, regardless of whether the MBLs’ broad substrate specificity was selected for or an emergent
property, it is evident that it is now an ingrained trait that is present throughout the B1-MBL family and will likely be a defining feature in the clinical landscape for the foreseeable future.
Chapter 5: The experimental genetic drift of VIM-2 reveals long term evolutionary trends

5.1 Summary

How do orthologs exhibit such striking sequence diversity while maintaining structural and functional similarity? I address this question by performing a long-term genetic drift experiment on a previously evolved, highly fit VIM-2 population at different functional thresholds. The fitness of the high functional threshold trajectory remained high throughout 50 rounds of genetic drift (as measured by EC$_{50}$ with ampicillin) whereas the fitness of the low functional threshold trajectory decreased 34-fold over the first 20 rounds, before remaining relatively constant for the last 30. A third trajectory, which saw an oscillation occur between the low and high thresholds after round 20 through a 12 round adaptive selection period, saw rapid restoration of fitness. On average, individual genes from the 40$^{th}$ round for the T10, TOsc, and T1000 trajectories had 17.5%, 15.9%, and 14.7% of their positions mutated, with 59% of all positions mutated at least once in at least one of the sequenced genes. This highlights the incredible flexibility of the MBL structure and explains the great diversity present in the B1-MBL family, despite its functional and structural homogeneity. Future work should use these genetic drift trajectories as a platform to explore how genetic drift can influence the phenotypic diversity within the family, while also determining the extent by which epistatic interactions can constrain MBL evolution and function.

5.2 Introduction

Some estimates suggest the biosphere contains up to one trillion different species (Locey & Lennon 2016). This remarkable biodiversity is a reflection of the power of evolution and the two
underlying processes by which variation becomes established in nature: natural selection, in which beneficial mutations providing a fitness advantage are fixed in a population, and genetic drift, in which neutral (or sometimes deleterious) mutations can spread through a population by chance. While genetic drift is thought to be responsible for the majority of genetic variation present in nature, few experiments have focused upon recreating and studying the effects of genetic drift at the protein level (Kimura 1968; Bershtein et al. 2006; Bershtein et al. 2008; Gupta & Tawfik 2008; Smith et al. 2011; Petrie & Joyce 2014; Rockah-Shmuel et al. 2015).

The B1 Metallo-β-Lactamase (MBL) family is an excellent example of the prevalence of genetic drift among proteins and provides a unique opportunity to study it in both nature and the laboratory. While I previously found that adaptation to an MBL’s host and altering the catalytic ability of the MBL were strong drivers for the fixation of mutations in an adaptive evolution experiment, the few mutations required to adapt to increase resistance before diminishing returns makes further improvements negligible (<10% of the amino acid sequence) pales in comparison to the full extent of genetic diversity present in the family. Indeed, despite striking phenotypic similarity in structure and function, these orthologs often share between 20-40% pairwise amino acid sequence identity, with only 16 residues (<7%) conserved between a small sampling of eight family members (Chapter 1, Figure 1.4). This suggests that genetic drift has played a significant role in the divergence of this family, leading to the diversity that is observed today.

Unfortunately, while the extent of the family’s genetic variation is well known, the mechanism by which genetic drift produced such diverse orthologs is unclear due to our limited view of the family and our inability to study the genetic ‘fossil record’. Many diverse MBLs and related isolates have been identified (i.e., variants differing by only a single mutation to up to 10% of the gene), but the gradual steps between such clusters remain missing. For example, the
next nearest relative of NDM-1, an MBL originally found in a clinical *K. pneumoniae* strain, is ElBla2, an ortholog from the marine bacteria *Erythrobacter litoralis* HTCC 2594, which shares only 56% amino acid identity (Jiang et al. 2018). This may be due to a bias in the sampling of clinical isolates compared to the less explored metagenome or it may be that these sequences are the only remaining extant branches today. Regardless, in the absence of the requisite data, we can recreate the genetic drift of the MBLs in the laboratory to better understand how genetic drift can create such divergent homologs with conserved structure and function. In doing so, I can investigate how genetic drift can influence the phenotypic diversity within the family and the emergence of new functions in the superfamily, while also determining the extent by which epistatic interactions can constrain MBL evolution and function.

Here, I established two long-term trajectories of genetic drift with a previously evolved VIM-2 population by mutagenizing and applying a purifying selection at a constant high and low threshold, in addition to a third oscillating trajectory that combines genetic drift with natural selection, and two shorter control and sub-inhibitory trajectories. These trajectories will be used to determine how genetic drift at various selective thresholds may have shaped the B1-MBL family, in addition to serving as a platform to better understand the role of epistasis in genetic drift and the emergence of novel enzymatic properties.

### 5.3 Methods

#### 5.3.1 Creation of mutagenized libraries

The 18th round of an adaptive evolution experiment in which VIM-2 was repeatedly mutagenized and selected for increasing resistance towards a β-lactam, ampicillin, was used at the starting
point for the genetic drift (Chapter 3). Randomly mutagenized libraries were created by application of the same protocol described in Chapter 3.

5.3.2 Application of purifying selection

The mutagenized libraries were electroporated with E. cloni® 10G E. coli cells (Lucigen), suspended in 1 mL of LB medium, and allowed to recover for 1 hour at 37°C. The transformations were plated on 15 cm LB agar plates containing 34 µg/mL of chloramphenicol and the requisite ampicillin for each trajectory (1000 µg/mL of the high threshold, 10 µg/mL for the low threshold, 1 µg/mL for the sub-inhibitory threshold, and 0 µg/mL for the control trajectory). The plates were incubated overnight at 37°C, yielding over $10^5$ bacterial colonies at the high threshold, and over $10^6$ bacterial colonies for the other trajectories. The colonies were collected and suspended in LB medium before centrifugation and plasmid extraction with the Mini-Prep Kit (Qiagen) and quantification. The extracted DNA was then used as the template in the next round of genetic drift. In total, 10 rounds of the control trajectory was performed (T0), 20 rounds of the sub-inhibitory trajectory (T1), and 60 rounds of both the high and low threshold trajectories (T1000 and T10 respectively). The TOsc trajectory was created from T10 after round 20 by applying increasing selection pressure from 10 µg/mL to 1000 µg/mL over the course of 8 rounds, holding it at that level for 4 rounds, before returning it to 10 µg/mL until round 50.

5.3.3 Determining the fitness of the drifted populations

To measure the approximate fitness of each trajectory, growth under selection by ampicillin was measured and compared to uninhibited growth for every other round of drift. 5 ng from every other library was transformed with E. cloni® 10G E. coli cells (Lucigen), suspended in 1 mL of
LB medium and recovered for 1 hour at 37°C. The suspension of cells was then added to 10 mL of LB medium with 34 µg/mL of chloramphenicol and incubated overnight at 30°C. The OD$_{600}$ for each culture was normalized to 0.5 and 20 µL of each was added in duplicate in a 96-well plate to 180 µL of LB medium with a varying concentration of ampicillin to create cultures with the following final ampicillin concentrations: 0 µg/mL, 10 µg/mL, 25 µg/mL, 50 µg/mL, 100 µg/mL, 250 µg/mL, 500 µg/mL, 1,000 µg/mL, 2,500 µg/mL, 5,000 µg/mL, 10,000 µg/mL, and 25,000 µg/mL. The plates containing the libraries were incubated for 6 hours at 37°C before a final OD$_{600}$ measurement was taken. The average of the duplicates was compared to the control culture to determine a percent fitness for that library at that concentration of ampicillin.

5.3.4 Sanger sequencing selected variants and the development of a python program for sequence data analysis

Between 5-10 variants from selected rounds of genetic drift were isolated and sent for Sanger sequencing (Genewiz). The sequence results were visually inspected in Geneious® (8.1.9) and complete sequences were used as input for MutationFinder3, a python program that was specifically built to identify both nucleotide and amino acid mutations in MBLs NDM-1, VIM-2, and IMP-1, as well as to produce a report outlining the mutations in the desired nomenclature while breaking down the number of transitions and transversions in each variant.

5.4 Results

5.4.1 Establishment of a comprehensive, long-term genetic drift experiment

Four genetic drift trajectories were created from a previously evolved VIM-2 population by separately mutagenizing and applying a purifying selection at four different thresholds: no
selective threshold (T0 trajectory, 0 µg/mL of ampicillin), a sub-inhibitory threshold (T1, 1 µg/mL of ampicillin), a low functional threshold (T10, 10 µg/mL of ampicillin), and a high functional threshold (T1000, 1000 µg/mL of ampicillin). The starting population used for the genetic drift was the 18th round of the previously described adaptive evolution experiment in which VIM-2 was evolved to provide higher resistance against ampicillin in E. coli (Chapter 2). This population was chosen as it provided the largest range between its ampicillin resistance and the host’s background resistance, thereby allowing for the largest range between the chosen high and low functional thresholds. The selection was performed by transforming E. coli cells with the mutagenized libraries, plating the transformants on an agar plate containing ampicillin at the desired selective concentration, and isolating the surviving plasmids for use in the next round of drift. The selective concentrations were chosen based on the ampicillin MIC of directly plated transformants from the starting population (approximately 8,192 µg/mL) and the host E. coli (4 µg/mL). The high functional threshold (1000 µg/mL) is close to the MIC of the starting population, whereas the low functional threshold (10 µg/mL) is slightly above the background resistance of the organism. This 100-fold difference allows for greater contrast between the two functional thresholds and may allow for the easier elucidation of the effects of varying the selective threshold. The sub-inhibitory threshold (1 µg/mL) is slightly below the background resistance of the organism, which allows for the investigation of weak, but non-lethal, selection pressure. The control trajectory was performed to determine mutation rates in the absence of selection and create a baseline from which the effects of the sub-inhibitory trajectory might be compared. The high and low functional threshold trajectories were carried forward for 60 rounds of genetic drift, while the trajectory with no selective threshold was halted at the 10th round and the sub-inhibitory was terminated at the 20th.
A fifth trajectory was created by diverging from the 20\textsuperscript{th} round of the low threshold trajectory and applying gradually increasing selective pressure over 12 rounds (from 21\textsuperscript{st}-32\textsuperscript{nd}) to raise the population’s resistance to the high threshold (1000 \(\mu\text{g/mL}\) of ampicillin). The selection was then reduced to the low threshold (10 \(\mu\text{g/mL}\) of ampicillin), which was again kept constant until the end of the trajectory. This trajectory was created to examine how the oscillation of selection pressure, which causes the mixing of adaptive selection and genetic drift, differs from genetic drift alone. This trajectory was carried forward to the 50\textsuperscript{th} round (Figure 5.1).

**Figure 5.1: The selection threshold applied to each round for the T0, T1, T10, T1000, and TOsc trajectories.** The concentration of ampicillin used for screening the T0 control (orange), T1 sub-inhibitory (yellow), T10 low functional threshold (green), T1000 high functional threshold (blue), and the TOsc oscillating (red) trajectories is denoted by circular data points for each round of the five trajectories. Two dotted black lines denote the ampicillin MICs for the previously evolved VIM-2 starting population and the *E. coli* host organism.
5.4.2 The fitness of each trajectory reflected the conditions of the purifying selection

To determine the effect that the genetic drift had on the overall fitness of each trajectory, the growth rate of every second round was assessed and compared at a large range of ampicillin concentrations in liquid culture. The OD$_{600}$ for each library was measured after 6 hours of growth at ampicillin concentrations ranging from 1 µg/mL to 25,000 µg/mL, and compared to its uninhibited growth to determine a relative ‘percent fitness’. The starting population (the 18$^{th}$ round of adaptive evolution for increased resistance towards ampicillin) was highly resistant, with growth only inhibited >85% at 25,000 µg/mL (Figure 5.2). As expected, the T1000 trajectory maintained a relatively consistent level of fitness, however, the fitness of the T10 trajectory gradually decreased over the first 20 rounds to reach a level of inhibition of >85% at 1,000 µg/mL, before stabilizing and remaining approximately constant for the remainder of the drift. In contrast to the T10 and T1000 trajectories, both the T0 and T1 trajectories exhibited a rapid decrease in fitness within the first 10 rounds of drift until the level of growth matched that of the background (growth inhibited >95% at 10 µg/mL). Interestingly, while the selection in T1 is not high enough to purge any bacteria harbouring nonfunctional MBLs, the low concentration of ampicillin did still appear to have an effect and slow the descent in the fitness of the trajectory. Indeed, whereas it took 8 rounds of genetic drift in the T1 trajectory for its fitness to fall below 20% at 10 µg/mL, it only took 4 rounds in the T0 trajectory.
Figure 5.2: The fitness of the drifted populations in the T1000 high functional threshold, T10 low functional threshold, T1 sub-inhibitory threshold, and T0 control trajectories. Every fourth round in the T1000 and T10 trajectories is visualized as a line with each data point representative of the growth of that population at the given concentration of ampicillin compared to its uninhibited growth in duplicate. Descending colours from dark blue to purple chart the progression from the initial starting population (R0) to round 48 (R48). Every second round in the T1 and T0 trajectories are visualized in the same way, with descending colours from dark blue charting the progression from the initial starting population (R0) to the final round for each trajectory (R10 for T0 and R20 for T1). The percent fitness for the host *E. coli* is denoted on each graph in black, while the fitness for the wild type VIM-2 is shown in red.

The oscillating trajectory, TOsc, quickly recovered its fitness after diverging from the T10 trajectory in round 20 and undergoing 12 rounds of adaptive selection towards improved
ampicillin resistance (Figure 5.3). Upon returning the selective pressure to the low threshold for round 33, the trajectory saw a substantial decrease in fitness that again, began to stabilize as the T10 trajectory had during the first 20 rounds.

**Figure 5.3: The fitness of the drifted populations in the TOsc oscillating trajectory.** Every second round in the TOsc trajectory is visualized as a line with each data point representative of the growth of that population at the given concentration of ampicillin compared to its uninhibited growth in duplicate. The trajectory is displayed in three panels to show the three major periods of the trajectory: the first 20 rounds (as T10) with low selection pressure, the second 12 rounds of natural selection for increased resistance, and the remaining 18 rounds with low selection pressure. The arrows indicate the general direction of the change in fitness over each period, whether fitness decreased (arrow to the left) or increased (arrow to the right). Descending colours from dark blue chart the progression from the starting population for each period of genetic drift or adaptive selection to the final round of that period. The percent fitness for wild type VIM-2 is denoted on each graph in red.
In order to compare the relative fitness between each trajectory, the EC$_{50}$ values were determined from the growth measurements by calculating the concentration of ampicillin required to inhibit growth by 50% (Figure 5.4). The T0 control and T1 sub-inhibitory trajectories rapidly became more sensitive to ampicillin and reached the EC$_{50}$ value of the host *E. coli* (2.5 µg/mL) by their respective 8$^{th}$ and 14$^{th}$ rounds. However, the T1 trajectory did so at a rate approximately 45% slower than the T0 trajectory, further suggesting that the sub-inhibitory concentration had an effect on the fitness of variants in the trajectory. The T1000 trajectory maintained a similar EC$_{50}$ throughout the genetic drift (around 10,000 to 15,000 µg/mL), while the T1 trajectory’s EC$_{50}$ value decreased over the first 20 rounds from 15,000 µg/mL to approximately 300 µg/mL, before stabilizing around this concentration for the remaining rounds. The T$_{Osc}$ trajectory, once diverged from the 20$^{th}$ round of T10, rapidly ascended from 500 µg/mL to 10,000 µg/mL over 8 rounds and maintained this value for an additional 4 rounds, before experiencing a gradual decrease back down to an EC$_{50}$ value of approximately 300 µg/mL when the low purifying threshold was reapplied in round 33. Interestingly, despite the relatively low purifying selection used in the T10 trajectory, the overall fitness appears to stabilize close to the fitness of the wild type VIM-2 gene (EC$_{50}$ value of 330 µg/mL). This suggests that a low functional threshold can supply enough selective pressure to maintain a population of antibiotic resistance genes at a fitness level equating to significant resistance.
Figure 5.4: The fitness of every second round in each of the five trajectories as measured by EC$_{50}$ value. Each data point is representative of the EC$_{50}$ value for ampicillin calculated from the growth measurements taken in duplicate. Each trajectory is represented by a different colour: T1000, blue; T10, green; TOsc, red; T1, yellow; and T0, orange. The upper black dotted line represents the EC$_{50}$ for the VIM-2 wild type MBL, while the background EC$_{50}$ for the host *E. coli* is denoted by the lower black dotted line.
5.4.3 The mutational load steadily increased in each trajectory over the course of the genetic drift

To determine how the mutation rate differs in each trajectory, 5-10 variants were isolated from selected rounds and sequenced (Appendix B, Figure B.5). As most mutations compromise protein stability and are therefore deleterious, it was expected that the accumulation of mutations through genetic drift would be, on average, deleterious, and would reduce the fitness of each trajectory until such a point where further deleterious mutations would drop the fitness of the MBLs below a functional threshold needed to survive. Consequently, it is expected that the lower the selective pressure applied is, the lower this threshold, and the greater the number of mutations that can be acquired.

The number of nucleotide and amino acid mutations per gene was determined in order to compare the rate of mutation acquisition across the five trajectories (Figure 5.5). The T10 and T1000 trajectories proceeded to acquire nucleotide mutations at approximately the same rate (1.3 and 1.1 nucleotide mutations per round respectively) but there is a slightly larger relative difference in the acquisition of amino acid mutations with T10 acquiring 0.6 mutations per round and T1000 acquiring 0.4. The T1 and T0 trajectories acquire mutations at a much higher rate; they acquire 3.0 and 4.2 nucleotide mutations per round and 2.1 and 2.9 amino acid mutations per round respectively. The higher number of nucleotide mutations is expected as redundancy in the genetic code makes non-synonymous mutations less likely to occur. Further, added functional constraints on amino acid mutations (i.e., protein folding and stability, catalytic function) makes the effect of non-synonymous mutations more deleterious on average, thereby limiting the number and the rate at which they can be acquired in comparison to the less constrained synonymous mutations. Again, while the T1 trajectory is exposed to a sub-inhibitory
concentration of ampicillin, the difference in mutation rate between it and the T0 trajectory with no selection is striking and an overall lower mutational load may explain the reduced rate at which its EC$_{50}$ values decrease when compared to the T0 trajectory. Taken together, these data demonstrate that the reduction or removal of purifying selection allows for the accumulation of more mutations.

**Figure 5.5:** The nucleotide and amino acid mutational rates for the five trajectories. Each data point is representative of the average of 5-10 genes that had been sequenced. Each trajectory is represented by a different colour: T1000, blue; T10, green; T0sc, red; T1, yellow; and T0, orange. The black dotted line represents the EC$_{50}$ for VIM-2 wild type.
The nucleotide mutations were further categorized into transition mutations, in which purine nucleotides mutated to other purine nucleotides or pyrimidine nucleotides mutated to other pyrimidine nucleotide, and transversion mutations, in which purines are mutated to pyrimidines and vice versa. These different types of mutations are produced in the mutagenesis by using different base analogues to create mismatches during the replication of the DNA molecules in PCR. 8-oxo-2’-deoxyguanosine-5’-triphosphate mismatches with adenine and is used to introduce transversion mutations, whereas 2’-deoxy-P-nucleoside-5’-triphosphate introduces transitions. The rate of acquisition for each type of mutation was then compared for each trajectory over the course of the genetic drift (Figure 5.6).

**Figure 5.6: The ratio of transitions to transversions for the five trajectories.** Each data point is representative of the average of 5-10 genes that had been sequenced. Each trajectory is represented by a different colour: T1000, blue; T10, green; TOsc, red; T1, yellow; and T0, orange. The black dotted line represents the EC$_{50}$ for VIM-2 wild type.
The T0 control trajectory exhibited acquired an average of 2.4 transitions and 2.0 transversions per round of drift (or 1.2 transitions per transversion). This suggests that the mutagenesis of the libraries is creating transition and transversion mutations at a relatively balanced rate. By contrast, the T1, T10, TOsc, and T1000 trajectories appear to all have a ratio of approximately 2 transitions per transversion. This suggests that selection, regardless of strength, favours transition mutations over transversions. As most non-synonymous mutations are deleterious, the fact that transition mutations are more likely to result in a synonymous codon change, and therefore avoid the majority deleterious effects, may provide for greater acquisition with a neutral result. Further increasing the number of samples sequenced and continuing the T0 control trajectory will allow for verification of this apparent trend.

5.4.4 The distribution of mutations reveals the adaptability of VIM-2

The overall difference in mutational patterns in terms of mutation rate in the three functional trajectories is subtle. To further characterize the difference in these trajectories, eight variants from the 40th round of each of the T10, TOsc, and T1000 trajectories were isolated and sequenced to determine the location of non-synonymous mutations accumulated through drift and selection. The positions that were mutated in those sequences were mapped on the primary sequence and secondary structure of VIM-2 (Figure 5.7). Of the 260 positions that were exposed to mutagenesis, 129 were mutated in at least one of the variants surveyed from the T10 trajectory (50%), 104 in the TOsc trajectory (40%), and 97 in the T1000 trajectory (37%) (Table 5.1). This is again expected as most mutations are deleterious and as the T10 trajectory has been exposed the longest to the lowest selection threshold, it could accumulate more deleterious mutations before its function falls below this threshold whereas the T1000 trajectory has a more stringent
threshold and many positions that could be mutated in the T10 trajectory would prove too deleterious to acquire. The TOsc trajectory was found to have slightly more positions mutated than the T1000 trajectory, which is consistent with having recently undergone an adaptive selection period (rounds 21 to 32) that likely purged many deleterious mutations.
Figure 5.7: The location of non-synonymous mutations acquired by the 40th round of T10, T1000, and TOsc mapped on VIM-2’s primary sequence and compared to secondary structure. Teal boxes represent a residue that was mutated in one of the sequenced variants of
round 40 from each trajectory while orange boxes represent residues that already had mutations that were fixed or in abundance in the adaptive evolution experiment (mutated in >50% of isolated variants that were sequenced from the 18th round of adaptive evolution). The grey boxes represent regions of the gene not exposed to mutagenesis, while the empty white boxes indicate that no mutations were observed at that location. Black highlighted residue numbers indicate the six metal binding residues (first zinc binding site: H116, H118, and H196; second zinc binding site: D120, C22, and H263). The secondary structure of VIM-2 is aligned above the primary sequence. The arrows designate β-strands, the cylinders denote α-helices, and the line represents loops, turns, and unstructured regions from the PDB ID 1KO3.
Table 5.1: Mutation location statistics produced from analysis of eight sequence variants from the 40th round of the T10, T0sc, and T1000 trajectories.

<table>
<thead>
<tr>
<th></th>
<th>T10</th>
<th>T0sc</th>
<th>T1000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positions</td>
<td>Total</td>
<td>Positions</td>
</tr>
<tr>
<td></td>
<td>Mutated</td>
<td>Residues</td>
<td>Mutated</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole Gene</td>
<td>129</td>
<td>260</td>
<td>50</td>
</tr>
<tr>
<td>Signal Peptide</td>
<td>33</td>
<td>41</td>
<td>80</td>
</tr>
<tr>
<td>Catalytic Domain</td>
<td>96</td>
<td>219</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface Accessible</td>
<td>82</td>
<td>156</td>
<td>53</td>
</tr>
<tr>
<td>Buried</td>
<td>14</td>
<td>63</td>
<td>22</td>
</tr>
<tr>
<td>Secondary Structure Elements</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-Helices</td>
<td>37</td>
<td>73</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>Surface Accessible</td>
<td>4</td>
<td>18</td>
<td>22</td>
</tr>
<tr>
<td>Buried</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-Strands</td>
<td>22</td>
<td>64</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>33</td>
<td>45</td>
</tr>
<tr>
<td>Surface Accessible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buried</td>
<td>7</td>
<td>31</td>
<td>23</td>
</tr>
<tr>
<td>Unstructured (Loops)</td>
<td>35</td>
<td>82</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>68</td>
<td>47</td>
</tr>
<tr>
<td>Buried</td>
<td>3</td>
<td>14</td>
<td>21</td>
</tr>
<tr>
<td>Active Site Residues</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residues 5Å from Zinc Ions</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Residues 6Å from Zinc Ions</td>
<td>2</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>Residues 8Å from Zinc Ions</td>
<td>7</td>
<td>28</td>
<td>25</td>
</tr>
</tbody>
</table>
A much higher percentage of the signal peptide was capable of acquiring mutations when compared to the catalytic domain across all three trajectories. For this analysis, I define the signal peptide as the unstructured region at the beginning of MBL from start codon to the first secondary structure element at Arg44. T10’s signal peptide showed the greatest diversity of all three trajectories with 80% of positions mutated, followed by TOsc with 68%, and T1000 with 59%. This is in stark contrast to the catalytic domain, where 44%, 35%, and 33% of positions were mutated respectively. This almost 2-fold higher amenability to mutation may be explained by the reduced structural constraints placed on the signal peptide region and is also reflected in the much greater diversity seen in signal peptides in nature. Between the closely related VIM-1 and VIM-2, 44% of the amino acid mutations occur within the signal peptide, despite it only accounting for ~10% of the entire protein. Within the MBL family, 44% of the mutations between the closely related VIM-1 and VIM-2 appear within the signal peptide. Within the entire B1-MBL family, no residue is fully conserved within the signal peptide.

When analyzing the catalytic domain, surface accessible positions were mutated at a much higher rate than those positions found buried in the structure (53% vs. 22% in T10, 43% vs. 14% in TOsc, and 42% vs. 13% in T1000). Again, this makes sense in light of the reduced structural constraints faced by surface accessible residues. However, while less restricted by interactions within the enzyme, these positions are still subject to constraints within their secondary structure element and therefore are not as freely mutable as the signal peptide.

Breaking down the analysis to secondary structure elements within the catalytic domain of the enzyme, it is clear that α-helices were the most mutable secondary structure element the percent of α-helical positions mutated is higher than both β-strand and loop regions in all three trajectories. For example, in T10, 51% of α-helical positions were mutated compared to 34% and
43% of β-strand and loop regions respectively. Interestingly, while the loop regions in the T10 trajectory were the second highest mutated element, this is not the case with TOsc and T1000, where the percent of positions mutated for each element in TOsc was 40%, 34%, and 19% for α-helices, β-strands, and loops respectively, and 45%, 28%, and 16% respectively for T1000. This suggests that the loops likely play a critical role in the maintenance of a high functioning MBL. This may reflect the structure of the MBL active site, which is composed primarily of loops and would be integral for catalysis and substrate specificity. Again, when comparing each secondary structure element by surface accessibility, it is evident that surface residues are far more amenable to mutation than residues buried within the MBL. For example, in the T10 trajectory, the ratio of mutated surface accessible positions to mutated buried residues was 2.7, 2.0, and 2.2 for α-helices, β-strands, and loops respectively.

Overall, the distribution of mutations in the genetic drift trajectories reveals the extent to which the MBL sequence and structure can accommodate mutations, and demonstrates how long term genetic drift may lead to the level of genetic variation currently observed in the B1-MBL family.

The mutated positions for each trajectory were mapped onto the crystal structure model of VIM-2 (Figure 5.8). This highlighted that of the 59% of positions that were mutated in at least one of the three trajectories, the majority are on the surface of the protein, particularly in the outwards-facing sides of the α-helices. While residues in the interior of the MBLs were less likely to be mutated, they were still much more likely than the core active site architecture. Indeed, of all seven residues within 5Å of the two zinc ions, including the six metal-binding residues, none were ever found to mutate. Expanding outwards to the 11 residues within 6Å, only 2, 2, and 1 positions were mutated in the T10, TOsc, and T1000 trajectories respectively.
Indeed, only expanding further to consider the 28 residues within 8Å of the zinc ions, saw the percent of positions mutated rise to comparable levels as other segments of the enzyme. This reflects the natural diversity present within the B1-MBL family. Of the 16 positions fully conserved between a small sampling of MBLs, 11 of the 16 are within 8Å of the active-site zinc ions. Further, 13 of the 16 are conserved within the three trajectories, with only two positions found to be mutable in the T10 trajectory and one position in the TOsc trajectory.
Figure 5.8: The location of non-synonymous mutations acquired by the 40th round of T10, T1000, and TOsc mapped onto a 3D model of VIM-2’s crystal structure separately and combined. Each position found to be mutated in at least one gene of the 8 variants sequenced in each trajectory was mapped onto the structure of VIM-2 (PDB ID 1KO3). The purple spheres show the location of all positions mutated in all three trajectories, while the green highlights indicate positions in T10, the blue indicate positions in T1000, and the red indicates positions in TOsc.
As for the question to whether or not the TOsc trajectory would remain unique or if after it was adapted to a high threshold, eight unique mutations were located that were not present in either the T10 or T1000 trajectory (Figure 5.9). These mutations may be the product of epistasis, where normally they may be too deleterious to survive selection but other mutations have changed the genetic context and the interaction between each other now allows for its fixation. Alternatively, their uniqueness may be a result of small sample size and further sampling of the T10 and T1000 libraries will reveal that these mutations are present there as well.

**Figure 5.9: The location of unique non-synonymous mutations acquired by the 40th round of the TOsc trajectory, not present in the T10 and T1000 trajectories, mapped onto a 3D model of VIM-2’s crystal structure.** Each position found to be mutated in at least one gene of the 8 variants sequenced from the TOsc trajectory that was not present in any of the variants from the T10 and T1000 trajectories were mapped onto the structure of VIM-2 as black spheres (PDB ID 1KO3).
5.4.5 Broad β-lactamase activity is maintained by selection with a single antibiotic

To determine the effect of the accumulated genetic variation on the phenotypic properties of the MBLs, I sought to characterize changes in the substrate specificity for the T10 and T1000 trajectories. The cefotaxime, meropenem, and ampicillin MICs were determined for 92 variants from the 60th round of each trajectory by measuring the growth of every variant on agar plates supplemented with each increasing concentration of each antibiotic (Figure 5.10). The behaviour within each trajectory was similar: Variants in the T1000 trajectory maintained a relatively high level of resistance that did not vary greatly for all three antibiotics, with only a few outliers, while the variants in the T10 trajectory had a lower, but more varied distribution in their ability to provide resistance. Specifically, for the T1000 trajectory, the ranges were: 512 to 4096-fold higher than the background for ampicillin; 512 to 8192-fold higher, with one outlier at 64-fold for cefotaxime; and 64 to 512-fold higher, with one outlier at 16-fold for meropenem. For the T10 trajectory, the ranges were: 1 to 512-fold, 1 to 256-fold, and 2 to 64-fold higher than the background MICs for ampicillin, cefotaxime, and meropenem respectively. This supports our previous conclusions that the different functional thresholds result in phenotypic variation consistent with the level of selective pressure.
Figure 5.10: The MICs of cefotaxime, meropenem and ampicillin for variants from the 60th round of the T10 and T1000 genetic drift trajectories. (A) The MIC of each antibiotic over the MIC of the host organism. Each data point represents a single variant screened from the two trajectories: green represents a T10 variant, blue represents a T1000 variant, and black represents the wild type VIM-2. (B) The fold-change in the MIC for each variant in comparison to the original starting population.
Interestingly, despite the 100-fold difference between the functional thresholds for each trajectory, the variants with the highest resistance in the T10 trajectory still overlapped with those with the lowest resistance in the T1000 trajectory for all three antibiotics. Of greater interest however, is that while the wild type VIM-2’s cefotaxime MIC was equal to the highest T10 variants, approximately half of the T10 variants surveyed provided higher ampicillin and meropenem resistance. This not only supports the hypothesis that low levels of selection can sustain high levels of antibiotic resistance over long evolutionary time, but that this selection may also preserve effective resistance to other antibiotics not used in selection.

5.5 Discussion

The long-term genetic drift of VIM-2 populations at various selective thresholds revealed the extent to which MBLs can accommodate mutations without the loss of function at different selection thresholds. The incredible flexibility of the MBL structure was particularly highlighted by the substantial accumulation of mutations by the 40th round of the T10, TOsc, and T1000 trajectories, which had, on average, 17.5%, 15.9%, and 14.7% of their respective amino acid positions mutated, with 59% of all positions mutated at least once in one of the sequences. Despite this heavy mutational load, the relative fitness of the trajectories subjected to a functional selective threshold remained high, with even the lowest functional threshold trajectory, T10, providing a similar level of ampicillin resistance as VIM-2 wild type.

This stands in contrast to the T0 control and T1 sub-inhibitory trajectories, where fitness rapidly decreased. While the T1 trajectory descended slower than the T0 trajectory, suggesting that the sub-inhibitory concentration of ampicillin was affecting the selection of the population,
both trajectories decreased to the same level of fitness as the background by the 8th and 14th rounds respectively, and contained many non-functional MBLs (e.g., stop codons inserted early in the sequence or metal-binding residues mutated to non-metal-binding residues). Without the overriding factor of survival, the burden that the expression of these genes places upon the host may be the dominant factor in their evolution (Kafri et al. 2016). Regardless, it is clear that the utility of an MBL to the host organism is questionable in the absence of functional selection pressure, and that the MBL genes rapidly degrade to pseudogenes during genetic drift. This raises an important question about the continual rise of antibiotic resistance genes from environmental reservoirs: If these genes quickly become pseudogenes in the absence of selective pressure, then how can so many functional resistance genes emerge so quickly and are readily found in the environment?

One possible explanation is that the pseudogenes are frequently ‘re-functionalized’ (Vazquez et al. 2017). While re-functionalization of pseudogenes has most certainly occurred in nature at some point, it is extremely unlikely to be the driving force between the emergence of resistance genes to the extent and scale observed. However, if the assumption that there is no selective pressure applied in nature is wrong, then three more possible explanations arise: First, a growing number of genes have been found to ‘moonlight’ and maintain a secondary physiological function (Jeffery 2014; Copley 2014). If the MBLs performed another cellular role in nature not recapitulated in this experiment, the loss of selection on both the primary and secondary function might explain the rapid loss here in contrast to their continued maintenance in nature. However, this again is unlikely because no biological secondary function has ever been identified for the MBLs and they undergo extensive horizontal gene transfer to diverse bacteria wherein a secondary function is less likely to be maintained. Second, low concentrations of
antibiotics are often found in the environment around human settlements and these may maintain the selective pressure at a level required for maintenance (Xi et al. 2009). Indeed, human contamination of the environment with antibiotics is well known: as of 2018, even UBC’s Risk Management Services had both ampicillin and cefotaxime on their list of chemicals that are safe for disposal down the drain. As a result of actions like these, resistant strains have been found in the sewage water of a hospital in Spain, polluted rivers in Nigeria, and wastewater treatment plants in China in recent years (Scotta et al. 2011; Luo et al. 2013). While a logical explanation at the local level, it fails to explain the isolation of fully functional MBLs from remote areas far from dense human settlements where the chance of synthesized antibiotics being found at higher than trace levels is extremely low (e.g., the North Sea) (Jiang et al. 2018; Gudeta et al. 2016; Delgado-Gardea et al. 2016). The third and most likely possibility is that the continual use of natural antibiotics in bacterial warfare has cultivated and continues to maintain a large reservoir of antibiotic resistance determinants in the environment. Indeed, most antibiotic classes, including the β-lactams, were originally isolated from microorganisms and recent reports have emphasized the important role that antibiotic can play in microbial attack and defense (Abrudan et al. 2015). This has troubling implications for the clinical resistance against antibiotics: any new antibiotic developed based upon a natural compound is likely to prompt a swift response from their bacterial targets, even if a resistance mechanism has yet to be characterized.

Of greater concern however, is the finding that broad β-lactamase resistance is maintained despite only a low level of selection with a single antibiotic. After 60 rounds of genetic drift with ampicillin selection, the majority of the variants in the T10 low threshold trajectory maintained a relatively high level of resistance towards both cefotaxime and
meropenem, comparable to that of VIM-2 wild type. It stands to reason that many of these variants would still be able to provide effective resistance in a clinical setting, and if not, may be able to swiftly adapt to do so. This finding undercuts a new treatment strategy that has gained popularity in recent years: the revival of old antibiotics. With growing resistance to modern drugs, clinicians are looking to older antibiotics, such as the β-lactams temocillin and pivmecillinam, that fell out of use in the early antibiotic era or have been passed upon in favour of other candidates that were more promising at the time (Zayyad et al. 2017; Theuretzbacher et al. 2015; Giske 2015). The ‘long memory’ of the MBLs may cast doubt as to the overall efficacy of this strategy, as effective resistance against unselected antibiotics may persist for long periods of evolutionary time.

Overall, it is clear that the substantial genetic variation present with the B1-MBL family could very well be a product of genetic drift alone and that their broad β-lactamase activity is generally maintained even with low levels of selection with a single antibiotic over long evolutionary time periods. However, a more likely scenario is that the genetic diversity present within the B1 MBL family is a product of both genetic drift and adaptive evolution. Indeed, MBLs have likely been subjected to long periods of genetic drift while under purifying selection for some basal level of activity during natural bacterial warfare, followed by short periods of intense selection for activity. These adaptive periods could take the form of a horizontal gene transfer event into a new host (as detailed in Chapter 3) or exposure to strong, novel antibiotics in the clinic. This would be consistent with the “rapid bursts and slow declines” enzyme evolution model recently described (Newton, Arcus, et al. 2015). The extent to which these ‘rapid bursts’ and ‘slow declines’ have occurred will likely remain unknown, but the interplay between the two and their effect on genetic diversity will continue to be investigated as the
genetic drift experiment proceeds. Beyond this, the genetic drift trajectories may serve as a platform to investigate four deeper evolutionary questions detailed below.

First, if the genetic variation introduced through genetic drift does not manifest phenotypically in terms of catalytic activity, could it emerge in different ways? I have previously shown that phenotypic variation in the B1-MBL family is pervasive, and largely manifests through variation in the level of their expression in different host organisms (Chapter 2). I also found that adaptive mutations rapidly accumulate, altering the MBL’s host specificity, and potentially lead to ‘domestication’ within their host organism (Chapter 3). Mutations that are considered relatively neutral and accumulate in the genetic drift trajectories may not be neutral in other organisms, and prove deleterious. As such, genetic drift may also be contributing to the domestication of the MBLs in the host, even without strong selection pressure. However, the extent of domestication and the mechanisms by which it occurs are unknown. Variants from selected rounds in each of the trajectories can be tested for fitness and changes in expression in *P. aeruginosa* and *K. pneumoniae* to determine answers for this question.

Second, how does epistasis promote or constrain the genetic and phenotypic diversity present in the B1-MBL family over long periods of genetic drift? A large part of determining how evolutionary forces shaped the family also include understanding the role that epistatic interactions may have played in promoting this diversity. Epistasis is a powerful influence in evolution, as the interactions between residues in a protein not only define the current structure and function of an enzyme, but also constrain and modulate its evolutionary potential (Breen et al. 2012; Starr & Thornton 2016). Of particular interest is the most dramatic type of epistasis, sign epistasis, in which deleterious mutations become beneficial (or vice versa). How can we determine to what extent epistasis has influenced the MBLs? The genetic drift experiment
provides an excellent platform to explore protein epistasis. Deep mutational scanning of final round variants from each of the trajectories could be performed by creating comprehensive codon mutagenesis libraries, functionally screening, and sequencing the screened libraries to determine what mutations are enriched in certain conditions (Araya & Fowler 2011; Fowler & Fields 2014). These variants could be compared to the wild type VIM-2 and genetic drift starting variants. Specifically, comparison of the enrichment values at each position should indicate what mutations in the final round of genetic drift are a result of sign epistasis, as those that are should be deleterious in VIM-2 and the starting population (and therefore not enriched). While this would provide a high-level view as to the extent of sign epistasis in genetic drift, such mutations may also be further characterized to determine how they occurred.

Third, will the growing genetic diversity in the genetic drift trajectories increase the evolutionary potential of the population by opening up new evolutionary pathways for the MBLs to provide higher resistance against β-lactam antibiotics? Previous genetic drift experiments have shown that the probability of new functions emerging in the population increases as the added genetic diversity results in the accumulation of ‘global suppressor’ mutations that stabilize the protein and allow for new function-changing mutations that would otherwise prove too deleterious to accommodate (Bershtein et al. 2008; Bershtein & Tawfik 2008). Further, ancestral reconstruction of ancient steroid receptors has shown how these ‘permissive’ epistatic mutations, which have no immediate effect individually, have resulted in major functional changes in the glucocorticoid receptor’s evolution (Ortlund et al. 2007; Harms & Thornton 2013). If such mutations can have such a large effect on the transition of one function to another, they may also permit new evolutionary pathways leading to higher resistance that were previously inaccessible. Taking selected rounds from the genetic drift trajectories and using them as the template for a
new evolutionary adaptation experiment could test this. Comparing the rate at which resistance increases as well as the level of resistance at which the trajectories plateau with the original VIM-2 wild type gene could show the extent of this effect (Chapter 3). This could have major implications for the development of clinical resistance, as low selection pressure over long period of time could promote the diversity that would allow for stronger clinical resistance in the future.

And fourth, a central question in evolutionary biochemistry is how has the sequence and functional diversity of enzyme superfamilies expanded from a common ancestral protein to what we observe today? The prevailing theory is that multifunctional or promiscuous enzymes underwent gene duplication, selection, and divergence towards difference functions (Jensen 1976; Jacob 1977; Baier et al. 2016). As the genes diverged, new promiscuous functions could emerge and act as the basis for further selection and divergence. Several studies have characterized the functional diversity present within superfamilies (Baier & Tokuriki 2014; Bastard et al. 2014; Huang et al. 2015), and more have selected for and evolved promiscuous functions (Gould & Tawfik 2005; Khersonsky et al. 2006; Tokuriki et al. 2012), however, none have actually replicated the process by which new functions emerge. While this work largely speaks to the genetic diversity within the structurally and functionally homogenous B1-MBL family, this genetic drift experiment can also be used to study greater sequence and functional diversity at the superfamily level. The genetic drift trajectories provide an excellent opportunity to determine how long-term protein evolution affects the emergence of new functions and how it may alter already established promiscuous functions. Previously, the functional diversity in the MBL superfamily was assessed by characterizing the level of activity of 24 MBL superfamily members for 10 catalytically distinct reactions (Baier & Tokuriki 2014). It was discovered that
VIM-2 has weak phosphodiesterase, phosphotriesterase, and phosphonatase promiscuous activities. The genetic drift trajectories could be probed for these and other common MBL activities and compared to wild-type VIM-2 to determine how robust promiscuous activities are and how easily new activities can emerge through genetic drift.

In the future, the long-term genetic drift trajectories will serve as a platform to investigate how genetic drift may alter the phenotypic diversity of the MBLs, how epistasis can promote or constrain diversity during genetic drift, how genetic drift may lead to greater antibiotic resistance, and genetic drift may promote the emergence of new enzymatic activities and lead to the divergence of superfamilies, in addition to continuing to explore how genetic drift promotes genetic and phenotypic diversity within the B1-MBL family. This will lead to a better understanding of how genotype affects phenotype and reveal long term trends in the development of antibiotic resistance that have yet to be systematically investigated.
Chapter 6: Conclusion and Future Outlook

6.1 General summary and conclusion

The primary goal of this thesis was to employ directed evolution to investigate the B1 MBLs, leading to a better understanding of their origin, current structure and function, and evolutionary potential. Specifically, I sought to identify how genetic variation in the B1-MBL family manifests phenotypically, and to replicate the evolutionary processes that produced that variation in order to identify and predict future evolutionary trends. In addition to the clinically relevant conclusions that can come from a better understanding of the nature of β-lactam antibiotic resistance, this work also addresses a fundamental question concerning molecular evolution and sequence diversity within protein families and the divergence of orthologs: How can proteins genetically diverge large distances while maintaining the same function?

To achieve this, I conducted a detailed biochemical survey of eight B1-MBL family members, performed directed evolution on three of these family members to increase their resistance against three different β-lactam antibiotics, and initiated a long-term genetic drift experiment from one of these evolved trajectories. These studies have successfully replicated the evolutionary processes that produced genetic variation in the B1-MBL family, while providing a wealth of biochemical information on the structure and function of the MBLs.

Specifically, in Chapter 2, I discovered how genetic variation in the B1-MBL family can result in phenotypic variation that only emerges upon expression in different host organisms. Further, our work suggests that the host-specific expression processes can potentially constrain the dissemination of the MBLs by limiting their effective deployment in certain organisms after horizontal gene transfer. In Chapter 3, I investigated how these host-specific constraints could be
overcome through adaptive evolution. Both Chapters 2 and 3 revealed the important role that the signal peptide plays in the dissemination and evolution of the MBL family.

In Chapter 4, I examined the broad substrate specificity featured by the MBLs and determined that while it is relatively robust, the MBLs’ broad specificity can sometimes be narrowed. However, the exchangeability of the active site residues allows for this to be readily reversed upon changing selection pressure, suggesting that broad substrate specificity is an ingrained trait that is hard to modify. Both Chapters 3 and 4 also demonstrated that MBLs cannot readily increase their ability to catalyze the hydrolysis of the MBLs, suggesting that these enzymes may occupy a fitness peak that is inescapable while under selection for β-lactamase activity. In Chapter 5, I determined how genetic drift could contribute to the striking genetic diversity within the B1-MBL family, while discovering that the MBLs’ broad substrate specificity could be maintained at a clinically relevant level over long periods of evolutionary time with only a very low level of selection with a single antibiotic.

In effect, Chapters 2 and 3 are concerned with identifying the molecular barriers that limit MBL expression and determining the adaptive mechanisms by which they may be overcome, while Chapters 4 and 5 convey a detailed investigation into the nature of the MBLs’ broad substrate specificity. Together, these four chapters present a detailed evolutionary perspective on the structure and function of the MBLs, and provide insight into their evolutionary potential.
6.2 Future outlook

Despite our growing knowledge of MBL-mediated antibiotic resistance, there are still many unanswered questions. Here, I detail multiple paths forward for the further investigation of the B1-MBL family.

6.2.1 Investigating phenotypic variation in regulatory elements

In Chapter 2, I performed a survey of eight B1-MBL family members and I determined that hidden phenotypic variation can constrain the effective dissemination of MBLs to different host organisms. However, I was unable to recapitulate all of the trends observed in nature. For example, NDM-1 is largely confined to clinical variants of *K. pneumoniae* and *E. coli*, whereas we found it still provided effective resistance in *P. aeruginosa*. Further, I was also unable to identify any difference between acquired and chromosomal MBLs. Indeed, IND-1, a non-mobile, chromosomal MBL, provided the highest level of resistance of all eight MBLs surveyed in all three organisms. Beyond the hidden phenotypic variation of each of the enzymes, what factors influence their dissemination and mobility patterns? Our experiment only analyzed the protein-coding region of each gene, while neglecting the native regulatory elements that constitute the entirety of the natural, full gene. This includes elements like the ribosome-binding site, promoter, and potentially, the origin for plasmid-borne variants. It is likely that the same type of host-specific phenotypic variation observed in the protein-coding regions is also pervasive in these regulatory elements, and that variation in these elements may even have a greater effect on expression than that of the coding region. To address these questions, elements from the full native genes should be assembled and tested both individually and in combination in different host organisms to determine the extent of their phenotypic variation and their effects on host
specificity. In line with these experiments, the methods in this study could be expanded to include other antibiotic resistance genes (beyond the B1-MBLs) and also include more organisms. Whereas, I have studied eight representative MBL enzymes, much wider-scale studies of which resistance genes are compatible with which organisms would allow for the identification of more nuanced trends. Additionally, an alternate line of inquiry includes investigating the effect of population-level factors such as population structure, community composition, environment, geography, and migration that may also influence dissemination. Taken together, these actions may lead to a better understanding of MBL dissemination patterns and allow for the prediction of their future spread. This in turn could facilitate the implementation of targeted treatment strategies that could be more effective at preventing their continued dissemination.

6.2.2 Examining the key features of evolutionary optimization processes

In Chapter 3, I performed 18 rounds of directed evolution on NDM-1, VIM-2, and IMP-1, and subsequently determined that the ability of the MBLs to confer ampicillin resistance to their host organism could be swiftly improved through adaptive evolution, largely by overcoming barriers that limit their functional periplasmic expression (i.e., cytoplasmic expression, translocation). While most of the mutations driving these changes appear to be in the signal peptide (as replacement of the signal peptide with a standardized PelB leader sequence negated most of the improvement), I did not definitively identify the beneficial effects that these mutations had on the level of resistance. I am currently determining the overall effect of these signal peptide mutations by creating hybrid constructs (i.e., wild type signal peptide with evolved catalytic domain and evolved signal peptide with wild type) and measuring their change in fitness.
However, these constructs have been difficult to create and I had hypothesized that they may be toxic to their host. If correct, the incompatibility between the catalytic domain and signal peptide region would suggest that there are epistatic interactions between the two segments: mutations in the evolved catalytic domain may compensate those in the signal peptide that would prove deleterious independently, and vice versa. In the absence of the compensatory mutations, the subsequent protein may decrease the fitness of the cell (*i.e.*, by aggregating or interfering in other cellular processes). The extent of this phenomenon should be investigated, as it may reveal yet another constraint placed on the structure and function of the MBLs, and thereby affect their evolutionary trajectory.

As previously discussed in Chapter 3, evolutionary optimization processes are subjected to the trade-offs, in which the gain of one function comes at the expense of another, and diminishing returns, in which the largest improvements occur first. I saw both features to some degree within our trajectories: I observed host specificity appear to weakly trade-off, with large improvements lessened or negated in other organisms, and I noted the diminishing improvement in resistance throughout each trajectory, which eventually plateaued entirely. However, this directed evolution experiment provides an excellent opportunity to delve further into both optimization features than anyone has previously reported. To further investigate trade-offs in host specificity, the directed evolution should be repeated in another organism, such as *P. aeruginosa*. Comparison between the complete *E. coli* trajectories with *P. aeruginosa* trajectories may reveal similar and contrasting trends in the adaptation of the MBLs. Further, these trajectories should be assessed in multiple bacteria that are evolutionarily further apart to better determine if there are significant trade offs in host specificity. This would better inform our ability to perform heterologous expression for biotechnological applications as well as further
support the concept that phenotypic variation can have a significant effect on the dissemination of antibiotic resistance determinants.

In terms of diminishing returns, we know that the largest improvements in resistance occur first, but we do not know specifically from where these improvements came and how that may differ between the MBLs. For example, it appears that the cytoplasmic expression of IMP-1 increased in the earlier rounds before the ratio of periplasmic to cytoplasmic expression increased, suggesting that expression was prioritized over translocation efficiency. However, VIM-2 maintained the same ratio of periplasmic to cytoplasmic expression throughout the entire trajectory, suggesting that expression was the greatest barrier to increased resistance. The individual mutations in each of the complete trajectories should be investigated individually and in the context in which they appeared to determine the order of priority for the optimization of resistance. This would reveal the relative importance of each of the barriers to expression that MBLs may face and help us to understand the challenges faced by heterologous proteins and identify mechanisms that could be employed to increase the success of recombinant protein expression.

Lastly, once again, I focused solely on the coding region of each MBL for this adaptation experiment. This neglects regulatory elements such as the promoter and ribosome-binding site. As I saw many mutations occurring in the signal peptide, which resulted in increased functional periplasmic expression, it stands to reason that including the promoter and ribosome-binding site will have a significant effect on the trajectory as their mutability may provide an easier pathway for increasing expression. This experiment would allow us to answer many questions: What dynamics would be observed between the different elements? Would the signal peptides be as heavily mutated if some of the mutational burden could be off-loaded onto the promoter? Would
the increase in resistance be more rapid than before or would the trajectory reach a higher resistance plateau? In contrast to the reductionist approach, it is often important to perform experiments with a holistic and inclusive view of the system being worked on to fully recapitulate what is happening in nature. Indeed, most directed evolution experiments with β-lactamases have even neglected the signal peptide and missed out on a wealth of clinically relevant data.

6.2.3 Thoroughly investigating the broad substrate specificity of the MBLs under different selection conditions

In Chapter 4, I performed 18 rounds of directed evolution on NDM-1, VIM-2, and IMP-1 using cefotaxime and meropenem for selection, and determined that the MBLs’ broad substrate specificity is not easily lost while under strong selection from any β-lactam antibiotic, with only two of seven completed trajectories exhibiting a substrate specificity shift (the IMP-1 cefotaxime and VIM-2 meropenem trajectories). Further, in these two trajectories, it appears that the substrate specificity shift (1), resulted from loss of catalytic activity with certain antibiotics rather than a gain for those under selection, (2), was largely driven by active site mutations, and (3), could be easily undone upon reversal of the selection conditions. However, our data is incomplete: the kinetic parameters for variants from round 18 of the IMP-1 cefotaxime trajectory must be acquired to confirm that our findings apply to both trajectories. Further, acquiring the kinetic parameters for those variants who have had their broad substrate specificity restored (round 21) will allow for elucidation of the mechanism by which this occurred.

Further, the effects of identified mutations should be validated experimentally and compared to the predictions made with the deep mutational scanning (DMS) data. These
mutations could be tested individually and in combination to not only test the accuracy of the DMS data, but to identify to what extent epistasis may influence the substrate specificity of the MBLs. To more confidently determine the robustness of the MBLs’ broad substrate specificity, a thorough review of the entire DMS dataset will reveal other residues that alter VIM-2’s substrate specificity. This may be compared to data from the DMS of other MBLs to be able to identify family-wide trends. We have created CCM libraries of NDM-1 and VIM-2, in which every codon position has been mutated to all 64 possible codons (including the wild type). Screening these libraries with different antibiotics at different selection thresholds and sequencing the surviving populations will allow us to determine relative enrichment values for each mutation in NDM-1, which can be compared across conditions. This will quickly identify all single codon mutations that result in a substrate specificity change, which can be followed up with more detailed characterization (i.e., resistance, kinetics), and will also allow for a comparison between NDM-1 and VIM-2 to identify common and contrasting trends.

6.2.4 Utilizing experimental genetic drift as a platform for long-term antibiotic resistance evolution

In Chapter 5, I determined that long-term genetic drift with a low level of selection could maintain an MBL population with a clinically-effective level of antibiotic resistance despite a substantial increase in genetic variation. However, it is unlikely that the level of selection experienced in nature would be static. To that end, we will complete the oscillation trajectory by performing the same number of rounds of mutagenesis and selection as the T10 low and T1000 high threshold trajectories. We will also attempt to rescue the T1 sub-inhibitory trajectory by adapting it for increased resistance
These trajectories will then be compared to the T10 and T1000 trajectories to determine what effect an oscillating level of selection has on the long-term evolution of the MBLs. Further, this trajectory will be taken into consideration with the future directions previously described in detail in Chapter 5: the genetic drift trajectories will be used as a platform to (1), quantify the effect that genetic drift has on phenotypic variation, (2), determine the extent of epistasis that emerges during genetic drift with deep mutational scanning, (3), investigate if genetic drift opens up new evolutionary pathways to the provision of higher resistance, and (4), determine if experimental genetic drift can lead to the emergence of novel catalytic activities.

6.2.5 Creating and screening indel libraries

Experimental evolution is often focused upon investigating the effects of single point mutations. This neglects an entire category of mutations: insertions and deletions (indels). These mutations may have significant effects on the structure, function, and evolution of proteins. Indeed, there are seven indel sites in the catalytic domains of the eight MBLs surveyed in Chapter 2 (Figure 1.4). To determine what effects indels might have on the current function and future evolution of the MBLs, one, two, and three-codon insertion and deletion libraries will be created using a transposon-based method to add or remove nucleotides to each MBL sequence. These indel mutants will be screened and sequenced to determine the relative fitness of each mutation for several β-lactam antibiotics (i.e., cefotaxime, meropenem, ampicillin, and aztreonam). In addition to demonstrating the overall receptiveness of the MBLs to accommodate these small indels and helping us understand the robustness of the MBL fold, these results will help to quickly identify any variants with large changes in phenotype. It is likely that indels allow for larger jumps in sequence space than single point mutations due to the fact that single point
mutations are limited in effects to changing a single codon, whereas indels provide a much larger disruption. This means that indels may result in more radical phenotypic jumps, such as converting a B1-MBL to a B2-MBL (one metal ion in the active site, and a preference for carbapenems) or allowing for the hydrolysis of aztreonam, a previously unreachable β-lactam. Any variant with a radically altered phenotype will be further characterized (i.e., measuring resistance, kinetics, thermostability, structures) to determine the mechanism by which the function was changed.

6.3 Conclusion

This thesis employed directed evolution and related techniques to investigate the B1-MBL family, resulting in a better understanding of their origin, current structure and function, and potential future evolution. The knowledge gained from this work will aid in combatting the rapid emergence of multi-drug resistant bacteria.
Bibliography


Bershtein, S. & Tawfik, D.S., 2008. Ohno's model revisited: measuring the frequency of potentially adaptive mutations under various mutational drifts. Molecular Biology and


Cameron, H.C., 1907. Lord Lister and the evolution of wound treatment during the last forty years: Being the James Watson Lectures delivered at the Faculty of Physicians and Surgeons


Hall, B.G., 2004a. In vitro evolution predicts that the IMP-1 metallo-beta-lactamase does not have the potential to evolve increased activity against imipenem. Antimicrobial Agents and Chemotherapy, 48(3), pp.1032–1033.


Smith, W.S., Hale, J.R. & Neylon, C., 2011. Applying neutral drift to the directed molecular evolution of a β-glucuronidase into a β-galactosidase: Two different evolutionary pathways...


Appendices

Appendix A Supplementary Figures and Tables

Appendix A contains supplementary figures and tables from chapters one through five.
Table A.1: The sequence and structural similarity data for all MBL variants examined in Chapter 2.

<table>
<thead>
<tr>
<th>Enzyme Similarity</th>
<th>BcII</th>
<th>IND-1</th>
<th>IMP-1</th>
<th>CcrA</th>
<th>VIM-1</th>
<th>VIM-2</th>
<th>NDM-1</th>
<th>SPM-1</th>
<th>Reference</th>
<th>GenBank Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>BcII</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
<td>M11189</td>
</tr>
<tr>
<td>IND-1</td>
<td>51%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EF394437</td>
<td></td>
</tr>
<tr>
<td>IMP-1</td>
<td>44%</td>
<td>53%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AY168635</td>
<td></td>
</tr>
<tr>
<td>CcrA</td>
<td>44%</td>
<td>48%</td>
<td>46%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M63556</td>
<td></td>
</tr>
<tr>
<td>VIM-1</td>
<td>44%</td>
<td>48%</td>
<td>46%</td>
<td>46%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td>Y18050</td>
<td></td>
</tr>
<tr>
<td>VIM-2</td>
<td>43%</td>
<td>47%</td>
<td>45%</td>
<td>45%</td>
<td>92%</td>
<td>100%</td>
<td></td>
<td></td>
<td>AF191564</td>
<td></td>
</tr>
<tr>
<td>NDM-1</td>
<td>39%</td>
<td>38%</td>
<td>40%</td>
<td>43%</td>
<td>48%</td>
<td>48%</td>
<td>100%</td>
<td></td>
<td>KX999121</td>
<td></td>
</tr>
<tr>
<td>SPM-1</td>
<td>40%</td>
<td>47%</td>
<td>44%</td>
<td>41%</td>
<td>44%</td>
<td>44%</td>
<td>38%</td>
<td>100%</td>
<td>AY341249</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Enzyme Similarity</th>
<th>BcII</th>
<th>IND-1</th>
<th>IMP-1</th>
<th>CcrA</th>
<th>VIM-1</th>
<th>VIM-2</th>
<th>NDM-1</th>
<th>SPM-1</th>
<th>Reference</th>
<th>GenBank Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>BcII</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AAA22276</td>
<td></td>
</tr>
<tr>
<td>IND-1</td>
<td>33%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ABO21411</td>
<td></td>
</tr>
<tr>
<td>IMP-1</td>
<td>33%</td>
<td>30%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AAN87168</td>
<td></td>
</tr>
<tr>
<td>CcrA</td>
<td>31%</td>
<td>28%</td>
<td>33%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AAA22904</td>
<td></td>
</tr>
<tr>
<td>VIM-1</td>
<td>36%</td>
<td>27%</td>
<td>30%</td>
<td>29%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td>CAE46717</td>
<td></td>
</tr>
<tr>
<td>VIM-2</td>
<td>36%</td>
<td>25%</td>
<td>32%</td>
<td>28%</td>
<td>90%</td>
<td>100%</td>
<td></td>
<td></td>
<td>AAK26253</td>
<td></td>
</tr>
<tr>
<td>NDM-1</td>
<td>28%</td>
<td>23%</td>
<td>31%</td>
<td>29%</td>
<td>35%</td>
<td>34%</td>
<td>100%</td>
<td></td>
<td>AF72857</td>
<td></td>
</tr>
<tr>
<td>SPM-1</td>
<td>26%</td>
<td>29%</td>
<td>32%</td>
<td>24%</td>
<td>26%</td>
<td>26%</td>
<td>23%</td>
<td>100%</td>
<td>AAR15341</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Enzyme Similarity</th>
<th>BcII</th>
<th>IND-1</th>
<th>IMP-1</th>
<th>CcrA</th>
<th>VIM-1</th>
<th>VIM-2</th>
<th>NDM-1</th>
<th>SPM-1</th>
<th>Reference</th>
<th>GenBank Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>BcII</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
<td>5FQA</td>
</tr>
<tr>
<td>IND-1</td>
<td>1.56</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3L6N</td>
<td></td>
</tr>
<tr>
<td>IMP-1</td>
<td>1.66</td>
<td>2.10</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5EV6</td>
<td></td>
</tr>
<tr>
<td>CcrA</td>
<td>1.20</td>
<td>1.70</td>
<td>1.57</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12NB</td>
<td></td>
</tr>
<tr>
<td>VIM-1</td>
<td>1.87</td>
<td>1.75</td>
<td>1.51</td>
<td>1.62</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td>4UWO</td>
<td></td>
</tr>
<tr>
<td>VIM-2</td>
<td>1.85</td>
<td>1.68</td>
<td>1.58</td>
<td>1.54</td>
<td>0.58</td>
<td>0.00</td>
<td></td>
<td></td>
<td>1KO3</td>
<td></td>
</tr>
<tr>
<td>NDM-1</td>
<td>1.75</td>
<td>2.14</td>
<td>1.93</td>
<td>1.82</td>
<td>2.07</td>
<td>2.00</td>
<td>0.00</td>
<td>0.00</td>
<td>3SPU</td>
<td></td>
</tr>
<tr>
<td>SPM-1</td>
<td>1.66</td>
<td>2.09</td>
<td>1.63</td>
<td>1.64</td>
<td>1.95</td>
<td>1.95</td>
<td>1.86</td>
<td>0.00</td>
<td>2FHX</td>
<td></td>
</tr>
</tbody>
</table>

1No crystal structures have yet been deposited to the PDB for IND-1 and VIM-1 so the closest related variant was used instead: VIM-26 (99.6% amino acid sequence identity to VIM-1) and IND-7 (100% amino acid sequence identity to IND-1).
## Table A.2: MIC measurements of all MBL variants for all tested antibiotics in three host organisms from Chapter 2.

<table>
<thead>
<tr>
<th>MIC (µg/mL)</th>
<th>5.9-BclI</th>
<th>5.9-IND-1</th>
<th>5.9-IMP-1</th>
<th>5.9-CcrA</th>
<th>5.9-VIM-1</th>
<th>5.9-VIM-2</th>
<th>5.9-NDM-1</th>
<th>5.9-NDM-1</th>
<th>5.9-NDM-1</th>
<th>5.9-NDM-1</th>
<th>5.9-NDM-1</th>
<th>Range Screened</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cefotaxime</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>0</td>
<td>512</td>
<td>256</td>
<td>64</td>
<td>16</td>
<td>0.125</td>
<td>0.032</td>
<td>0.032 - 4096</td>
<td></td>
</tr>
<tr>
<td>Ceftazidime</td>
<td>128</td>
<td>32</td>
<td>256</td>
<td>0</td>
<td>1024</td>
<td>16</td>
<td>1024</td>
<td>32</td>
<td>0.25</td>
<td>0.032</td>
<td>0.032 - 4096</td>
<td></td>
</tr>
<tr>
<td>Meropenem</td>
<td>16</td>
<td>128</td>
<td>8</td>
<td>0</td>
<td>16</td>
<td>8</td>
<td>4</td>
<td>0.25</td>
<td>0.064</td>
<td>0.016</td>
<td>0.016 - 64</td>
<td></td>
</tr>
<tr>
<td>Imipenem</td>
<td>8</td>
<td>8</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>0.032</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016 - 64</td>
<td></td>
</tr>
<tr>
<td>Ampicillin</td>
<td>4096</td>
<td>16384</td>
<td>256</td>
<td>0</td>
<td>16384</td>
<td>1024</td>
<td>1024</td>
<td>64</td>
<td>4</td>
<td>2 - 32768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Penicillin</td>
<td>8192</td>
<td>8192</td>
<td>512</td>
<td>0</td>
<td>4096</td>
<td>2048</td>
<td>512</td>
<td>256</td>
<td>32</td>
<td>2 - 32768</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>P. aeruginosa</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cefotaxime</td>
<td>1024</td>
<td>1024</td>
<td>256</td>
<td>8</td>
<td>2048</td>
<td>512</td>
<td>2048</td>
<td>1024</td>
<td>4</td>
<td>0.032</td>
<td>0.032 - 4096</td>
<td></td>
</tr>
<tr>
<td>Ceftazidime</td>
<td>64</td>
<td>16</td>
<td>128</td>
<td>2</td>
<td>1024</td>
<td>32</td>
<td>8192</td>
<td>1024</td>
<td>1</td>
<td>0.032</td>
<td>0.032 - 4096</td>
<td></td>
</tr>
<tr>
<td>Meropenem</td>
<td>&gt;64</td>
<td>&gt;64</td>
<td>64</td>
<td>4</td>
<td>&gt;64</td>
<td>&gt;64</td>
<td>&gt;64</td>
<td>&gt;64</td>
<td>0.5</td>
<td>0.016</td>
<td>0.016 - 64</td>
<td></td>
</tr>
<tr>
<td>Imipenem</td>
<td>&gt;64</td>
<td>&gt;64</td>
<td>8</td>
<td>2</td>
<td>32</td>
<td>64</td>
<td>64</td>
<td>16</td>
<td>2</td>
<td>0.016</td>
<td>0.016 - 64</td>
<td></td>
</tr>
<tr>
<td>Ampicillin</td>
<td>4096</td>
<td>16384</td>
<td>2048</td>
<td>256</td>
<td>16384</td>
<td>2048</td>
<td>4096</td>
<td>4096</td>
<td>512</td>
<td>2 - 32768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Penicillin</td>
<td>32768</td>
<td>8192</td>
<td>32768</td>
<td>8192</td>
<td>8192</td>
<td>16384</td>
<td>16384</td>
<td>8192</td>
<td>8192</td>
<td>2 - 32768</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>K. pneumoniae</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATCC 13883</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cefotaxime</td>
<td>512</td>
<td>256</td>
<td>128</td>
<td>32</td>
<td>128</td>
<td>128</td>
<td>512</td>
<td>32</td>
<td>0.125</td>
<td>0.032</td>
<td>0.032 - 4096</td>
<td></td>
</tr>
<tr>
<td>Ceftazidime</td>
<td>512</td>
<td>64</td>
<td>256</td>
<td>4</td>
<td>2048</td>
<td>8</td>
<td>1024</td>
<td>32</td>
<td>1</td>
<td>0.032</td>
<td>0.032 - 4096</td>
<td></td>
</tr>
<tr>
<td>Meropenem</td>
<td>64</td>
<td>64</td>
<td>16</td>
<td>16</td>
<td>64</td>
<td>8</td>
<td>16</td>
<td>16</td>
<td>0.064</td>
<td>0.016</td>
<td>0.016 - 64</td>
<td></td>
</tr>
<tr>
<td>Imipenem</td>
<td>64</td>
<td>64</td>
<td>8</td>
<td>4</td>
<td>32</td>
<td>8</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>0.016</td>
<td>0.016 - 64</td>
<td></td>
</tr>
<tr>
<td>Ampicillin</td>
<td>4096</td>
<td>8192</td>
<td>512</td>
<td>512</td>
<td>8192</td>
<td>512</td>
<td>512</td>
<td>256</td>
<td>256</td>
<td>2 - 32768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Penicillin</td>
<td>2048</td>
<td>2048</td>
<td>512</td>
<td>512</td>
<td>1024</td>
<td>512</td>
<td>512</td>
<td>256</td>
<td>256</td>
<td>2 - 32768</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure A.1: Measured minimum inhibitory concentrations (MICs) for the metallo-β-lactamases with representative β-lactam antibiotics in *E. coli*, *P. aeruginosa*, and *K. pneumoniae*. MICs displayed for cefotaxime (light blue), ceftazidime (dark blue), meropenem (light green), imipenem (dark green), ampicillin (light red), and penicillin (dark red) measured for each MBL (BclII, IND-1, IMP-1, CcrA, VIM-1, VIM-2, NDM-1, and SPM-1) in *E. coli* 10G, *P. aeruginosa* PA01, and *K. pneumoniae* ATCC 13883. MICs were determined with the concentration at which at least three of four replicates did not grow. The grey represents the background resistance of the organisms without MBL expression. The chemical structures of each antibiotic, as well as the concentrations screened to determine the MICs, are shown on the right.
Figure A.2: The initial rates of reaction for each enzyme-substrate pair over the testable substrate range for BcII, IND-1, IMP-1, CcrA, VIM-1, VIM-2, SPM-1, and NDM-1. Each data point represents the average initial rate of reaction of three replicates. Curves were determined by fitting the data to the Michaelis-Menten with KaleidaGraph (Synergy).
Table A.3: Kinetic parameters for all eight MBL variants for CENTA, cefotaxime, ceftazidime, meropenem, imipenem, ampicillin, and penicillin.

<table>
<thead>
<tr>
<th>Kinetics</th>
<th>BccII</th>
<th>IND-1</th>
<th>IMP-1</th>
<th>CcrA</th>
<th>VIM-1</th>
<th>VIM-2</th>
<th>NDM-1</th>
<th>SPM-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{cat}$ (s(^{-1}))</td>
<td>140 ± 10</td>
<td>150 ± 10</td>
<td>45 ± 2</td>
<td>55 ± 3</td>
<td>72 ± 2</td>
<td>33 ± 1</td>
<td>200 ± 10</td>
<td>30 ± 1</td>
</tr>
<tr>
<td>$K_m$ (µM)</td>
<td>74 ± 6</td>
<td>45 ± 5</td>
<td>2.2 ± 0.4</td>
<td>3.8 ± 0.8</td>
<td>54 ± 5</td>
<td>8.4 ± 0.9</td>
<td>8.0 ± 1.4</td>
<td>1.7 ± 0.2</td>
</tr>
<tr>
<td>$k_{cat}/K_m$ (M(^{-1})•s(^{-1}))</td>
<td>1.9 × 10(^6)</td>
<td>3.5 × 10(^6)</td>
<td>2.1 × 10(^7)</td>
<td>1.4 × 10(^8)</td>
<td>1.3 × 10(^8)</td>
<td>3.9 × 10(^8)</td>
<td>2.6 × 10(^8)</td>
<td>1.8 × 10(^8)</td>
</tr>
</tbody>
</table>

| Cefotaxime | $k_{cat}$ (s\(^{-1}\)) | 210 ± 10 | 130 ± 10 | 8.3 ± 0.9 | 33 ± 2 | 190 ± 10 | 38 ± 2 | 68 ± 2 | 36 ± 2 |
| $K_m$ (µM) | 58 ± 7 | 140 ± 20 | 10 ± 4 | 3.5 ± 0.6 | 73 ± 6 | 15 ± 3 | 7.2 ± 0.8 | 19 ± 2 |
| $k_{cat}/K_m$ (M\(^{-1}\)•s\(^{-1}\)) | 3.6 × 10\(^6\) | 9.5 × 10\(^5\) | 8.2 × 10\(^5\) | 9.3 × 10\(^5\) | 2.7 × 10\(^6\) | 2.5 × 10\(^6\) | 9.5 × 10\(^5\) | 1.9 × 10\(^6\) |

| Ceftazidime | $k_{cat}$ (s\(^{-1}\)) | 5.7 ± 0.8 | n.d. | 5.5 ± 0.4 | 36 ± 3 | 46 ± 4 | 1.1 ± 0.1 | 97 ± 2 | 12 ± 1 |
| $K_m$ (µM) | 430 ± 90 | n.d. | 37 ± 8 | 310 ± 40 | 270 ± 40 | 66 ± 8 | 51 ± 2 | 55 ± 7 |
| $k_{cat}/K_m$ (M\(^{-1}\)•s\(^{-1}\)) | 1.3 × 10\(^4\) | 1.5 × 10\(^4\) | 1.5 × 10\(^5\) | 1.2 × 10\(^5\) | 1.7 × 10\(^5\) | 1.6 × 10\(^4\) | 1.9 × 10\(^5\) | 2.1 × 10\(^5\) |

| Meropenem | $k_{cat}$ (s\(^{-1}\)) | n.d. | n.d. | 140 ± 10 | 110 ± 10 | 19 ± 1 | 0.90 ± 0.03 | 26 ± 1 | 57 ± 3 |
| $K_m$ (µM) | n.d. | n.d. | 30 ± 5 | 140 ± 20 | 140 ± 10 | 7.4 ± 1.0 | 21 ± 2 | 150 ± 20 |
| $k_{cat}/K_m$ (M\(^{-1}\)•s\(^{-1}\)) | 2.7 × 10\(^5\) | 1.8 × 10\(^5\) | 4.9 × 10\(^5\) | 7.9 × 10\(^5\) | 1.4 × 10\(^5\) | 1.2 × 10\(^5\) | 1.2 × 10\(^5\) | 3.8 × 10\(^5\) |

| Imipenem | $k_{cat}$ (s\(^{-1}\)) | 120 ± 10 | n.d. | 30 ± 5 | 110 ± 10 | 6.8 ± 0.3 | 8.7 ± 0.3 | 123 ± 3 | 12 ± 1 |
| $K_m$ (µM) | 190 ± 20 | n.d. | 100 ± 10 | 270 ± 40 | 24 ± 4 | 8.4 ± 1.1 | 56 ± 4 | 59 ± 7 |
| $k_{cat}/K_m$ (M\(^{-1}\)•s\(^{-1}\)) | 6.1 × 10\(^5\) | 7.8 × 10\(^5\) | 3.6 × 10\(^5\) | 4.1 × 10\(^5\) | 2.8 × 10\(^5\) | 1.0 × 10\(^5\) | 2.1 × 10\(^5\) | 2.0 × 10\(^5\) |

| Ampicillin | $k_{cat}$ (s\(^{-1}\)) | 290 ± 20 | 130 ± 10 | 95 ± 8 | 200 ± 10 | 130 ± 10 | 23 ± 1 | 720 ± 30 | 1500 ± 250 |
| $K_m$ (µM) | 1030 ± 120 | 140 ± 30 | 610 ± 110 | 150 ± 30 | 140 ± 30 | 83 ± 15 | 86 ± 15 | 550 ± 180 |
| $k_{cat}/K_m$ (M\(^{-1}\)•s\(^{-1}\)) | 2.8 × 10\(^5\) | 9.2 × 10\(^5\) | 1.5 × 10\(^6\) | 1.3 × 10\(^6\) | 9.2 × 10\(^5\) | 2.7 × 10\(^5\) | 8.4 × 10\(^5\) | 2.7 × 10\(^5\) |

| Penicillin | $k_{cat}$ (s\(^{-1}\)) | 290 ± 20 | 35 ± 2 | 410 ± 10 | 130 ± 10 | 580 ± 30 | 76 ± 3 | 320 ± 20 | 1300 ± 50 |
| $K_m$ (µM) | 200 ± 40 | 140 ± 20 | 480 ± 20 | 36 ± 9 | 170 ± 40 | 51 ± 10 | 61 ± 11 | 92 ± 14 |
| $k_{cat}/K_m$ (M\(^{-1}\)•s\(^{-1}\)) | 1.4 × 10\(^6\) | 2.5 × 10\(^5\) | 8.5 × 10\(^5\) | 3.8 × 10\(^6\) | 3.5 × 10\(^6\) | 1.5 × 10\(^6\) | 5.3 × 10\(^5\) | 1.4 × 10\(^7\) |

n.d. not determined
Figure A.3: Relationship between $k_{\text{cat}}/K_M$ and selected β-lactam antibiotic MICs for the metallo-β-lactamases in the three organisms. The measured cefotaxime, ceftazidime, meropenem, imipenem, ampicillin, and penicillin MIC values for each MBL in relation to their corresponding $k_{\text{cat}}/K_M$ for each organism are shown with the background resistance for each organism denoted by the grey box.
Table A.4: Measured copies of functional enzyme in the periplasm and cytoplasmic fractions for each variant in each organism.

<table>
<thead>
<tr>
<th>Enzyme Copy Number</th>
<th>E. coli 10G</th>
<th></th>
<th>P. aeruginosa PAO1</th>
<th></th>
<th>K. pneumoniae ATCC13883</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Periplasm</td>
<td>Cytoplasm</td>
<td>Periplasm</td>
<td>Cytoplasm</td>
<td>Periplasm</td>
<td>Cytoplasm</td>
</tr>
<tr>
<td>BclI</td>
<td>460 ± 50</td>
<td>5900 ± 300</td>
<td>230 ± 20</td>
<td>2100 ± 200</td>
<td>260 ± 20</td>
<td>560 ± 80</td>
</tr>
<tr>
<td>IND-1</td>
<td>6.0 ± 1.2 *</td>
<td>500 ± 20</td>
<td>3.7 ± 0.8 *</td>
<td>37 ± 6</td>
<td>5.8 ± 0.6 *</td>
<td>14 ± 2</td>
</tr>
<tr>
<td>IMP-1</td>
<td>83 ± 4</td>
<td>2600 ± 100</td>
<td>100 ± 10</td>
<td>730 ± 100</td>
<td>50 ± 3</td>
<td>61 ± 5</td>
</tr>
<tr>
<td>CcrA</td>
<td>54 ± 1</td>
<td>35 ± 5</td>
<td>31 ± 2</td>
<td>170 ± 70</td>
<td>6.9 ± 0.7</td>
<td>16 ± 5</td>
</tr>
<tr>
<td>VIM-1</td>
<td>8400 ± 700</td>
<td>16000 ± 1000</td>
<td>9600 ± 1100</td>
<td>13000 ± 1000</td>
<td>1900 ± 500</td>
<td>1100 ± 300</td>
</tr>
<tr>
<td>VIM-2</td>
<td>1500 ± 200</td>
<td>6900 ± 700</td>
<td>4500 ± 200</td>
<td>3300 ± 400</td>
<td>280 ± 30</td>
<td>190 ± 20</td>
</tr>
<tr>
<td>NDM-1</td>
<td>4.6 ± 0.8 **</td>
<td>230 ± 10</td>
<td>36 ± 2 **</td>
<td>890 ± 110</td>
<td>2.5 ± 0.2 **</td>
<td>30 ± 5</td>
</tr>
<tr>
<td>SPM-1</td>
<td>13 ± 1</td>
<td>260 ± 30</td>
<td>29 ± 1</td>
<td>910 ± 70</td>
<td>25 ± 2</td>
<td>25 ± 1</td>
</tr>
</tbody>
</table>

* IND-1 appears to be inactive in this particular assay for periplasmic activity

** NDM-1 has been found to be bound to the outer membrane, and was therefore unable to be accurately quantified in the periplasm with this method
Figure A.4: The variation within MIC, $k_{\text{cat}}/K_M$, and $[E_p]$ for each antibiotic, organism, and MBL. Each data point represents the fold-difference between the values for each MBL divided by the lowest value present for the group. The black line represents the average for each category.
Figure A.5: Relationship between MIC and the product of $k_{\text{cat}}/K_M$ and $[E_p]$ for each MBL in the three organisms. The correlation between the product of $k_{\text{cat}}/K_M$ and $[E_p]$ with MIC. NDM-1 and IND-1 were not included in the fit, as NDM-1 is known to be bound to the outer membrane and could therefore not be accurately measured with the assay, while IND-1 activity was not detectable in the assay.
Table A.5: Measured amounts of DNA, RNA, and protein in both the cytoplasm and periplasm for the eight MBL variants investigated in Chapter 2.

<table>
<thead>
<tr>
<th>Variant</th>
<th>DNA (ng/µl/OD)</th>
<th>RNA (relative)</th>
<th>Cytoplasm (µg/ul)</th>
<th>Percent in Periplasm (%)</th>
<th>Periplasm (µg/ul)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BclI</td>
<td>4.7 ± 2.7</td>
<td>580 ± 50</td>
<td>5900 ± 300</td>
<td>7</td>
<td>460 ± 50</td>
</tr>
<tr>
<td>IND-1</td>
<td>10.7 ± 3.5</td>
<td>390 ± 20</td>
<td>500 ± 20</td>
<td>1</td>
<td>n.d.</td>
</tr>
<tr>
<td>IMP-1</td>
<td>14.8 ± 4.7</td>
<td>150 ± 20</td>
<td>2600 ± 100</td>
<td>3</td>
<td>83 ± 4</td>
</tr>
<tr>
<td>CcrA</td>
<td>10.6 ± 3.8</td>
<td>620 ± 100</td>
<td>35 ± 5</td>
<td>61</td>
<td>54 ± 1</td>
</tr>
<tr>
<td>VIM-1</td>
<td>12.2 ± 4.7</td>
<td>450 ± 20</td>
<td>16000 ± 1000</td>
<td>34</td>
<td>8400 ± 700</td>
</tr>
<tr>
<td>VIM-2</td>
<td>6.4 ± 4.7</td>
<td>210 ± 40</td>
<td>6900 ± 700</td>
<td>18</td>
<td>1500 ± 200</td>
</tr>
<tr>
<td>NDM-1</td>
<td>4.8 ± 2.4</td>
<td>32 ± 7</td>
<td>230 ± 10</td>
<td>2</td>
<td>n.d.</td>
</tr>
<tr>
<td>SPM-1</td>
<td>5.4 ± 4.2</td>
<td>110 ± 60</td>
<td>260 ± 30</td>
<td>5</td>
<td>13 ± 1</td>
</tr>
</tbody>
</table>

n.d. not determined
Table A.6: MIC measurements of all MBL variants with the PelB leader sequence for all tested antibiotics in three different host organisms from Chapter 2.

<table>
<thead>
<tr>
<th>MIC (µg/mL)</th>
<th>5.9-Bcll</th>
<th>5.9-IND-1</th>
<th>5.9-IMP-1</th>
<th>5.9-CcrA</th>
<th>5.9-VIM-1</th>
<th>5.9-VIM-2</th>
<th>5.9-NDM-1</th>
<th>5.9-SPM-1</th>
<th>5.9 Plasmid</th>
<th>Range Screened</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli 10G</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cefotaxime</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>0</td>
<td>128</td>
<td>4</td>
<td>512</td>
<td>128</td>
<td>0.125</td>
<td>0.032 - 4096</td>
</tr>
<tr>
<td>Ceftazidime</td>
<td>16</td>
<td>8</td>
<td>128</td>
<td>0</td>
<td>32</td>
<td>1</td>
<td>2048</td>
<td>128</td>
<td>0.25</td>
<td>0.032 - 4096</td>
</tr>
<tr>
<td>Meropenem</td>
<td>1</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>0.064</td>
<td>0.016 - 64</td>
</tr>
<tr>
<td>Imipenem</td>
<td>0.25</td>
<td>0.25</td>
<td>1</td>
<td>0</td>
<td>0.25</td>
<td>0.032</td>
<td>1</td>
<td>0.125</td>
<td>0.016</td>
<td>0.016 - 64</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>128</td>
<td>512</td>
<td>128</td>
<td>0</td>
<td>128</td>
<td>32</td>
<td>4096</td>
<td>128</td>
<td>4</td>
<td>2 - 32768</td>
</tr>
<tr>
<td>Penicillin</td>
<td>256</td>
<td>512</td>
<td>512</td>
<td>0</td>
<td>256</td>
<td>128</td>
<td>2048</td>
<td>512</td>
<td>32</td>
<td>2 - 32768</td>
</tr>
<tr>
<td><strong>P. aeruginosa PA01</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cefotaxime</td>
<td>1024</td>
<td>256</td>
<td>512</td>
<td>8</td>
<td>1024</td>
<td>128</td>
<td>2048</td>
<td>1024</td>
<td>4</td>
<td>0.032 - 4096</td>
</tr>
<tr>
<td>Ceftazidime</td>
<td>16</td>
<td>4</td>
<td>128</td>
<td>2</td>
<td>256</td>
<td>16</td>
<td>4096</td>
<td>1024</td>
<td>1</td>
<td>0.032 - 4096</td>
</tr>
<tr>
<td>Meropenem</td>
<td>&gt;64</td>
<td>32</td>
<td>&gt;64</td>
<td>4</td>
<td>&gt;64</td>
<td>64</td>
<td>&gt;64</td>
<td>&gt;64</td>
<td>0.5</td>
<td>0.016 - 64</td>
</tr>
<tr>
<td>Imipenem</td>
<td>32</td>
<td>32</td>
<td>8</td>
<td>2</td>
<td>32</td>
<td>32</td>
<td>&gt;64</td>
<td>16</td>
<td>2</td>
<td>0.016 - 64</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>4096</td>
<td>4096</td>
<td>2048</td>
<td>256</td>
<td>2048</td>
<td>2048</td>
<td>4096</td>
<td>8192</td>
<td>512</td>
<td>2 - 32768</td>
</tr>
<tr>
<td>Penicillin</td>
<td>32768</td>
<td>8192</td>
<td>16384</td>
<td>8192</td>
<td>8192</td>
<td>16384</td>
<td>8192</td>
<td>8192</td>
<td>2 - 32768</td>
<td></td>
</tr>
<tr>
<td><strong>K. pneumoniae ATCC 13883</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cefotaxime</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>32</td>
<td>32</td>
<td>8</td>
<td>128</td>
<td>32</td>
<td>0.125</td>
<td>0.032 - 4096</td>
</tr>
<tr>
<td>Ceftazidime</td>
<td>8</td>
<td>16</td>
<td>256</td>
<td>4</td>
<td>32</td>
<td>2</td>
<td>1024</td>
<td>128</td>
<td>1</td>
<td>0.032 - 4096</td>
</tr>
<tr>
<td>Meropenem</td>
<td>2</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>16</td>
<td>1</td>
<td>0.064</td>
<td>0.016 - 64</td>
</tr>
<tr>
<td>Imipenem</td>
<td>4</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>0.016 - 64</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>256</td>
<td>2 - 32768</td>
<td></td>
</tr>
<tr>
<td>Penicillin</td>
<td>512</td>
<td>2048</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>256</td>
<td>512</td>
<td>512</td>
<td>2 - 32768</td>
<td></td>
</tr>
</tbody>
</table>
Figure A.6: Relationship between $k_{\text{cat}}/K_M$ and selected β-lactam antibiotic MICs for the three organisms harbouring the PelB-metallo-β-lactamase constructs. The measured cefotaxime, ceftazidime, meropenem, imipenem, ampicillin, and penicillin MIC values for each MBL in relation to their corresponding $k_{\text{cat}}/K_M$ for each organism are shown with the background resistance for each organism denoted by the grey box.
Figure A.7: The initial rates of reaction for each enzyme-substrate pair over the testable substrate range for NDM-1, VIM-2, IMP-1, and a selected variant from the 18th round of each ampicillin trajectory, NA, VA, and IA. Each data point represents the average initial rate of reaction of three replicates. Curves were determined by fitting the data to the Michaelis-Menten with KaleidaGraph (Synergy).
Figure A.8: The initial rates of reaction for each MBL-substrate pair. Each data point reflects the average of three replicates. For those pairs where substrate saturation of the enzyme was not possible, the linear portion of the Michaelis-Menten plot was used to determine the $k_{\text{cat}}/K_M$ values.
Appendix B  Sequencing Data

Appendix B contains supplementary data for the thesis, including the sequences of variants from the directed evolution trajectories in Chapters 3 and 4, the complete deep mutational scanning dataset that is referenced in Chapter 4, and the sequences of variants from the genetic drift experiment in Chapter 5.
Figure B.1: A map showing the location of each mutation acquired in the NDM-1 directed evolution trajectories. Each mutation is highlighted in green. Black represents portions of the variant not sequenced.
Figure B.2: A map showing the location of each mutation acquired in the VIM-2 directed evolution trajectories. Each mutation is highlighted in green. Black represents portions of the variant not sequenced.
Figure B.3: A map showing the location of each mutation acquired in the IMP-1 directed evolution trajectories. Each mutation is highlighted in green. Black represents portions of the variant not sequenced.
Figure B.4: A heat map displaying the relative enrichment of each possible amino acid mutation at each position of VIM-2 under selection from three different β-lactam antibiotics. Relative enrichment is shown by shading and colour: dark teal indicates that the mutation is beneficial compared to wild type, whereas dark brown indicates that the mutation is deleterious compared to the wild type.
Figure B.5: A map showing the location of mutations identified in the sequenced variants of selected rounds from each genetic drift trajectory. Each mutation is highlighted in purple. Grey represents portions of the gene not subjected to mutation.