Author:
Yang, Zongfan

Title:
Routes to inhibitors of colistin antibiotic resistance mediated by bacterial phosphoethanolamine transferases
computational and microbiological investigations

University of BRISTOL

Routes to inhibitors of Colistin Antibiotic Resistance mediated by Bacterial Phosphoethanolamine Transferases: Computational and Microbiological Investigations

By

Zongfan Yang

A dissertation submitted to the University of Bristol in accordance with the requirements for the award of the degree of Doctor of Philosophy in Cellular and Molecular Medicine in the School of Cellular and Molecular Medicine, Faculty of Life Sciences.

December 2021

Author’s Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE:.....

COVID-19 Impact Statement

The COVID-19 pandemic led to the closure of the laboratory for three and a half months in 2020 spring (March to June). After re-opening, the laboratory occupancy was strictly limited until summer 2021 due to social distancing rules. My access to the lab was limited due to these unexpected but essential actions. Access to essential shared equipment and facilities housed in other laboratories has been similarly restricted. Relevant affected facilities include equipment for protein purification, such as the cell disruptor and ultracentrifuge. Limited access to the lab and equipment reduced the amount of laboratory work possible, meanwhile, my lab work was also impacted by the disruption of reagent and chemicals supplies. In some cases it took us months to receive the order rather than the few weeks common before the pandemic.

Acknowledgements

I would like to express my sincere gratitude to my primary supervisor Prof. James Spencer for the continuous support during my PhD study. He is patient and cares the well-being of his students. He guided me throughout the project, and provided me with valuable advice in the writing of this thesis. Besides my supervisor, I would like to thank my secondary supervisor Prof. Adrian J. Mulholland for his enthusiasm, humour and knowledge that inspired me during my research.

I also would like to give my thanks to the two members of the annual progress monitor panel: Prof. Matthew Avison and Prof. Natalie Fey who provided me excellent views that helped improve my research. My thanks also go to the staff, particularly Dr. Samantha Southern, and my cohort of South West Biosciences Doctoral Training Partnership (SWBio) for providing me excellent opportunities to expand my PhD experience.

I would like to thank my colleagues of the D60 research community (Dr. Philip Hinchliffe, Dr. Catherine Tooke, Dr. Charlotte K. Colenso, Dr. Yuiko Takebayashi, Dr. Punyawee Dulyayangkul and Dr. Kirsty Goudar, etc.) for all the generous help and advice provided during this project. I also would like to express my thanks to the Mulholland group (especially Dr. Reynier Suardiaz, Dr. Eric Lang and Dr. Rebecca M. Twidale) for the value advice on computational modelling. I'd like to thank my MCR-1 cohort Dr. Emily Lythell and Dr. Roberto Perez Chavarria, for all the help and advice provided regarding the purification and mechanism of MCR-1 protein. I also would like to say thanks to John Shaw and the D-floor technical team for their fantastic support on laboratory supply and the technical aspects with various equipment used in this work.

I sincerely thank my family for supporting my study and especially my parents and grandparents for helping me digest negative emotions. Besides, I would like to thank to my friends near and far, old and new, for all their support, encouragement and being reliable during these years.

Last but not least, I would like to thank the staff in the University of Bristol for the administrative support and China Scholarship Council (CSC) for the financial support provided to accomplish this project.

Abstract

As one of the few remaining reliable antibiotics, colistin retains effectiveness against most multi-drug resistant (MDR) Gram-negative strains and is considered as the last resort to treat infections by these organisms. The mobile colistin resistance gene *mcr-1* was identified in 2015 in *Escherichia coli* and encodes a zinc-dependent phosphoethanolamine (PEtN) transferase that modifies lipid A of the bacterial outer membrane to reduce colistin affinity. One option to overcome colistin resistance is to combine colistin with inhibitors of colistin resistance. This project aims to use a combination of computational and experimental approaches to identify MCR-1 inhibitors to restore sensitivity of producer strains to colistin and related agents.

Virtual screening against the MCR-1 catalytic domain was carried out using the Bristol University Docking Engine (BUDE). Biochemical and biophysical assays were developed to investigate protein:ligand interactions and demonstrated interaction of the small thiol, thioglycolic acid (TGA) with the MCR-1 catalytic domain. 35 candidate compounds including top ranked ligands from BUDE and reported inhibitors of other zinc metalloproteins were tested in these and colistin susceptibility assays. Although in vitro assays did not give clear results, compound C4 (4-iodoisoxazole) was identified as reducing the colistin minimum inhibitory concentration (MIC) of MCR-1 producing *E. coli* and *Klebsiella pneumoniae* strains. Compound C4 also potentiated colistin activity against ArnT mediated colistin-resistant *K. pneumoniae* and meropenem activity against carbapenem-resistant *E.coli* producing the metallo-β-lactamase IMP-1. Compound C4 did not affect the cell membrane integrity/permeability of *E.coli* cells, but induced a colistin heteroresistance phenotype in *K. pneumoniae* strains with ArnT mediated colistin resistance.

A multiscale workflow to model zinc metalloprotein:ligand complex structures, including molecular mechanics (MM) molecular dynamics (MD) and quantum mechanics (QM)/MM MD simulations, was established. The workflow was extensively tested and accurately modelled a variety of tested complexes (Sfh-I, HDAC2, IMP-1, etc.) and applied to MCR-like complex systems including the MCR-1:TGA complex. Modelling results were consistent with current experimental findings, indicating low affinity of TGA for the mono-zinc form of MCR-1.

This project developed an experimental and computational pipeline to screen and validate inhibitors and study their interactions with MCR-1. An inhibitor of MCR-1-mediated colistin resistance, compound C4, was identified and investigated, and a modelling workflow developed with potential general application to zinc metalloproteins. These findings provide tools and knowledge useful to understanding of the MCR-1 zinc site and discovery of inhibitors for MCR-1 like PEtN transferases able to preserve colistin efficacy against MDR Gram-negative bacterial infections.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

BSA	**Bovine Serum Albumin**
NPN	**1-N-phenylnapthylamin**
PNP	**4-Nitrophenol**
PNP-PEtN	**4-Nitrophenol-Phosphoethanolamine**
BUDE	**Bristol University Docking Engine**
CLSI	**Clinical and Laboratory Standards Institute**
COL	**Colistin**
C4	**Compound 4 (4-iodoisoxazole)**
DFT	**density functional theory**
DPA	**Dipicolinic Acid**
EUCAST	**European Committee on Antimicrobial Susceptibility Testing**
GNB	**Gram-Negative Bacteria**
IPTG	**Isopropyl β-D-1-Thiogalactopyranoside**
KDa	**Kilo Daltons**
MEM	**Meropenem**
MBL	**Metallo-β-Lactamase**
MIC	**Minimum Inhibitory Concentration**
MCR	**Mobile Colistin Resistance**
MD	**Molecular Dynamics**
MM	**Molecular Mechanics**
MDR	**Multidrug-Resistant**
Ni-NTA	**Nickel-Charged Affinity Resin**
OST	**Osthole**
PAβN	**Phe-Arg-β-naphthylamide**
PEtN	**Phosphoethanolamine**
PI	**Propidium iodide**
PDB	**Protein Data Bank**
PT	**Pterostilbene**
QM	**Quantum Mechanics**

RPM	**Revolutions Per Minute.**
SBL	**Serine β-Lactamases**
TGA	**Thioglycolic Acid**
MMTZ	**Thiol-based Mercaptomethyl Thiazolidine**
FDA	**U.S. Food and Drug Administration**
VAN	**Vancomycin**
WHO	**World Health Organization**

Chapter 1. General Introduction

1.1 Antimicrobial resistance and colistin

Antibiotics (e.g., the well-known penicillins) are drugs used to kill bacteria or inhibit bacterial growth/replication. They are widely used in livestock farming, veterinary medicine and as medical treatments for bacterial infections in human beings[1,2]. While 'antibiotics' only refers to substances that are produced in the natural environment, the broader term 'antibacterial' refers to all kinds of substances that can be used to kill bacteria, including both natural and human-synthesized products. 'Antimicrobials' is a broader term which further includes medicines used to treat infections caused by other microbes such as viruses, parasites and fungi. Recently, due to rapid and extensive dissemination of multidrug resistance (MDR[3]) and the failure to identify novel antimicrobials over past decades, antimicrobial resistance (AMR) has become a worldwide health threat[4,5]. (MDR is defined as acquired non-susceptibility of microorganisms to at least one agent in three or more antimicrobial categories[6]). *Salmonella* and *Escherichia coli* (*E.coli*) strains with resistance to ampicillin, quinolones, tetracyclines, and sulphonamides have been frequently isolated from poultry and meat samples[7]. Moreover, the emergence of broad-spectrum β-lactamases (e.g., NDM-1), *Klebsiella pneumoniae* carbapenemase (KPC) and colistin resistance have also been reported in recent years exacerbating the current AMR situation [8,9]. Antimicrobial resistance can arise not only from genetic mutations inside the bacteria but can also be gained through transmission of resistance between bacterial strains, which makes it easy to spread and evolve[10]. Consumption of antimicrobials over past decades has led to widespread antimicrobial resistance in food-producing animals, which eventually increases incidence of antimicrobial resistant strains associated with human infections[11]. Antimicrobial resistant strains like MRSA (methicillin-resistant *Staphylococcus aureus*) dramatically erodes therapeutic options for clinical treatment of bacterial infection, leading to rising morbidity and mortality of vulnerable patients, such as surgery patients[12]. The global figure suggest over 700,000 people lost their lives every year due to antibiotics ineffectiveness, and this figure is estimated to reach 10 million per year by 2050[13,14].

The Enterobacterales (e.g., *Escherichia coli* and *Klebsiella pneumoniae*) are versatile Gram negative bacteria (GNB) responsible for various opportunistic infections, for instance, bloodstream, urinary tract and respiratory infections[15]. As well as being universally found as a commensal colonist of the human gastrointestinal tract, *E coli* is a leading human pathogen responsible for both community acquired and healthcare associated infections[16]. *E. coli* is the leading cause of bloodstream infections in the UK and the surveillance of *E.coli* infections has been made mandatory since 2011 by Department of Health[17]. Antimicrobial resistance in *E.coli* has been widely reported and the spread of MDR *E.coli* strains has increased in the last 20 years. The lack of treatment options with existing medicines for MDR GNB infections and the dearth of new antibiotic reaching the market highlight the threaten of GNB infections to human health[13].

Polymyxins, a class of polycationic lipopeptide antibiotics, are classified by the World Health Organization (WHO) as Critically Important Antimicrobials (CIA) for human medicine[18] with highest level of importance. Since they were discovered, polymyxins have been widely used in combatting GNB infections including those caused by most strains of Enterobacteriaceae[19,20]. Polymyxin B and Polymyxin E (colistin) are major members of the family used in clinical practice (**Figure 1-1**)[21]. However potential nephrotoxicity and neurotoxicity of colistin impeded its application in clinical treatment[22] and it was gradually replaced by other safer antimicrobials, such as cephalosporins and quinolones. Although direct usage of colistin as regular therapies for human GNB infections has been restricted, it has been constantly used in veterinary medicine [22]. Besides, owing to its strong antibiotic activity, colistin has been widely used as an in-feed growth promoter in intensive livestock farming (e.g., chicken and pigs) in some countries, for example India, China and Vietnam, to keep animals healthy and reduce the risk of lethal infections[20,23]. Nevertheless, interest in colistin has been renewed due to the recent emergence of multi-drug resistant bacterial infections and the failure of regular antibiotics, in particular for infections caused by carbapenemase-producing strains. As one of the few remaining reliable antimicrobials, colistin retains effectiveness against most MDR strains, including carbapenem-resistant isolates, thus being considered as a last-resort agent against bacterial infections with MDR. Colistin resistance was considered rare in past decades, particularly in *E. coli.* However, unregulated use of colistin for food-producing animals has become a driving force leading to global emergence and transmission of colistin resistance[24,25].

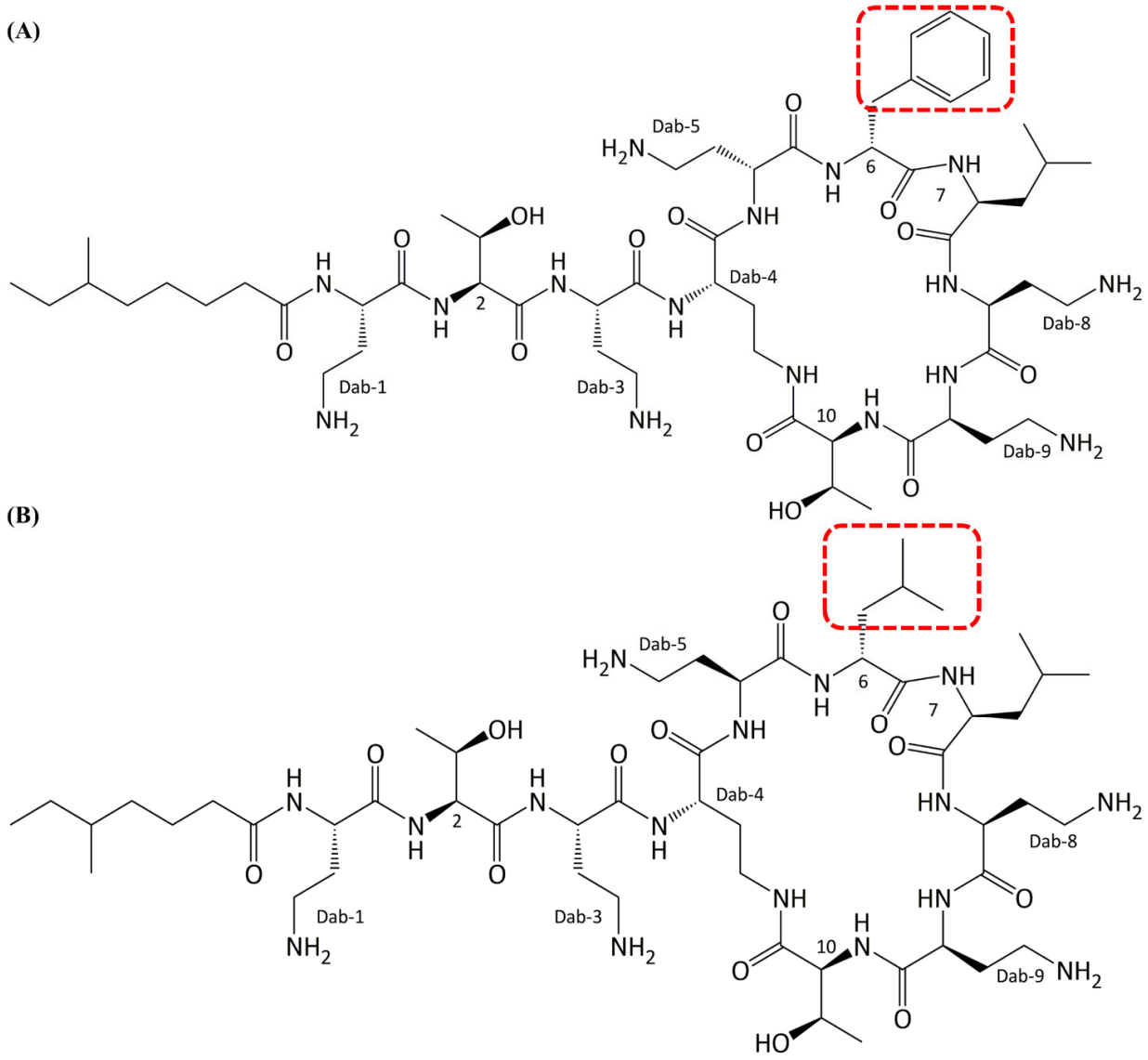


Figure 1-1. Structures of (A) Polymyxin B and (B) Polymyxin E, also known as colistin.
The major difference between polymyxin B and colistin is highlighted by the red box. Dab indicates the positive charged α,γ-diaminobutyric acid residue in polymyxins[26]. The figure is modified from Figure 3 in the dissertation of Roberto Pérez Chavarría[27]

1.2 Prevalence of transferable colistin resistance

The overall charge of the Gram-negative bacterial envelope is normally negative, owing to the phosphate groups on the lipid A moiety anchored to the lipopolysaccharide (LPS) of the outer membrane. As colistin is cationic, electrostatic interaction between colistin and the bacterial surface enables initial binding of colistin to the cell[28]. After binding, colistin penetrates the bacterial outer cell membrane and forms pores in the inner membrane, which eventually causes lysis of the bacteria[22,28]. The exact mechanism by which colistin kills cells after binding to the bacterial outer membrane is still unclear. A recent study claimed that colistin also targets

LPS in the inner membrane of bacteria and disrupts the cation bridges between these LPS molecules resulting in membrane permeabilization and finally bacterial lysis[29].

LPS is the primary target for colistin binding, colistin resistance in *Enterobacterales* is often associated with alteration of LPS[30]. Modification of the lipid A moiety of LPS is the most common mechanism, which reduces the net negative charge of LPS, impeding the binding of colistin[30,31]. These modifications can be achieved by adding phosphoethanolamine (PEtN) and 4-amino-4-deoxy-L-arabinose (L-Ara4N) to the phosphate group of lipid A (**Figure 1-2**). The PmrAB and PhoPQ two-component systems are chromosomal regulators that are responsible for these lipid A modifications via genes encoding the PEtN transferase PmrC (EptA) and the Arn proteins (ArnBCADTEF) in *K. pneumoniae*[21,25,30,32]. The PhoPQ two-component system mediates adaptions of bacteria to low Mg^{2+} environments and regulates the expression of other genes in some Gram-negative species[33]. The PmrAB two-component system responds to Fe^{3+} rich environments to protect cells from iron-mediated killing[34]. The expression of the PmrAB system also can be triggered by acidic pH[35]. Divalent cations like Mg^{2+} and Ca^{2+} are important for bridging adjacent negatively charged LPS molecules, thus stabilizing LPS and the outer membrane of bacteria. The addition of PEtN or L-Ara4N to LPS substitutes for the need for divalent cations thus allowing limiting Mg^{2+} levels to be used in more important activities of cells[33]. The addition of cationic PEtN or L-Ara4N to LPS also increase the net charge of cell envelope which reduces iron binding to the bacterial cell in high Fe^{3+} concentration environments[34]. The CrrAB two-component system can modulate the expression of the PmrAB component and PhoPQ component is modulated by the MgrB regulator. The inactivation of *mgrB* gene upregulates the expression of PhoPQ component conferring colistin resistance. PmrAB can also be also activated by PhoPQ via the expression of pmrD. Chromosomal mutations in genes related to the PmrAB and PhoPQ two-component system often cause colistin resistance[21,25,30]. Resistance to colistin is also reported to be caused by the complete loss of LPS on the cell outer membrane due to mutations in the LPS synthesis genes (e.g., *lpxA, lpxB* and *lpxC*) in *Acinetobacter baumannii*, which prevents the binding of colistin to the cell membrane[21]. In addition to the alternation of LPS, efflux pumps (e.g., Sap proteins system) and capsule formation in bacteria (e.g., *K. pneumoniae*) are other strategies for bacteria to compromise colistin activity[21,25,30].

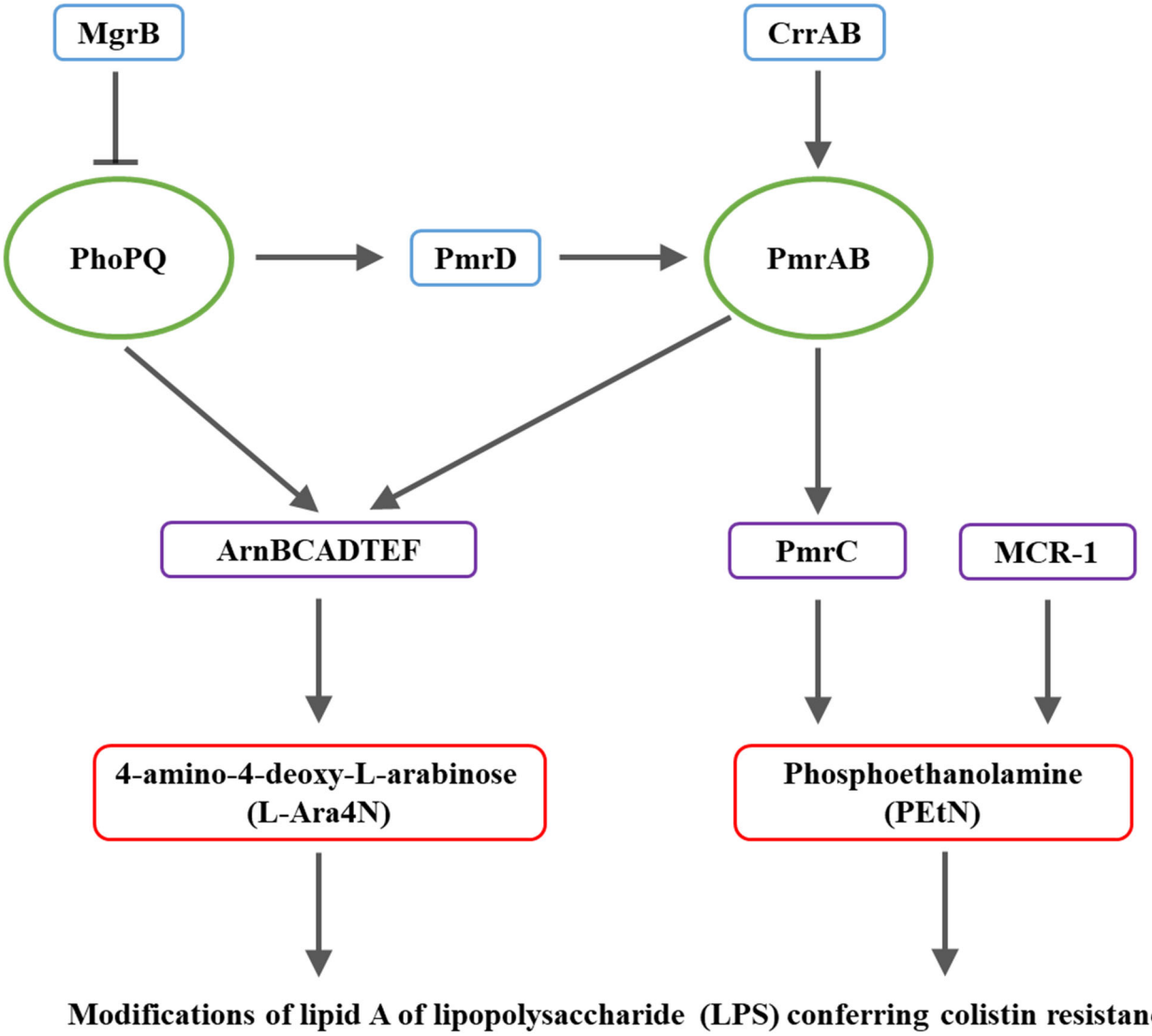


Figure 1-2. Pathways of LPS modifications in K. pneumoniae leading to colistin resistance.

Activation of the PhoPQ or PmrAB two-component systems are commonly observed by chromosomal mutations. The PhoPQ and PmrAB two-component systems activates the overexpression of Arn proteins and PmrC that are responsible for the addition of 4-amino-4-deoxy-l-arabinose (L-Ara4N) and phosphoethanolamine (PEtN) to the lipid A motif of LPS, respectively. PmrAB can also be activated by PhoPQ via the pmrD. The disruption of *mgrB* activates the PhoPQ system while mutations in *CrrAB* activates PmrAB but not PhoPQ. *mcr-1* is a recently identified mobile colistin resistance gene that mediates the addition of PEtN to lipid A of LPS conferring colistin resistance[21,25,27,30].

Until the identification of the *mcr-1* gene in 2015[20], almost all colistin resistance was believed to be chromosomally mediated, and plasmid mediated resistance acquisition was never reported[21,30]. The rise of the number of reports for colistin resistance may be due to the long-term use of colistin in veterinary medicine and agriculture. The continued use of colistin may lead to contamination of the environment (e.g., soil) resulting in selective pressure on microorganisms living in the environment. Colistin resistance may be subsequently mobilised into faecal bacteria of farmed animals being treated with colistin. Together with global trade and travel, these factors may finally lead to the dissemination of colistin resistance in human society[36].

Intrinsic resistance is chromosomally encoded and naturally occurring, for instance, the *Neisseria meningitidis*

EptA (enzyme lipid A PEA transferase A) gene[37,38]. The mobilized colistin resistance determinant *mcr-1* was first identified from an *E.coli* strain, SHP45, isolated from a pig in China[20]. Carried by plasmids, the *mcr-1* gene is able to mobilize to other strains, resulting in mobile colistin resistance (MCR). Soon after the first report, *mcr-1* carriage was also found in *E. coli* and *K. pneumoniae* isolates from other food producing animals (e.g., chicken), retail meat and unsurprisingly, human beings[20]. Subsequently, reports of the presence of *mcr-1* continued to emerge, either from comparison of existing sequences or by experimental characterization of new and existing strain collections from many countries, including Denmark, Germany, France, Canada and the United Kingdom, etc[39]. The *mcr-1* harbouring Gram negative pathogens have been detected over 40 countries from 5 continents[40] suggesting *mcr-1* has already disseminated worldwide [8]. Compared to the chromosomally encoded EptA, MCR-1 is more efficient in conferring colistin resistance leading to higher values of MIC (minimum inhibitory concentration) for *E.coli* expressing wild-type MCR-1[10]. MCR-2, a homologue of MCR-1, was detected in porcine and bovine colistin-resistant *E.coli* isolates in Belgium in 2016[9,41]. MCR-2 shares over 80% amino acid similarity with MCR-1, and gives significantly more robust growth in expressing *E.coli strains* compared to those expressing MCR-1[10]. MCR-1 (541 aa) and MCR-2 (538 aa) exhibit sequence identities of 34.4% and 34.5%, respectively, compared to the chromosomally encoded *Neisseria meningitidis* EptA (544 aa)[10]. In addition to *mcr-1*, more genes in the MCR gene family have been reported in *E. coli* and/or *Salmonella*, including *mcr-2 to mcr-9*[42]. Among these genes, *mcr-1* and *mcr-9* are the most disseminated[42]. However, these MCR-like enzymes show relatively low sequence identity to MCR-1 and MCR-2. Furthermore, variants of MCR-1 with genetic point mutations were also reported [43]. Although rare, recently some *mcr-1* and *mcr-2* variants have also been detected on the chromosomes of *Moraxella species* (e.g., ICR-Mo)[43,44]. The prevalence of these plasmid-borne transferable colistin determinants has potential to greatly threaten the effectiveness of colistin as a clinical therapy. Some MCR producing strains possess multi-resistance plasmids in addition to colistin resistance, making them effectively untreatable with available antibiotics[45–47]. In the absence of new antimicrobials effective against MDR Gram-negative strains, retaining the efficacy of colistin is of great therapeutic importance[48,49]. Therefore, understanding MCR-mediated colistin resistance and identifying methods of countering it is then of potential clinical importance.

1.3 MCR-1 mediated colistin resistance

MCR-1 is a membrane-associated phosphoethanolamine (PEA) transferase belonging to the alkaline phosphatase (AP) super-family[50]. Like the best studied PEA transferase in the AP family---*N. meningitidis* EptA, MCR-1 catalyses the addition of PEA from the physiological substrate phosphatidylethanolamine (PE)[10] to the 1' or 4' position of the LPS-Lipid A phosphate group (**Figure 1-3**). This modification of the LPS-lipid A reduces the net electronegative charge of the bacterial outer membrane[15,43], which consequently impairs the affinity of colistin when binding to the bacterial surface. Bacterial resistance to colistin as well as other polymyxins is therefore established. However, the catalytic mechanism of MCR-1 remains to be established.

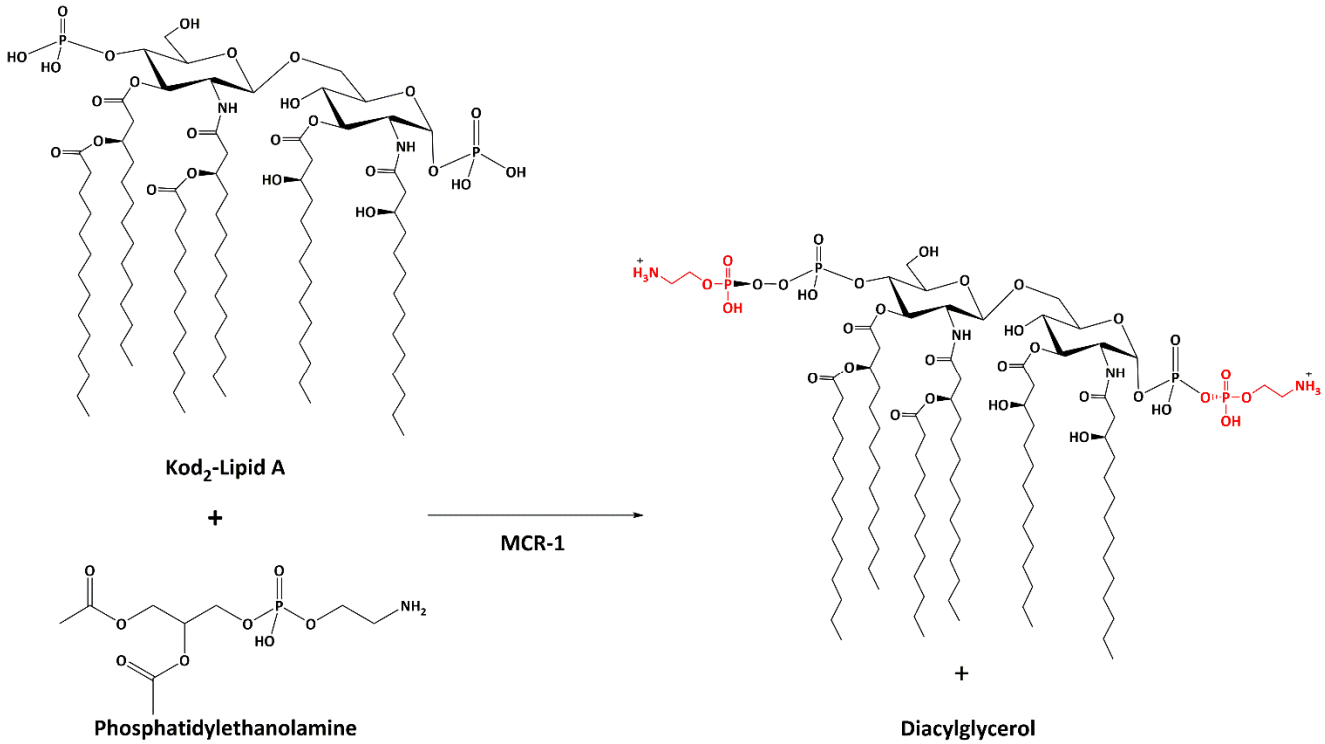


Figure 1-3. Proposed phosphoethanolamine transfer reaction catalysed by MCR-1.

Phosphoethanolamine (PEA) is transferred to the 1' or the 4' (as shown in red) positions of LPS-Lipid A phosphate group catalysed by MCR-1. (The figure cites Figure 6 in the dissertation of Roberto Pérez Chavarría[27])

1.4 Structural characterisation of MCR-1

Like other similar bacterial PEA transferases, MCR-1 is predicted, based on the protein sequence, to be an integral membrane protein [20]. The architecture of full-length MCR-1 is believed to consist of two separately folded domains: An N-terminal transmembrane (TM) domain with five putative transmembrane α -helices and a C-terminal catalytic domain facing the periplasm. Hinchliffe et al. and other research groups have expressed and purified the soluble MCR-1 catalytic domain (MCR-1^{CD}) and determined its crystal structure[15,51,52]. A conserved binding pocket including at least one zinc ion adjacent to phosphorylated Thr285 , which can also be found in EptA[28] and MCR-2[53,54], was identified in this structure, implying that MCR-1 is a zinc metalloenzyme[10]. The phosphorylation of Thr285 appears to happen in the bacterial cell during expression of MCR-1. The presence of a second zinc ion in the catalytic domain of MCR-1 is still in debate. Crystal structures of both mono-zinc and di-zinc MCR-1^{CD} have been reported by Hinchliffe *et al*[15] (**Figure 1-4**).

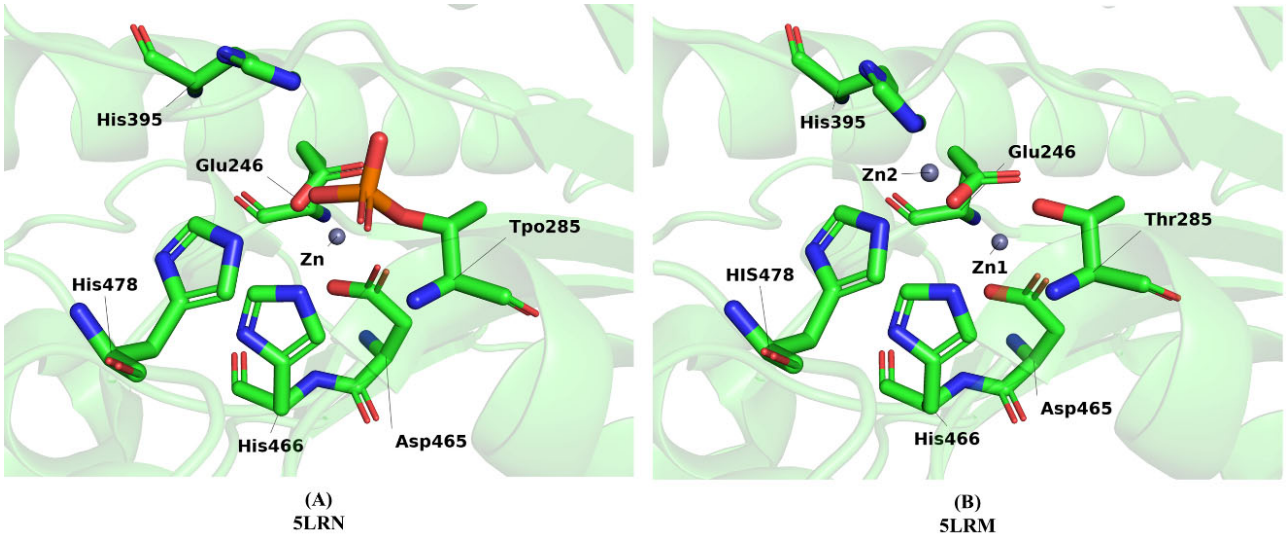


Figure 1-4. The zinc site of the MCR-1 catalytic domain.

Zinc sites in the catalytic domain of (A) mono-zinc (PDB code: 5LRN) and (B) di-zinc (PDB code: 5LRN) MCR-1. Six key residues related to MCR-1 activity are annotated (green for carbon atoms). Tpo285 represents the phosphorylated Thr285 residue. Blue represents nitrogen and red represents oxygen atoms, grey balls represent zinc ions.

Further microbiological experiments, in which the colistin MIC values of EDTA-treated MCR-1-producing *E. coli* strains were appreciably reduced from 2 mg/L to 0.25 mg/L (that of a vector-only negative control) confirmed the requirement of zinc for enzymatic activity. In the pocket, five zinc coordinated residues (Glu246, Thr285, Asp465, His466 and His478) were also identified as being critical for conferring colistin resistance. Site-directed mutagenesis was applied to create a collection of mutations at these residues to investigate their roles in MCR-1 activity[55]. MIC assays showed that none of these mutants can survive at 0.25 mg/L colistin, while *E.coli* strains carrying wild-type MCR-1 can tolerate 4 mg/L colistin. Further investigation using MALDI-TOF MS analysis demonstrated the failure of PEA transferase function of the aforementioned mutants[55]. Similar scenarios, involving loss of function, were observed for catalytic domain-deficient MCR-2 and EptA, suggesting that the five residues coordinating zinc are crucial for the enzymatic activity of both MCR-1/2 and EptA[53,54]. Unlike most characterised AP enzymes[56], it has been suggested that a single active site zinc ion may be sufficient for MCR-1 to catalyse PEA transfer to lipid A. However, comparing with other solved MCR-1 structures (e.g., PDB 5K4P[52]) and known bacterial PEA transferase structures (e.g., EptA[28] and EptC[57]), the possibility of a second zinc site (or another equivalent metal ion) cannot be ruled out. A second zinc ion is suggested to be required to assist incoming substrate to bind to the nucleophile (MCR-1 Thr285), and stabilise the resulting transition state[15]. In addition to the catalytic domain, the TM domain is also proposed to be important for full enzymatic function by playing a role in positioning MCR-1 on the periplasmic face of the cytoplasmic membrane[10]. The TM domain also likely contributes to binding of lipidated substrates (phosphaptidylPEA and lipid A) to MCR-1. However, the TM domain of MCR-1 remains poorly understood.

1.5 Inhibitors of resistance as part of combination therapies

Without obtaining new antibiotics, improving the effectiveness of currently used antibiotics is of great clinical importance. One strategy is to directly improve the safety and effectiveness of antimicrobial agents by chemical modifications, for instance, increasing the binding affinity of vancomycin against GNB[58]. Another strategy that uses a combination of an antibiotic and inhibitors of resistance is of more success in recent years and has been applied by the pharmaceutical industry[59]. In this combined therapy, the inhibitor was developed against the resistance mechanism of bacteria, thus restoring the efficacy of the antibiotic. The use of inhibitors avoided the challenge of developing novel antibiotics and re-activated currently available antimicrobial drugs that have been clinically proven to be safe and effective in treating microbial infections. A successful example is the development of β-lactamase inhibitors. β-Lactamases are enzymes that hydrolyze the β-lactam amide of the β-lactam antimicrobials (penicillins, cephalosporins, carbapenems and monobactams) thus conferring β-lactam resistance[60]. β-Lactamases are categorised as four classes (Ambler): A,B,C and D. The three classes including A, C and D are serine β-Lactamases (SBLs) and class B consists of zinc-dependent metallo-β-lactamases (MBLs). β-lactams including clavulanic acid and tazobactam act as mechanism-based inhibitors of class A β-Lactamases and are widely used in the clinic. Avibactam, based on a bicyclic scaffold, is a potent inhibitor for classes A and C SBLs[60]. The combined use of avibactam and the cephalosporin ceftazidime has been approved in clinical use for complicated infections[61]. Vaborbactam was developed based on the cyclic boronate scaffold to target SBLs of classes A and C. The therapy of vaborbactam combined with the carbapenem meropenem has been approved by the FDA (U.S. Food and Drug Administration) for complicated urinary tract infections[62]. The thiol group is known to be a zinc chelator thus on thiol-based structures are considered to be potential inhibitors for MBLs. A study repurposed approved thiol relevant drugs as inhibitors for class B MBLs including the inhibitor for angiotensin-converting enzyme (ACE), captopril, the antidiarrheal drug thiorphan and tiopronin that is normally used to treat arsenic, cadmium, and mercury poisoning[63]. These thiol-based drugs exhibit inhibition of MBLs *in vitro* with the usual dosage, even though they cannot restore the activity of imipenem[59]. To inhibit both SBLs and MBLs simultaneously, the cyclic boron-based compound taniborbactam is promising. Taniborbactam is effective against four classes of β-lactamases (broad spectrum) and is in clinical development (phase III) as a combination with the cephalosporin cefepime for use against infections caused by MDR Gram-negative strains[64].

Several strategies were investigated to overcome colistin resistance. First, new antibiotics effective against *mcr* positive strains were in development, for example, artilysi, eravacycline and plazomicin[65]. Efforts to identify bacteriophage effective against *mcr* producing strains were also made[66]. Another strategy is to develop medicines/tools to specifically reduce the expression of *mcr* gene (e.g., using CRISPR/Cas9 system to remove *mcr-1* harbouring plasmid from bacteria[67]). The combined usage of colistin and additional agents to produce synergistic effects is an option too. These agents could be antibiotics normally used for Gram-positive strains, MCR inhibitors and nature products that disturb LPS/outer membrane[65]. Currently, only a few compounds have been reported to be able to inhibit MCR-1, such as pterostilbene[68,69] and osthole[70]. This study aims to use a combination of computational (molecular docking, molecular dynamics simulations) and experimental

(biochemical and biophysical assays of protein:ligand interactions) approaches to identify possible inhibitors that could restore the activity of colistin against bacterial strains producing MCR-1 and other colistin resistance factors.

1.6 *In silico* molecular docking

Molecular docking refers to a computational simulation method used to predict the appropriate pose of a ligand molecule bound to its targeted receptor (i.e. the conformation and orientation of the ligand relative to its binding site)[71]. Starting in the 1980s, docking is now established as a standard tool in structure-based drug design[72]. Docking is capable of screening and virtually testing millions of ligands at acceptable speed without the demand of physically obtaining samples of these ligands. Only the top-scoring compounds are purchased for experimental validation[73]. It provides an alternative method for screening of ligand libraries at low cost and fast speed. Complemented with experimental techniques (e.g., validation assays), computational docking provides a cost-effective method to investigate new ligands[73], which have the potential to be active drugs. Consequently, molecular docking has become a key and widely used method in drug discovery at stages including hit identification, lead optimization, structure-based drug design and even drug metabolism analysis[74]. For instance, a selective inhibitor of HIV-1 (Human Immunodeficiency Virus 1) protease was identified through computational docking in 1990[75].

A typical computational docking process comprises two major stages: posing and scoring. Posing involves search methods used to handle ligand flexibility, and looks for the pose of a ligand (conformation and orientation) bound to its receptor in order to identify likely ligand-receptor binding configurations. Search strategies can be categorised into three basic methods:

- *Systematic Methods*

These sampling algorithms (incremental construction, conformational search, databases) are designed to treat all degrees of freedom in a molecule. However, systematic investigation eventually faces the problem of combinatorial explosion. Huge numbers of possible combinations will be generated when the search goes for many steps, resulting in increased computational time and/or resource requirement. To solve the problem, incremental construction strategies (stepwise docking process) are usually involved in these methods. In addition, using a library of pre-generated conformations is another solution[74].

- *Stochastic Methods*

These algorithms work by making random modifications to the configuration of a single ligand or a group of ligands. Newly modified ligands are evaluated based on pre-defined probability functions. The advantage of these algorithms is their fast computational speed, while the corresponding disadvantage is the lack of a full search of the conformational space[76]. Commonly used random approaches include Monte Carlo, genetic algorithms, and tabu search[74,76].

- *Simulation Methods*

Molecular dynamics and energy minimization are two major approaches used in simulation methods. Molecular dynamics is usually used for generating multiple sample conformations while energy minimization aims to optimize a single structure. In addition, molecular dynamics can also be used as an postprocessing tool in docking to refine docked ligand-protein complexes[77]. Energy minimization is often used complemented with other search algorithms such as Monte Carlo[74,78].

The scoring stage is designed to evaluate ligand poses by using a scoring function to estimate the binding affinity of poses to the receptor. By quantifying the binding interactions between ligand and receptor, the scoring function sorts out incorrect poses and ranks correct ones in order of scores (i.e., affinities). Early scoring functions evaluating ligand binding primarily rely on calculating steric and electrostatic complementarities. Nowadays, more detailed electrostatic interactions, and some more complicated treatments, have been incorporated into scoring functions including treatment of van der Waals force, rational entropic effects and hydrogen bonds between ligand and receptor[74,79,80]. Scoring functions implemented in docking can be classified into three categories:

- *Force-field-based scoring function*

A force-field here refers to a set of formulae and parameters used to describe the potential energies of intramolecular and intermolecular interactions. The values of these functions and settings is generally derived from experimental data. A basic form of a force-field contains two terms: bonded terms, including bond stretching, angle bending and bond rotation for links via covalent bonds; and non-bonded interactions usually described using electrostatic and van der Waals forces[81]. Further developments of forced-field based functions in packages such as AutoDock[82] and Gold[74] are extended to include ligand-protein hydrogen-bonding interactions.

- *Empirical-based scoring function*

Parameters for this type of scoring function are obtained (e.g., using regression analysis) from a set of experimentally determined ligand-protein binding energies, which also potentially limits the usage of empirical-based functions with systems outside the test set[74]. The basis of empirical-based functions is that binding energies can be approximated by analysis of a collection of individual unrelated terms. Four major energy terms constitute a basic empirical-based scoring function: ionic interactions, hydrophobic effects, hydrogen-bonding terms and binding entropy[83]. PLANTS[84,85] and Glide[86,87] are examples of docking applications implementing empirical-based scoring functions. In this project we used BUDE, a docking tool utilizing an empirical-based scoring function developed at the University of Bristol[88].

- *Knowledge-based scoring function*

Knowledge-based scoring functions use relatively simple atomic pair interaction potentials, derived from protein-ligand crystal structures, to evaluate the ligand-receptor complex[74]. The assumption of these approaches is that protein-ligand interactions that are statistically more common are replicated in favorable interactions in similar systems[76]. This scoring system is computationally simple, and not reliant on experimental binding data; it permits efficient screening of libraries containing a large number of compounds. An implementation based on this function type is DSX[89] derived from DrugScore[90].

1.7 Molecular interactions

Noncovalent molecular interactions are often electrostatic, between (partially) oppositely charged atoms in different molecules [91]. Although noncovalent interactions are relatively weak, and compared to covalent bonds are easily disrupted, they play a significant role in maintaining the structures and functions of large biomolecules such as proteins and nucleic acids[92]. Noncovalent interactions include ionic interactions, van der Waals forces, hydrogen bonds, and hydrophobic effects. Ionic interactions happen between oppositely charged atoms or groups (e.g., sidechains of amino acids). Van der Waals force is a non-specific transitory electrostatic interaction formed by any two closely approaching atoms, which is relatively weak. Hydrogen bonds are crucial for stabilising the three-dimensional structures of proteins. A hydrogen bond is formed when the covalent bond of a donor atom and a hydrogen atom is sufficiently polar such that the hydrogen nucleus is weakly attracted to a nearby electronegative acceptor atom. Donor atoms are often oxygen, nitrogen or sulphur, and lone electron pairs of these donors can make them serve as acceptor atoms. Directionality is a major feature of hydrogen bond interactions. In context of protein secondary structure, the mean donor-acceptor distance is close to 3.0 Å[93]. This value is helpful for seeking potential hydrogen bonds *in silico.* Hydrophobic effects describe the phenomenon that nonpolar molecules are excluded by surrounding water molecules because they cannot form hydrogen bonds. Therefore, these molecules are insoluble or almost insoluble in water, and associate together. Hydrophobic interactions are vital for protein folding and contribute to the binding to proteins of many small molecule ligands[91].

1.8 Molecular dynamics

To investigate the potential mechanism of binding of candidate inhibitors to MCR-1, molecular dynamics (MD) techniques were used. MD is a computational approach that simulates the movements of atoms using one of a range of possible algorithms. MD simulations are important tools to investigate e.g. the dynamics of protein conformations, enzymatic reactions and protein:ligand interactions[94,95]. MD simulations can also be used to further sample the poses of docked ligands and their dynamics after *in silico* molecular docking[96]. Molecular mechanics (MM) and quantum mechanics (QM) are two commonly used treatments in MD simulations. MM simulations use Newton's equations to describe the motions of molecules, while QM

treatment describes the motions of molecules by modelling the distributions of electrons[97]. The parameters of MM treatment are pre-defined and often called 'forcefields'. Amber[98–101], CHARMM[102] and GROMOS[103] are commonly used forcefields in MM MD simulations. The three types of forcefield are different in their method of parameterization but generally give similar results[104]. Compared to the QM treatment, MM simulations are often more computationally efficient but are less accurate. The computational efficiency of MM methods allows simulations with macromolecules consisting of thousands of atoms to be performed on the nanosecond or even microsecond time scale. QM calculations can usually provide more accurate descriptions than MM calculations, but with increased computational cost. Their accuracy is dependent on the level of QM theory. High-level QM calculations (e.g., *ab initio* QM) are often more accurate than calculations using lower levels of QM theory (e.g., semi-empirical QM) but also demand higher computational cost. The application of QM treatment in MD simulations is usually limited by the computational cost[105]. To combine the strengths of MM and QM approaches, a hybrid QM/MM method has been established. QM/MM has been widely used in recent years for biomolecular systems, including metal binding proteins[106–114].

In QM/MM calculations, the key region of the simulation system (e.g., the ligand binding site of a protein) is treated by QM approaches for maximise accuracy, while the rest of the system is treated by MM approaches for computational efficiency. QM/MM geometry optimization is often used to obtain the energy-minimised structure of interest of a biomolecular system (e.g., the binding conformation of an enzyme-inhibitor complex). QM/MM MD simulations can also be performed with low levels of QM theory. Apart from starting QM/MM calculations directly from crystal structures, QM/MM calculations starting from MM MD snapshots are becoming common[97]. These enable QM/MM calculations to refine the region of interest in structures obtained by MM MD, while retaining the opportunity to investigate the motions of the system (e.g., conformational changes of proteins) over larger time scales via MM MD. Several MD programs are commonly used in MD simulations, including AMBER[98–101], GROMACS[115], CHARMM[102] and NAMD[116], etc[94]. AMBER (https://ambermd.org/) is a one of the widely used MD packages with parallel calculation and GPU acceleration support[98–101]. Amber has incorporated some semi-empirical QM theories, and has an integrated interface for external QM programs to enable simulation data to be acquired at higher levels of QM theory. Gaussian is a well-known program developed for chemistry QM calculations and is one external QM program supported by Amber. The multiscale QM/MM framework ChemShell[117,118] is also a good option for QM/MM calculations and provides rich customized controls. Because Amber is user friendly, which is important for inexperienced users, this is used as the MD program in our project. Gaussian 16[119] is used as the external QM package.

1.9 Modelling zinc-containing protein:ligand complexes

In silico methods play an increasingly prominent role in ligand and drug discovery, driven in part by hardware and software innovations and by the growing availability of high-resolution crystal structures for many

biologically and/or pharmacologically important protein targets. Their application to zinc metalloproteins is then a desirable goal given the abundance of zinc-containing systems in the proteomes of many species, including humans[120–122]. Many properties of zinc, including the use of multiple ligands (Nitrogen, Oxygen and Sulphur), flexibility of zinc coordination geometry, polarization effects on the active site, and the existence of single and multi-nuclear sites, enable it to play a diverse range of roles in biological systems, but also complicate the application of computational approaches to protein zinc centers and their complexes.

To investigate protein:ligand interactions of zinc containing proteins, in particular MCR-1, MD is a useful tool. The bonded model[123], non-bonded model[124] and cationic dummy atom model[125] are three common models for modelling metal ions implemented in the AMBER MD package. The bonded model does not allow for ligand exchange or other changes in ligation, while the cationic dummy model requires a pre-defined zinc coordination geometry. Nonbonded models, as typified by the Lennard-Jones (LJ) models, that allow changes of zinc coordinating ligands, are then widely used owing to their simplicity and transferability[126]. The Lennard-Jones (LJ) 12-6 nonbonded model is a typical nonbonded model. It is widely used owing to its simple form and great transferability. The LJ12-6-4 non-bonded model was proposed and parameterized for divalent metal ions by Li and Merz in 2014[127]. A C4 term was added to the typical LJ12-6 model to represent the impact of ion-induced dipole interaction. Li and Merz claimed that the LJ12-6-4 model can simultaneously reproduce the hydration free energy (HFE), ion-oxygen distance (IOD) and coordination number, while in their tests the LJ12-6 model can only reproduce one or two these experimental values at a time[127–129]. To achieve the best description of a certain experimental term with the LJ12-6 model, users need to pick the most fitted parameter set from three: the IOD set specifically designed for reproduce ion-oxygen distance, the HFE set specifically designed for reproduce hydration free energy and the CM set designed as the general set with compromised performance. Since the LJ12-6-4 model can fulfil the performance requirement with one parameter set, it is thereby much easier to use in practice.

Following MM MD simulations, QM/MM calculations are used in our project to further improve the zinc geometries. Two levels of QM theory were adopted for the QM calculations: SCC-DFTB (self-consistent charge-density functional tight binding, abbreviated as DFTB)[130,131] and DFT (density functional theory). DFTB is a semi-empirical quantum chemical method derived from the DFT method by approximation. DFTB3 is at present the most recent version of the DFTB method and includes a complete third-order expansion of the DFT total energy around a reference density and an improved description of Coulomb potential. DFTB3 allows us to obtain reasonable accuracy at an acceptable computational cost and speed. DFT is a quantum mechanical modelling method based on Hohenburg-Kohn theorem and Kohn-Sham orbitals formulation that describes a many-body system by using functionals of electron destiny of the system. DFT calculations, e.g., with hybrid functionals such as B3LYP provide a good balance between accuracy and computational cost in describing the structures of transition metal complexes[132–134] and have been widely used for studies of zinc metalloproteins[15,135,136]. Grimme's empirical dispersion corrections with Becke-Johnson damping[137,138] which

are important for simulating proteins, are often used in DFT calculations. Detailed background and information of MD simulations performed in this project will be described in the corresponding results chapter (Chapter 4). QM/MM calculations have been applied in many zinc containing systems, such as IMP-1[139], L1 MBL[140] and HDACs[141], etc. For MCR-1, Suardiaz et al[142] have investigated the catalytic mechanism of MCR-1 by QM cluster modelling of the zinc site. Their simulations suggested that a single zinc ion is adequate for the step of transferring the phosphoethanolamine group from PEtN to Thr285 of MCR-1, but a second zinc ion (or another divalent metal ion) may be required for the transferring the phosphoethanolamine group from phosphorylated Thr285 to lipid A.

The aim of this PhD project is to seek potential inhibitors of MCR-1 and investigate interactions between MCR-1 and ligands. A robust and reliable platform that can be used to screen and evaluate various candidate inhibitors is key to achieving this aim. For example, a fluorescence-based high-throughput screen was used to identify and optimise indole carboxylates as broad-spectrum MBL inhibitors[143]. Cyclic boronates inspired by mimicking high-affinity reaction intermediates were identified as potent inhibitors of both MBLs and SBLs[144]. However, robust research tools and underlying knowledge of MCR-1 is lacking by far compared to other zinc-dependent systems such as MBLs. For instance, the chromogenic β-lactamase substrate nitrocefin[145] is available for rapid assays used to examine inhibition of β-lactamases while a robust reporter like nitrocefin is not available for MCR-1. Experimentally determined structures of the most important β-lactamases and their complexes with antibiotics/inhibitors are also available. Taking metallo β-lactamase NDM-1 as an example, the structure of NDM-1 (e.g., PDB code: 3S0Z[146]) and structures of NDM-1 complexed with either inhibitors (e.g., NDM1 bound to L-captopril, PDB code: 4EXS[147]) or hydrolysed antibiotics (e.g., NDM-1 with hydrolysed meropenem, PDB code: 4EYL[147]) are widely available. Unlike β-lactamases, limited crystal structures of MCR-1 are available and the structural understanding of MCR-1 (e.g., the number of zinc ions present) is still under discussion. In terms of antibiotic susceptibility testing, the issue of adhesion of colistin to plasticware raises difficulties for the reproducibility and accuracy of colistin susceptibility assays. Therefore, the combination of experimental and computational methods is applied to achieve our goal. We aim to use multiple source (e.g., virtual screen) to select candidate compounds and test the inhibitory effect of these compounds on MCR-1 by laboratory experiments. Molecular dynamics will be used to investigate the interactions between MCR-1 and its complexed ligand. The generality of these computational approaches by including other Zn-containing systems will also be explored in our investigations. These simulations provide useful information of the ligand binding mechanism of MCR-1 which may be applied in MCR-1 inhibitor discovery.

Chapter 2. Screening, assay development and experimental validation of potential MCR-1 inhibitors

2.1 Introduction

The prevalence of colistin resistance and other AMR mechanisms greatly threatens the availability of treatment options effective against severe bacterial infections by MDR strains[6,50]. It is of clinical significance to secure the effectiveness of antibiotics that represent the last line of defence (e.g., colistin) against multi-resistant bacteria. To ultimately overcome colistin resistance, a combined usage of inhibitors of the colistin resistance determinant (e.g., MCR-1) and colistin is an option. This chapter aims to use a combination of computational and experimental approaches to identify possible inhibitors that could restore sensitivity of MCR-1 producer strains to colistin.

Here, we first performed a virtual screening for potential MCR-1 inhibitors. We chose a mono-zinc MCR-1 catalytic domain (MCR-1^{CD}) structure (PDB: 5LRN) as our docking receptor (**Figure 2-1**) and ran a docking program named BUDE (Bristol University Docking Engine) to investigate potential MCR-1 inhibitors. BUDE is developed at the University of Bristol by Dr Richard Sessions and his colleagues using an empirical-based free energy scoring forcefield cooperating with a genetic-like Evolutionary Monte Carlo search algorithm to predict and rank binding affinity of receptor-ligand complexes[88,148]. BUDE energies approximate binding free energy in units of KJ/mol and the calculation is described as below:

$$E_{complex} = E_{steric} + E_{electrostatic} + E_{desolvation}$$

Here E_{steric} is a steric repulsion of atoms, $E_{electrostatic}$ is the electrostatic energy from charge–charge interactions and $E_{desolvation}$ is derived empirically for each amino acid from experimentally determined solvation energies.

Parameters of scoring function implemented in BUDE are derived from experimental data and have been optimized to approximate ligand poses and affinity predictions accurately[88,149]. The EMC method enables BUDE to find most favourable pose of each ligand close to the optimal via several 'evolution phases', where successive generations of pose candidates (children) are created from the best candidates from previous generations (parents). Specifically, the pose descriptor in a BUDE simulation describes the location of the ligand on the grid within the six degrees of freedom (3 translational and 3 rotational) of a -ligand relative to its receptor. BUDE utilises a notion of inheritance whereby children retain up to 2 of the positional descriptors from the parent and the other 1 to 4 descriptors are mutated at random, thereby sampling configurational space. Furthermore, BUDE is written in C++, OpenMP and OpenCL, enabling BUDE to use GPU (Graphic Processing Unit) acceleration, which provides BUDE with extremely fast calculation speed. In addition, no licence fees and expertise at hand are two other advantages, collectively making BUDE an ideal choice for

our project. ZINC is a free of charge public database and tool box aiming to provide ready access to ligands for virtual screening[150]. The majority of ligands included in the ZINC database are commercially available and the information barrier for the non-specialist has been continuously narrowed by development of various suites (e.g. ligand annotation, biology association tools, etc)[151152]. BUDE was utilized to screen the ZINC8 ligand library[151,152], which contains approximately 8 million molecules, to find potential inhibitors that interact with the MCR-1 catalytic site. The computational screening was performed on Blue Crystal Phase 4, which is one of the fastest and most advanced supercomputing facilities in the UK, capable of up to 600 trillion calculations per second[153]. A typical workflow of BUDE is shown in **Figure 2-1**. The BUDE screening of 8 million candidate compounds from the ZINC8 ligand library[151,152] was performed on Blue Crystal Phase 4 and the 28 top ranked ligands were analysed.

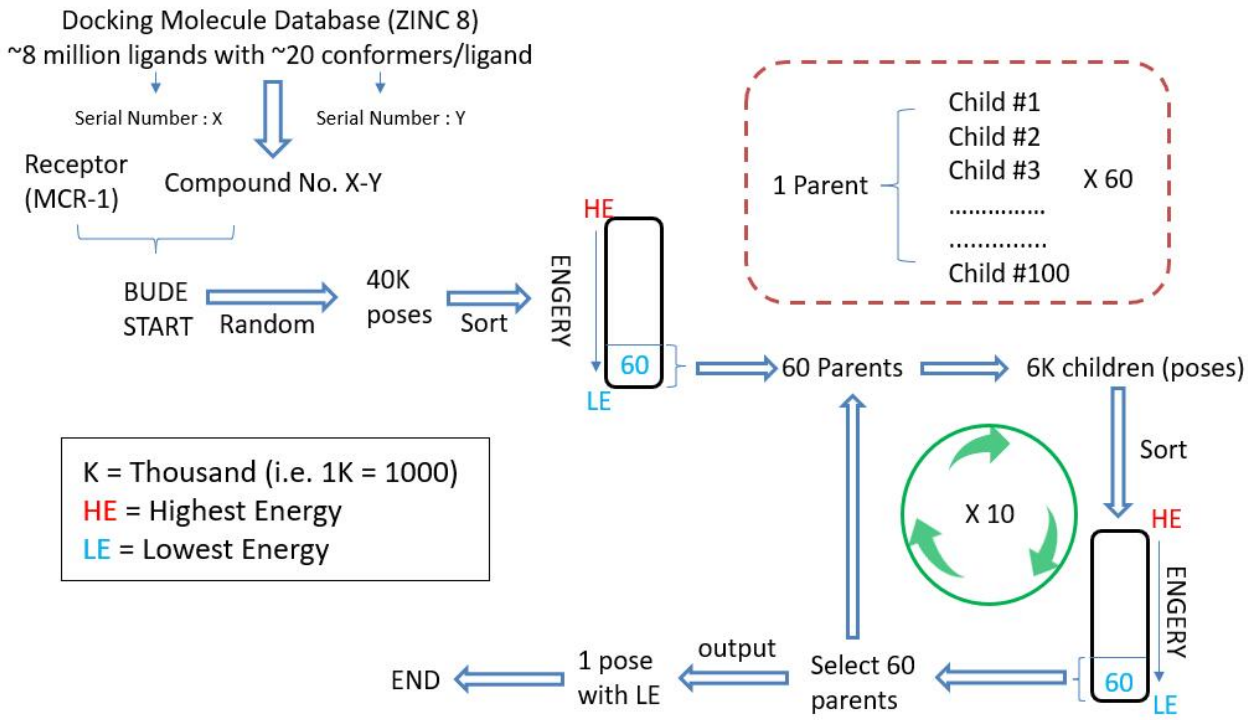


Figure 2-1. Typical BUDE workflow for ligand screening.

40 thousand calculations for generating initial random poses and 10 evolution cycles with 6 thousand calculations for creating 'children' poses per cycle, totalling 100 thousand calculations, are performed for a single ligand.

Because routine functional assays for MCR-1 activity are not currently available, two biochemical assays were established in our lab to assess MCR-1 activity: the intrinsic tryptophan fluorescence assay and chromogenic phosphoethanolamine transfer assay. In this work, these assays were applied to the purified recombinant MCR-1 catalytic domain (MCR-1^{CD}) which was produced after optimisation of bacterial growth medium for high yield expression in *E.coli*. The basic principle of the intrinsic tryptophan fluorescence assay is to utilize the intrinsic fluorescence of tryptophan residues within the MCR-1 protein, as such fluorescence is sensitive

to the conformation of the protein (i.e. the surrounding environment of tryptophan residues)[154]. The method was validated with the ligand thioglycolic acid. 4-Nitrophenyl phosphate (PNP) is a chromogenic substrate that has been widely used as molecular probe to assess alkaline phosphatase activity and in inhibitor screening[155–157]. The compound PNP is colourless but can be converted after hydrolysis to yellow 4-nitrophenolate which is detectable by absorbance at 405nm[158]. The PNP assay is generally quick and easy to carry out, hence setting up an assay based on the principle of the PNP assay (using 4-Nitrophenyl-PEtN [PNP-PEtN] as substrate instead) is an ideal choice to test the inhibitory ability of identified ligands toward MCR-1 activity. Phosphoethanolamine (PEA), also called PEtN, in its lipidated form ([phosphatidyl]phosphoethanolamine) is the substrate for the MCR-like PEA transferases which add PEA to the catalytic nucleophile in the first stage of the reaction. A preliminary experiment using the substrate-based (PNP-PEtN) chromogenic assay was successfully conducted with the ligand thioglycolic acid. Further validation assays of candidate compounds identified from virtual BUDE screening and the literature were performed to identify MCR-1^{CD} inhibitors.

In addition to the biochemical assays working on MCR-1^{CD} proteins, colistin susceptibility assays were established and used to screen potential MCR-1 inhibitors on the cellular level. The susceptibility of bacteria to colistin was evaluated by measuring minimum inhibitory concentration (MIC) values. The MIC of an antibiotic or compound is the lowest concentration of that chemical which prevents visible growth of a bacterium within the assay period. The broth microdilution method recommended by both the European Committee on Antimicrobial Susceptibility Testing (EUCAST) and the Clinical and Laboratory Standards Institute (CLSI) was adopted for the colistin MIC experiment. The colistin MIC assay was first validated with previously reported MCR-1 inhibitors[69,70,159], then used to screen our compounds. One limitation for colistin MIC assays is the loss of colistin due to the adhesion of colistin to surfaces consisting of plastic materials, which leads to poor reproducibility of MIC values between assays[160–162]. The colistin MIC assay protocol of was therefore optimized to solve this issue.

2.2 Results

2.2.1 Docking-based inhibitor screening

A crystal structure of the MCR-1 catalytic domain in the mono-zinc form (PDB code: 5LRN[15]) was chosen to generate our docking receptor (**Figure 2-2**). The zinc active site of the 5LRN structure was used. To generate the receptor pocket, a centre point (e.g., a crystallographic water molecule) was required to be defined in proximity to the zinc ion. However, none of the water molecules in the crystal structure was sufficiently close to the zinc ion. Hence, we aligned our MCR-1^{CD} structure to the previously described structure of the EptA PEtN transferase (PDB: 5FGN[38]) from *Neisseria meningitidis* to seek for a suitable water molecule that could be used as the docking centre point. Superposition of the 5FGN and 5LRN structures identified a crystallographic water molecule (HOH760A) in the EptA structure as sufficiently close to the active site zinc

ion, thereby a water molecule was placed in the equivalent position in MCR-1^{CD} and used to be the centre of the pocket. A receptor pocket including the active site zinc centre of MCR-1^{CD} was then generated by selecting all atoms within 20Å of the centre point. Modifications were made to the ZINC8 ligand library[151,152] to build the screening library: approximate 20 conformers for each ligand in the ZINC8 library of about 8 million compounds were generated and included into the screening library by our colleague Dr Richard Sessions. Each conformer was treated as an individual candidate thus in total 160 million ligands were evaluated by BUDE. In our simulation setup, the ligand was permitted to move in 1 Å increments (x: +7 to -7; y: +7 to -7; z: +7 to-7) and rotate in 10° increments (x: +170 to -170; y: +170 to -170; z: +170 to -170) within a 15 Å^3 search box centred around the water molecule as described above. Docking calculations were performed on Blue Crystal Phase 4, the high-performance computing facility of the University of Bristol. After weeks of GPU calculation, output files (raw data) from BUDE simulations were saved in 'tar.gz' format totalling 155 GB of data. The unzipping and sorting of these files took roughly a week, yielding a total of approximately 2.7 terabytes of individual results files (corresponding to the pose and energy for each conformer) and an 18 Gigabyte file listing the BUDE output filename (.mol2 format), ZINC ID, BUDE lowest energy score, heavy atom count and ligand efficiency for each conformer (total 166,752,212 conformers). To refine our data, different conformers of each original ligand were sorted by predicted binding energy as calculated by BUDE. The conformer with the lowest energy was kept as a representative of that ligand to yield a non-redundant data set (total 7,653,020 ligands). Then, the top 100,000 compounds having the most negative ligand efficiencies were selected for subsequent analysis to further reduce the size of the dataset.

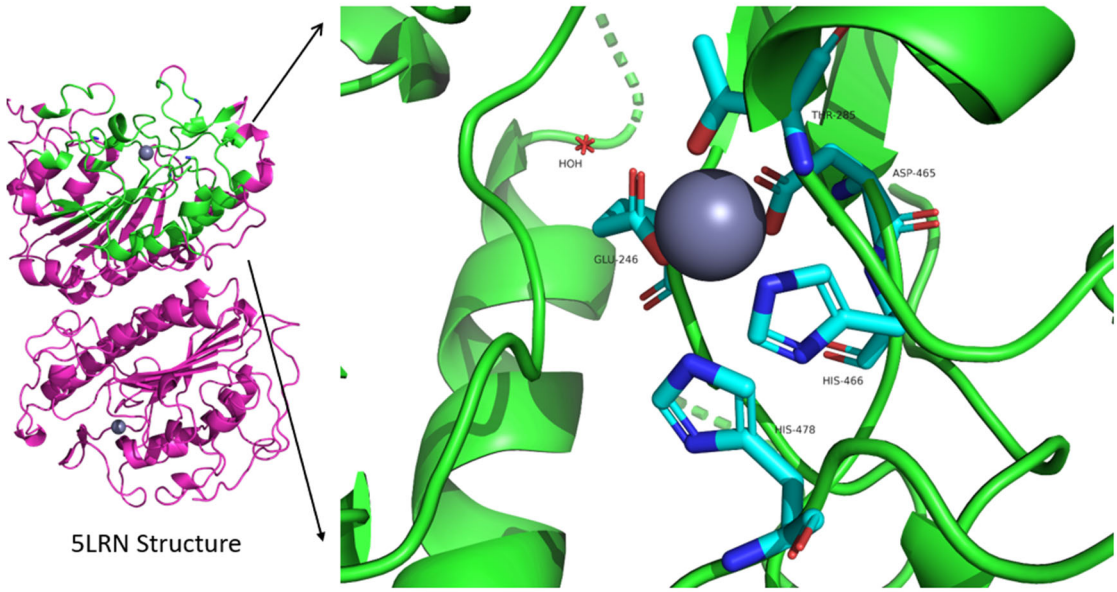


Figure 2-2. MCR-1 docking pocket used for BUDE ligand screening.

The pocket highlighted in green is a part of the catalytic domain structure of MCR-1 (PDB: 5LRN) and five key residues for MCR-1 activity are annotated (cyan for carbon atoms). Blue represents nitrogen, red represents oxygen and the grey sphere represents the zinc ion. The artificially introduced water molecule (HOH) represents the centre of the pocket and was deleted before screening.

The desired ligand-receptor binding information including ZINC ID (i.e., ligand ID), the path of mol2 file,

lowest Gibbs free energy of binding, heavy atom counts and ligand efficiency, was abstracted from raw data. Previous studies have indicated that the free energy of binding increases with number of heavy atom (nonhydrogen atoms) counts of a ligand[163]. To avoid analysis bias, a conception of ligand efficiency (ratio of Gibbs free energy to the heavy atom count (HA) of the compound) was proposed to evaluate the binding affinity fairly by measuring the energy per atom of a ligand to its receptor. The criteria of identifying a good candidate inhibitor for MCR-1 activity was based on ligand efficiency[164,165]. The abstracted data was subsequently sorted in order of ligand efficiency (i.e. the more negative the better). Analysis of the dataset suggested the best ligand efficiency for each heavy atom count gradually increases with heavy atom numbers.

Consequently, the top 30 compounds with the most favourable ligand efficiencies were small molecules, of which heavy atom numbers were under 16 (**Figure 2-3**). To ensure a diverse set of compounds in the final 'hit-list' we selected the compound with the lowest ligand efficiency for each heavy atom count. Since the heavy atom count ranged from 6 to 33 heavy atoms, 28 compounds representing the lowest ligand efficiency for each heavy atom number were picked to form our candidate pool of potential MCR-1 inhibitors (**Figure 2-4**).

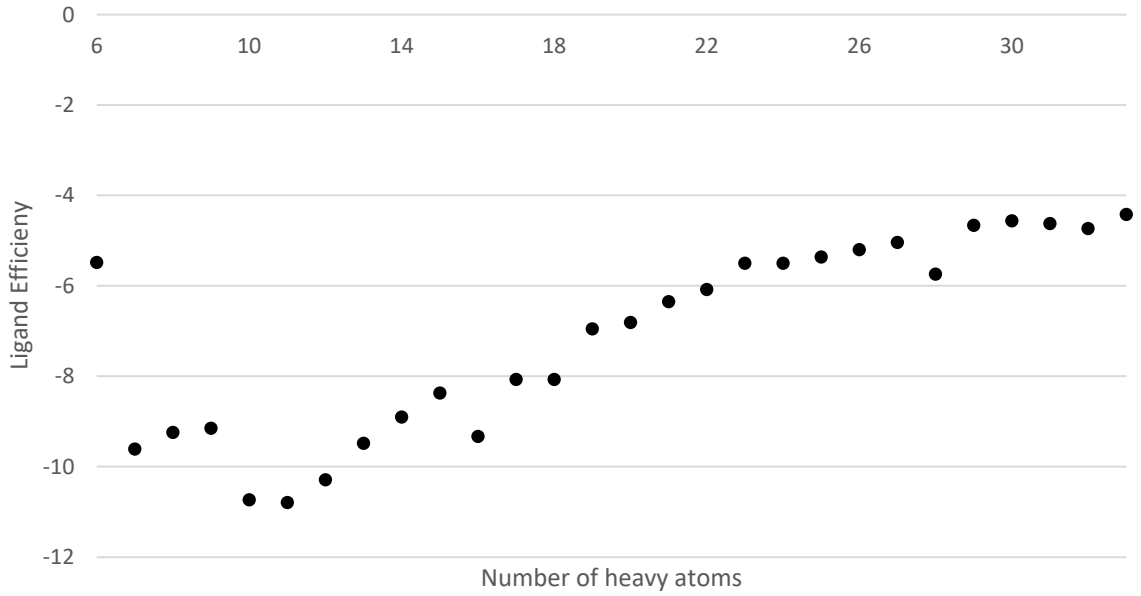


Figure 2-3. Plot of ligand efficiency against heavy atom number for top-scoring hit compounds.
The plot shows overall decrease in ligand efficiency with increasing heavy atom count.

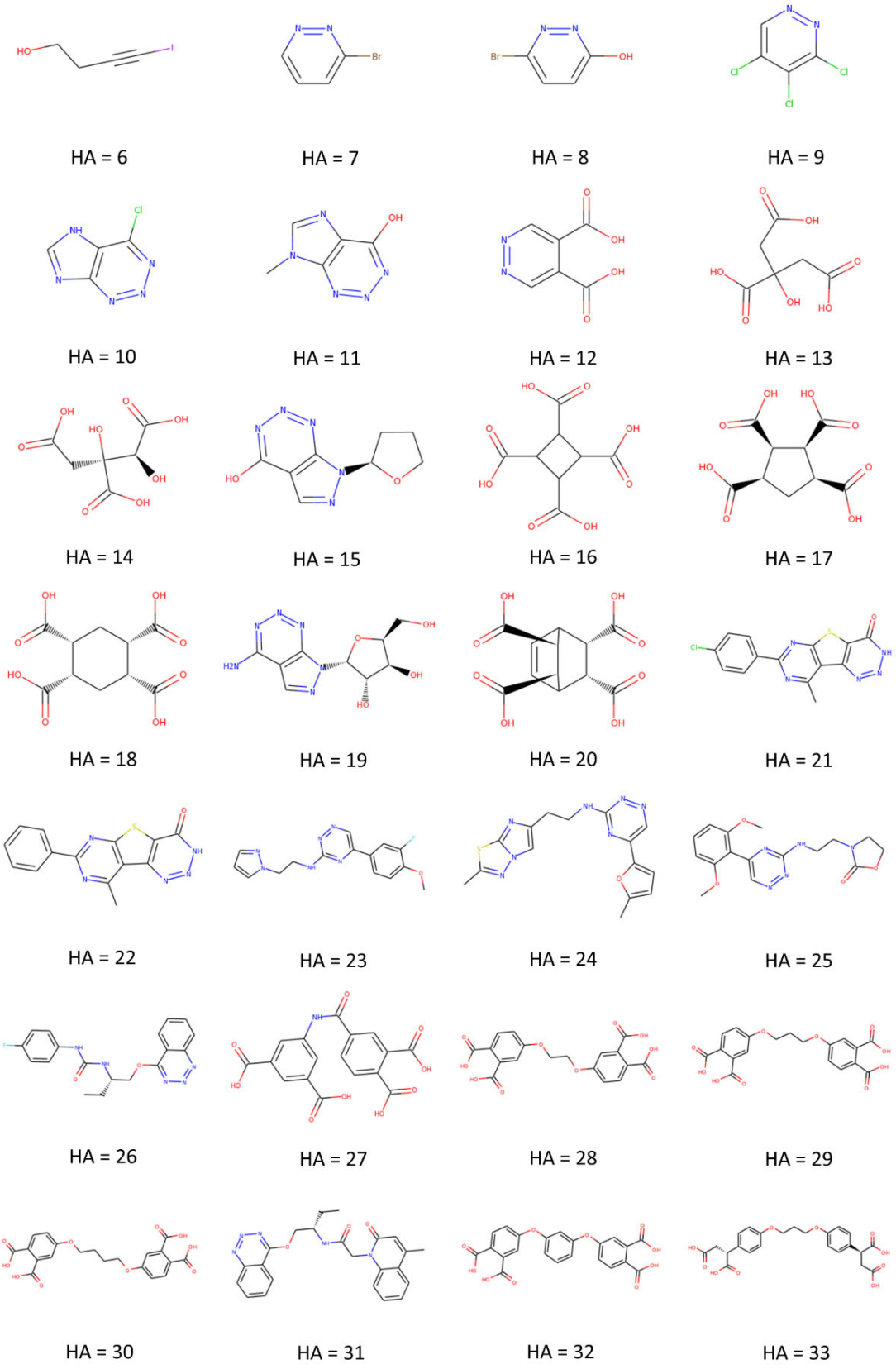


Figure 2-4. Structures of 28 selected ligands.

Structures shown are highest scoring ligands (according to BUDE ligand efficiency) with heavy atom counts between 6 and 33. Images of structures were obtained from ZINC15 database (https://zinc15.docking.org/) and are sorted in order of heavy atom number.

2.2.2 Assessment of selected ligands

Detailed information on the 28 selected compounds including solubility, approximate pH, net charges at physiological pH, numbers of hydrogen bond (H-bond) donors and acceptors and availability (for purchase) was obtained from the ZINC15 database (Table 2-1). Availability (Y) here was defined as the compound being commercially available at reasonable cost and capable of being delivered by the manufacturer within a reasonable time. In contrast, availability (N) indicates that a compound was not commercially accessible (e.g., too expensive or with too long a waiting time).

Table 2-1. Summary information for best BUDE screening hits for each heavy atom count

ZINC ID	HA	LE	logP	Net Charge	H-bond donors	H-bond acceptors	Availability Y/N
ZINC03164239	6	-5.48	0.765	0	1	1	Y
ZINC39088801	7	-9.61	1.239	0	0	2	Y
ZINC17003224	8	-9.24	0.945	0	1	2	Y
ZINC01235123	9	-9.15	2.437	0	0	2	Y
ZINC01409349	10	-10.73	0.401	0	1	4	Y
ZINC16980957	11	-10.79	-0.536	0	1	5	N
ZINC01529132	12	-10.29	-0.127	-2	0	0	Y
ZINC00895081	13	-9.48	-1.248	-3	1	7	Y
ZINC01656422	14	-8.90	-2.278	-3	2	8	Y
ZINC17176175	15	-8.37	0.236	-1	0	7	N
ZINC00161360	16	-9.33	-1.197	-4	0	8	Y
ZINC02077919	17	-8.07	-0.807	-4	0	8	Y
ZINC12343391	18	-8.07	-0.417	-4	0	8	Y
ZINC05053162	19	-6.95	-2.585	0	5	10	N
ZINC34241275	20	-6.81	-0.395	-4	0	8	Y
ZINC09484792	21	-6.35	2.298	-1	0	6	Y
ZINC09486575	22	-6.08	2.952	-1	0	6	Y
ZINC12716672	23	-5.50	1.995	0	1	7	N
ZINC11819514	24	-5.50	2.507	0	1	9	Y
ZINC12386452	25	-5.36	1.42	0	1	8	Y
ZINC21523364	26	-5.20	3.143	0	2	7	N
ZINC59367567	27	-5.04	1.732	-4	1	10	N
ZINC03010225	28	-5.74	1.937	-4	0	10	Y
ZINC03010227	29	-4.66	2.327	-4	0	10	Y
ZINC59502838	30	-4.56	2.717	-4	0	10	N
ZINC06747019	31	-4.62	2.622	0	1	7	Y
ZINC03147034	32	-4.73	4.064	-4	0	10	Y
ZINC13510463	33	-4.42	2.82	-4	0	10	Y

HA: Heavy atom counts, LE: Ligand Efficiency

The binding modes of all these compounds within the docking pocket (i.e., catalytic domain of MCR-1) generated by BUDE docking were viewed using PyMOL to identify compound orientation within the active

site and functional groups interacting with zinc. This showed that half of the identified compounds coordinate zinc via nitrogen atoms (contained within triazine or pyridazine groups), 12 of 28 ligands use carboxylate groups to interact with zinc, and 2 smaller compounds interact with zinc through hydroxyl groups. In addition, the potential of docked compounds to interact with the side chains of active site residues close to the zinc ion was also assessed. Evaluation primarily relied on the probability of forming non-covalent bonds between a ligand and its adjacent residues of the pocket. Two major non-covalent interactions were considered here: hydrogen bonds and van der Waals' forces. Since van der Waals' force is relatively weak and ubiquitous between any two atoms within favourable distances, it was omitted during assessment. Hydrogen bonds can contribute significantly to the specificity of receptor–ligand interactions, and so identifying potential hydrogen bonds between the ligand and binding site was our first aim. Given the consideration that hydrogen atoms added using PyMOL are random in orientation, it is pointless to seriously consider the hydrogen-donor-acceptor angle. Accordingly, a simple donor-acceptor distance dependent criterion was set to guide this job. A potential hydrogen bond requires:

1. A set of available donor atom, hydrogen atom and acceptors.
2. Donor-acceptor distance should be within 3.6 Å, best to be around 3.0 Å[93].

The number of potential hydrogen bonds of every candidate was counted. In addition, by alignment with the 5LRM structure (a di-zinc MCR-1 structure), these compounds were checked to see whether they occupied the position of a possible second active site zinc ion. The ligand is considered to occupy the position of the 2nd zinc ion when the minimum distance between the ligand and the Zn2 ion is shorter than 1.5 Å. The cut-off of 1.5 Å is selected based on the R_{min} value of Zn^{2+} (1.454 Å) in the LJ12-6-4 parameter set[124], which is used here to approximate the radius of Zn^{2+} coordination sphere. The binding poses of most molecules (27 of 29) do not conflict with the presence of a second zinc ion. 14 ligands are predicted to bind between the two zinc ions; the zinc binding group of these ligands is often an azole. 13 compounds are close to the position of the $Zn2^{2+}$ ion but more than 2.5 Å away from the $Zn1^{2+}$ ion, which suggests that these compounds are not predicted to interact with the $Zn1^{2+}$ ion but could interact with the $Zn2^{2+}$ ion. Only 2 ligands (HA6 and HA18) partially occupied the position of the second zinc ion. 13 of 14 ligands that are close to or partially occupied the crystal position of $Zn2^{2+}$ are predicted to coordinate the $Zn1^{2+}$ ion via carboxylate or hydroxyl groups. Representative figures of ligand poses bound to the receptor are shown in **Figure 2-5**. It is suggested that the majority of inhibitors identified could still potentially bind to MCR-1 in the presence of 2 zinc ions, especially in the case of ligands that are close to the crystallographically defined position of the second zinc. These results indicate that a second zinc (or other metal) ion could assist the ligand binding process. The aforementioned results are summarised in **Table 2-2**.

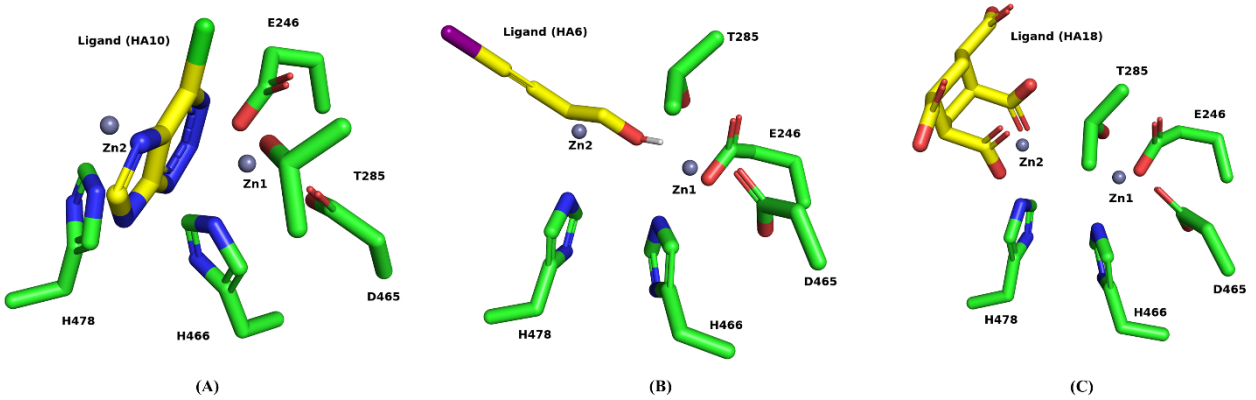


Figure 2-5. Representative positions of BUDE ligands relative to the di-zinc MCR-1 binding site.

Ligands identified by BUDE are (A) sitting between the two zinc ions (e.g., HA10) (B) fully or partially occupying the position of the 2nd zinc ion (e.g., HA6) or (C) only able to interact with the 2nd zinc ion (e.g. HA18). Green represents carbon atoms of 5 residues (Glu246, Thr285, Asp465, His466, His478) coordinating to zinc, yellow represents carbon atoms of the ligand. Blue represents nitrogen, red represents oxygen and grey balls represent the zinc ions.

Table 2-2. Position, functional groups and possible Hydrogen bonds of 28 ligands

ZINC ID	HA	LE	H-Bonds	Position	Function Groups	ZINC ID	HA	LE	H-Bonds	Position	Function Groups
ZINC03164239	6	-5.48	1	OC	-OH	ZINC34241275	20	-6.81	5	O	-COOH
ZINC39088801	7	-9.61	2	M	N	ZINC09484792	21	-6.35	4	M	N
ZINC17003224	8	-9.24	3	M	-OH	ZINC09486575	22	-6.08	4	M	N
ZINC01235123	9	-9.15	2	O	N	ZINC12716672	23	-5.50	3	M	N
ZINC01409349	10	-10.73	3	M	N	ZINC11819514	24	-5.50	3	M	N
ZINC16980957	11	-10.79	3	M	N	ZINC12386452	25	-5.36	3	M	N
ZINC01529132	12	-10.29	3	M	N	ZINC21523364	26	-5.20	3	M	N
ZINC00895081	13	-9.48	2	O	-COOH	ZINC59367567	27	-5.04	2	O	-COOH
ZINC01656422	14	-8.90	3	O	-COOH	ZINC03010225	28	-5.74	1	O	-COOH
ZINC17176175	15	-8.37	4	M	N	ZINC03010227	29	-4.66	2	O	-COOH
ZINC00161360	16	-9.33	3	O	-COOH	ZINC59502838	30	-4.56	3	O	-COOH
ZINC02077919	17	-8.07	4	O	-COOH	ZINC06747019	31	-4.62	4	M	N
ZINC12343391	18	-8.07	5	OC+O	-COOH	ZINC03147034	32	-4.73	1	O	-COOH
ZINC05053162	19	-6.95	3	M	N	ZINC13510463	33	-4.42	1	O	-COOH

HA: Heavy atom count; LE: Ligand Efficiency; OC: fully or partially occupying the position of the 2nd zinc atom; M: in the middle of two zinc ions; O: only interacting with the 2nd zinc atom.

2.2.3 Expression and purification of the MCR-1 catalytic domains

In order to investigate experimentally ligand interactions with the MCR-1 catalytic domain (MCR-1^{CD}), recombinant MCR-1^{CD} (36.1 kDa) was prepared. A recombinant pOPINF T7 plasmid[166] (kindly provided by Dr Philip Hinchliffe[15]) carrying the MCR-1 catalytic domain was transformed into competent SoluBL21(DE3) *E.coli* cells using the heat shock method. Cells were cultured with 2x YT medium then harvested and MCR-1^{CD} was purified by Ni-NTA affinity gravity column and size exclusion chromatography. The protein preparation protocol is based on the method described by Hinchliffe *et al* with minor adjustments[15]. An additional filtration step with a 0.45 μm syringe filter was introduced to remove remaining cell debris before loading into the gravity column. The protein quality and samples taken from the purification process were analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). A chromatogram showing the size exclusion chromatography step for a representative purification of MCR-1, and the associated gel images, is shown in **Figure 2-6**.

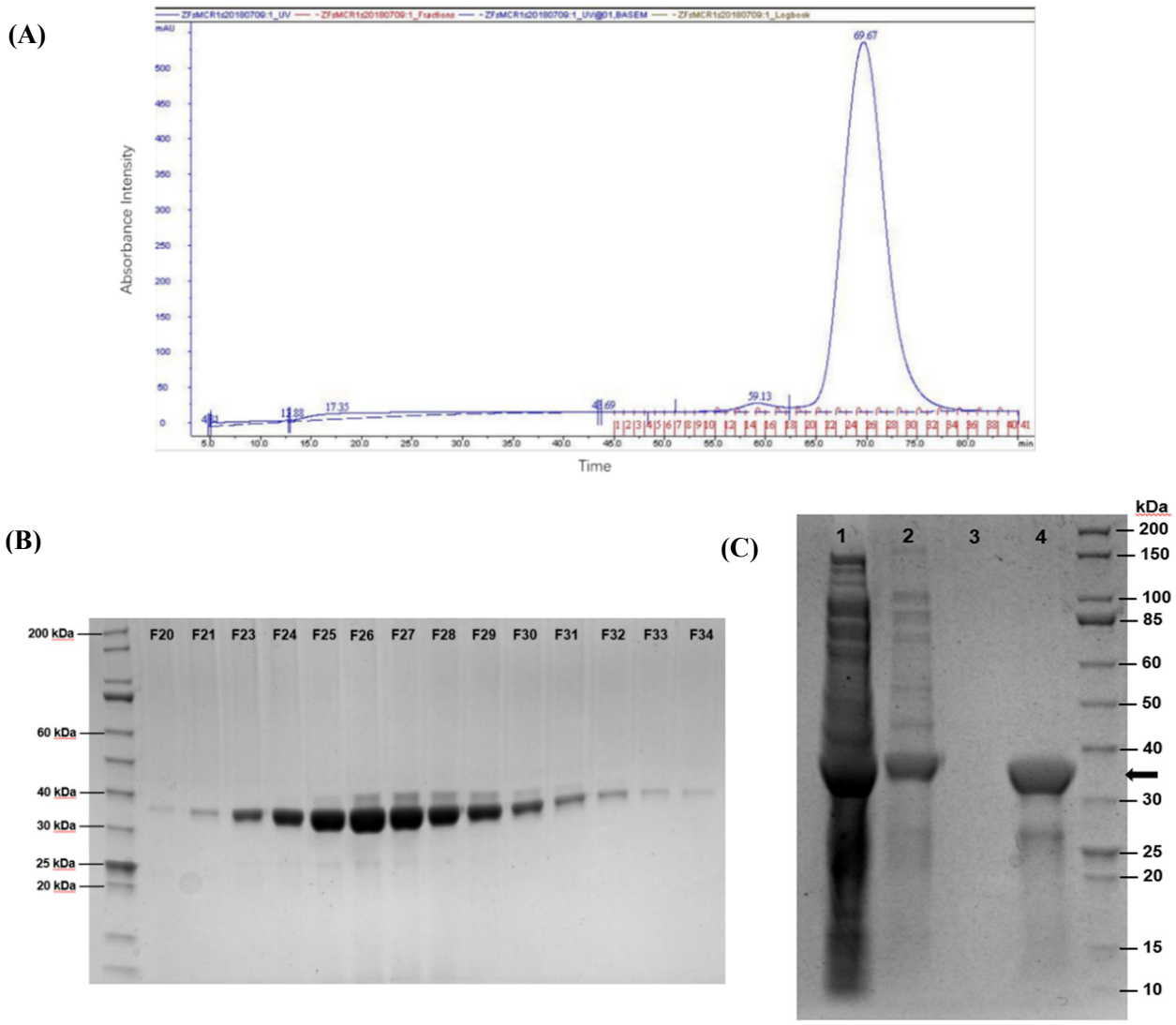


Figure 2-6. Representative purification of recombinant MCR-1 catalytic domain.

(A) Chromatogram of the size-exclusion column (Superdex 75, GE). The peak between fractions 20 and 34 (66 - 79 min) represents the elution of MCR-1^{CD} protein. (B) SDS-PAGE for fractions 20 to 34 collected from the size-exclusion chromatography. A band of the expected size of MCR-1^{CD} (36.1 kDa) is clearly present in fractions from 21 to 32. (C) Progress of the purification process as monitored by samples taken from several purification steps. Lane 1: supernatant before interaction with Ni-NTA resin. Lane 2: First wash step from Ni-NTA purification. Some loss of MCR-1^{CD} protein was detected. Lane 3: Flow-through buffer from concentration step. No residual protein was detected. Lane 4: Concentrated MCR-1^{CD} protein after purification on Ni-NTA resin. Arrow indicates the expected position of recombinant MCR-1^{CD} (36.1 kDa).

2.2.4 Preparation of candidate compounds

Candidate MCR inhibitors were purchased from different sources. Some compounds from the BUDE screen, particularly compounds with high heavy atom counts, often shared similar scaffolds and functional groups, for example compounds HA17 and HA18, compounds HA 28-33. To minimize the cost and improve efficiency, representative compounds and a few variants were ordered. In addition to the commercially available compounds identified by BUDE screening, some reported inhibitors of zinc enzymes such as histone deacetylases (HDACs), angiotensin-converting enzymes (ACEs) and Metallo-β-lactamases (MBLs) identified

from the literature were also purchased. These reported compounds include zinc binding groups such as benzamides, thiolates, hydroxamic and carboxylic acids, etc[167]. The ordered compounds are shown in **Figure 2-7** and details of the purchased compounds are shown in **Table 2-3**.

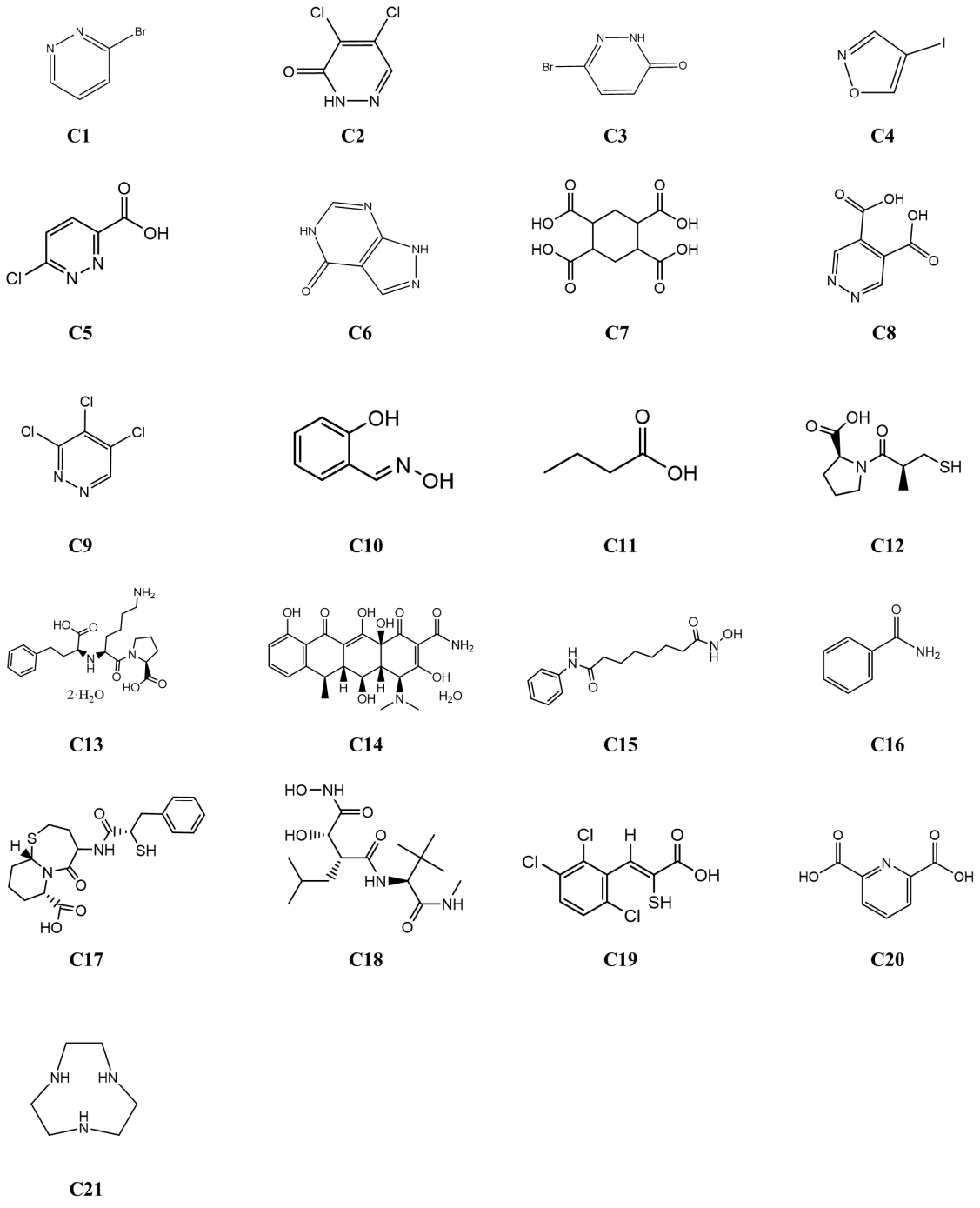


Figure 2-7. Compounds purchased according to BUDE screening and literature searches.

Compound C1-C10 were identified through BUDE screening, compounds C11-C21 were identified from the literature. Compounds from the literature include reported inhibitors of zinc metalloproteins that have similar zinc sites to MCR-1.

Table 2-3. Information of compound C1-C21

No.	MW	Solubility	CAS	Name
1	158.98	Slightly in water <10 mg/mL DMSO	88491-61-6	3-Bromopyridazine
2	164.98	DMSO: 10 mg/mL	932-22-9	4,5-Dichloro-2H-pyridazin-3-one
3	174.98	DMSO: 20 mg/mL	51355-94-3	6-bromo-3-pyridazinol
4	194.96	Slightly in water DMSO: 20 mg/mL	847490-69-1	4-Iodoisoxazole
5	158.54	DMSO: 20 mg/mL	5096-73-1	6-Chloropyridazine-3-carboxylic acid
6	136.11	DMSO: 20 mg/mL	315-30-0	1,5-Dihydro-4H-pyrazolo[3,4-d]pyrimidin-4-one
7	260.20	Water	15383-49-0	Cyclohexane-1,2,4,5-tetracarboxylic acid
8	168.11	DMSO: 10 mg/ml (heat)	59648-14-5	Pyridazine-4,5-dicarboxylic acid
9	183.42	Slightly in water DMSO: 20 mg/mL	14161-11-6	3,4,5-Trichloropyridazine
10	137.14	DMSO: 20 mg/mL	94-67-7	Salicylaldoxime[168]
11	88.11	Liquid	107-92-6	Butyric acid[167,169]
12	217.29	Water	62571-86-2	Captopril[63,170,171]
13	441.52	DMSO: 5 mg/mL Water: 10 mg/mL	83915-83-7	Lisinopril dihydrate[171,172]
14	462.45	ETOH: sparingly soluble DMSO: 20 mg/mL	17086-28-1	Doxycycline monohydrate[173]
15	264.32	DMSO: 20 mg/mL	149647-78-9	Vorinostat (SAHA)[167]
16	121.14	DMSO: 20 mg/mL	55-21-0	Benzamide[167]
17	408.53	DMSO: 20 mg/mL	167305-00-2	Omapatrilat[174]
18	331.40	DMSO: 30 mg/mL	154039-60-8	Marimastat[173,175]
19	283.56	DMSO: 20 mg/mL	NA	ML302F[176,177]
20	167.12	ETOH 25mg/mL	499-83-2	Dipicolinic Acid (2,6-Pyridinedicarboxylic acid)[159,178]
21	129.20	ETOH 50mg/mL	4730-54-5	1,4,7-Triazacyclononane (TACN)[179,180]

2.2.5 Intrinsic tryptophan fluorescence assay for inhibitor binding

Apart from the chromogenic substrate assay, intrinsic tryptophan fluorescence assays were developed to test potential ligands of cMCR-1 identified from docking. The catalytic domain of MCR-1 includes 3 tryptophan residues which could be used as an indicator of protein conformational change [181] e.g. on binding ligands. Bovine serum albumin (BSA) containing 4 tryptophan residues was used as the negative control. The hypothesis was that the ligand-protein interaction caused conformation change of MCR-1^{CD} resulting in variation of intrinsic tryptophan fluorescence. Thioglycolic acid (TGA), a simple compound with a thiol group and a carboxyl group, was identified as a candidate inhibitor according to our knowledge of the thiol part as a zinc-binding group[167,173]. The hypothesis was that the thiol group can interact with the zinc in the active site of the MCR-1 catalytic domain leading to the impairment of enzymatic function. The tryptophan fluorescence reading of BSA is not affected by the addition of TGA. A reduction of the intrinsic tryptophan fluorescence of MCR-1^{CD} was observed on the addition of thioglycolic acid which is distinguishable from the behaviour of the negative control (**Figure 2-8**). Observing fluorescence at Ex280 nm/ Em324 nm of purified MCR-1^{CD} with a set of TGA concentrations produced a binding isotherm which could be fitted (equation 1) to yield a dissociation constant (K_d) of 210 μM. This preliminary test also indicated that 50 μM is an appropriate MCR-1^{CD} concentration for this assay. The system test using thioglycolic acid demonstrated that this assay can be used to detect binding interactions between the ligand and MCR-1^{CD}.

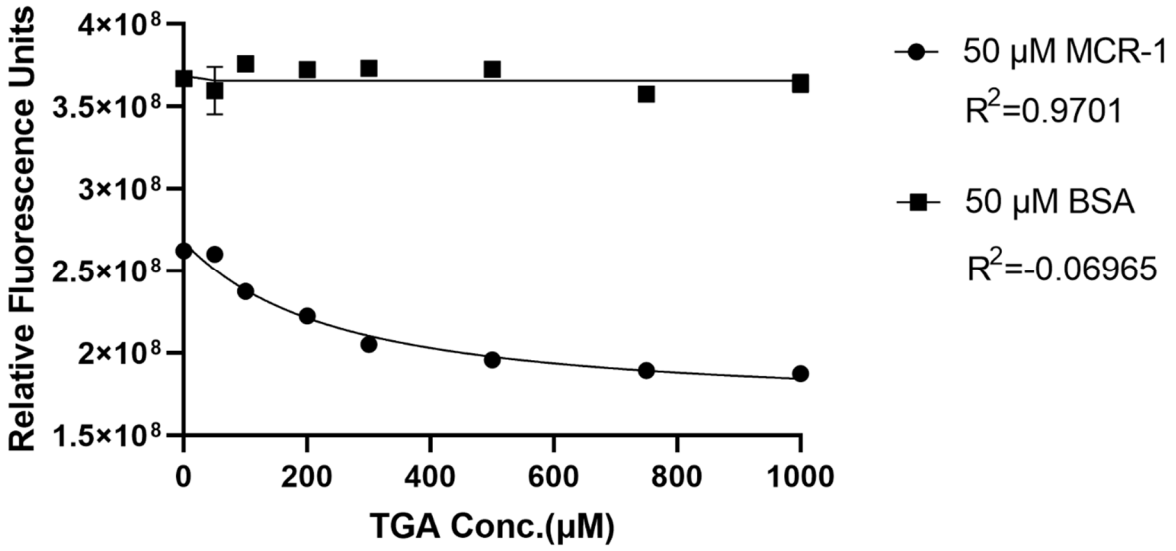


Figure 2-8. Assay validation with thioglycolic acid (TGA).

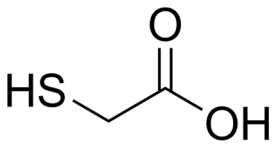
Intrinsic tryptophan fluorescence of 50 μM MCR-1 catalytic domain and BSA was plotted as a function of TGA concentration. The solid line is fitted to Equation 1 giving a dissociation constant of 210 μM for MCR-1 and no binding for BSA. Each concentration point prepared in triplicate and measured with excitation wavelength at 280 nm and emission wavelength at 324 nm. The point in the figure is the mean of three readings of each point with error bars. The R square values of fitted binding curves are shown.

2.2.6 Substrate-based chromogenic assay for inhibitor validation

4-Nitrophenyl phosphate (PNP) is a chromogenic substrate that has been previously used as a molecular probe to assess alkaline phosphatase activity and in inhibitor screening [155–157]. The compound PNP is colourless but can be converted after hydrolysis to yellow 4-nitrophenolate which is detectable by absorbance at 405nm [158]. The PNP assay is generally quick and easy to carry out, hence setting up an assay based on the principle of the PNP assay is a good choice to test the inhibitory activity of identified ligands toward MCR-1. By coupling the 4-nitrophenyl (PnP) and phosphoethanolamine (PEA, also named PEtN), 4-Nitrophenyl-PEtN (PNP-PEtN) was used as a soluble chromogenic substrate to detect the inhibitory ability of identified ligands[182]. Due to its commercial unavailability, the substrate was synthesized by the School of Chemistry, University of Bristol. The ligand, protein and its substrate were mixed and continuously incubated at room temperature (25°C). The catalytic domain of MCR-1 is expected to cleave the PEA portion from the PNP-PEtN chromogenic substrate releasing 4-nitrophenolate. The release of 4-nitrophenolate leads to the colour change resulting to an increase of absorbance at 405 nm. As described above, TGA was selected, alongside the zinc chelating agent EDTA, as a candidate inhibitor to test the chromogenic assay.

Five sets of controls were used in the assay: (1) 500 μM PNP-PEtN only; (2) 500 μM PNP-PEtN with 50 μM MCR-1 ^{CD}; (3) 500 μM PNP-PEtN with 1mM TGA; (4) 500 μM PNP-PEtN with 50 μM EDTA-treated MCR-1^{CD} (Zn ion deprived); (5) buffer only as blank. Two sets of thioglycolic acid concentrations: 1 mM and 0.1 mM were tested in reactions with 500 μM PNP-PEtN and 50 μM MCR-1^{CD}. The assay was conducted in duplicate and absorbance at 405 nm was continuously recorded every 30 seconds for 16 hours. The result is shown in **Figure 2-9B**.

(A)



(B)

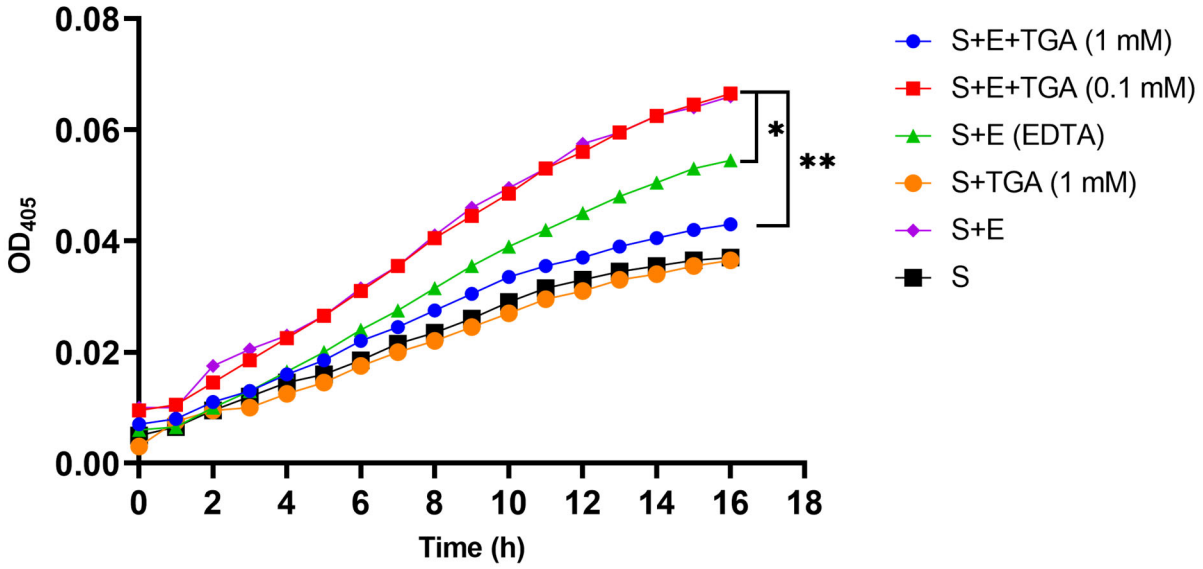


Figure 2-9. Inhibition of MCR-1 Activity by Thioglycolic Acid.

(A) The structure of thioglycolic acid. (B) Plot of absorbance (405 nm) against time for reaction of MCR-1^{CD} (E) with PNP-PEtN (S). Reduction in absorbance increase on treatment with 1 mM TGA compared to no-inhibitor control (MCR-1+ PNP-PEtN) shows a clear inhibition of MCR-1^{CD} activity caused by TGA when the ratio of TGA to MCR-1^{CD} is 20:1. Curves shown are after subtraction of buffer blank reading. The assay was conducted in at minimum two independent experiments in duplicate. T-tests were conducted using the endpoint OD_{405} values (after 16 hours) to demonstrate the difference between 1mM TGA experiment group (S+E+TGA) and the control (S+E) group ($P < 0.01$, n=2) as well as EDTA-treated group ($P < 0.05$, n=2).

The absorbance increase of the chromogenic substrate observed in the absence of enzyme is due to the slow spontaneous hydrolysis of the substrate over time. The degradation rate of the substrate is affected by the assay buffer, particularly depending on the pH of the buffer, according to the study of my colleague Dr Roberto Perez Chavarria[27]. A significant reduction in MCR-1^{CD} activity was observed with 1 mM thioglycolic acid treatment, suggesting that thioglycolic acid can inhibit MCR-1^{CD} activity. However, treatment using 0.1 mM thioglycolic acid did not show evidence of inhibition of MCR-1^{CD} activity. EDTA-treated MCR-1^{CD} that was used as a control, exhibited a detectable, but moderate, decrease in enzyme activity, suggesting that the zinc ions were not fully removed from the MCR-1 catalytic domain. A higher concentration of EDTA is likely to be required for complete removal of Zn. However, as is shown in Figure2-9B, the 16-hour endpoint absorbance of the sample 'PNP-PEtN and MCR-1^{CD}' is 0.066 which is a less than 2-fold increase over the absorbance of PNP-PEtN alone (0.037). Although the overall signal is low, the samples can still be distinguished in this

preliminary experiment, demonstrating the sensitivity and feasibility of this substrate-based chromogenic assay for validation of the activity of MCR-1 inhibitors.

2.2.7 Evaluating candidate compounds using the chromogenic assay

Although the MCR-1^{CD} protein could be purified according to the above protocol, the activity of purified proteins varied among batches. The inconsistency of protein activity and slight differences in spontaneous substrate hydrolysis rate make the inter-batch results of chromogenic substrate assays not directly comparable. To eliminate this inconsistency, the protein activity is evaluated by its enzymatic reaction rate. The equation is described as below:

$$\text{Protein activity} = \frac{\text{Absorbance}_{S+E}}{\text{Absorbance}_{S}}$$

where S is the chromogenic substrate PNP-PEtN and E represents the enzyme MCR-1^{CD}. The end-point absorbance values of the chromogenic assays were used. High-activity samples have high ratio values while low activity samples have ratio values equal to or close to 1. Lower protein activity can also make the assessment of candidate ligands harder due to the narrow enzyme activity assessment window.

Candidate compounds C1 - C12 and C14 - C21 were then evaluated against MCR-1^{CD} using the chromogenic assay (compound C13 was not delivered at the time of assay). Among these compounds, compound C20 is dipicolinic acid (DPA), one of the previously reported inhibitors for MCR-1[159]. To eliminate the inconsistency caused by the difference in enzymatic activity and substrate spontaneous hydrolysis among assays, the inhibitory effect of the compound was assessed as following:

$$\text{Inhibitory score} = \frac{\text{Absorbance}_{S+E+I}}{\text{Absorbance}_{S+I}} \times \frac{\text{Absorbance}_{S}}{\text{Absorbance}_{S+E}}$$

Here S is the chromogenic substrate PNP-PEtN, E represents the enzyme MCR-1^{CD}, and I is the tested candidate compound. The 18-hour end-point absorbance values of the chromogenic assays were used. The inhibitory score is defined as the ratio of the enzymatic activity of MCR-1^{CD} with compound and without compound.

For each compound, two conditions were tested: (1) the compound control (substrate with 1 mM compound); (2) the treatment group (1 mM compound reacted with substrate and MCR-1^{CD}). For each 96-well plate, two control groups were used: (1) the substrate control (substrate only); (2) the MCR-1^{CD} control (substrate with MCR-1^{CD}). The concentration of the substrate and MCR-1^{CD} was 500 μM and 50 μM respectively. The evaluation results of the inhibitory scores of different compounds are shown in **Figure 2-10**.

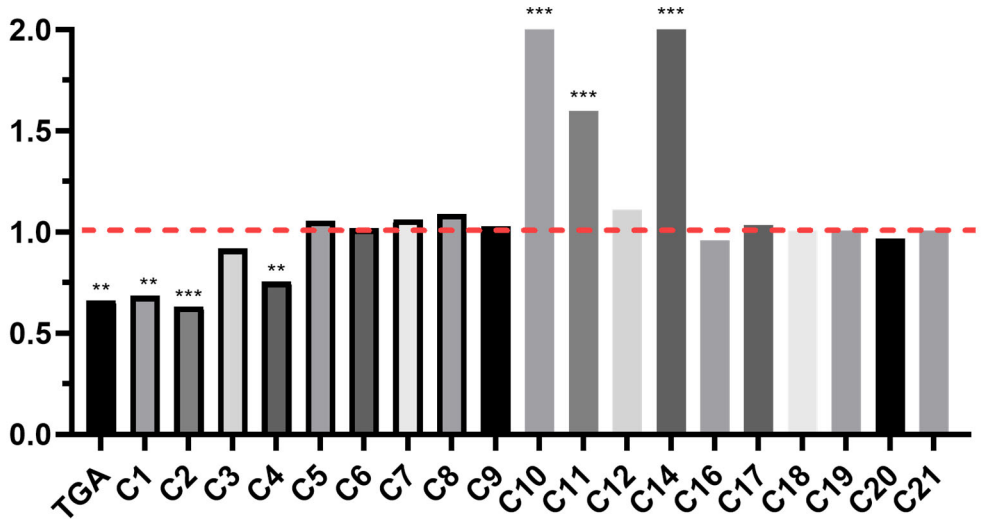


Figure 2-10. Inhibitory scores of different compounds evaluated by the PNP-PEtN chromogenic assay.

The red dashed line is equal to 1.0 representing the enzymatic activity of MCR-1^{CD} without the presence of a compound. Inhibitory scores close to 1.0 indicate that the compound did not inhibit the enzymatic activity of MCR-1^{CD}. Inhibitory scores significantly higher than 1.0 indicate that the compound may promote the enzymatic activity of MCR-1^{CD}. Inhibitory scores significantly lower than 1.0 indicate that the compound may inhibit the enzymatic activity of MCR-1^{CD}. The assay was conducted in duplicate. The aggregated results used a combination of 3 batches of MCR-1 protein, the coefficient of variation in the enzyme activity of these individual batches is 11%. T-test analyses were conducted to evaluate the effects of these scores compared to the score value 1.0 (n=2). Significance found were shown in the figure (**: $p < 0.01$, ***: $p < 0.001$).

We found that compounds C1, C2 and C4 have scores below 0.75 while C10, C11 and C14 have scores above 1.5. However, inspection of the individual absorbance curves identified that, for C1, the background signal caused by the compound is higher (>10 fold) than the hydrolysis of substrate, while the results for C2 were affected by high levels of noise in the compound control group. The absorbance of the MCR-1 treatment (S+E+I) and MCR-1 control (S+E) groups were consistent in the test suggesting C2 did not inhibit the activity of MCR-1^{CD}. Similarly, for C4, the signal increased in the presence of the compound, but no difference was found between the MCR-1^{CD} treatment group and the MCR-1^{CD} control group. The presence of compounds C11 and C12 increased the signal of the MCR-1^{CD} treatment group while the signal of the compound control group was lower than the substrate control group. It is possible that C11 and C12 may facilitate the activity of MCR-1^{CD}. In contrast, compound C14 formed a yellow-coloured solution with significant high signal, Compound C20 (DPA, a known but weak MCR-1 inhibitor) did not show clear inhibitory effects on the activity of MCR-1^{CD} in our test.

The above experiments suggest that using the chromogenic assay to screen candidate compounds may not be efficient. The PNP-PEtN assay has a high demand of active MCR-1^{CD} protein, but the activity of MCR-1^{CD} varies among batches. The substrate PNP-PEtN is not commercially available thus taking effort to synthesize. Besides, the candidate compounds may have an intrinsic absorbance signal at 405nm which may mask the

signal produced by PNP-PEtN hydrolysis (e.g., compounds C1 and C14). The presence of candidate compound may also interfere with the signal stability during the assay, even if the compound does not have background signal (e.g., compound C2, C3, C8 and C17), making it difficult to evaluate the results. The most important issue is that, while the assay appeared to confirm the ability of TGA to interact with MCR-1^{CD}, the results yielded from the PNP-PEtN assay do not necessarily represent the inhibitory activity of the tested compounds (e.g., as seen for compound C20). Altogether, under these conditions the PNP-PEtN assay using the MCR-1^{CD} may not be an efficient platform to screen candidate inhibitors of MCR-1. It is possible that, in combination with improved production methods yielding better levels of active protein, it may still be a good tool to investigate the binding of certain ligands to the enzyme.

2.2.8 Colistin susceptibility test

The substrate-based chromogenic assay and intrinsic tryptophan fluorescence measurements are biochemical assays that monitor direct interactions of ligands with, or inhibition of activity of, the MCR-1 protein. However, due to the unavailability of purified, stable, full-length MCR-1 protein, only the catalytic domain of MCR-1 (MCR-1^{CD}) can be used. As the transmembrane domain of MCR-1 may also be involved in the activity of MCR-1[9,43,55,183,18410], biochemical assays using MCR-1^{CD} may not correctly present the effects of potential inhibitors. Moreover, the two assays require relatively large amounts of purified MCR-1^{CD} protein, which makes both assays more suitable for the study of protein:ligand interactions rather than fast screening of candidate inhibitors for MCR-1. To efficiently identify potential inhibitors for MCR-1 activity in the bacterial cell, colistin MIC (Minimum inhibitory concentration) assays were performed to measure the effect of compounds on colistin susceptibility of MCR-1 producing bacteria.

The colistin MIC assay referred to the recommendation by the joint CLSI-EUCAST polymyxin breakpoint working group (2016)[185] and the broth dilution method was used. The known non-specific zinc chelator dipicolinic acid (DPA)[159], the natural products pterostilbene (PT)[69] and osthole (OST)[70] are three reported inhibitors for MCR-1 (**Figure 2-11**). The three inhibitors were used as positive controls to validate the colistin MIC assay. In previous studies by others the colistin MIC values in the presence of DPA, pterostilbene and osthole were reported to be 900, 32 and 32 mg/L for most *E. coli* strains tested. The EUCAST breakpoints for colistin (for *E. coli* and *K.pneumoniae* strains) are listed as the following: sensitive: ≤2 mg/L; resistant: >2 mg/L. Wild-type *E. coli* strain ATCC-25922 was used as the colistin-susceptible strain for quality control, for which the target colistin MIC is 0.5-1 mg/L and the acceptable range is: 0.25-2 mg/L. The colistin MIC assay was first validated using previously reported inhibitors against laboratory *mcr-1* positive BL21(DE3) *E. coli* strain and clinical *mcr-1* carrying *K.pneumoniae* strain C180. The *mcr-1* gene carried on the plasmid pET24a was delivered into the BL21 (DE3) cell using the heat shock method. The results are presented in **Table 2-4**. MIC values that were not determined are shown as ND (not determined).

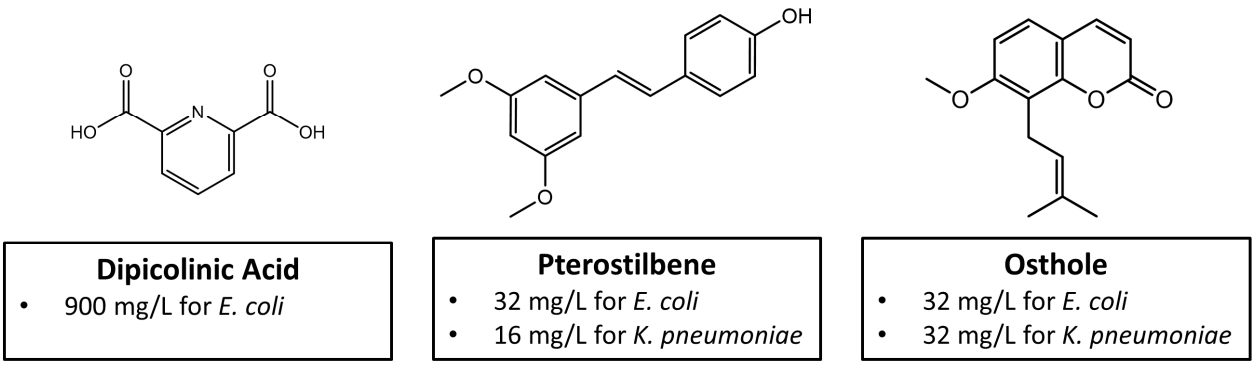


Figure 2-11. Structures of three reported inhibitors of MCR-1.

The reported concentration of these inhibitors (Dipicolinic acid[159], pterostilbene[69], osthole[70]) for reversing colistin resistance were listed in the information box.

Table 2-4. Colistin MICs against BL21 (*mcr-1*) *E. coli* and C180 *K.pneumoniae.*

	Control	Pterostilbene	Osthole	DPA
BL21 (*mcr-1*)	8	1	1	1
C180	8-16	2	ND	ND

After successfully setting up the colistin MIC assays in the lab the effects on colistin susceptibility of candidate compounds C1-C21 at 100 mg/L were tested by MIC assays. The colistin MIC assays were performed against *E. coli* BL21 (DE3) (*mcr-1*) with *K.pneumoniae* C180 included in some assays. The Results are presented in **Table 2-5**.

Table 2-5. Colistin MICs with the presence of 100 mg/L candidate compounds (C1-C21) against MCR-1 producing BL21(DE3) *E. coli* and C180 *K.pneumoniae.*

Compound	BL21 (DE3)	C180	Compound	BL21 (DE3)	C180
C1	ND	ND	C12	8	8
C2	8	4	C13	8/4	8/2
C3	8	8	C14	0-0.25	0.25-2
C4	1	1	C15	4	4
C5	16	8	C16	8	8
C6	8	8	C17	8	ND
C7	8	8	C18	8	ND
C8	8	8	C19	8	ND
C9	4	2	C20	8	ND
C10	4	2	C21	0-0.25	8/4
C11	ND	ND	Without compound	8	8-16

Note: The MIC values were determined in at minimum two independent experiments in duplicate. ND in the table suggests the MIC value is not determined.

We found that, in combination with colistin, two compounds (C4 and C14) effectively inhibited the growth of both tested strains and three compounds (C9, C10 and C21) effectively inhibited the growth of one test strain (*E. coli* BL21 (DE3) (*mcr-1*) or *K. pneumoniae* C180). The growth inhibition caused by these compounds was confirmed at minimum in two independent experiments in duplicate.

C14 (doxycycline monohydrate) is a broad-spectrum tetracycline-class antibiotic which totally inhibited the growth of the tested strains even without the presence of colistin. Compounds C9 and C10 potentiated colistin activity against the MCR-1-producing clinical *K. pneumoniae* strain C180 and the effects were clearly concentration dependent. 50 mg/mL C9 and C10 reduced colistin MICs against C180 cells from 8 mg/L to 2 mg/L making the cells colistin-susceptible, while 200 mg/L C9 and C10 further reduced the colistin MIC values to 0.5 mg/L. Compounds C9 and C10 also showed concentration-related inhibitory effects on the growth of cells when colistin was not present. The two compounds also showed inhibitory effects on the growth of MCR-1 BL21 *E.coli* cells but did not change colistin MICs against the strain. Compound C21 worked the other way around. It potentiated colistin activity against the MCR-1-producing *E.coli* BL21 (DE3) strain and exhibited strong inhibition on the growth of the strain when the concentration of C21 was equal to or more than 50 mg/L. Tested BL21 (DE3) cells did not survive exposure to more than 100 mg/L C21 even without colistin. However, compound C21 only showed minor inhibitory effects on the growth of *K. pneumoniae* C180 cells (i.e., reduction in maximal absorbance at colistin concentrations ranging from 0-2 mg/L) and did not affect the colistin MIC values against this strain (**Figure 2-12**).

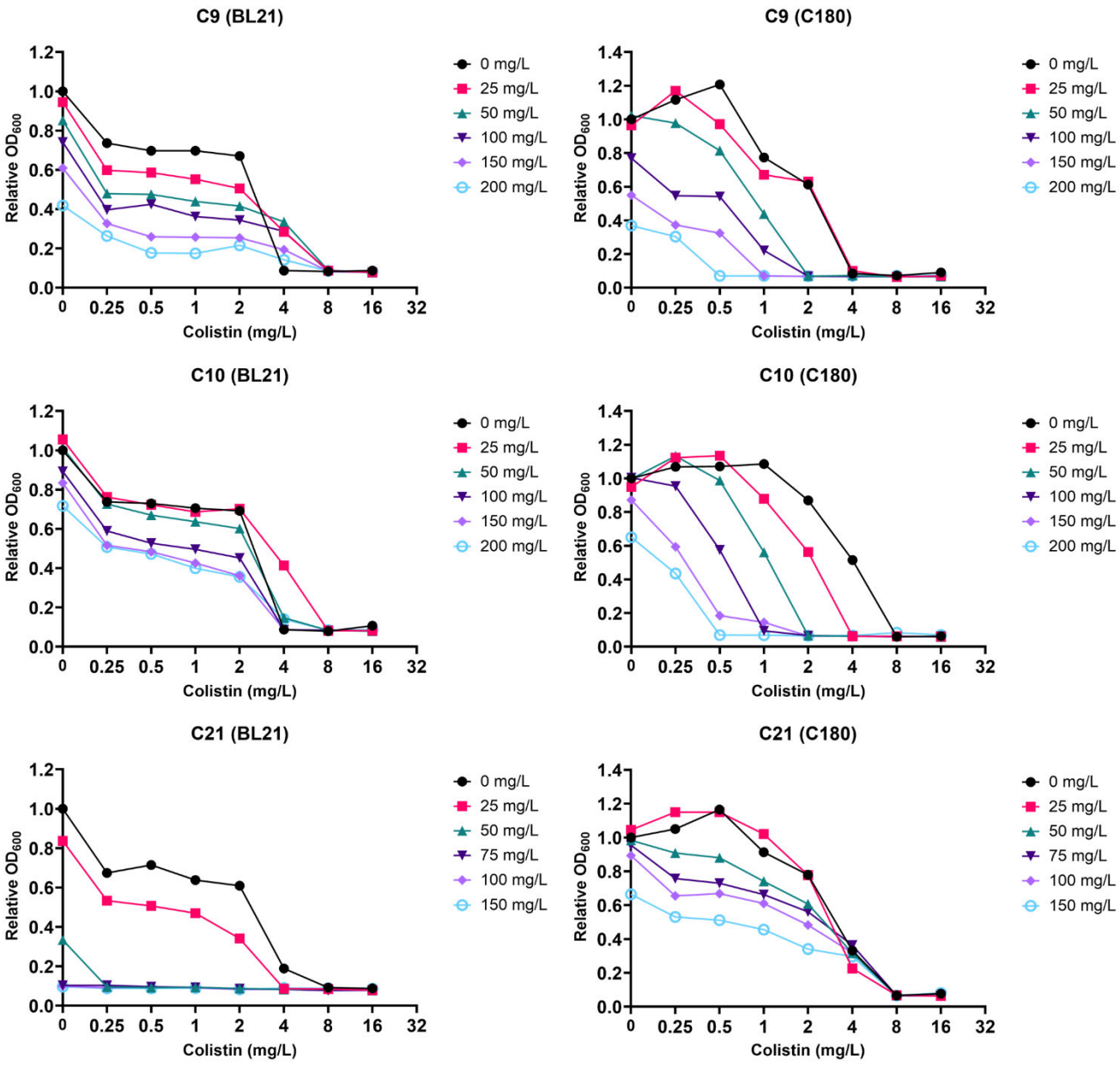


Figure 2-12. Colistin combined with various concentrations of compound C9, C10 and C21 against BL21 (DE3) (*mcr-1*) *E.coli* and MCR-1 clinical *K.pneumoniae* C180.

The OD_{600} values were obtained from the endpoint of colistin MIC assays. For each subplot, relative OD_{600} values were obtained by comparing to the OD_{600} values of the colistin only control. Cells were grown once in duplicate, the mean of each data point is presented in the plot (error bars are not shown for clarity).

Compound C4 (4-Iodoisoxazole) is related to the BUDE compound HA-7 (i.e., compound C1) and was selected additional halogen substituted heterocyclic aromatic compound that was readily commercially available. Compound C4 was identified as a potential inhibitor as it potentiated colistin activity, reducing colistin MICs against both MCR-1-producing *E.coli* and *K.pneumoniae* strains. Due to the simplicity of the molecular scaffold of compound C4, we sought to investigate whether it could be used as a scaffold for obtaining chemically modified inhibitors with higher potency. To investigate the nature of the functional group(s) of the compound C4 responsible for inhibitory effects, compounds similar to C4 (compounds C22-

C26) were purchased and tested in colistin MIC assays. In addition, because my colleague Emily Lythell serendipitously identified bound boric acid in one of her crystal structures of MCR-1^{CD}, we also screened some boron-containing compounds (C27-C32) which could mimic transient tetrahedral species as zinc ions[59,60] to look for other potential inhibitors of MCR-1. The MCR-1 inhibitory effect of compounds showing similarities to phosphoethanolamine (C33-C34), and of 4-mercaptobenzoic acid (C35) that like TGA contains both a thiol and carboxyl group were also tested. The properties (**Table 2-6**) and structures (**Figure 2-13**) of these additional compounds are presented below:

Table 2-6. Information of compound C22-C35.

No.	MW	Tested Solubility	CAS	Name
22	147.96	DMSO 20 mg/mL	97925-43-4	4-Bromoisoxazole
23	69.06	DMSO 20 mg/mL Density: 1.078	288-14-2	Isoxazole
24	83.09	DMSO 20 mg/mL Density: 1.016	6454-84-8	4-Methylisoxazole
25	208.99	DMSO 10 mg/mL	7064-38-2	4-Iodo-5-methylisoxazole
26	223.01	DMSO 20 mg/mL	10557-85-4	3,5-Dimethyl-4-iodoisoxazole
27	112.88	DMSO 20 mg/mL	1008139-25-0	Isoxazole-4-boronic acid
28	247.83	DMSO 20 mg/mL	5122-99-6	4-Iodophenylboronic acid
29	165.94	Water 20 mg/mL (heating needed)	14047-29-1	4-Carboxyphenylboronic acid
30	101.94	DMSO 20 mg/mL	84110-40-7	(2-Methylpropyl) boronic acid
31	121.93	DMSO 20 mg/mL	98-80-6	Phenylboronic acid
32	61.83	Water 20 mg/mL	10043-35-3	Boric Acid
33	141.15	Water 20 mg/mL (heating needed)	926-39-6	2-aminoethyl hydrogen sulfate
34	169.07	Water 6 mg/mL	1071-83-6	N-(Phosphonomethyl) glycine
35	154.19	DMSO 20 mg/mL	1074-36-8	4-Mercaptobenzoic acid

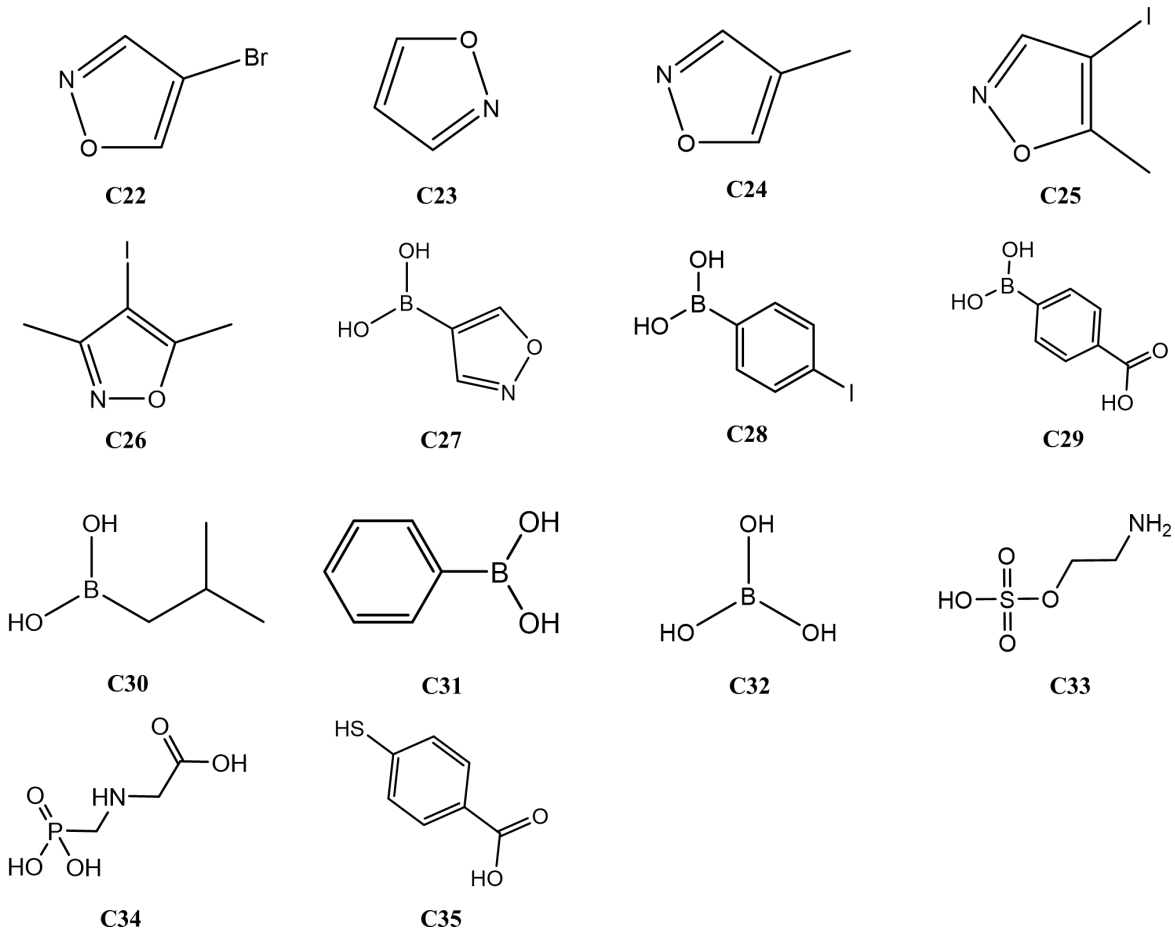


Figure 2-13. Structures of compound C22-C35.

Compound C22-C26 are C4 similarities; compound C27-C32 are boron related compounds; compound C33-C34 are similarities of phosphoethanolamine and compound C35 with a thiol and carboxyl group.

The colistin MICs for *E. coli* BL21 (DE3) (*mcr-1*) and *K. pneumoniae* C180 on addition of 100 mg/L compounds C22-C35 are shown in **Table 2-7**. None of the compounds potentiated the activity of colistin against either of the two tested strains. The ineffectiveness of compounds similar to C4 suggested that chemical modifications at any position of the isoxazole ring of compound C4 abolished the inhibitory activity of C4 against MCR-1, or the ability of C4 to penetrate the bacterial cell membrane. Replacing the iodine of C4 with bromine also abolished activity, suggesting an essential role for iodine.

Table 2-7. Colistin MICs with the presence of 100 mg/L candidate compounds (C22-C35) against MCR-1 producing BL21(DE3) *E. coli* and C180 *K.pneumoniae.*

Compound	BL21 (DE3)	C180	Compound	BL21 (DE3)	C180
C22	16	ND	C29	16	ND
C23	8	16	C30	16	ND
C24	8	16	C31	16	ND
C25	8	16	C32	8	16
C26	8	16	C33	16	ND
C27	8	16	C34	16	ND
C28	8	ND	C35	8	ND

Note: The MIC values were determined at minimum in two independent experiments in duplicate. ND in the table suggests the MIC value is not determined.

In total, 35 compounds were screened in colistin MIC assays. These compounds included 10 compounds from BUDE screening, 11 compounds identified from literature searches of inhibitors of zinc-dependent enzymes, 5 compounds with similarities to compound C4, 6 boron-containing compounds and 3 others that do not fall into these categories. One compound named C4 clearly potentiated colistin in the assays

2.3 Discussion

In this chapter, extensive docking calculations using BUDE and the isolated catalytic domain of MCR-1 have been applied to identify potential inhibitors of MCR-1. To test these compounds in the lab, we purified recombinant MCR-1^{CD} from recombinant *E.coli* BL21(DE3). Two biochemical assays respectively using the chromogenic substrate of MCR-1 PNP-PEtN and the intrinsic fluorescence of tryptophan residues of MCR-1^{CD} were developed for candidate inhibitor validation. Colistin susceptibility assays were also performed against MCR-1 expressing *E.coli* and *K.pneumoniae* cells. Compound C4 (4-iodoisoxazole) was finally identified as an inhibitor for MCR-1.

The compounds identified by BUDE normally have hydroxyl groups (individually or as part of carboxyl groups) and/or triazine or pyridazine structures. It is also worth noting that most hydrogen bonds formed in the predicted protein-ligand complexes are associated with histidine residues (His395/His466/His478) indicating that these conserved active site histidine residues may be important to stable binding. In terms of interacting with the catalytic zinc ion(s), some of the highest-scoring ligands identified by BUDE screening against the mono-zinc MCR-1^{CD} structure did not interact closely with the zinc ion (> 3.5 Å) leaving a space for the involvement of a second zinc ion. In future experiments virtual screening with BUDE may then also be performed for the di-zinc MCR-1^{CD} structure, as well for full length MCR-1. In the absence of a crystal structure screening against the full-length enzyme will require homology models as investigated by my

colleague Emily Lythell. We will then be interested to know whether the BUDE screen results for the di-zinc MCR-1^{CD} structure share common features with the screening results for the mono-zinc MCR-1^{CD} structure. It will also be interesting to know whether the inclusion of the transmembrane domain will significantly change the nature of ligands predicted by BUDE to interact with the MCR-1 active site.

One of the main tasks of this chapter was to investigate assays that can be used to experimentally detect MCR-1 inhibitors. However, a major limitation was the time spent on preparing MCR-1^{CD} proteins, for which it was hard to obtain consistent levels of activity. There are several possible reasons for this. Although the addition of extra nutrients[27] to MCR-1 cultures improved the yield nearly threefold, the condition of cells was consistently unhealthy at the time of harvest indicating reduced effectiveness of *E. coli* SoluBL21(DE3) cells for MCR-1^{CD} production. Other group members have since shown that replacement of *E. coli* SoluBL21 (DE3) cells with *E. coli* BL21(DE3) increased cell numbers and improved cell condition, suggesting that further optimisation of expression conditions might help to increase the yield of active MCR-1^{CD} (Dr. Philip Hinchliffe personal communication). In the purification process MCR-1^{CD} precipitated during the overnight cleavage of His-tag with 3C proteinase, potentially due to insufficient mixing using the rotator or use of too high a concentration of MCR-1^{CD} with His-tag. In future experiments the concentration of MCR-1^{CD} protein should be kept at around 1-2 mg/mL to avoid protein precipitation. Most importantly, the activity of purified MCR-1^{CD} was not consistent between preparations. In some cases the activity of MCR-1^{CD} was lost after going through the size exclusion chromatography. It is possible that the addition of Zn^{2+} ions to the protein solution during purification may help preserve the activity. The variable activity of MCR-1^{CD} is most likely due to differing levels of phosphorylation of Thr285[15]. Because phosphate components were not involved in the purification procedure of MCR-1^{CD}, this phosphorylation is likely to take place during protein expression by the *E. coli* cells. Phosphorylation of Thr285 during recombinant production of both MCR-1^{CD} and full-length MCR-1 has been demonstrated by multiple groups[15,27,182]. It is currently unknown whether MCR-1 phosphorylated at Thr285 can show any catalytic activity. It has also been suggested that the protein expression efficiency (e.g. post-translation modification or protein folding) of MCR-1^{CD} is not as high as for the catalytic domain of the MCR-2 isoform (MCR-2^{CD}). It has been reported that growth of recombinant MCR-2^{CD} expressing clones can be significantly more robust than that of MCR-1^{CD}-producing clones [10]. In future it may be of interest to explore a wider range of MCR isoforms to identify the system in which assays are most robust.

Limitations are also present in inhibitor validation methods. The chromogenic assay is pH sensitive. Specifically, the spontaneous hydrolysis rate of PNP-PEtN increases dramatically when the pH is beyond 7.5[27]. Considering the physiological environment required by the MCR-1^{CD}, the assay pH range needs to be carefully controlled in the range between 7.0 and 7.5. The assay time is long (more than 12.5 hours) as well so it is possible that the activity of MCR-1^{CD} may vary over time. The long assay time is needed along with high amounts of protein. Besides, the activity of MCR-1^{CD} against PNP-PEtN is low in the chromogenic assay,

which is consistent with what others have reported for assays of phosphoethanolamine transferase activity using PNP-PEtN[28,186,187]. This may reflect the absence from the assay of the transmembrane domain, which is expected to contribute to binding of the substrate, and use of the non-lipidated substrate analogue PNP-PEtN rather than the true membrane-bound substrate. In addition, assays of many compounds were made more difficult by high levels of background absorbance, which could be substantially larger than the signal change.

The intrinsic fluorescence tryptophan assay was validated with TGA but not used to test candidate compounds. One limitation was that the magnitude of the fluorescence signal change in the presence of compound (i.e. TGA) was small. Another limitation was the evidential link between the fluorescence signal change and MCR-1^{CD} inhibition was weak. The decrease of fluorescence signal may not necessarily suggest that the compound was bound to MCR-1^{CD} protein. The presence of the compound in solution may just simply quench the signal produced byMCR-1^{CD} thus reducing the signal. It is also possible that the compound may interact with MCR-1^{CD} in a non-specific fashion, or at a site other than the active site, leading to slight conformational change of the protein, and change in fluorescence, without any inhibition of the activity of MCR-1^{CD}. Equally, the intrinsic fluorescence from compounds containing e.g. a benzene or naphthalene structure may interfere with readings of tryptophan fluorescence, as the optimal excitation/emission wavelengths for these and tryptophan residues are similar. To test this we mixed compounds C5 and C7 with MCR-1^{CD} and then tested the fluorescence signal of the mixture. The results showed the fluorescence signal was reduced in the presence of either compound C5 or C7. Therefore, we considered that the intrinsic tryptophan fluorescence assay was not suitable for screening inhibitors of MCR-1^{CD}. The limitations with both the intrinsic fluorescence and chromogenic substrate assays show the need for better assays of MCR activity (e.g., using SPR as discussed in Chapter 5) to simplify the process of identifying inhibitors.

The colistin susceptibility test was currently the most efficient tool for inhibitor screening, but the assay cannot examine molecular interactions between the tested compound and MCR-1^{CD}. Although my colleague Emily Lythell has seen evidence of boric acid interacting with the zinc ion of MCR-1^{CD}, none of the tested boron-containing compounds potentiated colistin activity against *E.coli* cells. It might be because these compounds cannot get through the *E.coli* outer membrane, and are thus unable to interact with the catalytic domain of MCR-1 located in the bacterial periplasm. Experiments with the addition of a membrane permeabilizing agent (potentially Phe-Arg-β-naphthylamide [PAβN][188]) may be interesting to perform to investigate these boron related compounds. It was interesting that compounds C9, C10 and C21 were able to potentiate colistin activity against MCR-1-producing bacteria, but the effect depended on the tested strain. Differences between *E.coli* and *K. pneumoniae* in outer membrane structure may be the reason for this.

Although biochemical assays of MCR binding and inhibition proved difficult, use of MIC assays did provide a platform for testing candidate inhibitors, leading to the identification of compound C4 (4-iodoisoxazole) as a candidate inhibitor. Experiments on a limited selection of analogues showed compound C4 to be intolerant

of substitutions. It would be interesting to see if further substitutions (e.g., of 4-iodine with 4-methyl-iodine) allowed activity to be retained. In general, it would then be of interest to test a wider range of compounds, including those identified from BUDE screening, in MIC assays, particularly representative compounds with higher HA counts such as HA-25, HA-31 and HA-32. Disruptions to the experimental programme prevented this during the lifetime of this project.

In this chapter, we virtually screened the ZINC8 ligand library using BUDE to seek potential inhibitors of MCR-1. Top ranked ligands from BUDE, known inhibitors of zinc containing enzymes from the literature and potential inhibitor compounds identified based on our knowledge/experience were physically tested in the lab. The chromogenic substrate (i.e., PNP-PEtN) and intrinsic tryptophan fluorescence assays were developed to investigate the binding of ligand to MCR-1. Colistin susceptibility tests were developed and used to screen all purchased compounds. Compound C4 was identified as a potential inhibitor of MCR-1 via colistin susceptibility assays. The potential activity and the mechanism of action of compound C4 will be investigated in Chapter 3.

2.4 Methods and materials

2.4.1 BUDE simulation setup

PDB files 5LRN[15] and 5FGN[38] were downloaded from the RCSB PDB Protein Database (https://www.rcsb.org/). Manipulations of the 3D structures of 5LRN and 5FGN were performed using PyMOL (Version 2.0.6, www.pymol.org). A water molecule (HOH760) from pdb 5FGN was selected as a centre point for docking experiments by aligning structure 5LRN to 5FGN and then added to molecule 5LRN. The resulting PDB file was cleaned, only coordinate information for protein atoms, the zinc ion and HOH760 were kept. PDB files were edited using Notepad++ (Version 7.54). A pocket with radius of 20 Å centred on the water molecule HOH760 was then created in PyMoL, extracted and saved as a PDB file, and sorted on a Linux system machine with essential BUDE packages for input to BUDE. The bude_centre function was used to convert the coordinates of atoms in the pocket and the water molecule HOH760 was set as an origin (coordinates: 0,0,0). Following successful configuration, HOH760 was removed from the pocket PDB file and this file was reordered to fit BUDE input requirements using the pdbfix command. The PDB format was converted to mol2 format to be compatible with the BUDE input system. The BUDE program and ZINC8 compound library had been previously installed on the Blue Crystal Phase 4 High Performance Computing Facility of the University of Bristol Advanced Computing Research Centre (https://www.acrc.bris.ac.uk/acrc/phase4.htm).

BUDE screening utilised a modified version of the ZINC8 ligand library[6,7] constructed by Dr Richard Sessions (School of Biochemistry, University of Bristol): approximately 20 conformers were generated for each ligand in the ZINC8 library of about 8 million compounds and included in the screening library, with each conformer treated as an individual candidate ligand to give a total of 160 million ligands for evaluation by BUDE. The 5LRN pocket mol2 file was uploaded to the high-performance computer and the docking target was set by modifying the receptor configuration file of BUDE. Bash scripts (written to be compatible with the queueing system of the computational platform) were utilized to automatically submit BUDE jobs. The ZINC8 database was split into 362 sections, with each section comprising 25 serials, where a serial contains over 15000 candidates. In the BUDE docking setup, the ligand was permitted to move in 1 Å increments (x: +7 to -7; y: +7 to -7; z: +7 to-7) and rotate in 10° increments (x: +170 to -170; y: +170 to -170; z: +170 to -170) within a 15 $Å^3$ search box centred around the position of the HOH760 water molecule as described above. Formal docking using the ZINC8 database was performed after successfully running a test simulation with 4 serials. BUDE is accessible by contacting Dr Richard Sessions (R.Sessions@bristol.ac.uk) and the ZINC database and libraries can be obtained from https://zinc15.docking.org/.

2.4.2 Analysis of BUDE output

Output files from BUDE docking runs were compressed as tgz formatted files hence unzipping process was required before analysis. The location, identification, simulated free energy of binding and ligand efficiency records for each candidate were extracted from the unzipped raw data and put in the same dataset. Afterwards, conformers belonging to the same ligand were compared, with only the one with the lowest calculated free energy retained in order to represent that ligand. Free energies approximating to binding free energy in units of KJ/mol are calculated by BUDE as described below:

$$E_{complex} = E_{steric} + E_{electrostatic} + E_{desolvation}$$

where E_{steric} is a repulsion between atoms, $E_{electrostatic}$ is the electrostatic energy and $E_{desolvation}$ is derived empirically for each amino acid from experimentally determined solvation energies

The size of the dataset was 20 times smaller after this comparison. Ligands in the tidied dataset were further sorted based on calculated ligand efficiency (LE), defined as the ratio of Gibbs free energy (ΔG) to the heavy atom count (HA) of the compound:

$$LE = (\Delta G)/HA$$

and best hits for each heavy atom number were kept.

Information for these ligands including purchasability, logP value, pH range, net charge, donor and acceptor counts for hydrogen bonds, and chemical structures, was accessed from the ZINC15 web database using ZINC IDs. These ligands were then loaded into PyMOL together with the docking pocket to visualise their binding poses and assess the potential for hydrogen bonding interactions with protein sidechains. Hydrogens atoms were added to the sidechains of amino acids in the binding pocket via the automatic hydrogen addition function in PyMOL, and the distances between atoms were determined using PyMOL measurement tools. Images of ligand-protein binding were generated using PyMOL.

2.4.3 Heat shock transformation

Except where specified, all reagents used in the experiments were purchased from Sigma-Aldrich. The pOPINF T7 expression vector[189] carrying the MCR-1 (wild-type) catalytic domain[15] was kindly provided by Dr Philip Hinchliffe (University of Bristol School of Cellular and Molecular Medicine) and stored at -20 °C. To start the transformation, 0.5 µL of pOPINF plasmid was added to 25 µL SoluBL21 (DE3) competent *E. coli* cells (Genlantis). The mixture was first placed on ice for 30 minutes and then went through heat-shock at 42 °C for exactly 45 seconds. The cells were then quickly transferred onto ice for 5 minutes cooling. After the addition of 250 µL S.O.C medium, the cells were incubated for 1 hour in the shaking incubator (37 °C, 180 rpm). Afterwards, the liquid culture was centrifuged at 4000 rpm for 60 seconds (VWR Micro Star 17R). 250 µL supernatant was removed and the pellet was resuspended in the remaining supernatant. Subsequently, the liquid culture was spread on an LB agar plate with 50 µg/ml ampicillin and incubated at 37 °C (180 rpm)

overnight. A single colony on the selection plate was picked for overnight LB broth culture at the same condition above.

2.4.4 Growth of MCR-1^{CD} producing *E. coli.*

The overnight broth culture was transferred into a 2l flask containing 2X YT medium with 50 mg/L ampicillin (5 ml culture per 500 ml medium). The flasks were then placed in the shaking incubator (37 °C, 180 rpm) and samples were taken every 20 minutes after 3 hours incubation to measure cell density via absorbance at 600 nm (OD_{600}, Amersham Biosciences Ultrospec 2100 pro) until the reading reached 0.6. IPTG (isopropyl-β-D-thioga-lactopyranoside) was subsequently added into each flask to a final concentration of 500 μM. Following an 18-hour incubation at 18 °C (180 rpm), the cell culture was centrifuged at 4°C, 6500 g for 10 minutes (Beckman Coulter Avanti J-26XP). The supernatant was then discarded, and cells were harvested into 50 mL Falcon Tubes and frozen at -80 °C until needed.

2.4.5 Protein purification of MCR-1 catalytic domain

The cells were resuspended in 50mM HEPES (pH 8), 500 mM NaCl, 2 mM β-mercaptoethanol, 1x EDTA-free protease inhibitor tablet (Roche), homogenized manually and broken by twice passing through the cell cracker (Constant Systems T5 Model) at 25 kpsi. The lysate was spun down at 38000 rpm in Ti70 rotor (Beckman Coulter Optima L-80 XP Ultracentrifuge) for 1 hour at 4 °C to remove cell debris. Subsequently, the supernatant was passed through a 0.45 μm syringe filter and 10 mM imidazole and Ni-NTA Agarose (QIAGEN, 1 mL for 1 L prep) were then added to the filtered supernatant. After a 2-hour incubation at 4 °C with constant rotation, the supernatant along with suspended Ni-NTA resin was loaded into a gravity flow column and excess liquid allowed to elute. The Ni-NTA resin in the column was first washed in 50mM HEPES (pH 7.5), 400 mM NaCl, 10 mM Imidazole, 1 mM TCEP, 0.1 mM $ZnCl_2$. Then the resin beads were sequentially washed in 50mM HEPES (pH 7.5), 400 mM NaCl, 10 mM Imidazole, 1 mM TCEP, 0.1 mM ZnCl2, 0.1% TX-100 and 50 mM HEPES (pH 7.5), 200 mM NaCl, 15 mM Imidazole, 1 mM TCEP, 0.1 mM $ZnCl_2$. After washing, the bound MCR-1 catalytic domain (MCR-1^{CD}) protein was eluted in 50 mM HEPES (pH 7.5), 200 mM NaCl, 400 mM Imidazole, 1 mM TCEP. The protein solution was concentrated using 10K MWCO concentrators (Sartorius Vivaspin 20) and diluted in 50 mM HEPES (pH 7.5), 150 mM NaCl, 1 mM TCEP until the concentration of Imidazole was less than 10 mM. After that, recombinant 6His-tagged 3C proteinase (prepared and stored at -80 °C by colleagues in the lab) was added and the solution was rotated at 4°C for 16 hours to cut off the his-tags on the MCR-1^{CD} protein. The Ni-NTA resin was then added to the protein solution for a 30-minute incubation at 4°C to capture cleaved His tags from the solution. After flowing through the gravity column, MCR-1^{CD} protein (catalytic domain) was collected and subsequently loaded onto a Superdex 75 size-exclusion chromatography column (GE Healthcare) mounted on an AKTA

chromatography system (GE Healthcare) equilibrated in 50 mM HEPES (pH 7.5), 150 mM NaCl, 1 mM TCEP, 0.1 mM $ZnCl_2$. Peak fractions corresponding to the MCR-1 protein from the size-exclusion chromatography were collected. After concentration, purified MCR-1^{CD} protein was stored at 4 °C. Protein concentration was determined by measuring absorbance at 280 nm using a Lambda 35 spectrophotometer (Perkin Elmer) and subsequently checked using a NanoDrop Lite spectrophotometer (Thermo Fisher Scientific). SDS-PAGE was used to analyse the purity of the final products as well as samples from purification steps and fractions. Precast nUView Tris-Glycine gels (NuSep) were used in conjunction with a 10 - 200 kDa unstained protein standard (New England Biolab) used as a protein ladder. After gel staining with instant blue (Expedeon), images were taken with the gel doc and imaging systems (Syngene).

2.4.6 Substrate-based PNP-PEtN chromogenic assay

Assay validation

Zn^{2+} ions in the purified MCR-1 catalytic domain and buffer were removed by serial concentration and dilution followed by an overnight dialysis (GeBAflex-Mini Dialysis Tubes 8K MWCO) in assay buffer (50 mM HEPES, 180 mM NaCl, pH 7) with addition of 1 mM EDTA. The Zn-depleted MCR-1^{CD} protein was used as a control. All reagents were dissolved in the assay buffer and five sets of controls used in the assay: (1) substrate only; (2) substrate with 50 μM MCR-1^{CD}; (3) substrate with 1mM inhibitor; (4) substrate with EDTA-treated MCR-1^{CD}; (5) buffer-only blank. Two concentrations of ligand were tested: 1mM and 100 μM. Each concentration of ligand was reacted with 500 μM substrate and 50 μM MCR-1^{CD}. Samples were gently mixed and equilibrated for 1 min prior to loading. After loading samples into 96-well microplates (Corning Costar 3595) with duplication (110 μL per well), the plate was incubated at 25 °C inside the SpectraMax iD5 microplate reader (Molecular Devices) recording the absorbance at 405 nm of each well every 30 seconds for a total of 16 hours.

Compound screening

100 mM stock solution of each compound was first prepared in DMSO or water to according to the solubility of the compound. The stock solution was then mixed with other substances in the assay buffer (50 mM HEPES, 180 mM NaCl, pH 7) to the final concentration of 1 mM. For each compound, two conditions were tested: (1) the compound control (substrate with compound); (2) the treatment group (compound reacted with substrate and MCR-1^{CD}). For each 96-well plate, two control groups were used: (1) the substrate control (substrate only); (2) the MCR-1^{CD} control (substrate with MCR-1^{CD}). The concentration of the substrate and MCR-1^{CD} was 500 μM and 50 μM respectively. DMSO was added into the two control groups at final concentration of 1%. Samples were gently mixed and equilibrated for 1 min prior to loading. After loading samples into the 96-well microplate (Corning Costar 3595) with duplication (110 μL per well) the plate was incubated at 25°C inside

the CLARIOstar microplate reader (BMG LABTECH) recording the absorbance at 405 nm of each well every 45 seconds for a total of 12.5 hours.

2.4.7 Intrinsic tryptophan fluorescence assay

To selectively excite tryptophan, the excitation and emission spectra were first scanned at 1 nm intervals using SpectraMax iD5 microplate reader (Molecular Devices). The excitation wavelength was set at 280 nm and emission wavelength was set at 324 nm. 50 μM MCR-1^{CD} protein or BSA were mixed with a range of concentrations of TGA (50 μM, 100 μM, 200 μM, 300 μM, 500 μM,750 μM and 1 mM) in the assay buffer (50 mM HEPES, 150 mM NaCl, 1 mM TCEP). Samples were gently mixed and equilibrated for 5 min in 1.5 mL microcentrifuge tubes prior to transferring into a 96-well microplate (Corning Costar 3595), the tryptophan fluorescence intensity was read by the microplate reader at Ex280 nm/Em324 nm. Data were corrected for dilution and fitted to the tight binding quadratic equation[190]:

$$\mathbf{F} = \boldsymbol{F_{ini}} - \Delta\boldsymbol{F} * \frac{([\boldsymbol{E}]+[\boldsymbol{L}]+\boldsymbol{K_d})-\sqrt{([\boldsymbol{E}]+[\boldsymbol{L}]+\boldsymbol{K_d})^2-4*[\boldsymbol{E}]*[\boldsymbol{L}]}}{2*[\boldsymbol{E}]} \qquad \text{(Equation 1)}$$

Where and F, F_{ini} and ΔF denote fluorescence, maximal fluorescence and maximal change in fluorescence respectively; [E] is the concentration of protein (i.e. cMCR-1 or BSA), [L] is the concentration of ligand (i.e. TGA) and K_d is the dissociation constant.

2.4.8 Colistin susceptibility testing

The colistin susceptibility testing method followed ECUAST guidelines for antimicrobial susceptibility testing (https://www.eucast.org/ast_of_bacteria/). A colistin concentration series (Sigma-Aldrich) was prepared in cation-adjusted Mueller-Hinton (MH) broth. Broth or PBS suspension of the test organisms from an overnight agar plate or liquid culture was made and adjusted to OD_{600} 0.08-0.1 (equivalent to 0.5 McFarland standard). This suspension should contain approximately 1 to 2 x 10^8 CFU/ml. Then the suspension was diluted 1:100 with MH broth. The colistin solution and bacterial suspension were then assembled in the 96-well plate. For each 96-well plate, one row of antimicrobial dilutions was set as a negative control with no bacterial inoculation to ensure that observed growth was not due to contamination. 75 μl of each colistin dilution was transferred into the 96 well plate. For wells containing bacteria, 75 μl of the 1:100 diluted bacterial suspension was added into each well. This yielded a final bacterial concentration of approximately 5 x 10^5 CFU/ml. For the control row with no bacterial inoculation, 75 μl of MH broth was added. The 96-well plate was then incubated at 37 °C for 16 - 20 h and the OD_{600} of each well in the plate was read by a plate reader. For assays involving MCR-1 inhibitors, the inhibitors/compounds were added in advance to the MH broth at a concentration of twice the final concentration.

Chapter 3. Microbiological investigation of interactions of compound C4 with *E.coli* and *Klebsiella pneumoniae*

3.1 Introduction

The compound C4 (4-iodoisoxazole) was identified as an MCR-1 inhibitor in via the experimental verification in Chapter 2. In this chapter, we aimed to further investigate compound C4 and try to understand its mechanism of action. We first evaluated the compound C4 by comparing with three known inhibitors of MCR-1: pterostilbene (PT)[69], osthole (OST)[70] and dipicolinic acid (DPA)[159] in terms of effectiveness, solubility and toxicity. Extensive colistin susceptibility assays, also known as minimum inhibitory concentration (MIC) assays, were performed in the presence of compound C4 to verify its effectiveness on MCR-1-producing clinical isolates and ArnT-mediated colistin-resistant *K. pneumoniae* strains. Unlike *Escherichia coli,* where colistin resistance is caused by the transferable *mcr* gene, colistin resistance in *Klebsiella pneumoniae* is mainly caused by chromosomal mutations which often increase the expression of the Arn protein, particularly ArnT, and/or the PEtN transferase PmrC (**Figure 1-2, Chapter 1**). The expression of the ArnT protein is mediated by the *arn* (also known as *pbg* or *pmrF*) operon that can be independently activated by the PhoPQ and PmrAB pathways[32]. The produced zinc-containing ArnT protein modifies lipid A in lipopolysaccharide by adding 4-amino-4-deoxy-L-arabinose to the core disaccharide of lipid A at the 1' and/or 4' position, with the modification of phosphate groups increasing the net charge of the cell envelope and causing colistin resistance[191–193]. Activation of PhoPQ can be triggered by loss of expression of MgrB, a membrane-bound regulator of PhoPQ expression[194]. The activation of the PmrAB pathway also leads to the upregulation of Arn protein expression, thus conferring colistin resistance, but disruption of *pmrA* only slightly reduced Arn protein production in *mgrB*-silenced mutants. We tested two colistin resistant *K.pneumoniae* strains, P23 and P23 (*pmrA*), that have been reported in a recent publication[32]. P23 is the *mgrB* gene deactivated mutant of the parent Ecl8[195] *K.pneumoniae* strain, and P23 (*pmrA*) is the *pmrA* gene-disrupted derivative of the P23 strain. Ideally an inhibitor of colistin resistance will act not just upon MCR-producing *E. coli* but also on other species and potentially against other resistance mechanisms. Therefore, the potency of these inhibitors of colistin resistance (compound C4 and comparators) was tested against the ArnT-producing P23 and P23 (*pmrA*) *K.pneumoniae* strains.

The mechanism of action of these candidate MCR-1 inhibitors is not fully understood, one possibility is that they may increase the permeability of the outer membrane. The effect of compound C4 and pterostilbene on the permeability of the cell membrane of *E.coli* was then investigated. 1-N-phenylnapthylamin (NPN) is a fluorescent probe that has been widely used to investigate the permeability of the outer membrane of *E. coli*[58,196–199]. NPN is a small molecule that exhibits strong fluorescence when it binds to the phospholipid layer of the cell membrane. The outer membrane of Gram-negative bacteria is a permeability barrier that prevents

the easy uptake of hydrophobic molecules and large molecules such as the glycopeptide antibiotic vancomycin. As NPN is hydrophobic, it normally cannot cross the outer membrane efficiently. Once the outer membrane has become more permeable, for instance through being damaged, NPN can cross the barrier and bind to phospholipids inside the cell, resulting in significantly increased fluorescence. Therefore, NPN can be used an indicator of permeabilization of the outer membrane. In the presence of NPN, cells with an intact outer membrane show weak fluorescence, while cells with a compromised outer membrane show enhanced fluorescence. Of note, NPN is a probe of outer, but not inner, membrane permeability.

Propidium iodide (PI) is a fluorescent stain for DNA that is normally used to evaluate cell viability in flow cytometry[200]. It exhibits strong fluorescence when it binds to DNA. Similarly to NPN, PI is normally excluded by cells with intact cell membranes. Compromised outer and inner membranes allow the efficient uptake of PI enabling it to bind to cytoplasmic DNA. Therefore, PI was used as a probe to assess the permeability of both the outer and inner cell membranes[58,201]. In addition to the assays using small fluorescent probes, MICs of the glycopeptide antibiotic vancomycin were determined and used to evaluate the effects of compounds on cell membrane integrity. Vancomycin effectively inhibits the growth of Gram-positive bacteria by impeding cell-wall synthesis[202]. However, due to the presence of the outer membrane of Gram-negative cells, the large vancomycin scaffold (~1450 Daltons) cannot easily cross the outer barrier and is thus usually not effective against Gram-negative strains[203]. Vancomycin has previously been used as an indicator in cell membrane permeabilization studies of *E.coli*[203,204] and *P. aeruginosa*[205].

Phe-Arg-β-naphthylamide (PAβN, also previously called MC-207,110) is a known efflux pump inhibitor that was reported to potentiate the activity of fluoroquinolone antibiotics against multidrug resistant *Pseudomonas aeruginosa* and some *Enterobacterales* species[206–208]. A study in 2011 showed that more than 4 mg/L PAβN permeabilized the outer membrane of wild-type *E. coli* MG1655 allowing accelerated nitrocefin hydrolysis[188]. Nitrocefin is a chromogenic β-lactam can be efficiently hydrolysed by periplasmic β-lactamases, and enters the cell and is hydrolysed more readily when the Gram-negative outer membrane is compromised. PAβN, at a concentration of more than 10 mg/L, was also reported to potentiate the activity of β-lactam antibiotics against *P. aeruginosa* and to permeabilize the *P. aeruginosa* outer membrane[205]. Considering the ability of PAβN to permeabilize the outer membrane of Gram-negative strains, we aimed to investigate how PAβN acts in combination with colistin/vancomycin on colistin resistant *E.coli* NCTC-13846 (MCR-1 producing) in comparison with compound C4. NPN uptake assays, PI uptake assays, vancomycin MIC assays and growth curves were performed to investigate the effects of compound C4 and pterostilbene on the cell membrane of colistin resistant *E.coli* cells. PAβN was included in the study to extend the comparison. The results indicated that these compounds did not affect the cell membrane integrity or permeability.

In addition to reversing colistin resistance, pterostilbene was reported to restore activity of the carbapenem (β-lactam) antibiotic meropenem against New Delhi metallo-β-lactamase (NDM-1) producing *E.coli* and

K.pneumoniae isolates[209]. This suggests that pterostilbene may have multiple activities on other zinc-dependent enzymes in addition to MCR-1. As compound C4 shares similar inhibitory effects to pterostilbene upon MCR-1 producing bacteria, the meropenem susceptibility of strains producing the metallo-β-lactamases NDM-1 and IMP-1, and of strains producing the serine-β-lactamase KPC-2, was also assessed in the presence of compound C4. By assessing compound C4 and its comparator pterostilbene, we found that these compounds inhibited colistin resistance caused by MCR-1and ArnT, and meropenem resistance caused by some, but not all, metallo-β-lactamases. Although a definitive mechanism of action for these compounds was not uncovered, our data indicate that both compound C4 and pterostilbene are capable of reversing colistin resistance in MCR-producing *E. coli*, including both laboratory and clinical strains, but that both have activities upon the bacterial cell that extend beyond inhibition of MCR-1.

3.2 Results

3.2.1 Protocol optimization for colistin MIC experiments

After screening candidate compounds against MCR-1 producing *E. coli* using colistin MIC assays as described in Chapter 2, we found that colistin activity was not always consistent between assays. We considered that this was due to the uncertainty of colistin loss in the experiment through e.g. binding to the plastic surfaces[160]. This unstable colistin activity often resulted in 2-fold differences in colistin MIC values in replicate experiments. Although a 2-fold difference is generally considered acceptable in MIC assays, it may cause ambiguity when MIC values are around the colistin breakpoint (2 mg/L) and bring about difficulties for repeated measurements. As a result, colistin-sensitive strains may exhibit MIC values representative of colistin resistance in some experiments.

To improve the reproducibility of the colistin MIC experiment, we optimized the protocol. The assay medium remained as cation-adjusted Mueller-Hinton broth (CAMHB) which is recommended by the EUCAST/CLSI protocol[210,211]. A colistin-resistant (*mcr-1* positive) reference *E.coli* strain NCTC-13846 was introduced. The strain is recommended by EUCAST's guidance for colistin broth microdilution testing[212]. The NCTC-13846 strain has a colistin MIC target value of 4 mg/L. The colistin MIC value of NCTC-13846 strain should usually be 4mg/L and only occasionally 2 or 8 mg/L. Addition of the NCTC-13846 strain was used to ensure the colistin activity at 4 mg/L. The combined use of NCTC-13846 with the quality control (QC) colistin-sensitive *E.coli* ATCC-25922 strain (for which colistin MIC values should mostly be 0.5-1 mg/L, but may occasionally be 0.25mg/L or 2 mg/L) should ensure that the activity of colistin is consistent at concentrations between 1 - 4 mg/L. In addition, the brand and type of all consumables was fixed so that the loss of colistin in individual assays could be consistent.

The incremental dilution method was used to first dilute colistin from stock solution to the target concentration (i.e., two times the highest concentration of colistin serial used in the MIC assay). The number of steps and

containers contacting colistin were minimized because colistin may be lost quicker with increasing numbers of dilution steps[160]. The initial concentration of colistin was finally adjusted referring to the quantitative analysis by Karvanen *et al*[160] that also showed that colistin was also lost from solution over time. In that work, most colistin loss took place in the first 6 hours, with 25 to 80 % of the expected concentrations left after 24 hours. In their tests, the lower the concentration of colistin, the higher percentage of colistin loss. Around 20 % to 30 % of colistin may be lost for concentrations of 2 mg/L and 1 mg/L colistin contacting polystyrene material (e.g. MIC assay plates) over 24 hours. For 0.5 mg/L colistin, only half of the initial colistin concentration was left after 24 hours. To compensate for this inevitable loss of colistin, we increased the initial concentration of colistin based on the target concentration. For example, to meet a goal of 2 mg/L colistin in the assay, 2.6 mg/L colistin was prepared. Colistin MIC assays were performed with different initial concentrations of colistin against colistin-resistant NCTC-13846 and colistin-sensitive ATCC-25922 *E.coli* strains to examine whether the expected colistin MICs for both strains could be stably achieved (4 mg/L for NCTC-13846, 0.5-1 mg/L for ATCC-25922). After confirming this to be true, the amplification factor was set to be 1.3. The increased initial concentration of colistin also compensates for the loss of colistin during serial dilutions, ensuring that the concentration of colistin in the MIC experiment hits the target value.

Experiments using the optimized colistin MIC protocol yielded more consistent results between different assays. The optimized protocol enabled us to test compounds at low colistin concentrations (e.g., 0.5 mg/L) with confidence, which was particularly important when testing colistin-sensitive strains. The optimized MIC assay was then used to investigate the mechanism of action of compound C4 in the remaining parts of this Chapter.

3.2.2 Comparing compound C4 with known MCR-1 inhibitors

To investigate the activity of compound C4 (4-iodoisoxazole) against that of previously reported MCR-1 inhibitors, the compound C4 was compared with the reported inhibitors pterostilbene (PT)[69], osthole (OST) and dipicolinic acid (DPA) using checkerboard MIC assays. The effects of the four compounds on colistin activity against laboratory-made *mcr-1*-positive *E.coli* BL21 (DE3) cells were evaluated using checkerboard MIC assays. The same assays (except DPA) were also performed against the colistin-resistant EUCAST reference *E.coli* NCTC-13846 strain. The results are presented in **Figure 3-1.**

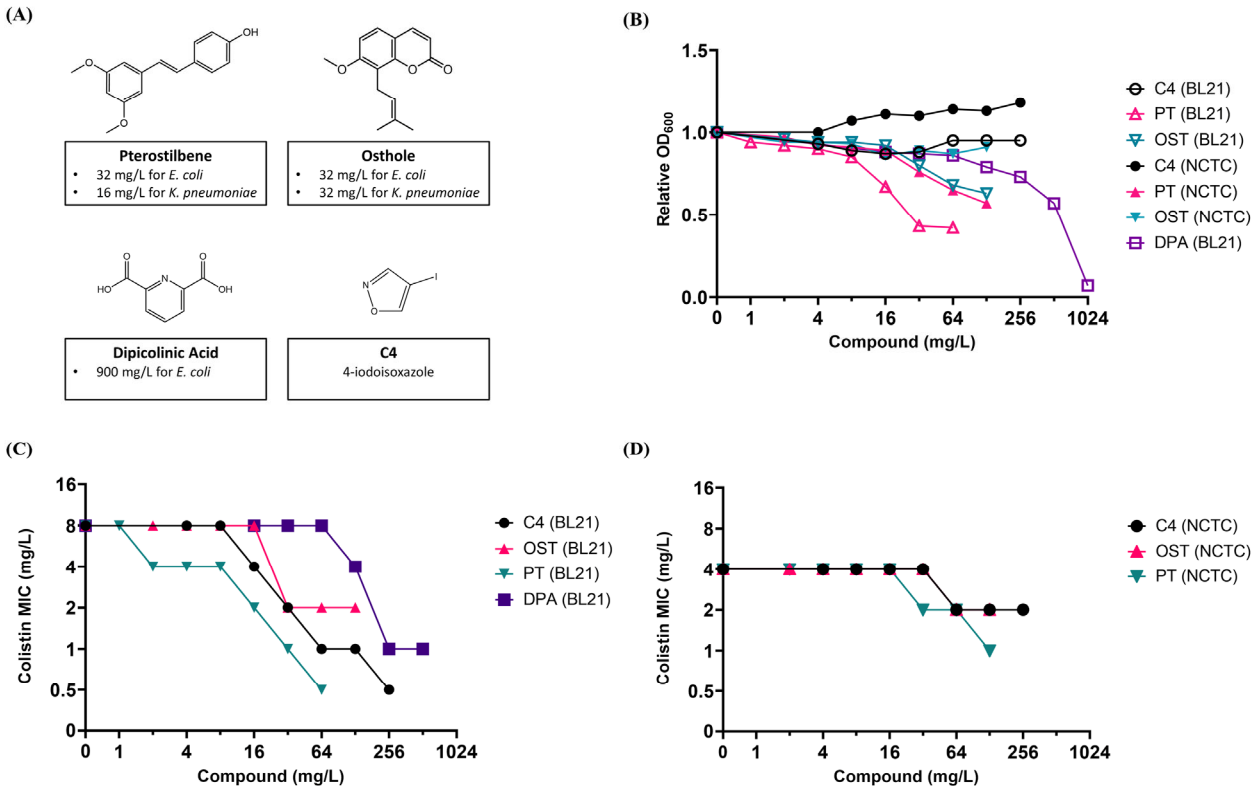


Figure 3-1. Comparing C4 with three known MCR-1 inhibitors via checkerboard MIC assays.

(A) Structures of DPA, PT, OST and C4. (B) The effects of C4, PT, OST, DPA at different concentrations on the growth of MCR-1 *E.coli* BL21 (DE3) and NCTC-13846 cells without colistin after 18 hours incubation. (C) Effects of C4, PT, OST and DPA on colistin activity against BL21 cells. (D) Effects of C4, PT, OST on colistin activity against NCTC-13846 cells. Cells were grown in triplicate, the mean of each data point is presented in the plot (error bars are not shown for clarity). The checkerboard assay was independently done twice.

The colistin breakpoints for *E.coli* and *K.pneumoniae* are 2 mg/L (sensitive: ≤ 2 mg/L, resistant: > 2 mg/L) according to the EUCAST and CLSI breakpoint tables[210,211]. To reduce the colistin MIC against MCR-1 producing *E. coli* BL21 (DE3) to 2 mg/L, 32 mg/L C4, 16 mg/L pterostilbene, 32 mg/L osthole or 256 mg/L DPA were required. To decrease the colistin MIC of *E. coli* NCTC-13846 to 2 mg/L, 64 mg/L C4, 32 mg/L pterostilbene or 64 mg/L osthole were required. There is a two-fold difference in the concentration of each compound that was required to sufficiently potentiate colistin activity against BL21 (DE3) and NCTC-13846 cells to restore susceptibility. The difference indicated that these agents each affect *E.coli* BL21 (DE3) more readily than the NCTC-13846 strain. The potency of the compounds in potentiating colistin activity was evaluated according to the lowest concentration of each compound that is required to reverse colistin resistance. Pterostilbene has the strongest effect among the four compounds as the required concentration of pterostilbene for reversing colistin resistance was the lowest. Compound C4 and osthole have similar performance in reversing colistin resistance, while DPA was the least effective.

Absorbance measurements after 18 hours of bacterial growth in the absence of colistin (**Figure 2-1B**) suggested that ≥16 mg/L pterostilbene inhibited the growth of both BL21 (DE3) and NCTC-13846 strains. Osthole at a concentration of ≥ 32 mg/L inhibited the growth of MCR-1-expressing *E. coli* BL21 (DE3), but did not inhibit that of the NCTC-13846 strain. 1024 mg/L DPA totally prevented the growth of BL21 cells leading to no cell growth at the end of the experiment. However, compound C4 (up to 256 mg/L) showed no inhibition on the growth of cells indicating that it possesses no intrinsic antibacterial activity. Compound C4 even slightly promoted the growth of NCTC-13846 cells. To further investigate the effects of C4 on the growth of *E.coli* cells, 1024 mg/L C4 was used to treat ATCC-25922, NCTC-13846 and MCR-1 producing BL21 (DE3) (**Figure 3-2**).

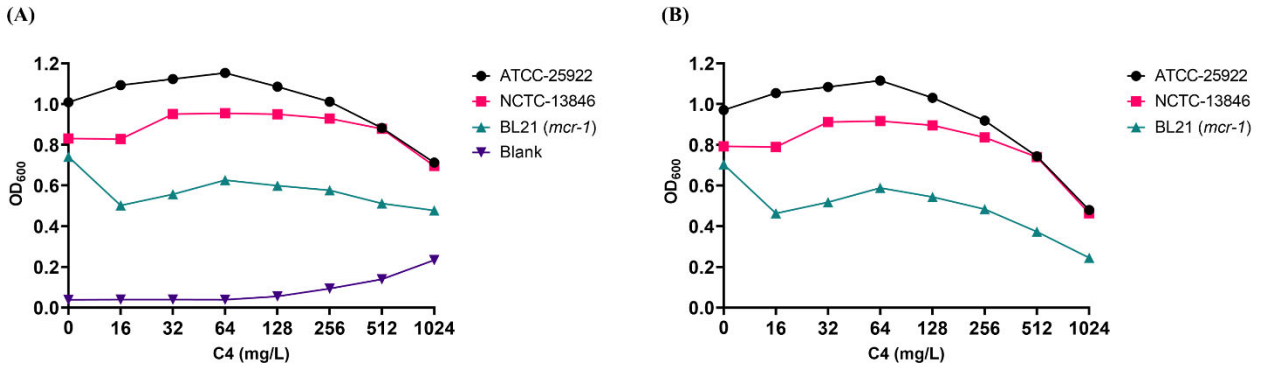


Figure 3-2. OD_{600} of *E.coli* strains incubated with compound C4 after 18 hours.

(A) The raw OD_{600} values with blank (CAMHB + C4). (B) The blank subtracted OD_{600} values. Cells were grown once in triplicate, the mean of each data point is presented in the plot (error bars are not shown for clarity).

Result for the blank group (C4 alone in CAMHB) suggested that C4 at concentrations higher than 256 mg/L increased the absorbance signal at 600 nm. The blank-subtracted results showed that compound C4 began to exhibit growth inhibition effects on the *E.coli* strains in a concentration-dependent manner when the concentration of C4 was higher than 256 mg/L. The results suggested the maximum concentration of compound C4 in colistin susceptibility assays should be 256 mg/L. Although 32 mg/L C4 can decrease the colistin MIC of *E.coli* BL21 (DE3) to 2 mg/L, more than 64 mg/L C4 is required to reduce the colistin MIC of *E. coli* NCTC13846 to 2 mg/L. It is suggested that 64 mg/L C4 may be the optimal concentration for future assays.

3.2.3 Effects of compound C4 against MCR-1 producing clinical strains

To verify whether compound C4 has an effect on the colistin susceptibility of clinical *mcr-1* positive isolates, 21 randomly picked clinical isolates (provided by Dr Jon Tyrrell, University of Bristol School of Cellular and Molecular Medicine) were tested. The effects of compound C4 against these cells were evaluated by colistin MIC assays. Strain C179 and C180 are *mcr-1* positive *K.pneumoniae* isolates and the remaining 19 strains are

clinical *mcr-1* positive *E. coli* isolates. Pterostilbene was used as a positive control. Compound C4 at a concentration of 64 mg/L clearly improved colistin activity against MCR-1-expressing clinical isolates, including *E.coli* and *K.pneumoniae* strains (**Table 3-1)**.

Table 3-1. Colistin MIC determination of compound C4 against clinical *mcr-1* positive isolates.

Strain	Control	64 mg/L C4	32 mg/L PT	Strain	Control	64 mg/L C4	32 mg/L PT
C179	16	2	1	C219	4	2	2
C180	16	2	2	C222	8	2	1
C193	8	2	1	C224	8	1	1
C195	8	2	1	C225	8	2	1
C197	8	2	1	C228	8	2	2
C203	8	1	2	C229	8	2	2
C204	8	1	2	C232	8	2	2
C211	8	1	1	C239	8	2	2
C215	8	2	2	C243	4	2	1
C217	4	2	1	C245	8	1	2
C218	4	2	1				

Colistin concentration range: 0.25-16 µg/ml; MIC unit: µg/mL
C4: 64 µg/mL C4 + colistin; Pterostilbene: 32 µg/mL pterostilbene + colistin; Control: colistin only.
MIC values were determined in duplicate.

3.2.4 Compound C4 potentiated colistin activity against ArnT-mediated colistin resistant *K. pneumoniae*

In addition to the MCR-1 producing strains, the colistin potentiating effect of compound C4 was tested against the laboratory-generated colistin-resistant *K.pneumoniae* strains P23 and P23 (*pmrA*) in which colistin resistance is primarily mediated by ArnT overexpression[32]. Pterostilbene was used as a comparator for potentiating colistin activity, although its activity against ArnT producing strains was not previously reported. We found that compound C4 improved colistin potency against all tested colistin-resistant strains. The colistin MICs for *K.pneumoniae* P23 and P23 (*pmrA*) strains are 64 and 128 mg/L respectively. It was surprising that, in the presence of 128 mg/L compound C4, the colistin MICs for the two *K. pneumoniae* strains were reduced to 32 mg/L. The colistin MICs for P23 and P23 (*pmrA*) strains could be further decreased with increased C4 concentration. We also found that both of the reported MCR-1 inhibitors pterostilbene and osthole also reduced colistin MICs against the P23 and P23 (*pmrA*) *K. pneumoniae* strains that do not express MCR-1. Indeed, the effect of both of these compounds on colistin susceptibility of *K. pneumoniae* P23 and P23 (*pmrA*) was more pronounced than that of compound C4. Notably, however, white cell deposits appeared in the bottom

of individual wells in the MIC plate containing intermediate concentrations of colistin and compound C4. This will be discussed further below (section 3.2.7). A summary of colistin MICs measured against the different strains is shown in **Table 3-2**. Colistin MIC values for another MCR-1 expressing *E.coli* strain, C43 (DE3)[213], as well for control experiments using the colistin-sensitive *E.coli* BL21 (DE3) and C43 (DE3) strains carrying the empty pET24a vector were also included.

Table 3-2. Colistin MICs with C4, PT and OST against lab and reference strains.

		C4				Pterostilbene	Osthole		
	Control	**16**	**32**	**64**	**128**	**32**	**32**	**64**	**128**
Ec NCTC-13846	4	ND	4	2	ND	2	2	2	ND
Ec ATCC-25922	0.5-1	ND	ND	0.25-0.5	0.25-0.5	0.5	ND	ND	ND
Ec BL21(*mcr-1*)	8	ND	1	0.5-1	ND	0.5-1	2	ND	ND
Ec BL21(pET24a)	1	ND	0.5	ND	ND	0.25	ND	ND	ND
Ec C43(*mcr-1*)	8	4	0.5-2	0.5*	0.5*	0.25	ND	ND	ND
Ec C43(pET24a)	0.5-4	0.5	0.25-0.5	0.5*	0.5*	0.25	ND	ND	ND
Kp P23	64-128	ND	ND	16-64**	8-32**	2	4	4	4
Kp P23(*pmrA*)	64-128	ND	ND	16-64**	8-32**	2	4	4	4
Kp Ecl8	0.5*	ND	ND	0.5	ND	0.5	ND	ND	ND

* Colistin concentration lower than the values shown were not tested.
** Values were determined without shaking and cell deposits were observed in the plate.
MIC values were determined in triplicate. ND indicates the MIC value was not determined.

3.2.5 Growth curves of *E.coli* cells in the presence of compound C4 and colistin

To investigate the effects of C4 and pterostilbene on *E.coli* cells in more detail, we grew wild-type *E.coli* ATCC-25922, laboratory-generated *mcr-1* positive *E.coli* BL21 (DE3) and the EUCAST reference *mcr-1* positive *E.coli* NCTC-13846 strain in CAMHB with various compound combinations (**Figure 3-3**). For ATCC-25922, the presence of C4 dramatically promoted growth at stationary phase with the degree of improvement dependent on the concentration of C4 **(Figure 3-3A).** The growth-promoting effect of C4 was less pronounced for NCTC-13846 cells, and not observed for *mcr-1* expressing BL21 (DE3) cells when colistin was not present **(Figure 3-3B).** The growth curves also suggested that both 16 mg/ L and 32 mg/L pterostilbene exhibited inhibitory effects on cell growth of both the *mcr-1*-positive BL21 (DE3) and NCTC-13846 strains when colistin was absent, with the effect correlating with the concentration of pterostilbene (**Figure 3-3C, E**). In the presence of a sub-MIC concentration of colistin (2 mg/L), both C4 and pterostilbene at all test concentrations showed inhibitory effects on the cell growth **(Figure 3-3B)**. The laboratory-generated

mcr-1-positive BL21 (DE3) strain was less robust compared to the reference *mcr-1* positive strain NCTC-13846. Specifically, 32 mg/L C4 combined with 2 mg/L colistin fully inhibited the growth of *mcr-1*-positive BL21 (DE3) cells but only extended the lag phase of NCTC-13846 cells **(Figure 3-3D, F)**.

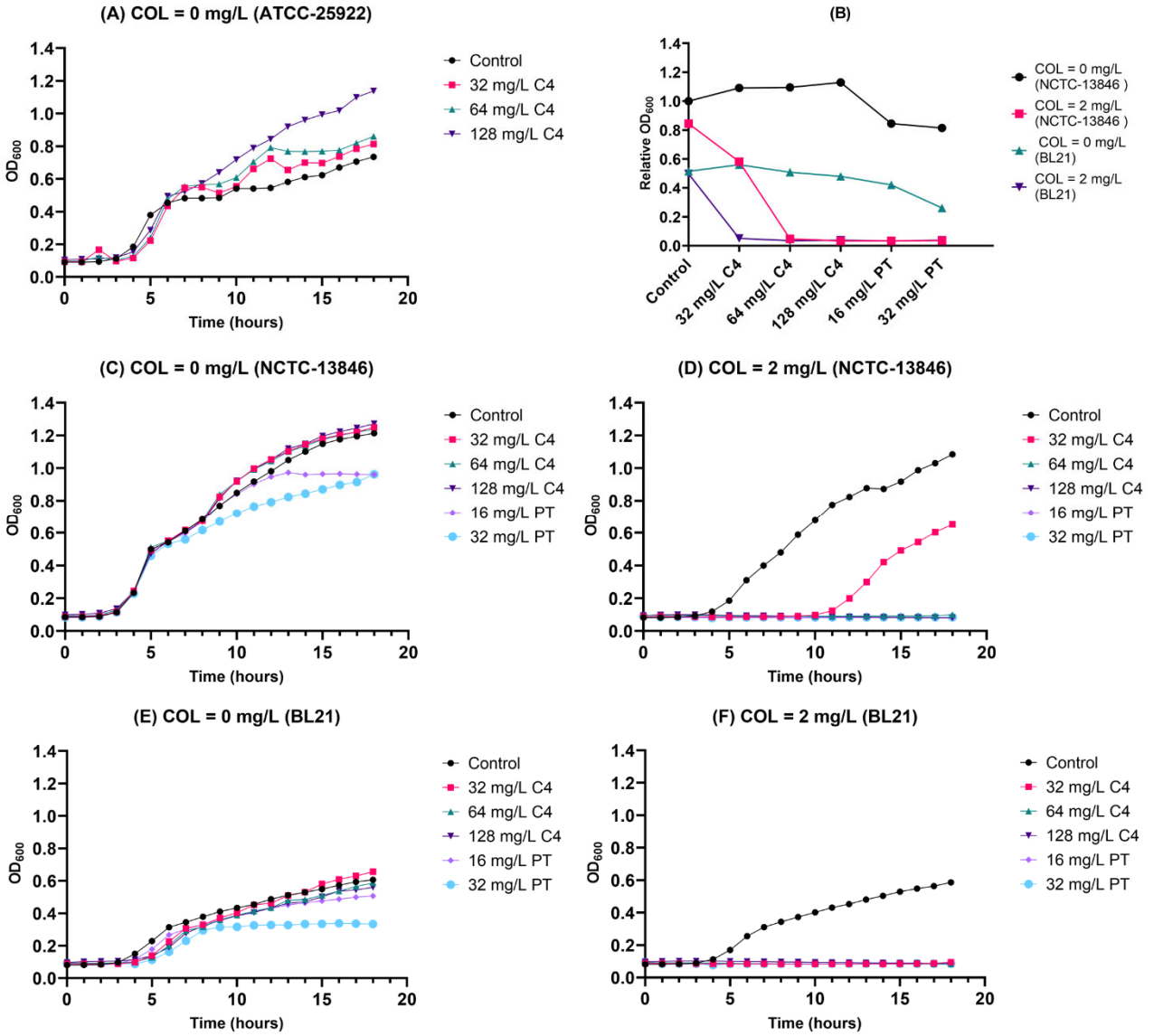


Figure 3-3. Growth of *E.coli* ATCC-25922, NCTC-13846 and BL2 (*mcr-1*) cells in CAMHB with C4, pterostilbene and colistin.

ATCC-25922 cells were incubated with 0, 32, 64, 128 mg/L C4 without colistin. NCTC-13846 and BL21(*mcr-1*) cells were incubated with CAMHB, 32, 64, 128 mg/L C4, 16 and 32 mg/L PT with 2 mg/L colistin or without colistin respectively. Cells were incubated at 37°C for 18 hours and the growth data were collected every 1 hour. Cells were grown in triplicate; the mean of each data point is presented in the plot. Relative OD_{600} values (panel B) were calculated by dividing readings by the OD_{600} values for the colistin free group after 18 hours incubation. Cells were grown in triplicate, the mean of each data point is presented in the plot (error bars are not shown for clarity). The assay was independently done at minimum two times.

3.2.6 Investigating the effects of compound C4 on the cell membrane

NPN uptake assay

The possibility was considered that compound C4 and pterostilbene might act to enhance the effect of colistin on bacterial membranes. The effect of colistin, compound C4 and pterostilbene upon bacterial membranes was investigated. NPN uptake assays were carried out to investigate the effect of colistin on the outer cell membrane integrity of the NCTC-13846 *E.coli* strain. The experiment was first set up according to the published studies[58,196–199] and validated by using colistin as the positive control. The final concentration of NPN used in the assay was 10 μM and the final OD_{600} value of the cell inoculum was 0.5. The fluorescence signal was monitored for 10 min after adding colistin/compound. The NPN fluorescent intensity was clearly promoted by the addition of colistin and a colistin concentration-dependent increase in the intensity of NPN fluorescence was observed (**Figure 3-4**). The effect of different concentrations of cell inoculum on NPN fluorescence intensity was also tested. The results suggested that cells with a relatively lower OD_{600} value were able to cause a higher fluorescence response using our current assay conditions (**Figure 3-4B**). The NPN validation assay was then carried out using cells at a final OD_{600}=0.1. Results suggested that, even though the absolute fluorescence intensity was higher for cells inoculated at OD_{600}=0.1, the relative change in fluorescence intensity after adding colistin is similar between cells inoculated at final OD_{600}=0.1 and OD_{600}=0.5 (Compare **Figure 3-4A and C**). Therefore, an inoculum with a final OD_{600} between 0.1 and 0.5 can be used in the NPN assay, as long as the inoculum was the same in each experiment.

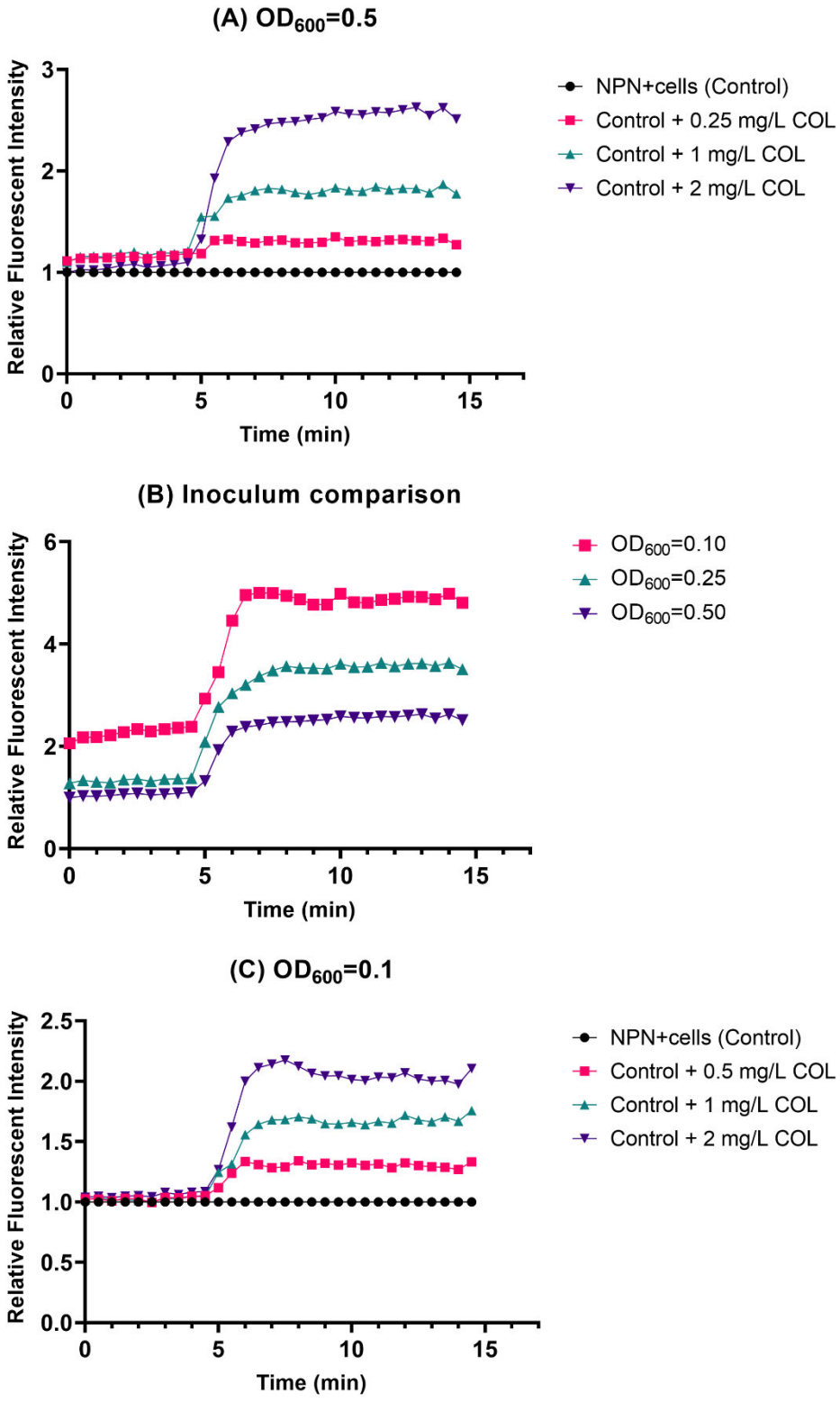


Figure 3-4. Validation of the NPN uptake assay.

Colistin was added at 4.5 min. (A) NPN assay with inoculum of NCTC-13846 cells with final OD_{600}=0.5. (B) NPN assay with different inoculum of NCTC-13846 cells with 2 mg/L colistin. (C) NPN assay with inoculum of NCTC-13846 cells with final OD_{600}=0.1. The assay was conducted twice in triplicate, the mean of each data point is presented in the plot.

After successful validation of the NPN assay, the effects of C4 and pterostilbene on outer membrane permeability were investigated using the NPN assay platform. The results are presented in **Figure 2-5**. Although C4 alone at concentrations between 32 - 128 mg/L did not have any fluorescence signal at the Ex/Em settings used for NPN (the same intensity values as the assay buffer), it reduced the fluorescence signal of NPN in the presence (**Figure 3-5A**) and absence (**Figure 3-5C**) of bacterial cells. The addition of colistin (up to 2 mg/L) cannot reverse this decrease (**Figure 3-5D)**. Therefore, testing the effect of C4 upon the outer membrane using the NPN uptake assay resulted in a decreased fluorescence signal, which made the assay not feasible for monitoring the effect of C4.

The NPN uptake assay was also used to investigate the mechanism of action of pterostilbene on the *E. coli* outer membrane. Unfortunately, using the Ex/Em settings for NPN, pterostilbene in the assay buffer showed a stronger fluorescence signal (above the maximum reading of the plate reader used) than does NPN in the presence (**Figure 3-5B**) and absence (**Figure 3-5C**) of bacterial cells. Due to the unfavourable effects of both C4 and pterostilbene upon NPN fluorescence, it is then hard to directly use NPN as a probe to investigate the effect of these two compounds on the outer membrane integrity of Gram-negative cells.

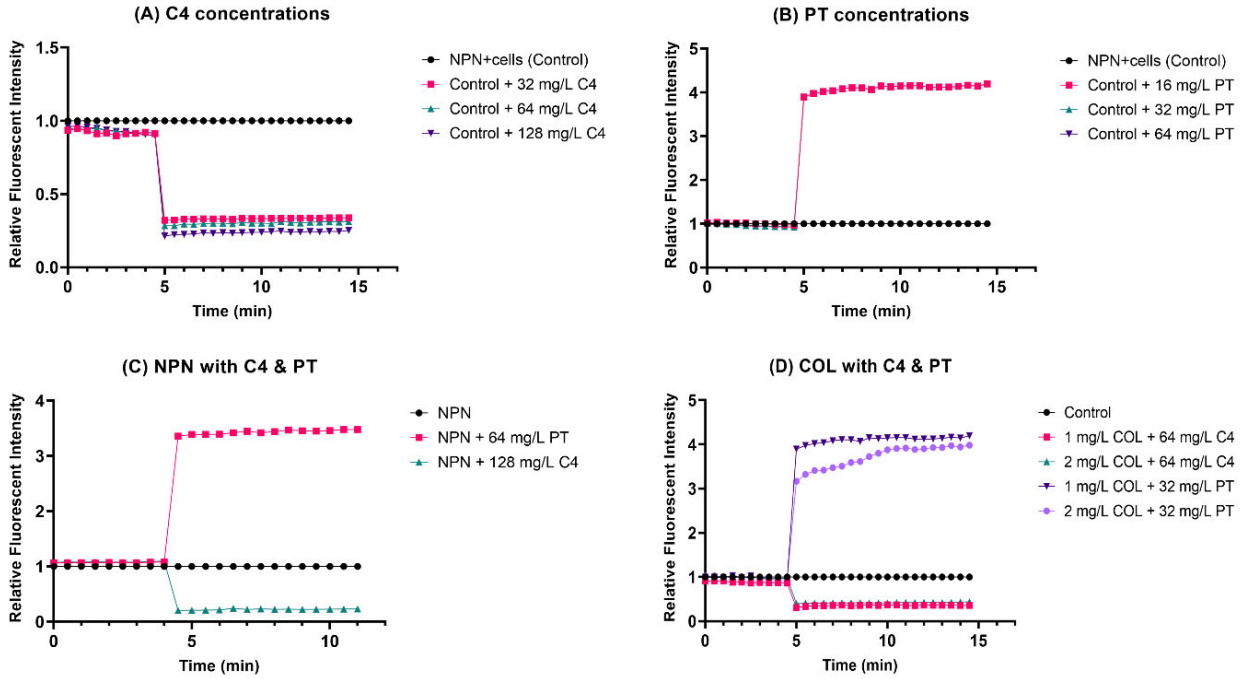


Figure 3-5. NPN uptake assays for C4 and pterostilbene.

Compounds were added at 4.5 min. The effect of (A) C4 and (B) pterostilbene on *E. coli* NCTC-13846 cells. (C) NPN fluorescent intensity for C4 and pterostilbene without cells. (D) NPN fluorescence intensity for the combination of colistin and C4/pterostilbene with NCTC-13846 cells. The assay was conducted twice in triplicate, the mean of each data point was presented in the plot.

To avoid direct contact between C4 and NPN in solution and possible resulting of loss of NPN fluorescence, we attempted to optimise the sample preparation process. Instead of adding C4 and bacterial cells together into the assay buffer, *E. coli* NCTC-13846 cells were first co-cultured with C4 or colistin and then washed in

assay buffer before carrying out the NPN uptake assay. The results of NPN uptake assays using washed cells (**Figure 3-6**) showed that cells co-cultured with 2 mg/L colistin still showed significantly higher NPN fluorescence intensity compared to the cells cultured in CAMHB alone. It is suggested that these cells have damaged cell membranes as a result of colistin exposure which they remained unable to repair after washing in buffer. In contrast, no difference was seen between the cells co-cultured with CAMHB alone and with 64 mg/L C4, suggesting that C4 did not permeabilize the outer membrane of the bacteria. These data suggest that the effect of compound C4 on the bacterial outer membrane is not the same as that of colistin. In addition, the optimized protocol may be used to investigate the effect of C4 and colistin in combination as well as pterostilbene using NPN assays in the future.

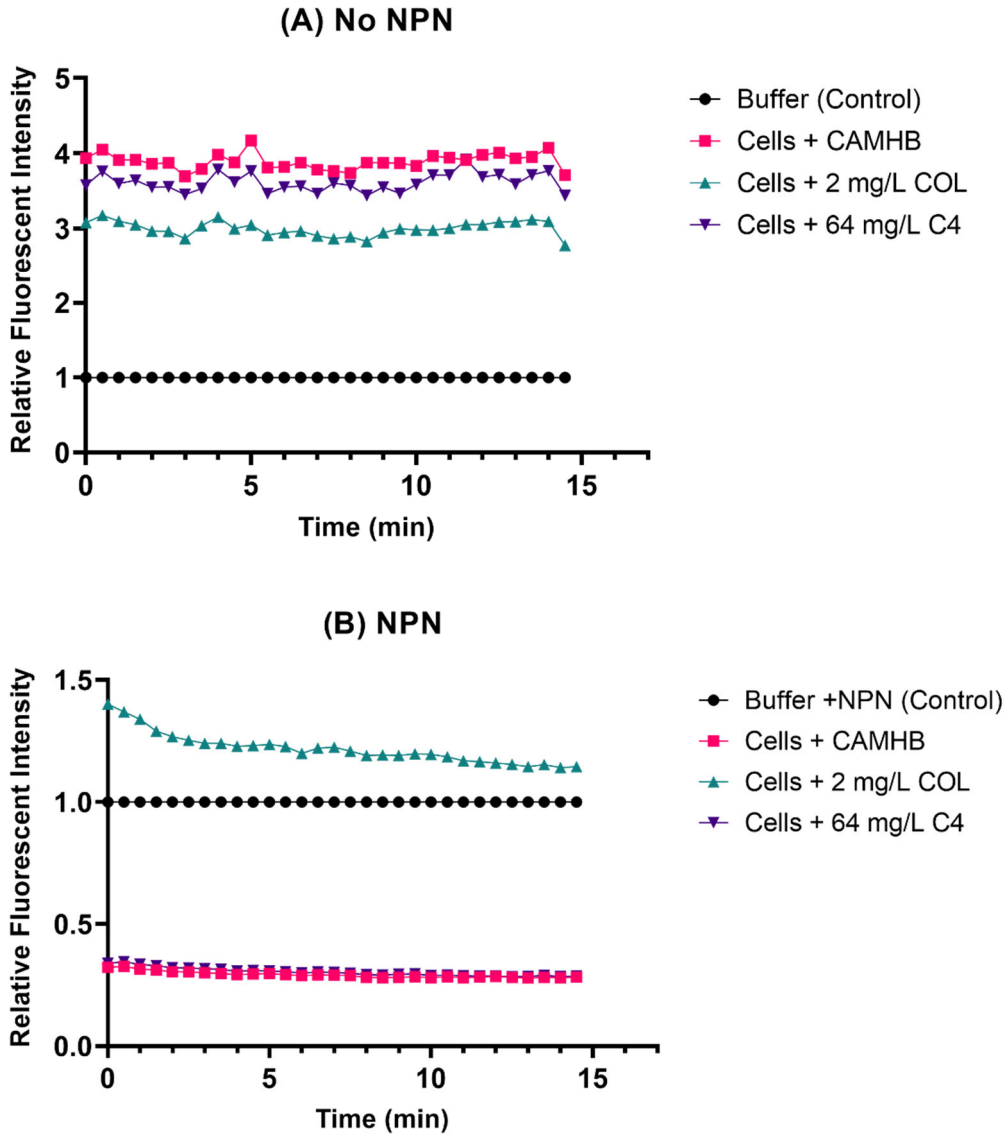


Figure 3-6. NPN uptake assays with buffer washed cells.

Assays without (A) and with (B) NPN using *E. coli* NCTC-13846 cells. Cells were co-cultured with CAMHB, 2 mg/L colistin (COL) and 64 mg/L C4 respectively. Relative fluorescence intensity of samples was calculated by referring to the cell free control group in each set-up respectively. The assay was conducted twice in triplicate, the mean of each data point was presented in the plot.

Propidium iodide (PI) uptake assay

We also used the probe propidium iodide (PI) to investigate the effect of compounds on the cell membrane integrity (including both outer and inner membrane) of the *E.coli* NCTC-13846 strain. The experiment was first set up following methods described in previously published studies[58,200,201] and validated by using colistin as the positive control for membrane permeabilization. The final concentration of PI used in the assay was 10 μM and the final OD_{600} value of the cell inoculum was 0.5. The fluorescence signal was monitored for 10 min after adding colistin/compound. The PI fluorescence intensity was clearly promoted by the addition of colistin and the increase was colistin concentration-dependent (**Figure 3-7**). In contrast, the addition of C4 or pterostilbene does not have any effect on the fluorescence intensity of PI. These results suggested that C4 and pterostilbene could not permeabilize the cell inner membrane. However, the addition of C4 (64 μg/ml) slightly reduced the fluorescence increase caused by colistin. Because the presence of C4 alone did not reduce the fluorescence of PI, it may be that C4 interferes with the ability of colistin to interact with the inner membrane, rather than affecting the permeability of the membrane directly. Unlike compound C4, the addition of pterostilbene (32 μg/ml) did not affect the fluorescence increase caused by colistin.

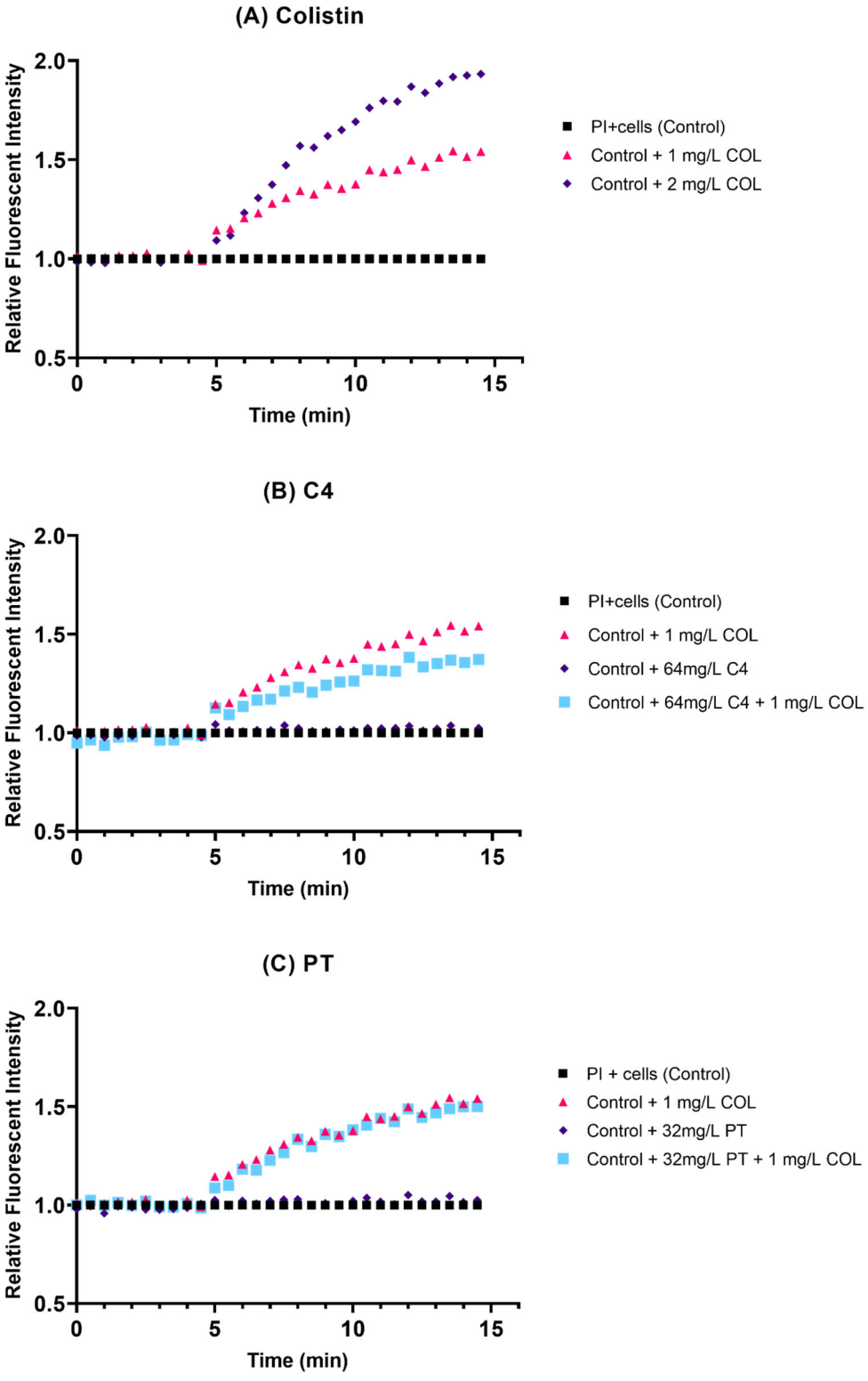


Figure 3-7. PI uptake assays for colistin, C4 and pterostilbene.

Compounds were added at 4.5 min. The effect of (A) colistin, (B) C4 and (C) pterostilbene on NCTC-13846 cells. The assay was conducted once in triplicate, the mean of each data point was presented in the plot.

Vancomycin susceptibility testing

Vancomycin (VAN) inhibits cell-wall synthesis in Gram-positive bacteria but is normally ineffective against Gram-negative bacteria because vancomycin is unable to penetrate the outer membrane of Gram-negative cells.[58,214] We used vancomycin MICs as an indicator of membrane integrity for Gram-negative strains at 37°C.[215] PAβN is a known inhibitor of efflux pumps and has been reported to permeabilize the outer cell membranes of Gram-negative bacteria. So PAβN was considered as a potential membrane permeabilizing agent[188,205]. Vancomycin MICs were determined to be 256 mg/L for *E.coli* NCTC-13846 and 512 mg/L for the *E.coli* ATCC-25922 strains. 32 mg/L pterostilbene or 128 mg/L C4 did not potentiate vancomycin activity, while 2 mg/L colistin reduced vancomycin MICs at least 2-fold. This may be indirect evidence suggesting that C4 and pterostilbene do not affect the outer membrane integrity of Gram-negative strains. However, the presence of PAβN (up to 50mg/L) similarly did not potentiate vancomycin activity against MCR-1-producing *E.coli* NCTC-13846 cells. The vancomycin MIC values were summarized in **Table 3-3**.

Table 3-3. Vancomycin MIC determination with C4, PAβN, PT and COL.

Compounds added with VAN	ATCC-25922	NCTC-13846
Control (CAMHB)	512	512
32 mg/L C4	ND	512
64 mg/L C4	ND	256
128 mg/L C4	ND	512
12.5 mg/L PAβN	512	>128
25 mg/L PAβN	512	>128
50 mg/L PAβN	512	>128
32 mg/L PT	ND	>128
2 mg/L COL	ND	128

MIC values were determined in triplicate. ND indicated the MIC value is not determined.

E. coli growth curves in the presence of Vancomycin

To further investigate the effects of compound C4, pterostilbene and PAβN on vancomycin potency against the *E.coli* strains ATCC-25922 and NCTC-13846, we grew the cells in CAMHB with different combinations of compounds. The effect of PAβN on vancomycin potency against the *E.coli* type strain ATCC-25922 was first evaluated. The effects of PAβN, C4, PT and colistin (COL) on vancomycin potency against the *mcr-1* positive colistin-resistant EUCAST control *E.coli* strain NCTC-13846 were then evaluated (**Figure 3-8**).

The results showed that, in the presence of vancomycin, PAβN has an inhibitory effect on the growth of the ATCC-25922 strain (**Figure 3-8A**). For the NCTC-13846 strain an effect of PAβN was evident even in the absence of vancomycin, and growth was more evidently inhibited in the presence of vancomycin (**Figure 3-8B**). However, the vancomycin MIC did not change for either of the two strains. The data showed that the growth inhibitory effect caused by 25 or 50 mg/L PAβN is greater than that caused by 12.5 mg/L PAβN, while the difference between 25 and 50 mg/L is less clear (**Figure 3-8**).

Data for the growth of cells against time were also collected (**Figure 3-9**). The presence of PAβN changed the shape of the growth curve of *E. coli* ATCC-25922 in the vancomycin control group (Vancomycin = 0 mg/L). The growth curve of the vancomycin control group has one exponential phase, but two exponential phases of the growth curve were shown when the cells were treated with PAβN (**Figure 3-9**). Growth curve data for NCTC-13846 cells showed similar results to those obtained with the ATCC-25922 cells (**Figure 3-10**). The shapes of the growth curves of the vancomycin control group of NCTC-13846 cells were also changed. However, PAβN has little effect on vancomycin susceptibility, indicating that it is not permeabilizing either *mcr-1* negative or positive *E. coli* to vancomycin in our experiments.

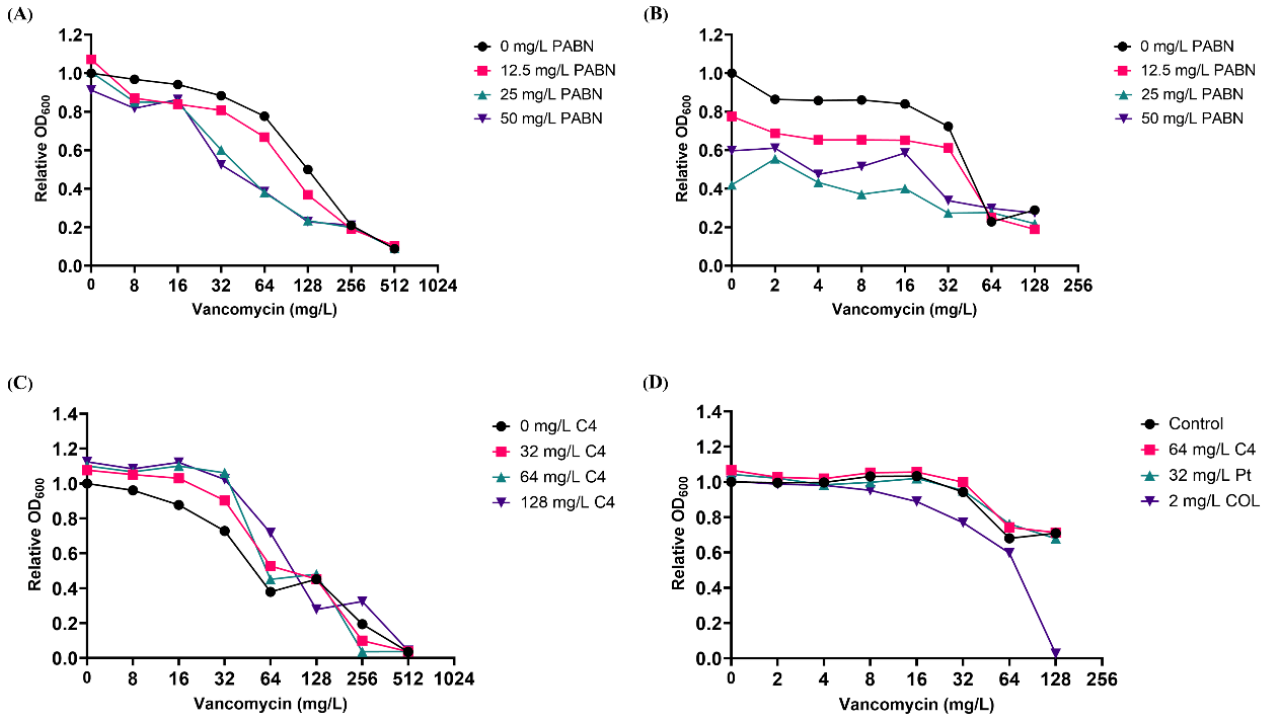


Figure 3-8. Potency analysis of vancomycin against *E.* coli strains treated with compounds at various concentrations.

The relative growth status of (A) ATCC-25922 and (B) NCTC-13846 cells incubated with vancomycin and PAβN at 37°C after 18 hours (quoted OD600 values are divided by the readings for the PaβN-free controls). (C) The relative growth status of NCTC-13846 cells incubated with vancomycin and C4 at 37°C after 18 hours (quoted OD600 values are divided by the reading for the 0 mg/L C4 group). (D) The relative growth status of NCTC-13846 cells incubated with vancomycin, C4, PT and COL at 37°C after 18 hours (quoted OD600 values are divided by the reading for the vancomycin only control group). Cells were grown in triplicate, the mean of each data point is presented in the plot (error bars are not shown for clarity). The assay was independently done two times. VAN = vancomycin.

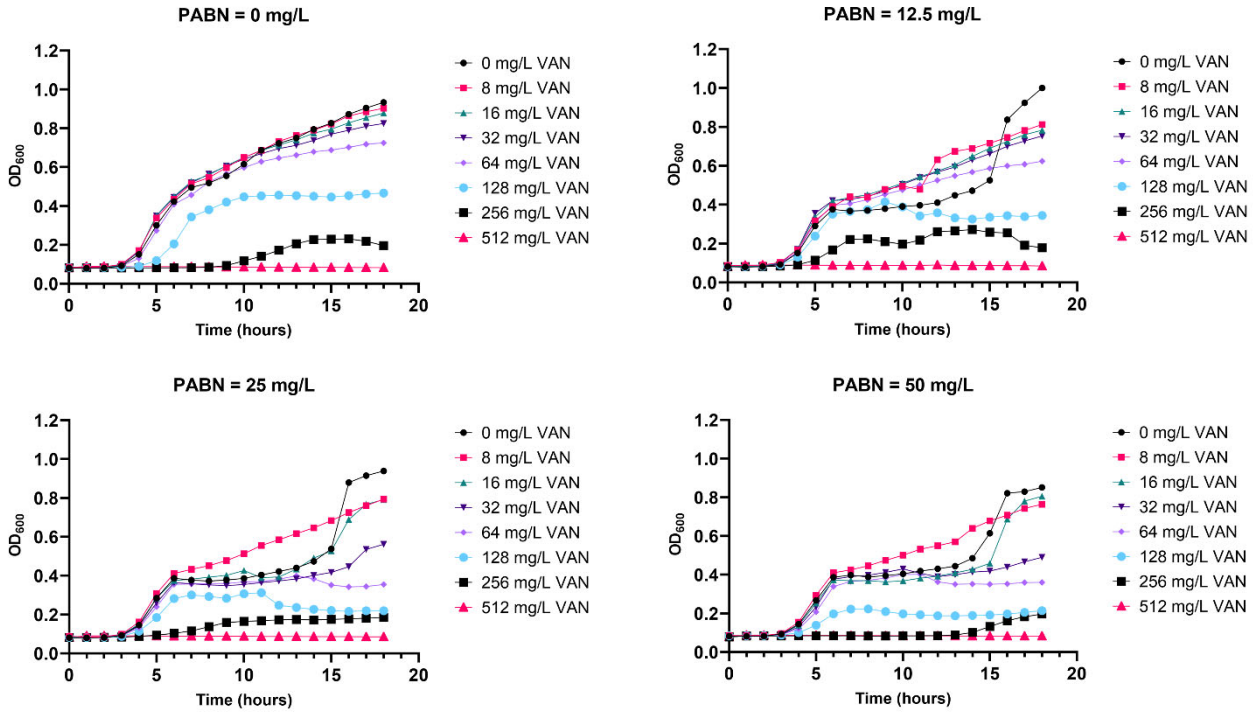


Figure 3-9. Growth curves of ATCC-25922 cells incubated with PAβN and vancomycin.

Cells were incubated with 0, 8, 16, 32, 64, 128, 256, 512 mg/L vancomycin companied with 0,12.5,25 and 50 mg/L of PAβN at 37°C respectively. The growth data were collected every hour for 18 hours. Cells were grown once in triplicate, the mean of each data point is presented in the plot (error bars are not shown for clarity). VAN = vancomycin. Note: For the treatment of ATCC cells with 12.5 mg/L PAβN, data points for cells incubated with Vancomycin at 512 mg/L are omitted at reading times between 10h and 17h incubation due to issues with condensation on the surface of the lid of the 96-well plate. Visual and OD_{600} checks with the lid removed showed no cell growth in the wells containing 512 mg/L vancomycin after 18h incubation.

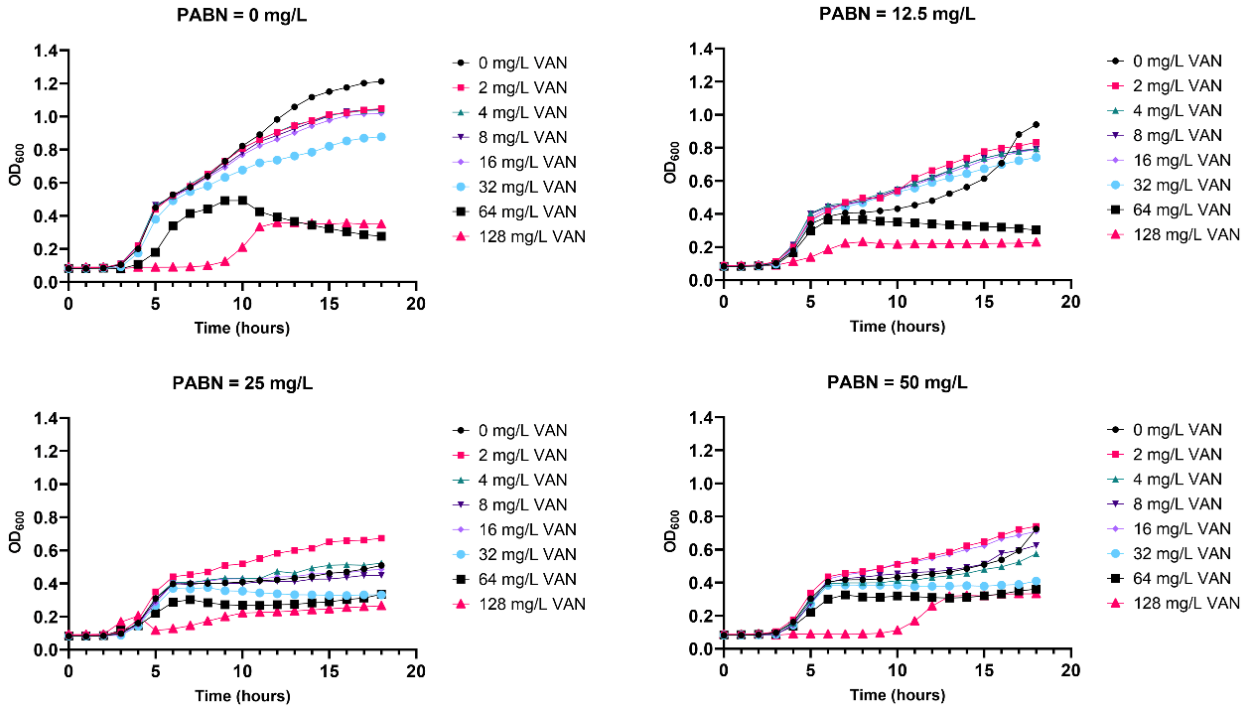


Figure 3-10. Growth curves of NCTC-13846 cells incubated with PAβN and vancomycin.

Cells were incubated with 0, 2, 4, 8, 16, 32, 64, 128 mg/L vancomycin companied with 0,12.5,25 and 50 mg/L of PAβN at 37°C respectively. The growth data were collected every hour for 18 hours. Cells were grown once in triplicate, the mean of each data point is presented in the plot (error bars are not shown for clarity). VAN = vancomycin.

We found that C4 at concentrations of 32, 64 and 128 mg/L modestly promoted the growth of NCTC-13846 *E.coli* cells (i.e., increased endpoint OD600 values were obtained compared to the C4 free control group) when the concentration of vancomycin was lower than 128 mg/L (**Figure 3-11**). The growth data of NCTC-13846 cells treated by C4 and vancomycin were collected. The data suggested that the presence of C4 did not have effect on the potency of vancomycin against these cells and the shape of curves (**Figure 3-11**).

To directly compare the effect of C4, pterostilbene and colistin (the membrane penetrating positive control) on vancomycin potency against NCTC-13846 cells, the minimum working concentration of C4 (64 mg/L) and pterostilbene (32 mg/L) as well as sub-MIC concentration of colistin (2 mg/L) were tested (**Figure 3-12**). We found that neither C4 nor pterostilbene have effects on the potency of vancomycin against NCTC-13846 cells, but colistin increased the potency of vancomycin, reducing vancomycin MICs to 128 mg/L (**Figure 3-12-D**). The time dependent growth data also showed that the presence of 2 mg/L colistin extended the lag phase of NCTC-13846 cells treated by sub-inhibitory (32 mg/L and 64 mg/L) concentrations of vancomycin. The results suggested that any effects of C4 or pterostilbene upon the integrity of the *E. coli* cell are not the same as those of colistin.

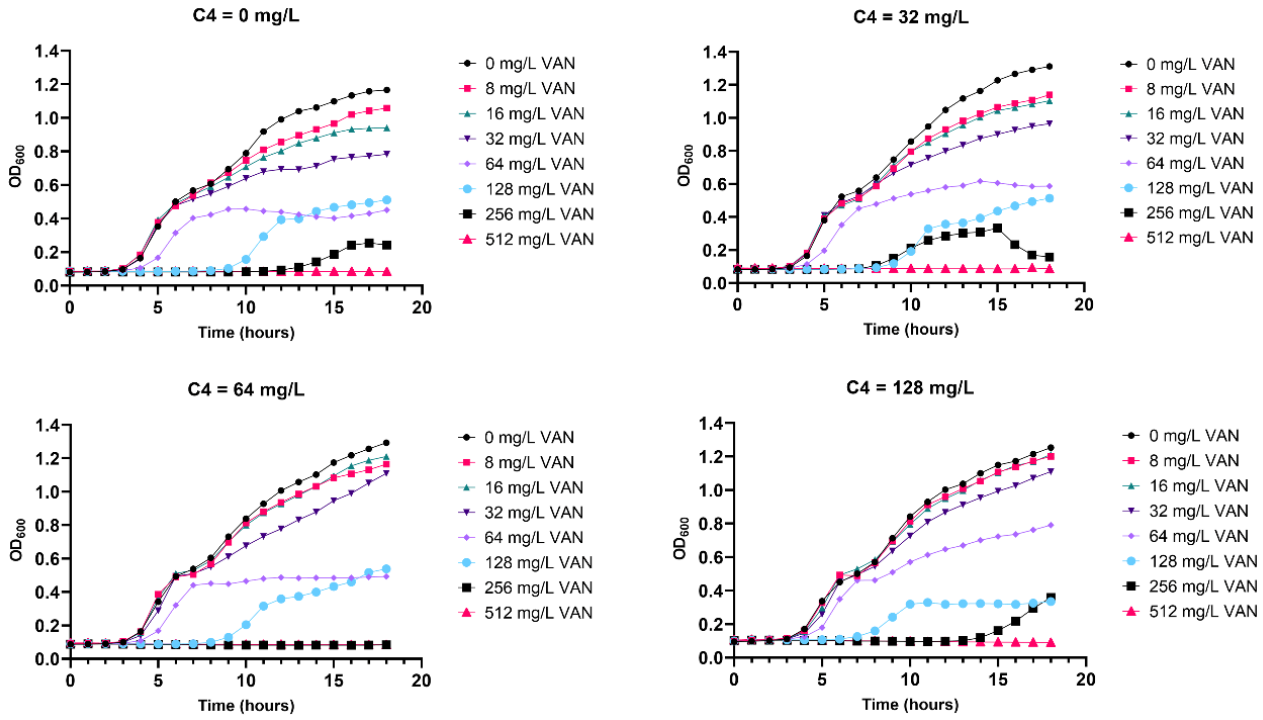


Figure 3-11. Growth curves of NCTC-13846 cells incubated with C4 and vancomycin.

Cells were incubated with 0, 8, 16, 32, 64, 128, 256, 512 mg/L vancomycin companied with 0, 32, 64 and 128 mg/L of C4 at 37°C respectively. The growth data were collected every hour for 18 hours. Cells were grown once in triplicate, the mean of each data point is presented in the plot (error bars are not shown for clarity). VAN = vancomycin.

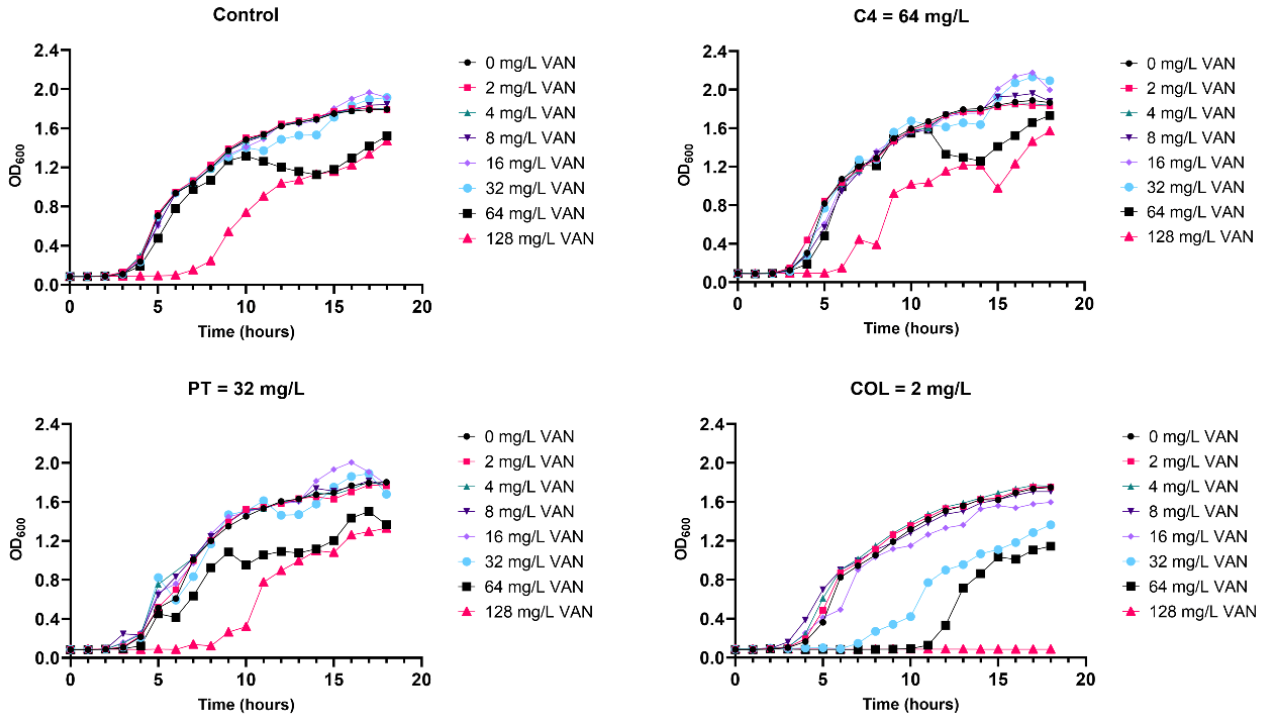


Figure 3-12. Growth curves of NCTC-13846 cells incubated with C4, PT, COL and vancomycin.

Cells were incubated with 0, 2, 4, 8, 16, 32, 64, 128 mg/L vancomycin companied with 64 mg/L of C4 or 32 mg/L PT or 2 mg/L colistin at 37°C respectively. The growth data were collected every hour for 18 hours. Cells were grown once in triplicate, the mean of each data point is presented in the plot (error bars are not shown for clarity). VAN = vancomycin.

3.2.7 Colistin heteroresistance of *K. pneumoniae* P23 and P23 (*pmrA*) strains

As described above (section 3.2.4) the effect of compound C4 on the *K. pneumoniae* P23 and P23 (*pmrA*) strains was to cause change in appearance in the MIC plates: cell deposition at the bottom of the 96-well microplate well. The cell deposit has a white appearance, is irregular in shape but is most often round. The cell deposit sits at the bottom of the well, immersed by the assay solution that consists of CAMHB, colistin and compound C4. The solution remains translucent in most cases when the cell deposit is observed (**Figure 3-13**).

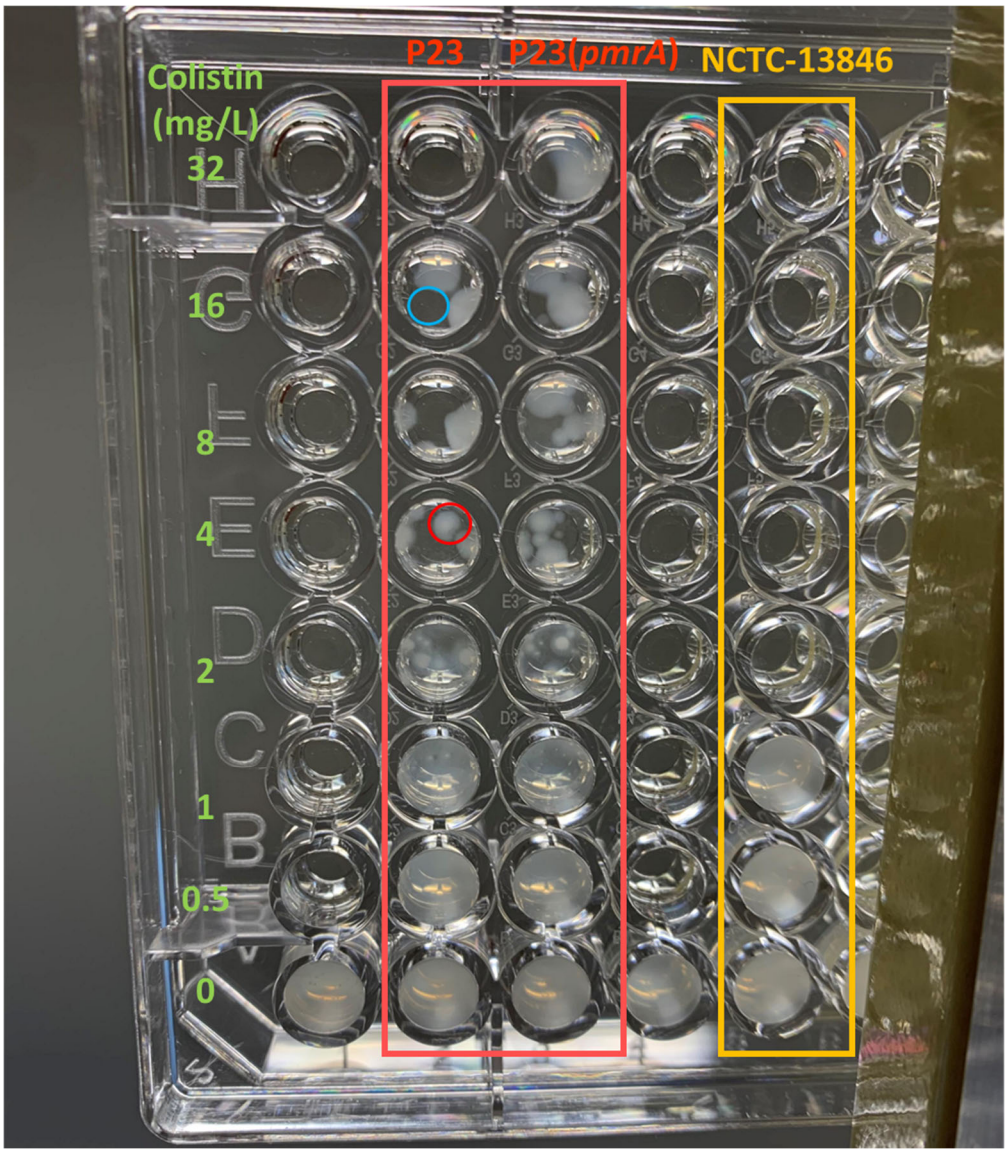


Figure 3-13. Representative image of white cell deposits of *K. pneumoniae* P23 and P23(*pmrA*) induced by the treatment of colistin and compound C4.

Red box: *K. pneumoniae* P23 and P23(*pmrA*) cells incubated for 18 hours at 37°C (static). Yellow box: *E.coli* NCTC-13846 cells under the same conditions. Red circles suggest the white cell deposit of *K. pneumoniae* cells. Blue circle shows that the medium solution of a well containing the white cell deposits remain transparent. Medium contains colistin serially diluted in CAMHB (concentrations in green (left) in the presence of 64 mg/L compound C4. Note: After a further 18 hours incubation the white cell deposits disappeared, and the medium was turbid and homogeneous.

On leaving the plate at room temperature after reading MIC values, the cell deposit gradually 'dissolved' in the solution and the translucent solution eventually turned turbid within 24 hours. This suggests that the cells of the white deposit are alive. To confirm this, we discarded the translucent assay buffer solution and then inoculated the cell deposit onto an LB agar plate with 2 mg/L colistin. The cells successfully grew and colonies were formed after overnight culture at 37°C. The cell deposits started to appear when the P23 or P23 (*pmrA*) cells were exposed to a concentration of colistin reaching the colistin MIC value for these strains. In the presence of 64 mg/L compound C4, it was most likely to see the cell deposits in wells of the serial dilution containing colistin at 8, 16 or 32 mg/L.

Because the plates for MIC assays are kept static during incubation, the growth curves of P23 and P23 (*pmrA*) cells were monitored using the same chemical combination as was used in MIC assays to find out if constant plate shaking could avoid the cell deposition phenomenon. The plate was monitored for 18 hours at 37°C with 100 rpm constant shaking, and the absorbance at 600 nm (OD_{600}) was recorded by the plate reader every 60 minutes. The growth curves of P23 and P23 (*pmrA*) cells with two concentrations (64 and 128 mg/L) of compound C4 were monitored. The growth curves are presented in **Figure 3-14**. With shaking, the frequency of cell deposit occurrence was greatly decreased, with only tiny dots observed in a few wells (<5 for each strain). In addition, due to the disappearance of the cell deposit, constant plate shaking at 100 rpm further reduced the colistin MIC values for the P23 and P23 (*pmrA*) strains to 2 and 4 mg/L respectively in the presence of 128 mg/L compound C4. These values are 4 – 8-fold lower compared to the MIC values measured in the static colistin MIC assays. It might be that constant shaking promotes the even distribution of colistin in the wells while also preventing cells from accumulating or attaching to the plastic surface of the well.

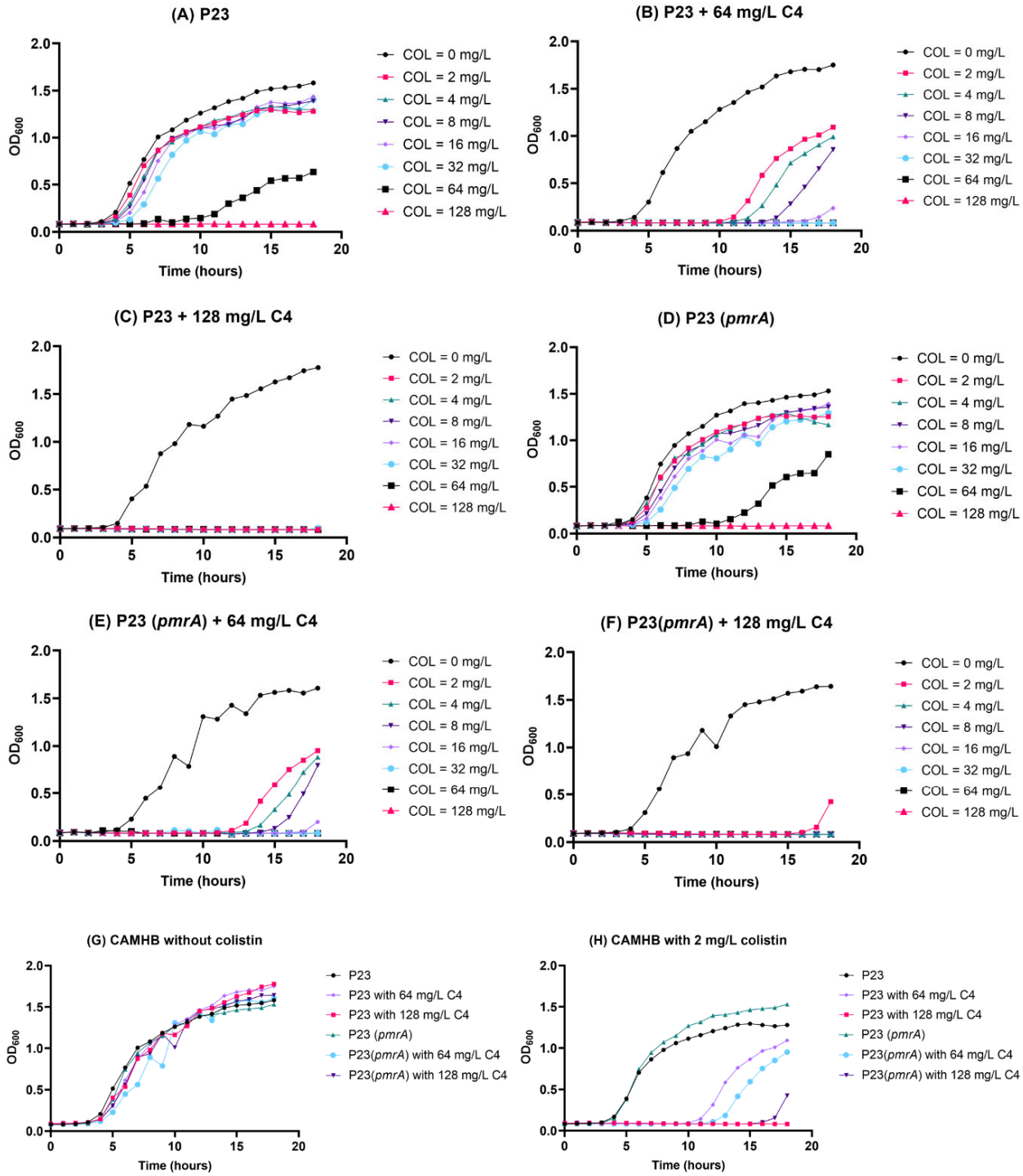


Figure 3-14. Growth curves of *K. pneumoniae* P23 and P23 (*pmrA*) cells incubated with C4 and colistin.

(A) P23 cells were incubated with 0, 2, 4, 8, 16, 32, 64, 128 mg/L colistin. (B-C) P23 cells were incubated with 0, 2, 4, 8, 16, 32, 64, 128 mg/L colistin supplemented with 64 and 128 mg/L C4 respectively. (D) P23(*pmrA*) cells were incubated with 0, 2, 4, 8, 16, 32, 64, 128 mg/L colistin. (E-F) P23 (*pmrA*) cells were incubated with 0, 2, 4, 8, 16, 32, 64, 128 mg/L colistin supplemented with 64 and 128 mg/L C4 respectively. Three treatments (CAMHB, 64 mg/L C4, 128 mg/L C4) on P23 and P23 (*pmrA*) cells without colistin (G) / with 2 mg/L colistin (H). The growth data were collected every 60 minutes for 18 hours at 37°C (100 rpm constant shaking). Cells were grown twice in triplicate, the mean of each data point is presented in the plot (error bars are not shown for clarity). COL= colistin.

To investigate whether the cell deposit was caused by heteroresistance (resistance emerging from a subpopulation of a susceptible strain), cell deposit samples were taken from the statically-grown 96-well plate used in the P23 and P23 (*pmrA*) colistin MIC assays . Cell deposits from wells with the chemical condition of 8 mg/L colistin and 64 mg/L C4 were collected for both the P23 and P23 (*pmrA*) strains. The cells were then streaked on LB agar plates containing 8 mg/L colistin for overnight culture. A colistin MIC assay using strains from these overnight LB agar plates as the inoculum was then carried out. The results showed that, in the presence of 64 mg/L C4, and when the plate was grown with shaking at 100 rpm, the colistin MIC value increased from 8 to 32 mg/L for the P23 strain and from 16 to 64 mg/L for the P23 (*pmrA*) strain, respectively. In both cases the cells grew better compared to P23 and P23 (*pmrA*) strains that had not been exposed to C4. Although white cell deposits could still be found in some wells where the colistin concentration was close to the MIC value, the size of these white deposits formed by the C4-exposed generation was dramatically larger than those obtained from the initial generation. In wells containing colistin at concentrations of less than 16 mg/L, the cells grew well, turning the broth in these wells turbid and cell deposits were not seen in these wells. These findings indicate that, compared to the initial generation of P23 and P23 (*pmrA*) cells, the new generation derived from the cell deposit shows better adaptation to an environment where both colistin and C4 are present. In addition, the colistin MIC values for the C4-exposed generation of P23 and P23 (*pmrA*) strains, measured when cells were grown, with shaking, in the absence of C4, were both 128 mg/L. The integrity of these cells was confirmed by gentamicin nonsusceptibility (at 5 mg/L) suggesting that the genetic characteristics of P23 and P23 (*pmrA*) strains still remained. To investigate the environmental adaption of these ArnT expressing *K. pneumoniae* cells, growth curves of the P23 (*pmrA*) cells deriving from the white deposit (P23(*pmrA*)-WD) were monitored (**Figure 3-15**). Consistent with the findings above, the results showed that previous exposure to compound C4 protects cells from colistin action when compound C4 is present (compare **Figure 3-14E and F** with **Figure 3-15B and C**) while having little effect when compound C4 is absent (compare **Figure 3-14D** with **Figure 3-15A**).

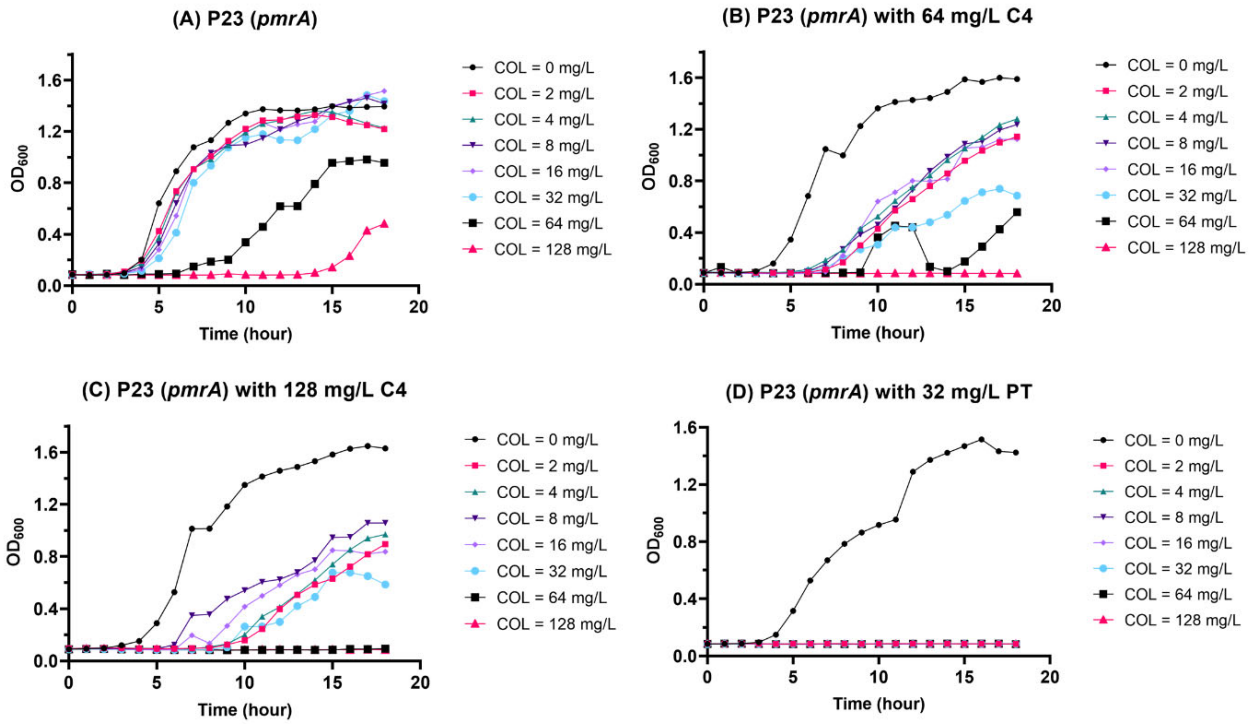


Figure 3-15. Growth curves of P23 (*pmrA*)-WD cells incubated with C4, PT and COL.

(A) P23 (*pmrA*)-WD cells were incubated with 0, 2, 4, 8, 16, 32, 64, 128 mg/L colistin. (B-C) P23 (*pmrA*)-WD cells were incubated with 0, 2, 4, 8, 16, 32, 64, 128 mg/L colistin companied with 64 and 128 mg/L C4 respectively. (D) P23 (*pmrA*)-WD cells were incubated with 0, 2, 4, 8, 16, 32, 64, 128 mg/L colistin companied with 32 mg/L pterostilbene. Cells were grown once, with shaking at 100 rpm, in triplicates, the mean of each data point is presented in the plot (error bars are not shown for clarity).

The colistin-susceptible *K. pneumoniae* Ecl8 strain is the parent for the P23 and P23 (*pmrA*) strains. We then investigated whether we could see a similar phenomenon when treating the Ecl8 strain with 64 mg/L compound C4 in the presence of serially diluted colistin (concentrations 0, 0.0625, 0.125, 0.25, 0.5,1,2,4 mg/L). The results showed that no cell deposit was observed under any condition, and the colistin MIC value for Ecl8 (1~2 mg/L) was not affected by the addition of compound C4. This result suggested that C4 cannot induce formation of cell deposits by the colistin-susceptible *K. pneumoniae* Ecl8 strain. However, use of low concentrations of colistin can be problematic because the issue of colistin loss becomes more significant as the colistin concentration becomes lower[160]. It might be possible that the cell deposits of the Ecl8 strain did not appear because the concentration of colistin is low. Taken together, these and previous data suggest that this may be an observation of colistin heteroresistance in the presence of compound C4 for both the P23 and P23 (*pmrA*) strains.

Further experiments were carried out to investigate whether the improved tolerance to the combination of C4 and colistin can be stably maintained by the *K. pneumoniae* strains. Specifically, the stability of the heteroresistance phenotype after storage was assessed using frozen stocks of the samples used in the experiments shown in **Figure 3-15**. In colistin MIC tests carried out (without shaking) in the presence and absence of compound C4 (64 mg/L), contrary to previous observations no difference was observed between

the colistin MICs of the parent P23 and P23 (*pmrA*) strains and cells taken from the cell deposits that been frozen after growth on LB agar without selection, or with 8 mg/L colistin in the absence or presence of 64 mg/L compound C4. This was true when frozen cells were cultured on LB agar either in the absence or presence of colistin at 16 mg/L before inoculating the MIC plate. In all cases cell deposits/heretoresistance were observed, rather than the turbid growth described above, suggesting that the frozen strains had reverted to the original phenotype. For cells that had been frozen after growth on colistin (8 mg/L) MIC measurements were also carried out on samples grown on LB agar containing colistin (16 mg/L) and compound C4 (64 mg/L) prior to inoculating the MIC plate. Again, no difference in colistin MIC was observed compared to the parent strains, and cell deposits/heteroresistance were evident across the MIC plate.

Elevated MICs and strong growth (absence of cell deposits/heteroresistance) were however observed when cells that had been frozen after growth on 8 mg/L colistin plus 64 mg/L compound C4 were used to inoculate the MIC plate. In this case colistin MICs of 16 mg/L (P23) and 64 mg/L (P23 (*pmrA*)) were measured in the presence of 64 mg/L compound C4, but, importantly, these experiments were distinguished by strong growth at sub-inhibitory concentrations of colistin and an absence of the cell deposit/heteroresistance phenotype. This experiment was carried out after taking the frozen strains directly into liquid culture and growing in LB broth (plus colistin 16 mg/L and compound C4 64 mg/L) for 1.5 hours (to an OD = 0.02-0.04) before inoculating the MIC plate. However, when the same frozen strains were plated onto LB agar plus 16 mg/L colistin and 64 mg/L compound C4 and then used to inoculate 96-well plates, colistin MICs in the presence of compound C4 (64 mg/L) decreased to 4 mg/L and 32 mg/L for the P23 and P23 (*pmrA*) strains, respectively. In both strains, cell deposit formation/heteroresistance was observed, but this was particularly noticeable for P23 (*pmrA*). These results indicate that cell adaptation to colistin/compound C4 on prior exposure to compound C4 is not a stable phenotype and can reverse on storage at – 80 ºC and after recovery on agar plates. Loss of polymyxin resistance on storage of resistant isolates has previously been described for several Gram-negative species[216].

We considered investigating possible changes in protein expression when cells were exposed to C4 using proteomics assays. To ascertain whether the C4-tolerant phenotype could be stably maintained, further experiments were carried out in which bacteria were grown on a larger scale in liquid culture. As cells that had been frozen after growth on 8 mg/L colistin plus 64 mg/L compound C4 had the highest tolerance to the combination of colistin and C4 in the experiments above, these were used for the following experiments.

Frozen stocks of the P23 and P23 (*pmrA*) strains (grown on on 8 mg/L colistin plus 64 mg/L compound C4 before freezing) were grown overnight on LB agar in the presence of 16 mg/L colistin and 64 mg/L C4. Three single colonies on the plate were selected for each strain for overnight culture in 10 mL CAMHB with 16 mg/L colistin and 64 mg/L C4. The overnight culture was then diluted 1:100 with CAMHB containing 16 mg/L colistin and 64 mg/L C4 in 50 mL sub-culture flasks in triplicate. Cells in the flasks were cultured at 37°C shaking at 180 rpm until the optical density (OD_{600}) reached 0.6 - 0.8. For each replicate, samples were taken out for colistin MIC testing and the rest of the cells harvested by centrifugation. Surprisingly, these

samples yielded colistin MIC values in the presence of 64 mg/L compound C4 of ≤ 2 mg/L (the lowest concentration used in the colistin serial dilution in this experiment). Control experiments on the same plate showed the samples to grow well in the absence of either compound C4 or colistin, and samples of the parent P23 and P23 (*pmrA*) strains yielded MIC values consistent with previous measurements.

As it was not possible to replicate in liquid culture the apparent heteroresistance phenotype induced by exposure of *K. pneumoniae* P23 and P23 (*pmrA*) to compound C4 using previously frozen cells grown on LB agar plates in the presence of colistin and compound C4, we attempted to do so using as inoculum strains grown directly in small scale liquid culture without previous recovery on agar plates. As described above, enhanced growth and decreased colistin susceptibility were observed when frozen cells, originating from cell deposits in the MIC experiments that were cultured on LB agar plus 16 mg/L colistin and 64 mg/L compound C4 before freezing, were inoculated directly into 10 mL LB broth and the resulting cultures used in MIC assays. Frozen stocks of P23 and P23 (*pmrA*) strains treated as described were inoculated directly into 10mL LB broth, in triplicate, until the optical density (OD_{600}) was between 0.6-0.8, samples were taken out for phenotypic (colistin MIC) testing and the rest of the cells were harvested by centrifugation. The resulting colistin MIC values for the P23 and P23 (*pmrA*) samples in the presence of 64 mg/L C4 were 2-8 mg/L and 64 mg/L respectively with cell deposits/heteroresistance found most noticeably in individual wells for the P23 (*pmrA*) strain. However, in comparator experiments using the parent (i.e not previously exposed to the colistin/compound C4 combination) P23 and P23 (*pmrA*) strains the equivalent MIC values were 32 mg/L and 64 mg/L respectively, with cell deposits/heteroresistance clearly in evidence. These data further support a loss of the colistin/compound C4 tolerant phenotype after storage/recovery, indicating that this is not a stable adaptation.

These experiments suggested that the improved tolerance of P23 and P23 (*pmrA*) to compound C4 observed after previous exposure was not stable. When the current generation of cells had been exposed to C4, the next generation may be either more resistant or susceptible to its effects, as measured by colistin MIC. It was concluded that these observations are most likely to be due to heteroresistance rather than genetic mutations. Under challenge by C4, the strain P23 (*pmrA*) in general grew better (reached higher OD600 values), and consistently acquired higher resistance to C4 in the next generation, than did P23. These data suggest that the *pmrA* mutation may reduce the susceptibility of the strain to the effects of C4. Although C4 did not inhibit growth of *K. pneumoniae*, it enhanced the activity of colistin against these cells, as measured by MIC, and appeared to induce colistin heteroresistance in P23 and P23 (*pmrA*) cells. Therefore, if appropriate conditions could be identified to maintain the cell deposit/heteroresistance phenotype, a proteomics experiment may be used in the future to investigate cells newly exposed to the combination of 32 mg/L colistin (sub-MIC of colistin for the P23 strain) and 64 mg/L compound C4.

3.2.8 Compound C4 with meropenem

The data above suggest that C4 can affect colistin susceptibility of both MCR-1-expressing *E. coli* and ArnT-expressing *K. pneumoniae*. As both MCR-1 and ArnT are zinc-dependent metalloproteins, it is possible that compound C4 can act by binding non-specifically to enzyme zinc centres. To investigate whether C4 has any effect upon the activity of other zinc metalloproteins, or of proteins without metal ions, that are involved in bacterial resistance to antibiotics, the metallo-β-lactamases NDM-1 and IMP-1 as well as class A serine-β-lactamase KPC-2 were studied. The effects of compound C4 against IMP-1, and NDM-1-producing *E.coli*, and KPC-2-producing *K.pneumoniae*, laboratory strains were investigated using MIC assays measuring susceptibility to the carbapenem (β-lactam) antibiotic meropenem (MEM) (**Table 3-4**). As pterostilbene has previously been shown to potentiate meropenem activity against metallo-β-lactamase-producing Gram-negative strains[209], this was included as a comparator. The MEM MIC values were 8 mg/L for the NDM-1-producing *E. coli* strain (MG1655 + pSU18), 4 mg/L for the NDM-1-producing *E. coli* strain (Top10 + pSU18) and 2 mg/L for the NDM-1-producing *E. coli* strain (DH5α + pUBYT) respectively. MEM MIC values for IMP-1-producing *E. coli* (MG1655 + pSU18) and KPC-2 producing *K. pneumoniae* (pUBYT) strains were 16 mg/L and 32 mg/L respectively. C4 (up to 128 mg/L) did not potentiate meropenem activity against NDM-1-producing *E.coli* strains while 32 mg/L pterostilbene restored susceptibility to the sensitive level. Either 128 mg/L C4 or 64 mg/L pterostilbene reduced MEM MIC values for the IMP-1-producing *E. coli* strain to 4 mg/L (4-fold reduction). For the KPC-2-producing *K. pneumoniae* strain, C4 has no effect on the MEM MIC value while both 32 mg/L and 64 mg/L Pterostilbene reduced the MIC to 16 mg/L (a 2-fold reduction, usually considered within the error of MIC experiments). In addition, the meropenem MICs for 4 NDM-1-producing clinical isolates (C107, C109, C114, C134) of four different Gram-negative species were tested. Neither the addition of pterostilbene nor C4 potentiated meropenem activity when treating any of the clinical isolates. Pterostilbene reduced the colistin MIC value for the *K. pneumoniae* C134 strain by 2-fold, but this was not regarded as significant. The MIC results for these four clinical isolates suggested that, in contrast to the results obtained for laboratory strains, pterostilbene did not have effects on NDM-1-producing clinical strains.

Table 3-4. Meropenem MICs with C4 and PT.

Strain	Meropenem	Meropenem + 64 mg/L C4	Meropenem + 128 mg/L C4	Meropenem + 32 mg/L PT
NDM-1 *E.coli* Top10 (pSU18)	4	4	4	2
NDM-1 *E.coli* DN5a (pUBYT)	2	2	2	1
NDM-1 *E.coli* MG1655 (pSU18)	8	8	8	2
IMP-1 *E.coli* MG1655 (pSU18)	16	8	4	8
pSU18 empty vector *E.coli* MG1655	0.25	0.25	0.25	0.25
KPC-2 *K. pneumoniae* (pUBYT)	32	32	32	16
pUBYT empty vector *K. pneumoniae* Ecl8	0.25	0.25	0.25	0.25
C107 (NDM-1 *E. cloaceae*)	4	4	4	4
C109 (NDM-1 *P. stuartii*)	16	16	16	32
C114 (NDM-1 *E. coli*)	8	8	8	8
C134 (NDM-1 *K. pneumoniae*)	4	4	4	2

3.2.9 Compound C4 with streptomycin and spectinomycin

The results above showed that compound C4 can potentiate colistin activity against ArnT-expressing *K. pneumoniae*, as well as reducing the carbapenem (meropenem) MIC value for IMP-1 (but not NDM-1) metallo-β-lactamase producing *E.coli*. To further investigate the possibility that compound C4 had a broader effect upon antimicrobial susceptibility, the effect of co-administration on the MICs of the aminoglycoside antibiotics streptomycin and spectinomycin was investigated. *K. pneumoniae* is known for being resistant to aminoglycoside antibiotics (e.g. streptomycin and spectinomycin)[217]. Aminoglycoside resistance is commonly mediated by aminoglycoside modifying enzymes or 16S ribosomal RNA methylases in *K. pneumoniae*[217], but can also be associated with efflux pump activity[218]. The streptomycin MIC against P23 and P23 (*pmrA*) cells was determined to be more than 128 mg/L, while the spectinomycin MIC against P23 and P23 (*pmrA*) was determined to be 8 mg/L. The addition of compound C4 (64 & 128 mg/L) or pterostilbene (32 mg/L) had no effect on either streptomycin MIC values nor spectinomycin MIC values for P23 and P23 (*pmrA*) cells. These results suggested that neither pterostilbene nor C4 affect aminoglycoside susceptibility in the two *K. pneumoniae* strains.

3.3 Discussions

In this chapter, we demonstrated that compound C4 potentiated colistin activity against MCR-1 mediated colistin-resistant *E.coli* and *K. pneumoniae* strains, as well as the ArnT-mediated colistin resistant *K. pneumoniae* strains P23 and P23 (*pmrA*). This evidence suggested that the compound C4 may not just be a specific inhibitor of MCR-1, but instead has broader capability to potentiate colistin activity against *E.coli* and *K. pneumoniae* strains where resistance is caused by other factors. Because the compound C4 did not have any inhibitory effects against bacterial growth, and the presence of C4 did not potentiate colistin activity against colistin-sensitive strains (e.g., *E. coli* ATCC-25922, *K. pneumoniae* Ecl8), the mechanism of action of compound C4 may be relevant to modifications causing colistin resistance. As most factors that cause colistin resistance involve modifications to the cell membrane, we investigated whether compound C4 could permeabilize the cell membrane of MCR-1-producing, colistin-resistant, *E. coli* NCTC-13846 cells. The NPN and PI uptake assays, that target the outer membrane and the inner membrane respectively, as well as vancomycin MIC assays were performed. The outcomes all suggested that the compound C4 did not permeabilize the cell membrane.

The PI assay had shown that the presence of the compound C4 did not promote the inner membrane permeability of colistin. Compound C4 contains iodide which is a known and effective fluorescence quencher[219,220]. The reduction of NPN fluorescence observed in the presence of C4 shows that C4 is able to quench NPN fluorescence. However, compound C4 did not quench the fluorescence of PI to the same extent as for NPN, while reducing the increase in PI fluorescence when colistin is present. This differs from the effect of pterostilbene, which showed no effect upon PI fluorescence in the presence of absence of colistin. These data suggest that the mechanisms of action of pterostilbene and C4 may not be the same. It is possible that C4 affects the ingress of PI (or colistin) into the cell, or that C4 can affect the binding of PI to the cytoplasmic DNA after colistin permeabilizes the membrane. It is still possible that the presence of the compound C4 improves outer membrane permeability of colistin. It will be interesting to verify this assumption in the future using an optimized NPN assay.

For the two ArnT-producing colistin resistant *K. pneumoniae* strains P23 and P23 (*pmrA*), we did not obtain significant differences between them in colistin MICs, although the colistin MIC value of P23 is in theory expected to be higher than that for P23 (*pmrA*)[32]. However, we found white colonies in the 96-well microplate when P23 or P23 (*pmrA*) cells were treated with the combination of colistin and the compound C4. Cells of these white colonies were viable. The cells kept growing and eventually turned the well turbid outside the time period for MIC determination. The cells selected from the white colonies usually showed adaption to the combined treatment of colistin and the compound C4, compared to the initial generation, as evidenced by an increase in the colistin MIC values. However, our investigations suggested that the improved adaption was not stably preserved between generations of cells. Sometimes, cells selected from the white colonies had similar, or even worse performance (as measured by colistin sensitivity in MIC assays), compared to the

behaviour of the parent generation. This indicated that the occurrence of white colonies of P23 or P23 (*pmrA*) strains was more likely to be caused by heteroresistance rather than gene mutations.

Heteroresistance in the antibiotic resistance context normally refers to the phenomenon where subpopulations of seemingly isogenic bacteria exhibit various susceptibilities to an antibiotic (e.g. colistin)[221]. Heteroresistance was observed in both Gram-negative and Gram-positive bacteria such as *Staphylococcus aureus, Acinetobacter baumannii* and *Klebsiella pneumoniae*[221]. Alteration of the PmrAB and PhoPQ two component regulatory system was identified as a common cause of colistin heteroresistance in *K. pneumoniae*[222–224]. However, these studies of colistin heteroresistance in *K. pneumoniae* focus on the variability of colistin MICs, and the genomic profile of the heteroresistance subpopulation, without describing the appearance of the heteroresistance subpopulation. Chemical communication of antibiotic resistance from the more resistant members of the population may protect less resistant bacteria[225], which may be one reason for formation of cell deposits. Although a few cell deposits remained when 100 rpm constant shaking was introduced, their size was much smaller indicating only a minority of cells were 'naturally' highly resistant. It is possible that colistin-susceptible cells in the centre of the cell deposit were protected from antibiotic challenge.

In addition to the heteroresistance triggered by genomic alterations in *K. pneumoniae,* differences in levels of biofilm formation may be a possible explanation[226] of the observed heteroresistance phenomenon in *K. pneumoniae* P23 and P23 (*pmrA*). Biofilms mainly consist of extracellular polymeric substances including polysaccharides, proteins, lipopeptides and DNA[227,228]. Communities of bacteria are enclosed in biofilms on a variety of surfaces inside the body and on invasive devices (e.g catheters)[229–231]. The formation of biofilms protects *K. pneumoniae* from antibiotics thereby reducing the effectiveness of antibiotic treatment[227–233]. Although *K. pneumoniae* can form biofilms, the biofilm producing capacity varies between strains ranging from nonbiofilm producers to strong biofilm producers[228]. The first step of forming biofilms is attaching to a surface[233]. Our results showed that the cell deposit phenomenon of P23 and P23 (*pmrA*) mainly occurred in static MIC assays but rarely occurred in experiments with constant shaking. The constant shaking may impede the formation of biofilms of *K. pneumoniae* thereby cell deposits did not occur and lower colistin MICs were obtained with the presence of compound C4 compared to static conditions. Besides, our attempts to isolate the isogenic subpopulation from cell deposits failed, suggesting that the phenomenon may not be caused by stable genomic alterations. It is possible that C4 selected *K. pneumoniae* P23 and P23 (*pmrA*) cells that capable of producing biofilms to a certain extent, or induced these cells to produce biofilms. Because the phenomenon of cell deposition was not observed with the combination of colistin and pterostilbene in either static or shaking conditions, heteroresistance and/or biofilm formation appears not to be a consequence of pterostilbene exposure.

The iron-binding protein lactoferrin is reported to interfere with biofilm formation by scavenging iron[234,235]. MIC experiments with lactoferrin, colistin and C4 may be helpful to investigate whether biofilm formation is

relevant to the cell deposit phenomenon. Reported biofilm inhibitors such as bromo-pyrrole, furanone, imidazole and indole, etc[233] may also be useful for this purpose. It is also notable that these reported biofilm inhibitors have molecular scaffolds that are similar in some respects to that of compound C4.

In addition to its effects upon colistin susceptibility, the compound C4 potentiated activity of the carbapenem antibiotic meropenem against *E.coli* carrying the IMP-1 metallo-β-lactamase. However, compound C4 did not potentiate meropenem activity against an *E.coli* strain carrying the NDM-1 metallo-β-lactamase, or a *K. pneumoniae* strain carrying the serine-β-lactamase KPC-2. Selective reduction in the meropenem MIC of the IMP-1-producing *E.coli* strain by the compound C4 is like the behaviour of pterostilbene, which is reported to inhibit NDM-1[209], and in our experiment did not show inhibition of IMP-1-producing *E.coli*. However, it is too early to conclude that the MEM MIC reduction is directly attributed to the inhibition of IMP-1. Further confirmation, with additional IMP-1 producing strains and biochemical assays of purified IMP-1 activity using nitrocefin, is needed to further provide information on the inhibitory effect of compound C4 on IMP-1, and to investigate whether compound C4 is a more general metallo-beta-lactamase inhibitor. To verify if the compound C4 can generally sensitise *K. pneumoniae* cells to antibiotics, streptomycin and spectinomycin susceptibility tests were performed on *K. pneumoniae* strains in the presence of the compound C4. The MIC values suggested that the compound C4 did not potentiate the activity of either streptomycin or spectinomycin.

In our study, compound C4 is a versatile compound which had similar effects to those of osthole on colistin resistant *E.coli* and *K. pneumoniae strains,* and like pterostilbene also improved meropenem activity against *E.coli* strains producing a metallo-β-lactamase, in this case IMP-1. It was less toxic to cells compared to the previously characterised compounds osthole, pterostilbene or the non-specific zinc chelator DPA. Moreover, compared to pterostilbene and osthole, compound C4 has better solubility in DMSO. Although our attempts to identify routes to chemical modification of the C4 small molecule scaffold that retain its activity against MCR-producing strains have so far been unsuccessful, possibilities for future exploration still remain. Even though the working mechanism of compound C4 remains not clear, some clues were obtained. Permeability assays with compound C4 have shown C4 cannot permeabilize the outer membrane of *E.coli* cells. Compound C4 is a hydrophobic molecule, is generally non-toxic to cells when colistin is not present and only shows effects in the presence of antibiotics, which indicates that a compromised outer membrane caused by colistin may be essential for C4 to get into the cells. Considering that compound C4 showed effects on MCR-1, ArnT and IMP-1 producing cells, one hypothesis is that compound C4 may bind to the zinc ions of these proteins. In the *E. coli* and *K. pneumoniae* strains investigated here, colistin resistance is caused by modifications to the lipid A phosphate. These modifications are mediated by MCR-1 and ArnT, that catalyse the addition of phosphoethanolamine and 4-amino-4-deoxy arabinose to lipid A, respectively. Because compound C4 works on colistin resistance conferred by both of these pathways, and does not potentiate colistin activity against colistin-susceptible strains, it may inhibit the activity of MCR-1 and ArnT via binding to the zinc in these

enzymes. If C4 does not bind to MCR-1 nor ArnT protein, it may act with some component of the cell membrane shared by all tested cells.

3.4 Methods and materials

3.4.1 Bacterial strains and chemicals

Chemicals are mostly purchased from Sigma-Aldrich unless otherwise specify. All clinical isolates including *E.coli* and *K.pneumoniae* are kindly provided by Dr. Jon Tyrrell from School of Cellular and Molecular Medicine (CMM), University of Bristol. The laboratory bacteria strains used in this study and chemicals that are not from Sigma-Aldrich are listed in the following **Table 3-6**.

Table 3-5. Laboratory bacteria strains, plasmids and chemicals used in this study.

Chemical	**Manufacturer**
Pterostilbene	Cayman Chemical, USA
Osthole	Santa Cruz Biotechnology, USA
4-iodoisoxazole	Fluorochem, UK
Meropenem	Zeneca, UK
Vancomycin HCl	Serva Electrophoresis, Germany

***E.coli* strains**	**Source**
MCR-1 BL21 (DE3) (pET24a) BL21 (DE3) (pET24a)	Genlantis, USA
MCR-1 C43 (DE3) (pET24a) C43 (DE3) (pET24a)	New England Biolabs, UK
VIM-1 MG1655 (pSU18) IMP-1 MG1655 (pSU18) NDM-1 MG1655 (pSU18)	Dr. Philip Hinchliffe (CMM), University of Bristol[236].
NDM-1 Top10 (pSU18) NDM1 DH5α (pUBYT)	Dr. Yuiko Takebayashi (CMM), University of Bristol[144,237].

***K.pneumoniae* strains**	**Source**
Ecl8 P23 P23 (*pmrA*)	The group of Prof. Matthew Avison (CMM), University of Bristol[32].
KPC-2 Ecl8 (pUBYT)	Dr. Catherine Tooke (CMM), University of Bristol[238].

3.4.2 Antimicrobial susceptibility testing (MIC determination)

The bacterial strains were cultured on Luria-Bertani (LB) agar plate at 37 °C for overnight. PBS or CAMHB suspension of the bacterial bacteria from an overnight agar plate or liquid culture was made and adjusted to OD_{600} 0.08-0.1 (equivalent to 0.5 McFarland standard). Then the suspension was diluted 1:100 with CAMHB. The antibiotics stock solution was prepared in the sterile 7 mL Bijou sample container (Greiner). Antibiotics (colistin sulfate, vancomycin hydrochloride and meropenem) concentration series solution was prepared in cation adjusted Mueller-Hinton broth (CAMHB) in the 96-well plate (Corning 3595). For assays including inhibitors/compounds, the inhibitors/compounds were assembled with the antibiotics before the addition of colistin. The final volume of each well was 75µL before bacterial inoculum and the concentration of chemicals should be twice the final concentration. For each assay, one row of antimicrobial dilutions was set as a negative control with no bacterial inoculation to ensure no contamination. For wells with bacteria, 75 µL of the diluted bacterial suspension was added into each well. This yielded a final bacterial concentration of approximately 5 x 10^5 CFU/ml. For the control without bacterial inoculation, 75 µL of CAMHB was added. The final volume of each well was 150 µL. The 96-well plate was then incubated at 37°C (with low evaporation lid on) for 16-20 hours and then inspected by the plate reader (OD_{600}) and visual inspection. The microdilution method followed the guideline of Clinical and Laboratory Standards Institute (CLSI) and ECUAST's for antimicrobial susceptibility testing. For colistin MIC assays, MIC determination against *E.coli* NCTC-13846 and ATCC-25922 strains were performed in each assay for the quality control purpose.

3.4.3 Checkerboard colistin MIC assay

The checkerboard method used was the broth microdilution assay performed, in triplicate, in a 96-well plate with the final volume of 150 µL, which is similar to the above antimicrobial susceptibility testing. The bacteria strains were cultured in Luria-Bertani (LB) agar plate at 37 °C for overnight. PBS suspension of the bacteria from the agar plate was made and adjusted to OD_{600} 0.08-0.1 (equivalent to 0.5 McFarland standard). Then the suspension was diluted 1:100 with CAMHB. The concentration serial of inhibitors was first prepared in the plate and the colistin concentration serial was prepared afterwards to minimize dilutions of colistin. 150 µL 2x the highest concentration of compounds was added into each well of column 8 (A-G) and 150 µL 4x the highest concentration of compounds was added in the well H8. The 2x serial dilution with CAMHB was performed to spread the compounds from column 8 to column 2 with a final volume of 75 µL per well. 75 µL CAMHB was added in column 1. Afterwards, 75µL 2x the highest concentration of colistin was added into each well of Row H. The 2x serial dilution was performed to spread the colistin from Row H to Row B with a final volume of 75 µL per well. Subsequently, 75 µL of the diluted bacterial suspension was added into each well. This yielded a final bacterial concentration of approximately 5 x 10^5 CFU/ml and a final volume of 150 µL per well. The 96-well plate was then incubated at 37°C (with low evaporation lid on) for 16-20 hours and then inspected by the plate reader (OD_{600}) and visual inspection. Colistin MIC determination against *E.coli*

NCTC-13846 and ATCC-25922 strains were performed in each assay for the quality control purpose. The relative OD_{600} values were calculated by the following equation:

$$\text{Relative } OD_{600} = \frac{OD_{600}(\text{strain incubated with the compound })}{OD_{600}(\text{strain alone })}$$

3.4.4 Growth curves of bacteria

The tested strains were cultured in Luria-Bertani (LB) agar plate at 37 °C for overnight. Broth or PBS suspension of the test strains from the overnight agar plate was prepared and adjusted to OD_{600} 0.08-0.1 (equivalent to 0.5 McFarland standard). This suspension should contain approximately 1 to 2 x 10^8 CFU/mL bacteria. Then the suspension was diluted 1:100 with cation-adjusted Mueller-Hinton broth (CAMHB). The concentration serial of antibiotics (colistin and vancomycin) with/without the presence of compounds (e.g., compound C4 and pterostilbene) were prepared in the 96-well plate (Corning 3595) to a final volume of 90 µL. The concentration of the prepared solution in each well should be twice the target value. 90 µL of prepared bacteria inoculum was then added into the 96-well plate to the final volume of 180 µL. The bacteria were cultured at 37 °C with shaking at 100 rpm (with low evaporation lid on), and the growth of the bacteria was estimated by measuring the OD_{600} every 60 minutes for 18 hours.

3.4.5 Large-scale growth of P23 and P23 (*pmrA*) samples

The P23 and P23 (*pmrA*) strains were grown overnight on the LB agar with the presence of 16 mg/L colistin and 64 mg/L C4. Three single colonies on the plate were selected for each strain and transferred respectively into 10 mL CAMHB with 16 mg/L colistin and 64 mg/L C4 for overnight culture. The overnight culture was then diluted in 1:100 with 10 mL CAMHB with 16 mg/L colistin and 64 mg/L C4 in 50 mL sub-culture flasks in triplicate. The cells in flasks were culture at 37°C at 180 rpm in the incubator till the optical density of cells reaching OD_{600} 0.6-0.8. For each replicate, samples were taken out for the phenotypical test (colistin MIC assay) and the rest of the cells were harvested by centrifuge.

3.4.6 NPN uptake assay

10 mM stock solution of NPN was prepared in ethanol and diluted in the assay buffer (5 mM HEPES, 5 mM glucose, pH 7.2) to a concentration of 20 µM.

Direct NPN uptake assay

The bacteria in Luria-Bertani (LB) broth were grown to an OD_{600} of 0.5–0.6. The cells were then harvested by centrifuge at 3000 RCF and resuspended in the assay buffer to a concentration of OD_{600}=1. Then, 99uL of

resuspended cells and 99uL assay buffer containing 20 μM NPN were mixed in the black-walled 96-well microplate. The final concentrations of NPN and testing cells in the plate were 10 μM and OD_{600}=0.5. The fluorescence was immediately monitored every 30 seconds for 5 minutes inside the CLARIOstar plate reader (BMG LABTECH) with excitation wavelength of 350 nm and emission wavelength of 420 nm at 25°C. Afterwards, either 2 μL of a chemical compound, or the corresponding solvent, was added to the corresponding wells and fluorescence was immediately monitored for another 10 minutes. The fluorescence was read from the top of the plate.

Indirect NPN uptake assay

The bacteria were grown to an OD_{600} of 0.5 – 0.6 in LB broth with the chemical compound (e.g. colistin/compound C4) at the target concentration. The cells were then harvested by centrifuge at 3000 RCF, washed and resuspended in the assay buffer to a concentration of OD_{600}=1. Then, 100 uL of resuspended cells and 100 uL assay buffer containing 20 μM NPN were mixed in the black-walled 96-well microplate. The final concentrations of NPN and testing cells in the plate were 10 μM and OD_{600}=0.5. The fluorescence was immediately monitored every 30 seconds for 15 minutes inside the CLARIOstar plate reader (BMG LABTECH) with excitation wavelength of 350 nm and emission wavelength of 420 nm at 25 °C. The fluorescence was read from the top of the plate.

3.4.7 PI uptake assay

5 mM stock solution of PI was prepared in DMSO and diluted in the assay buffer (5 mM HEPES, 5 mM glucose, pH 7.2) to a concentration of 20 μM. The bacteria in LB broth were grown to an OD_{600} of 0.5–0.6. The cells were then harvested by centrifuge at 3000 RCF and resuspended in the assay buffer to a concentration of OD_{600}=1.0. Then, 99uL of resuspended cells and 99uL assay buffer containing 20 μM PI were mixed in the black-walled 96-well microplate. The final concentrations of PI and testing cells in the plate were 10 μM and OD_{600}=0.5. The fluorescence was immediately monitored every 30 seconds for 5 minutes inside the CLARIOstar plate reader (BMG LABTECH) with excitation wavelength of 535 nm and emission wavelength of 617 nm at 25°C. Afterwards, either 2 μL of a chemical compound, or the corresponding solvent, was added to the corresponding wells and fluorescence was immediately monitored for another 10 minutes. The fluorescence was read from the top of the plate.

Chapter 4. Computational investigation of protein:ligand complexes of MCR-1 and other zinc metalloproteins

4.1 Introduction

Chapter two identified significant activity reduction of the MCR-1 catalytic domain (MCR-1^{CD}) in the presence of TGA using the chromogenic PnP-PEtN assay. In this chapter, computational approaches including molecular dynamics (MD) and molecular docking are used to investigate the possible binding mode of the inhibitor thioglycolic acid (TGA) to the catalytic domain of MCR-1 (MCR-1^{CD}). First, a computational workflow for modelling protein:ligand complexes of zinc containing proteins was developed and tested. This work has been published in the *Journal of Chemical Information and Modeling* (DOI: 10.1021/acs.jcim.1c01109). This manuscript is integrated as a whole in this chapter. The supporting information for the manuscript is also included. An introduction, and details of the simulation methods used in this Chapter are included in the manuscript.

The pipeline was then used to model complexes of MCR-1. Due to the lack of available crystal structures of MCR-1 complexes, a crystal structure of the non-covalent complex with phosphoethanolamine of ICR^{Mc} (the intrinsic colistin resistance enzyme from *Moraxella catarrhalis*) was first modelled. Although ICR is not a mobile colistin resistance determinant, it represents the closest known ortholog to MCR-1 and MCR-2. The key structural and functional characteristics of the catalytic domain of ICR^{Mc} are similar to those of MCR-1.[183] Because the crystal structure of the MCR-1^{CD}:TGA complex is not available, both molecular docking and MD simulations were used to predict the binding conformation of TGA to MCR-1^{CD}. This work, which is not included in the published manuscript, is described in section 4.3 of this Chapter.

4.2 Development of a multiscale workflow for modelling zinc containing protein-ligand complexes

Contribution statement

The work is published in the *Journal of Chemical Information and Modeling* (doi: 10.1021/acs.jcim.1c01109). All simulations and data analysis were done by Zongfan Yang. Zongfan Yang wrote the complete manuscript draft, with editing by Prof. Jim Spencer and Prof. Adrian Mulholland and comments and suggestions by all other authors. The author's version of the accepted manuscript and the supporting information are presented below.

A Multiscale Workflow for Modelling Ligand Complexes of Zinc Metalloproteins

Zongfan Yang[a,b], Rebecca M. Twidale[a], Silvia Gervasoni[a,c,d], Reynier Suardíaz[a,e], Charlie Colenso[a,b], Eric J. M. Lang[a], James Spencer[b] and Adrian J. Mulholland[a]*

[a]Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, BS8 1TH, U.K.

[b]School of Cellular and Molecular Medicine, University of Bristol, Bristol, BS8 1TD, U.K.

[c]Department of Pharmaceutical Sciences, University of Milan, Via Mangiagalli, 25, I-20133 Milano, Italy.

[d]Department of Physics, University of Cagliari, S.P. Monserrato-Sestu km 0.700 I-09042, Monserrato (CA), Italy (present address)

[e]Departamento de Química Física, Facultad de Química, Universidad Complutense, 28040 Madrid, Spain (present address)

KEYWORDS: Zinc, Metalloenzyme, Molecular Dynamics, DFTB3, QM/MM, Metallo-beta-lactamase

ABSTRACT: Zinc metalloproteins are ubiquitous, with protein zinc centers of structural and functional importance, involved in interactions with ligands and substrates and often of pharmacological interest. Biomolecular simulations are increasingly prominent in investigations of protein structure, dynamics, ligand interactions and catalysis, but zinc poses a particular challenge, in part because of its versatile, flexible coordination. A computational workflow generating reliable models of ligand complexes of biological zinc centers would find broad application. Here we evaluate the ability of alternative treatments, using (non-bonded) molecular mechanics (MM) and quantum mechanics/ molecular mechanics (QM/MM) at semiempirical (DFTB3) and density functional theory (DFT) levels of theory, to describe the zinc centers of ligand complexes of six metalloenzyme systems differing in coordination geometries, zinc stoichiometries (mono- and di-nuclear), and the nature of interacting groups (specifically the presence of zinc - sulfur interactions). MM molecular dynamics (MD) simulations can overfavor octahedral geometries, introducing additional water molecules to the zinc coordination shell, but this can be rectified by subsequent semiempirical (DFTB3) QM/MM MD simulations. B3LYP/MM geometry optimization further improved the accuracy of description of coordination distances, with the overall effectiveness of the approach depending upon factors including the presence of zinc – sulfur interactions that are less well described by semiempirical methods. We describe a workflow, comprising QM/MM MD using DFTB3 followed by QM/MM geometry optimization using DFT (e.g., B3LYP), that well describes our set of zinc metalloenzyme complexes and is likely to be suitable for creating accurate models of zinc protein complexes when structural information is more limited

Introduction

Zinc has an essential role in biology, with zinc sites contributing to the structural integrity, stability and catalytic activity of a wide range of proteins with highly diverse functions in both pro- and eukaryotic cells.[239] Many such proteins are actual or potential targets for pharmaceutical intervention, including by small molecule therapeutics used to treat conditions such as hypertension, cancer, infectious disease etc. Interaction of small molecule ligands with their protein targets frequently involves participation of bound zinc; in some cases, interactions involving zinc ions are essential to ligand binding and removal of zinc abolishes this.[240]

Molecular simulation methods play an increasingly prominent role in ligand and drug discovery, driven in part by hardware and software innovations and by the growing availability of high-resolution crystal structures for many biologically and/or pharmacologically important protein targets.[241,242] Their application to zinc metalloproteins is desirable, given the abundance of zinc-containing systems in the proteomes of many species, including humans. Unfortunately, many of the properties that enable zinc to play a diverse range of roles in biological systems make modelling protein zinc centers and their complexes challenging. These include the ability of zinc to coordinate different types of ligands, including N, O and S; flexibility of coordination geometry (zinc is 6-coordinated (octahedral) in aqueous solution but may be tetrahedral, 5- or, in some catalytic sites, 6-coordinated in proteins[243–246]);[247,248] the ability of coordinating water molecules to exchange with substrates or inhibitors during complex formation;[245,249] and the existence of single and multi-nuclear sites. As we have recently demonstrated,[139] treatment of zinc centers as point charges (as in for example many ligand docking methods) often leads to unrealistic coordination geometries.

A variety of treatments has been applied to zinc (and other metal) centers in proteins in molecular dynamics (MD) simulations. In the widely used AMBER MD package (https://ambermd.org/)[98–101] bonded,[123] non-bonded[124] and cationic dummy atom[125] approaches have all been implemented as molecular mechanics (MM) treatments for metal ions and their complexes. Bonded models do not allow for ligand exchange and/or changes in zinc coordination geometry, while cationic dummy atom approaches require a pre-defined zinc coordination geometry, limiting exploration of alternative ligation patterns, and are more challenging to set up. In contrast, nonbonded models, as typified by the Lennard-Jones (LJ) 12-6 model, are widely used owing to their simplicity and transferability.[126] The LJ12-6 model does, however, require user selection of the most appropriate parameter set: the IOD set, specifically designed to reproduce ion-oxygen distances, the HFE set specifically designed to reproduce hydration free energies, or the CM set designed as a compromise set for more general application.[124] More recently, the LJ12-6-4 non-bonded model, proposed and parameterized for divalent metal ions by Li and Merz,[127] includes an additional C4 term to represent the impact of ion-induced dipole interaction, and is claimed to simultaneously reproduce the hydration free energy (HFE), ion-oxygen distance (IOD) and coordination number, whereas the LJ12-6 model reproduced only one or two these experimental values in a given simulation.[127–129] Since the LJ12-6-4 model fulfils many of the performance requirements for MD simulations using a single parameter set, its ease of use in practice has led to its wide application in simulations of divalent metal cations (e.g., Zn^{2+}, Mg^{2+} and Mn^{2+}, etc.) involved in biological systems.[250–254]

Although the MM MD approach is advantageous in terms of computational efficiency, its accuracy is highly dependent on the predefined MM force fields. However, most current MM force fields do not accurately describe interactions of protein zinc centers, and force field parameter optimization may be required for a particular biological system.[105] The lack of any accurate description of polarization effects and the ability to simulate charge transfer, as well as the flexible coordination geometry of zinc, makes correct simulation of zinc ions in proteins using MM methods difficult. Quantum mechanics (QM) provides another approach to modelling interactions of zinc ions in proteins. QM methods can provide more accurate description of interactions involving zinc ions than MM methods, but at a computational cost that increases rapidly with increasing system size.[255] Accordingly, quantum mechanics/molecular mechanics (QM/MM) approaches provide a balance between calculation accuracy and efficiency. QM/MM approaches have been used in efforts to obtain more accurate descriptions of protein zinc centers in simulations.[110,256,257] Results can be dependent upon the appropriate partitioning of the system under investigation between the QM and MM regions.[258]

The computational requirements, and potential accuracy, can also be dependent upon the level of QM theory applied in QM/MM calculations. QM/MM calculations can be performed using semiempirical QM methods, ab initio QM or density functional theory (DFT) treatment. One of the most popular semiempirical QM methods is density functional tight binding (DFTB) approaches are derived in the framework of DFT.[259] DFTB methods have previously been used to simulate a variety of metalloprotein systems.[109,113,114,171,260–262] Higher levels of accuracy are in principle possible with ab initio QM techniques, but the computational costs of wave function optimization make ab initio QM/MM calculations impractical for routine application. DFT calculations, e.g. with hybrid functionals such as B3LYP provide a good balance between accuracy and computational cost in describing the structures of transition metal complexes[132–134] and have been widely used for zinc protein studies.[15,135,136] The choice of QM method then represents a compromise between accuracy and computational cost.

We have previously[263] developed a computational workflow, involving docking, MM and QM/MM simulations at two levels of QM theory, with which we have successfully reproduced crystal structures of

complexes of the zinc-dependent (metallo-) beta-lactamase (MBL) IMP-1 (imipenemase-1) with thiol-based mercaptomethyl thiazolidine (MMTZ) inhibitors.[264] In this work we develop and expand this work to test the ability of multiscale modelling to replicate crystal structures of a wider range of zinc metalloenzyme complexes varying in their zinc stoichiometry (mono- and dinuclear zinc sites), zinc coordination number (CN) or coordination geometry, and the identity of zinc ligating atoms from both protein and small molecule ligands. (Specifically, these include protein Asp, His and Cys ligands, representing Zn - N, Zn - O and Zn - S interactions, respectively; and thiolate, carboxylate and hydroxamate small molecule ligands). Our initial model system, the MBL Sfh-I in complex with an MMTZ inhibitor, features a typical tetrahedral zinc geometry with His, Asp, Cys and inhibitor thiolate ligands.[265] Subsequently we extend our investigations to a further five systems (**Figure 1**): Angiotensin-Converting Enzyme (ACE) complexed with the thiol inhibitor L-captopril (PDB: 2X8Z),[266] ACE-2 (the ACE isoform involved in viral spike protein processing during infection by SARS-CoV-2) complexed with the carboxylate inhibitor MLN-4760 (PDB: 1R4L),[267] histone deacetylase (HDAC) 2 complexed with the hydroxamate inhibitor SAHA (PDB: 4LXZ),[268] and the dinuclear IMP-1(PDB: 6JED)[269] and L1 MBLs (PDB: 7A63)[270] complexed with thioglycolic acid (TGA) and the hydrolyzed form of the penem antibiotic, faropenem, respectively. Collectively, these represent single-zinc systems with 4- and 5- coordination geometries and dinuclear systems with 4, 5 and 4, 6 coordination with and without sulfur ligands.

The results show that, while MM MD approaches can in some circumstances provide reasonable descriptions of zinc coordination distances, these usually introduce changes to zinc geometry that require rectification by semi-empirical QM/MM MD. The accuracy with which semi-empirical methods can describe protein zinc centers varies, with those involving Zn – S interactions among the most challenging. Our data suggest that a multiscale approach involving increasing levels of theory is necessary to obtain accurate models of zinc enzymes and their complexes, and identify a workflow that may be broadly applicable in such cases.

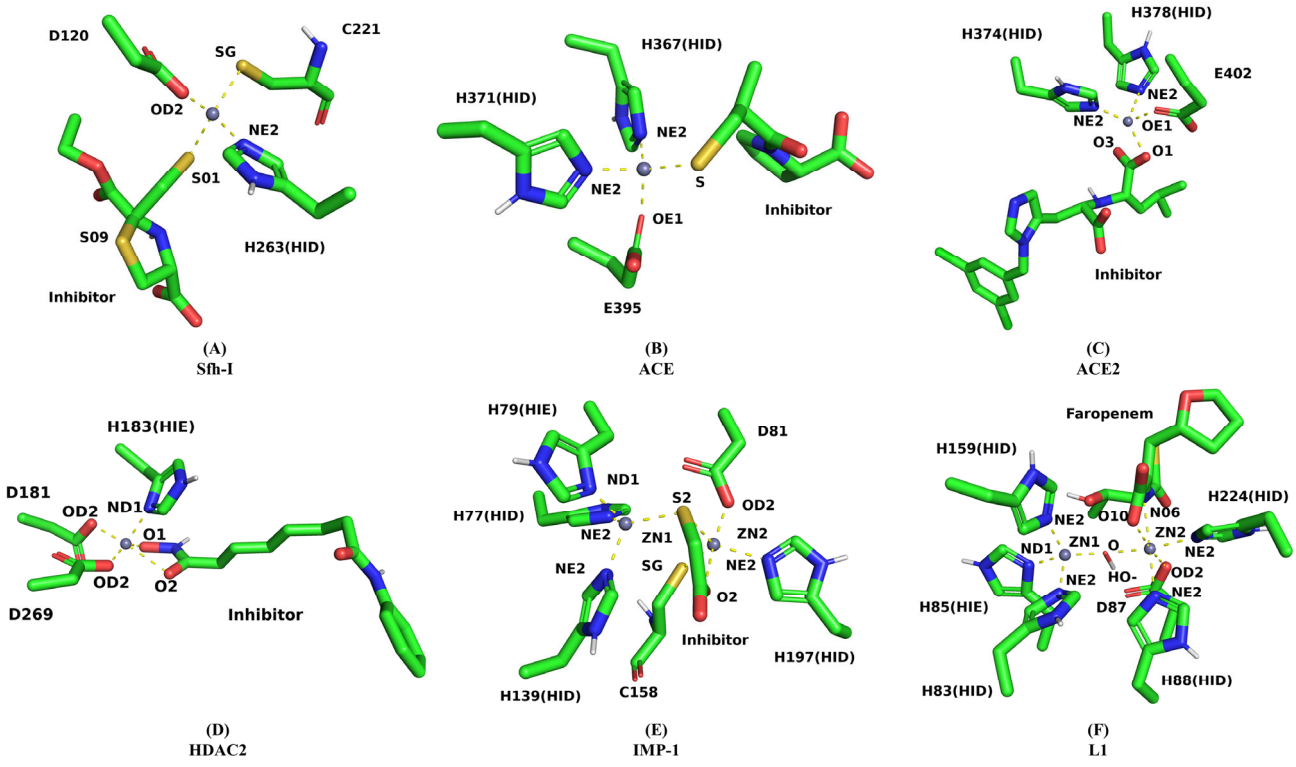


Figure 1. Zinc site geometries of the 6 protein-ligand systems included in this study. (A) Sfh-I MBL with MMTZ inhibitor L-anti-1a, coordination number (CN) = 4, PDB code: 7BJ9;[265] (B) ACE with the thiol inhibitor L-captopril, CN=4, PDB code: 2X8Z;[266] (C) ACE2 with the carboxylate inhibitor MLN-4760, CN=4, PDB code:1R4L;[267] (D) HDAC2 with the hydroxamate inhibitor SAHA, CN=5, PDB code : 4LXZ;[268] (E) IMP-1 MBL with thioglycolic acid (TGA), CN = 4,5, PDB code: 6JED;[269] (F) L1 MBL with hydrolyzed faropenem, CN=4,6, PDB code:7A63.[270] Carbon atoms are in green, nitrogen atoms are in blue, oxygen atoms are in red and sulfur atoms are in yellow, with zinc ions represented as gray balls.

Results

Our long-term goal is to develop a robust computational workflow able to generate realistic models of complexes of zinc metalloproteins. In this work we investigate the ability of different computational approaches to successfully maintain zinc centers in architectures consistent with starting crystal structure(s) for a range of representative complexes. As our aim is to develop a methodology that can be applied to suboptimal starting structures, derived from either poorer quality experimental data or models from e.g. docking experiments, and for which MM MD optimization might be necessary, we tested the inclusion of MM MD steps prior to any QM optimization.

Initial testing: complex of the MBL Sfh-I with the MMTZ inhibitor L-anti-1a.

Initial investigations were carried using the complex of the mono-zinc MBL Sfh-I (a carbapenem hydrolyzing beta-lactamase from the environmental bacterium and occasional pathogen *Serratia fonticola*[271]) with the MMTZ inhibitor L-anti-1a (PDB code: 7BJ9)[265] as a model system. This structure contains a single zinc ion in tetrahedral geometry coordinated by Asp, His and Cys residues and the inhibitor thiolate. Three different modelling methods were tested with the Sfh- I: L-anti-1a complex: MM MD (using a non-bonded model), QM/MM MD with a semi-empirical QM method (DFTB3) and QM/MM geometry optimization/ energy minimization with a DFT QM method (B3LYP).

Four different nonbonded models were applied to model the zinc center of the Sfh- I: L-anti-1a complex and tested in triplicate 100 ns MM MD simulations: the unrestrained LJ12-6 model (LJ12-6), the restrained LJ12-6 model (LJ12-6-R), the unrestrained LJ12-6-4 model (LJ12-6-4) and the restrained LJ12-6-4 model (LJ12-6-4-R). Atomic distance restraints when used were placed between the zinc ion and coordinating protein residue atoms during MD production runs in order to maintain crystallographically observed zinc coordination geometry and leave the fourth coordination site open for small molecule ligand binding. The RMSD plots (compared to the crystal structure) and representative zinc center geometries of individual models are presented in **Figures 2** and **3**, respectively.

Figure 2 shows the RMSD values of the Zn^{2+} binding site in the unrestrained LJ12-6-4 model to be consistently lower than those using the LJ12-6 model, whereas the LJ12-6 and LJ12-6-4 models have similar RMSD values when restraints are applied. Although the mean RMSD value for the MMTZ ligand is about 0.5 Å higher using the restrained LJ12-6-4 models than the restrained LJ12-6 models, the pose of the ligand in the former simulation is more stable. Taken together, the RMSD plots for the ligand and binding sites indicate that protein-ligand interactions are more stable over the duration of the simulation when the LJ12-6-4 model, rather than the LJ12-6 model, is used.

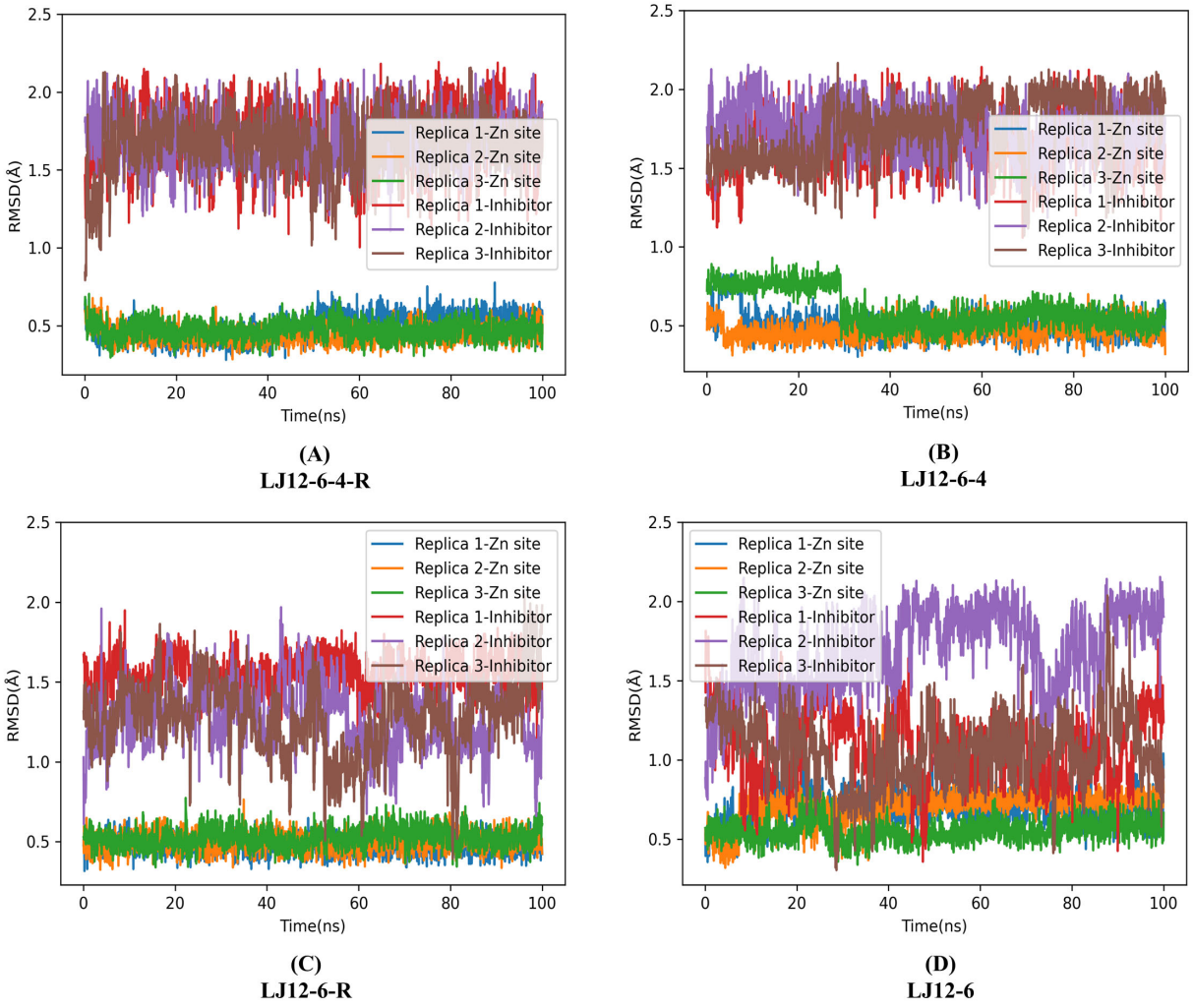


Figure 2. Time-dependence of RMSD values (compared to the crystal structure) for MM MD simulations of Sfh-I:L-anti-1a complex using non-bonded models. Three replicate simulations were performed for each model and each replicate is 100 ns. The 'Zn site' refers to zinc ions and zinc coordinating residues and 'Inhibitor' refers to the zinc bound compound.

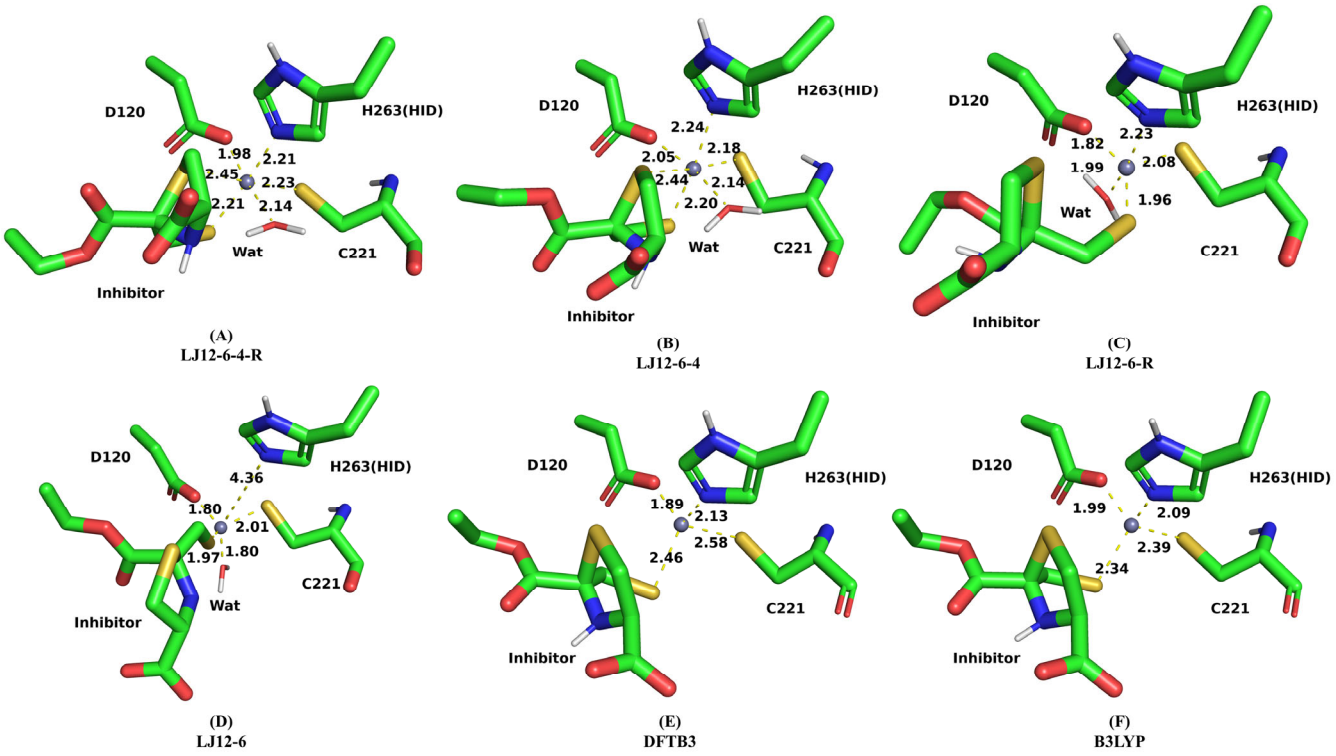


Figure 3. Zinc site geometries of Sfh-I:L-anti-1a complexes. (A-D) The representative zinc site geometry of the four non-bonded models after 100ns MM MD simulation. (E) The representative zinc geometry after 100 ps DFTB3/MM MD simulation. The extra coordination by an additional water molecule was removed. (F) The representative zinc geometry after B3LYP-D3BJ based QM/MM geometry optimization

A comparison of zinc coordination distances obtained from simulations using the four MM models is shown in **Table S1**. The performance of the two simulations using the LJ12-6-4 model is significantly better than that of those using the LJ12-6 model when considering the distances between the zinc ion and coordinating atoms. It is worth noting that during 100ns MD production runs using the unrestrained LJ12-6 model zinc coordination by the residue His263 was lost. Overall, the coordination distances predicted by the LJ12-6-4 models more closely approach the crystallographically determined values (the coordinate error of the crystal structure is 0.10 Å) and the similarity can be slightly improved with the use of distance restraints. In the majority of simulations using the restrained LJ12-6-4 model the distance restraint penalties were not triggered, demonstrating the improved ability of the LJ12-6-4 model, compared to the LJ12-6 model, to describe bond lengths. Our results suggest that the LJ12-6-4 models not only provide positional predictions for binding site residues closer to the crystal structure, but also can better reproduce coordination bond length values than the LJ12-6 model. However, when coordination geometry is considered, all four models showed a strong tendency to increase the coordination number of the Zn^{2+} ion to five or six, through the addition of an extra Zn-coordinating water molecule in the case of the LJ12-6 model, and with both a water molecule and the inhibitor thiazolidine sulfur atom forming additional bonds to Zn^{2+} in the LJ12-6-4 model. For the LJ12-6 model, application of distance restraints partially corrected this, resulting in improved performance yielding zinc coordination numbers (CN) closer to the experimental value.

Semi-empirical QM/MM MD. The data presented above demonstrate that, although the bond lengths between zinc and crystal coordinating ligands were largely consistent (~ 0.10 Å difference on average) with experimental values, the geometries were not. Accordingly, in an effort to improve the outcome, 100 ps DFTB3/MM MD was performed. As the MD trajectory and zinc site coordination were stable, the last frame of the production run of a LJ12-6-4 model was chosen as a typical snapshot after MD simulation for the following DFTB3/MM MD. The Zn^{2+} binding site (i.e., the zinc ion, the side chains of zinc coordinating residues and zinc coordinating water molecules added by MM MD) and the inhibitor were simulated at the semi-empirical level of QM theory using DFTB3 with the 3OB-3-1 parameter set,[131] while the rest of system was simulated using the Amber ff14SB forcefield.[272] The two additional coordination interactions introduced by the MM MD were not retained after the DFTB3 QM/MM dynamics simulation, reducing the Zn^{2+} coordination number to four, which restored the structure to the crystal geometry (**Figure 2(C)**). The Zn^{2+}-ligand coordination

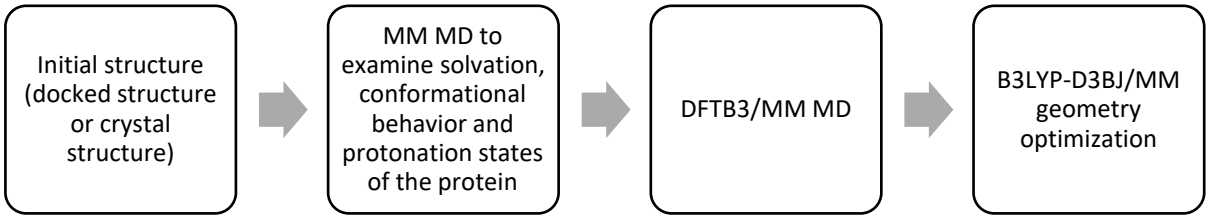
distance was also optimized, getting closer to the value of the crystal structure (**Table S2**). Inspection of the plot of Zn^{2+}-ligand coordination distance against simulation time (**Figure S1**) demonstrated that that the major shift in zinc geometry (specifically removal of 'inappropriate' coordinating atoms) happened at the beginning of the QM/MM MD simulation, with the distance between Zn^{2+} and the additional water molecule immediately rising to above 5 Å, indicating loss of coordination, and that the system was relatively stable after 100 ps. Therefore, a simulation time scale of 100 ps appears sufficient to obtain a stable zinc coordination geometry with DFTB3/MM MD.

B3LYP/MM optimization. Although DFTB3/MM MD treatment substantially improved the quality of the model for the Sfh-I:L-anti-1a complex, discrepancies remained with respect to the experimental crystal structure, specifically regarding the coordination distances for Zn-S interactions involving both the Cys221 ligand and the inhibitor thiolate. Hence a QM/MM energy minimization using higher level DFT theory was performed to further optimize the geometry of the zinc site. The hybrid functional B3LYP, widely used in QM/MM studies of enzymes including zinc-containing proteins[15,132,136,273,274], and the 6-31G(d) basis set were selected for the DFT calculations. Two parameters were tested in the DFT QM/MM geometry optimization: the diffuse function and the empirical dispersion correction. The B3LYP/MM optimization was carried out with and without diffuse functions for heavy atoms, and with and without the D3BJ (D3 version of Grimme's dispersion with Becke-Johnson damping) empirical dispersion correction,[137,138] to test their effects. The results of these comparisons are shown in **Table S2**.

Use of diffuse functions improved the accuracy of Zn^{2+} - ligand coordination distance predictions, particularly for electron-rich sulfur atoms. The absence of empirical dispersion corrections leads to longer Zn^{2+} - ligand bond lengths, reducing the accuracy with which the experimental geometry is reproduced. Accordingly, the results presented here are those obtained using B3LYP/MM optimization with diffuse functions for heavy atoms and GD3BJ empirical dispersion. Compared to the DFTB3-optimized starting structure, we found that this resulted in further improvement in prediction of Zn^{2+}-ligand coordination distances, especially those between the zinc ion and ligating sulfur atoms. Ultimately the values for DFT-refined Zn^{2+} coordination distances were almost identical to those observed in the crystal structure (the total absolute deviation of all zinc ligating distances was 0.05 Å). The geometry of the zinc center in the DFT - based QM/MM optimized structure is shown in **Figure 3**. We found that the atomic distance between the zinc ion and its ligands only changes slightly in the zinc site during the B3LYP geometry optimization (**Figure S2**). We also found no significant difference in accuracy between a fully converged B3LYP-D3BJ/MM geometry optimization, and one progressed over 250 steps. As B3LYP/MM geometry minimization normally takes around 500-750 steps to converge, use of the 250th step as the end point for geometry refinement is thereby much more computationally efficient. A comparison of the Zn^{2+} site geometry predicted from this simulation and the crystal structure is shown in **Figure 4 (A).** In additional experiments, we also attempted application of the B3LYP/MM geometry optimization process directly after the 100ns MM MD simulation, instead of starting from the DFTB3-optimized structure. The results showed that the QM/MM optimization process cannot remove the additional coordinating water molecules to the zinc site introduced by the MM treatment. Instead, B3LYP/MM optimization led to the stable incorporation of additional water molecules into the Sfh-I zinc site, increasing the discrepancy between the modelled zinc site and that observed in the crystal structure.

The results of simulations of the Sfh-I:L-anti-1a complex suggested the workflow shown in **Scheme 1** as an approach capable of accurately describing this tetrahedrally coordinated zinc center. Non-bonded LJ12-6-4 MM MD simulation was first used to examine the general motion of the complex (e.g., examine solvation, conformational behavior and protonation states) and establish the stability of the system over a relatively long timescale. DFTB3 QM/MM MD is then applied to better describe the zinc site, and to rectify changes in coordination introduced by MM treatment. Subsequent B3LYP-D3BJ /MM optimization can be used to further refine the geometry of the zinc site (e.g., atomic distances, particularly for Zn – S interactions).

Scheme 1. Schematic of the workflow for modelling zinc-containing protein-ligand complexes.



Workflow validation: Inhibitor complexes of ACE and ACE2

To assess the broader applicability of this approach, the pipeline described above was then tested with two additional tetrahedral zinc systems (complexes of angiotensin converting enzyme (ACE) and its ACE2 isoform with the inhibitors L-captopril and MLN-4760). Compared to Sfh-I, above, the protein zinc sites lack a Cys ligand, and the inhibitors are coordinated via thiolate and carboxylate groups respectively, so representing a wider range of enzyme zinc centers. In the L-captopril complex[266] the ACE zinc ion is coordinated by three protein ligands (His367, His371 and Glu399) and the captopril thiolate (**Figure 1(B)**). In the ACE2:MLN-4760 complex[267] (a structure determined at lower resolution (3.00 Å) enabling assessment of the performance of the modelling workflow with a lower-quality starting structure) the zinc ion is coordinated by the equivalent protein ligands (His374, His378, Glu402) and the inhibitor carboxylate (**Figure 1(C)**). Based upon results obtained for the Sfh-I:D-anti-1a complex (above) the LJ12-6 models were discarded for the MM MD simulations and the LJ12-6-4 and LJ12-6-4-R models only were used. RMSD plots (compared to the crystal structures) and representative geometries for the ACE and ACE2 zinc centers at different stages of the modelling workflow are presented in **Figures S3 - S6**.

For both the ACE and ACE2 models, the RMSD values of residues within the zinc binding site were consistently lower than 1 Å, suggesting that stable geometry is maintained across the 100 ns of the simulation. (Notably, although the simulations overall remained stable the inhibitor RMSD value of the bound ACE-2 inhibitor in one replicate using the LJ12-6-4 model increased suddenly from ~ 1 Å to 2.5 Å at around 35ns of the simulation and remained at ~ 2 Å for a further 30 ns before reverting to baseline). This restoration of inhibitor binding geometry, and the disappearance of other occasional RMSD peaks, suggested that the MD treatment using LJ12-6-4 models was able to handle incorrect ligand poses and form a reasonable binding geometry. Subsequent DFTB3/MM MD simulations were carried out starting from the last frame of the 100ns MM MD simulation, and were followed by a 250 step DFT-based QM/MM geometry optimization using the B3LYP functional with GD3BJ dispersion correction (B3LYP-D3BJ) and the 6-31+G(d) basis set. In the case of ACE, the DFTB3 MM MD step was trialed starting from the last frame of the unrestrained MM MD simulations using the LJ12-6-4 model, with negligible differences between the starting or end points. For ACE2 DFTB3 QM/MM MD began from the last frame of both the restrained and unrestrained MM MD simulation. Zn^{2+} coordination geometries from the various treatments are reported in **Table S3**. (Note that both oxygen atoms of the ACE2:MLN-4760 carboxylate were included to monitor the performance of the model on Zn^{2+}-carboxylate interactions).

In general, the QM/MM optimized structures were close to the crystal structures and, although the coordination distances were not perfectly predicted, the Zn^{2+} binding geometry in the crystal structure was successfully restored. Alignment of the ACE and ACE2 Zn^{2+} sites to the respective crystal structures is shown in **Figure 4(B) and (C)**. The results showed that the workflow developed using Sfh-I worked well for these additional systems with tetrahedral zinc binding centers that do not contain Cys, with both thiolate (ACE) and non-thiolate (ACE2) small-molecule ligands.

For the ACE2:MLN-4760 complex, **Table S3** also includes the interaction between the zinc ion and the inhibitor O3 atom, even though the crystallographically observed distance (2.62 Å) is beyond the 2.5 Å boundary considered to be the limit for a coordination bond. This is due to our wish to investigate whether the MM and QM/MM models can correctly handle this weaker interaction. The results showed that, during MM simulations using the LJ12-6-4 series models, the Zn^{2+} – O3 distance reduced to form an additional coordination bond not present in the crystal structure, while the QM/MM simulation correctly handles this interaction. The modelled ACE2 structure after QM/MM refinement showed an unexpectedly large (0.3 Å) deviation from the experimentally observed distance for the interaction between Zn and His378-NE2. As this is a low-resolution (3 Å) structure, the ACE2 zinc center was inspected using the CheckMyMetal server[275,276]. The results indicate that the experimentally determined value (2.31 Å) for this Zn^{2+} - N distance is likely to be an outlier, given that Zn^{2+} - N distances in protein zinc sites are mainly distributed in the range: ~1.9 - 2.2 Å, and that the gRMSD value (defined as the RMSD of the observed ligand-metal-ligand angles compared to their idealized values) (23.3°), is also an outlier. These findings, when considered together with the low resolution of this structure, may explain why the QM/MM optimized values in Table S1 lie relatively far from the experimental figures.

Extension to a 5-coordinate system: HDAC2 complex with SAHA

After testing the ability of this approach to model 4-coordinated Zn^{2+} centers, we then tested it on a Zn^{2+} center with 5-coordination. We selected the crystal structure of HDAC2 in complex with the hydroxamate inhibitor suberanilohydroxamic acid (SAHA, also known as Vorinostat and used clinically for treatment of cutaneous T cell lymphoma) solved at a resolution of 1.85 Å (PDB code: 4LXZ)[268] as a model system. In this structure the HDAC2 zinc ion is coordinated by three amino acids (Asp181, His183 and Asp269) and by two oxygen atoms of the SAHA hydroxamate head group (**Figure 1(D)**).

MM MD. The performance of MM MD simulations using nonbonded models was investigated first. Four models were tested: the LJ12-6 and LJ12-6-4 models, each with and without restraints. Three replicate simulations (100ns/replicate) were performed in each case, RMSD values, compared to the starting crystal structure, for the active site and bound SAHA are shown in **Figure S7**. Although RMSD values of the zinc binding site were consistent (~0.4 Å) across the LJ12-6 models and the LJ12-6-4 restrained model in general. the unrestrained

LJ12-6-4 model yielded higher RMSD values than the other models. However, compared to the relative stability of the zinc-coordinating residues, the RMSD values of the SAHA ligand show much greater fluctuation, with bound SAHA adopting a similar pose at the end of each simulation that is distinct from that observed in the crystal structure. Comparison of zinc coordination distances shown in **Table S4** indicates that the overall performance of the four models with respect to the zinc site is similar, although the accuracy with which coordination distances are predicted varies. In general, the LJ12-6-4 models yield Zn-O coordination distances that are more realistic, and more consistent with the crystallographically determined values, than those obtained from the LJ12-6 models. In contrast to the outcome from our earlier calculations with Sfh-I (above) application of distance restraints did not further improve the performance of either model, possibly indicating that the non-bonded MM models handle zinc binding sites that include only Zn-O and Zn-N interactions more easily than those that also include a Zn-S interaction. However, in all cases the zinc coordination number at the end of the simulation had increased to 6 as a result of a move to bi-, rather than monodentate coordination by Asp 181.

DFTB3/MM MD. 100ps of DFTB3/MM MD was then performed starting from the last frame of the 100 ns MM MD production run with the restrained LJ12-6-4 model, with the HDAC2 zinc site and SAHA ligand included in the QM region. After DFTB3 QM/MM MD simulation, zinc coordination by Asp 181 shifted from bi- to mono-dentate, returning the Zn coordination number to five, as observed in the crystal structure. However, we also found occasionally that zinc was in a tetrahedral geometry after DFTB3 QM/MM simulation, with the detachment of the SAHA O2 atom.

B3LYP/MM optimization. We then ran a B3LYP-D3BJ /MM geometry optimization after the DFTB3 treatment to further optimize the zinc-ligand coordination distance. The 6-31G(d) basis set was first applied, and the diffuse function subsequently added. The results of this further round of QM/MM optimization are shown in **Table S5**. In this case, incorporation of the diffuse function had no effect on the accuracy with which Zn^{2+} coordination distances were predicted. The results clearly show that in this case B3LYP-D3BJ/MM yielded a better result for Zn^{2+} coordination distance and the overall geometry closely resembles the crystal structure (**Figure 4(D)**). Moreover, as detailed for simulations of Sfh-I (above) we observed only slight improvement in prediction accuracy for structures that had undergone 542 steps (converged) of B3LYP-D3BJ/MM geometry optimization, compared to those that had only undergone 250 steps. This increases our confidence that 250 steps of B3LYP QM/MM geometry optimization should be sufficient to build accurate models of complexes of zinc metalloproteins.

Application to di-zinc systems: complexes of the IMP-1 and L1 metallo-beta-lactamases.

In addition to the wide range of zinc metalloproteins with mono-zinc centers, several enzyme classes of mechanistic and pharmacological and/or biotechnological interest possess dinuclear zinc centres.[277,278] Accordingly, we also investigated the application of the combined MM and QM/MM MD approach to model complexes of di-zinc metalloenzymes, specifically metallo-beta-lactamases (MBLs), enzymes that confer resistance to a broad range of beta-lactam antibiotics upon producer bacteria.[60,279] Two MBL model complexes were selected: the complex of IMP-1 with the small thiol thioglycolic acid (TGA, resolution 1.57 Å, PDB code: 6JED) [269] and the complex of the L1 MBL from *Stenotrophomonas maltophilia* with the hydrolyzed form of the penem antibiotic faropenem (resolution 1.57 Å, PDB code: 7A63). [270] Both complex structures contain two zinc ions that are close together in space (3.55 Å and 3.57 Å respectively for the IMP-1 and L1 complexes). In the IMP1:TGA complex one zinc ion (Zn1) is tetrahedrally coordinated by three histidine residues (His77, His79, His139) and the thiolate of TGA (**Figure 1(E)**), while the second (Zn2) is 5-coordinated by three amino acids (Asp81, Cys158 and His197) and the thiolate and carboxylate oxygen atom of TGA. In the L1:faropenem complex the equivalent zinc ions are respectively tetrahedrally coordinated by three histidine residues (His83, His85 and His159) and a water molecule (which bridges the two zinc ions); and octahedrally coordinated by three amino acids (Asp87, His 88, His224), the carboxylate oxygen and ring nitrogen of hydrolyzed faropenem, and the zinc-bridging water. The two systems thus resemble one another in some aspects (sharing a tetrahedrally coordinated zinc ion in a tri-histidine center and with a bridging, non-protein ligand); but differ in others (with the second zinc ion in IMP-1 5-, as opposed to 6-coordinated and with IMP-1 containing a Cys ligand).

Workflow validation for the di-zinc system using the IMP-1 complex

MM MD. MM MD simulations on the IMP-1:TGA complex were carried out as above, using both the LJ12-6 and LJ12-6-4 models with and without restraints. Of note, the GAFF2 (Ver 2.11) forcefield was used to parameterize the inhibitor. Three replicate simulations were run for each model; RMSD values for the zinc centers and bound inhibitor, compared to the crystal structure, are presented in **Figure S9**.

Inspection of **Figure S9** shows that in all cases bound TGA remains close to the crystallographically observed binding pose throughout the duration of the simulation, with RMSD values remaining stable and below 0.5 Å throughout the simulation. However, when the unrestrained LJ12-6 model was used RMSD values for the binding site residues increased dramatically, although these remained consistently low (~0.7 Å) when restraints were applied or the LJ12-6-4 model was used. Visual inspection of the simulation trajectory identified that in the case of the unrestrained LJ12-6 model zinc coordination was disrupted by replacement of coordinating histidine residues with water molecules.

Compared to the LJ12-6 models, the zinc coordination distances obtained from simulations using the LJ12-6-4 models were overall closer to the crystallographic values (**Table S6**). Accuracy also improved when restraints were applied, although the effect was far less pronounced than in the case of the LJ12-6 model. The distance between two zinc ions was however better predicted by the LJ12-6 series models, whereas this increased by ~0.4 Å when the LJ12-6-4 model was used. Overall, and consistent with our findings for the mono-zinc systems, we consider the LJ12-6-4 model to also be the first choice for this di-zinc system. However, the coordination number for each zinc ion increased to 6, as a result of the introduction of additional water molecules into the coordination shells of both zinc ions after LJ12-6-4 MM MD simulations.

DFTB3/MM MD. As with our previous approaches, 100 ps of QM/MM MD using DFTB3 was then performed starting from the last frame of the MM production run. The QM region included the inhibitor, the zinc ions and the side chains of their coordinating residues, and the zinc-bound water molecules. The additional water molecules were successfully removed by this treatment, restoring the total Zn^{2+} coordination number to nine. However, with the exception of the interaction with His197, the Zn^{2+}-ligand distances for Zn2 increased by 0.1 ~ 0.2 Å, and approached the coordination boundary of 2.5 Å. This was particularly the case for Zn-S interactions involving both Cys221 and the TGA thiolate.

B3LYP/MM optimization. BL3YP-D3BJ based QM/MM geometry optimization was carried out using the last frame from the 100ps DFTB3/MM MD. In the interests of computational efficiency, the 6-31G(d) basis set was applied first, and the diffuse function was added subsequently. The results of the DFTB3/MM MD calculation and DFT/MM optimization are shown in **Table S7**. After DFT refinement the Zn^{2+} coordination distances were very close to their values in the crystal structure. However, the incorporation of the diffuse function did not show a large improvement on the accuracy of predictions for Zn^{2+} coordination distances. We also found that there was no improvement in accuracy between structures that had met the convergence criteria (449 steps) for DFT/MM geometry optimization and structures that had only undergone 250 steps. Snapshots of the active site of the IMP-1:TGA complex at different stages of the simulation workflow are shown in **Figure S10** and alignment of Zn^{2+} site geometries obtained from simulations and the starting crystal structure, is shown in **Figure 4(E).**

Workflow implementation for L1 complex with hydrolyzed faropenem

After demonstrating successful treatment of the IMP-1:TGA complex, the pipeline was tested against the complex of the L1 metallo-beta-lactamase with the hydrolysis product of the penem antibiotic faropenem. As observed for IMP-1, MM MD simulations using the LJ12-6-4 model yield similar results regardless of the inclusion of restraints: the RMSD plots showed that the zinc center remained stable across the simulations while the geometry visualization showed that in each case both zinc ions were both octahedrally coordinated after 100 ns MM MD simulation (**Figures S11, S12**). MD treatment leads to coordination of both zinc ions by the faropenem C6 carboxylate group, and moves coordination of Zn2 by Asp87 from mono- to bidentate. In addition, the hydroxyl group connected to the faropenem C1 carbon moved to coordinate Zn1. As the above modes of coordination generated by simulations using MM non-bonded models are clearly unrealistic (compared to the crystal structure), DFTB3/MM MD was carried out from the last snapshot of a restrained LJ12-6-4 simulation to address this artificial Zn^{2+} coordination. The distorted zinc geometry was successfully restored to that observed in the crystal structure by this subsequent 100 ps DFTB3/MM MD step. The structure obtained from DFTB3 treatment was then further refined to optimize zinc coordination distances with the B3LYP-D3BJ based QM/MM minimization approach (**Figure S12, Table S**). The 6-31G(d) basis set was first applied and the diffuse function was then added. In this case, incorporation of the diffuse function slightly improves the prediction accuracy of the Zn^{2+} coordination distances but significantly increased the computational cost. The alignment of the Zn^{2+} site geometry predicted from these simulations to the crystal structure is shown in **Figure 4(F)**.

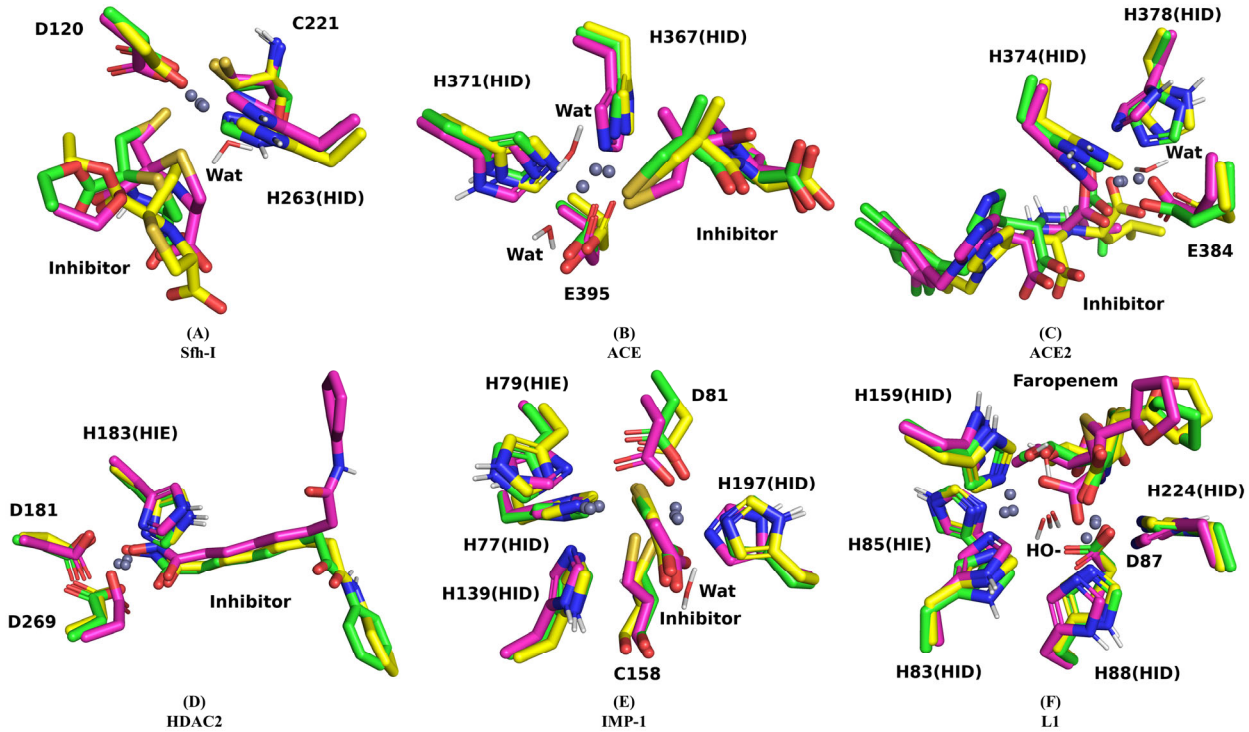


Figure 4. Zinc binding sites of the 6 protein-ligand systems included in the study after simulations at different levels of theory. Carbon atoms are in green/yellow/purple depending on the model: yellow represents the crystal structure; purple represents a typical structure after 100 ns MM simulation with LJ12-6-4 models; green represents a representative structure after DFTB3/MM MD and B3LYP-D3BJ based QM/MM geometry optimization. Nitro atoms are in blue, oxygen atoms are in red and sulfur atoms are in dark yellow, zinc ions are represented as gray balls.

DFTB3 calculations directly from crystal structures.

Comparison of the results obtained from the various systems identified that when DFTB3/MM treatment was used to rectify distortions (e.g. additional zinc-coordinating interactions, such as by water molecules, introduced during MM MD simulations) this often resulted in preferential production of 4-coordinate geometry for individual zinc ions. To investigate whether DFTB3 can maintain appropriate zinc coordination when applied to a more accurate starting structure, we ran 2ns DFTB3/MM MD simulations directly from the crystal structures of each system using the same QM regions as in the QM/MM calculations described above. The outcomes of these simulations are shown in **Figure 5** and **S13**. The results showed that some structures (the ACE, ACE2 and L1 complexes) were well described by direct application of the DFTB3 QM/MM approach, but that this was not the case for the Sfh- I, HDAC2 and IMP-1 complexes. For the Sfh-I:L-anti-1a complex, the inhibitor showed a relatively high RMSD value compared to the ligands in other simulations. The interaction of the inhibitor thiolate with Zn^{2+} was lost, with Zn^{2+} coordination substituted by a water molecule. This may reflect the relatively high importance of the interaction between the inhibitor thiolate and zinc ion to the affinity of this system, compared to the much less extensive interactions made with the remainder of the active site[265], and is also consistent with the relatively poor performance of our previous DFTB3-based simulations in describing zinc-sulfur interactions.

For the HDAC2:SAHA inhibitor system, although the SAHA binding pose remained similar to that observed in the crystal structure, the SAHA O2 atom sometimes lost its attachment to the zinc ion (with the Zn^{2+} – O2 distance increasing to > 2.5 Å) after DFTB3 QM/MM simulation, with the zinc ion adopting a tetrahedral geometry. This was the same outcome as observed on DFTB3 optimization after MM MD simulations of this system (see above). The situation was however improved when DFTB3 treatment was applied using a larger QM region, that included not only the zinc ion, the side chains of coordinating residues and the inhibitor, but also the side chains of residues that are not in the zinc site but may interact with the inhibitor. In the case of HDAC2, when Tyr297 and His135 were included in an enlarged QM region coordination between SAHA O2 and the zinc ion was retained. However, analysis of the distances between SAHA and interacting residues suggested that these were not affected by changes to the size of the QM region. In the case of IMP-1, the major inaccuracy when using DFTB3 treatment is the incorrect calculation of Zn^{2+}-inhibitor coordination distances, with both Zn^{2+} - S and Zn^{2+} - O coordination distances predicted to be greater than 2.5 Å. As in the case of Sfh-I, above, this may reflect the involvement of sulfur atoms in interactions with zinc by both the protein and small-molecule ligand.

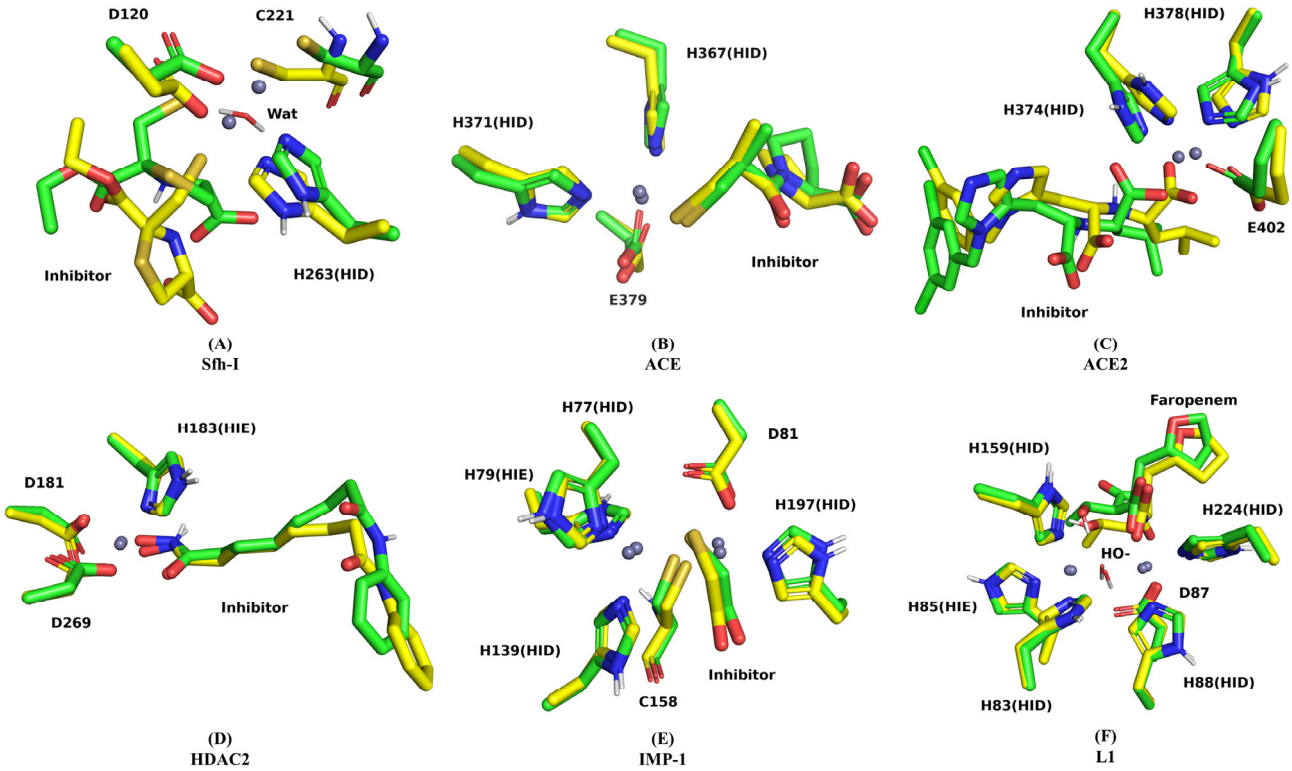


Figure 5. Zinc binding sites of the 6 protein-ligand systems included in the study after DFTB3/MM MD starting directly from the crystal structure. Carbon atoms are in green and yellow depending on the model: green represents a representative zinc site structure after 2ns DFTB3/MM MD starting directly from the crystal structure, yellow represents the crystal structure. Nitro atoms are in blue, oxygen atoms are in red and sulfur atoms are in dark yellow, zinc ions are represented as gray balls.

Discussion

The diversity and versatility of zinc coordinating interactions in zinc metalloproteins and their complexes with small molecule ligands makes accurate modelling of these systems an undertaking that is frequently challenging. In this work our aim was to develop a computational approach able to accurately describe ligand complexes of 6 diverse zinc metalloprotein systems, representing a range of coordination chemistries and geometries. Whilst we here focus on complexes of known crystal structure against which the accuracy of the various approaches can be assessed, in many cases crystal structures of zinc metalloprotein complexes with known or putative small molecule ligands will not be available. Hence we sought to develop a pipeline that is sufficiently robust to deal with more approximate starting structures. For this reason, our evaluation included MM MD methods (as might be applied to optimize structures of putative complexes generated from e.g. docking experiments) as well as QM/MM MD at both semi-empirical (DFTB3) and higher (B3LYP-D3BJ based DFT) levels of theory, with each approach representing increasing levels of accuracy in treatment of zinc and its interactions, at the expense of increasing demands upon computational resources.

For the 6 tested systems, non-bonded MM MD approaches using the unrestrained LJ12-6-4 and restrained (both LJ12-6-R and LJ12-6-4-R) models yielded RMSD values of whole protein backbone atoms, compared to the starting crystal structures, for the 6 systems that were consistent across the 4 tested non-bonded models. This suggests that the differences between the LJ12-6 and LJ12-6-4 models mainly occur at the zinc center. In general, the LJ12-6-4 nonbonded model outperformed the LJ12-6 models in terms of accuracy of Zn^{2+}-ligand coordination distance (smaller overall deviations from crystallographic values), consistency of simulation (lower standard error and more stable RMSD plots) and lower dependence on restraints. However, the LJ12-6-4 models show a much greater tendency than the LJ12-6 models to increase the zinc ion coordination number to six, usually by means of introduction of additional water molecules to the zinc site, resulting in octahedral coordination geometries (e.g., Sfh-I, ACE, ACE2, IMP-1, etc.). This difference likely reflects differences between the Zn^{2+} sphere radius used in the various models. Specifically, the R_{min} value for zinc (1.276Å) for the LJ12-6 CM parameter set is much smaller than that for the 12-6-4 parameter set (1.454 Å). Smaller sphere sizes (R_{min} values) reduce the space available for additional water molecules to "crowd" into the zinc site. Despite inclusion in the LJ12-6-4 model of a C4 term to represent dipoles, its performance still reflects design and validation based on aqueous Zn^{2+} complexes, in which the preference for octahedral

geometry is well established, compared to a relatively earlier stage of development for describing the behavior of zinc in a protein environment where a wider range of geometries are evident, with tetrahedral particularly common. Fortunately, however, our work here demonstrates that additional zinc interactions introduced in MM MD simulations using the LJ12-6-4 model are normally water molecules, and that these can usually be removed with subsequent QM/MM MD approaches. With this caveat, the LJ12-6-4 model is a good choice for modelling small molecule complexes of zinc containing proteins, being easy to be set up, delivering reproducible simulations and overall predicting reasonable values for zinc coordination distances, over a longer timescale than is possible with QM/MM MD approaches and at reasonable computational cost.

The use of distance restraints was essential for obtaining reasonable results with LJ12-6 models, as without these some ligand or protein side-chain interactions with the zinc could be lost (e.g., in the above IMP-1:TGA complex). Although the application of distance restraints did not affect the LJ12-6-4 model as much as the LJ12-6 model, these did improve the consistency between individual simulations in our sets of three replicates and contributed during the initialization of the system (energy minimization and equilibration). Although small molecule ligands were not restrained during either the equilibration or production MD processes, these did require restraining during energy minimization to retain or obtain a reasonable starting pose. Selection of appropriate restraints then represents an important determinant of the overall success of MM MD steps, as inappropriate choices may lead to artificial interactions resulting in unrealistic descriptions of zinc binding or detachment of the small molecule from the zinc ion(s); and requires input based on prior knowledge of common interaction patterns between zinc and typical zinc-binding groups (e.g. thiolates, carboxylates, hydroxamates).

An additional factor influencing the quality of the results obtained with MM MD simulations is the ligand parameters. GAFF and GAFF2 from the AMBER package are general forcefields for parameterizing small molecule ligands; in the work presented here GAFF generally performed well. (Difficulties with implementation in LJ12-6-4 models precluded more extensive use of GAFF2). Of the model systems investigated here, limitations to ligand parameterization may be apparent in MM MD simulations of Sfh-I, where the ligand pose is poorly replicated, possibly due to the presence of the thiazolidine ring system and the additional thiolate sulfur atom; and HDAC2, where the aromatic ring at the end of the SAHA alkyl tail is oriented oppositely to its position in the crystal structure. In both cases the accuracy with which the ligand position was modelled was improved by subsequent DFT/MM optimization. Nevertheless, the GAFF approach retains the major advantages of speed and ease of use, making it a good choice for fast model building or ligand screening tasks with acceptable levels of accuracy.

The DFTB3 and B3LYP-D3BJ methods were the two QM approaches investigated here. In our simulations, the DFTB3 approach showed a preference for formation of a 4-fold (tetrahedral) zinc coordination geometry, and was an effective tool for removing additional water molecules introduced during MM MD with LJ12-6-4 models. However, when both sulfur atoms and water molecules added during MM MD simulations are present in the zinc binding site, with DFTB3 the Zn - S contact is occasionally lost and zinc instead interacts with water molecules. In addition, the coordination distances for Zn - S interactions were often badly described with DFTB3 treatment (i.e., the Zn - S coordination distances were often close to 2.5 Å, substantially longer than the typical experimental values of 2.15 – 2.35 Å). Taken together, these findings suggest that the DFTB3 approach may be less effective in handling Zn - S interactions. Although DFTB3/MM MD simulations were run for 100ps, additional water molecules introduced in preceding MM MD simulations departed from the zinc center at the very beginning (i.e., first 10ps) of the run, suggesting that shorter timescales of 50ps or even 20ps may be sufficient for this purpose. Limiting the duration of DFTB3/MM MD simulations may also reduce the possible impact on other Zn^{2+} - ligand interactions, leaving final refinement of the zinc center for subsequent B3LYP QM/MM optimization.

Our results demonstrate that removal of additional zinc-coordinating water molecules requires DFTB3/MM MD simulations. In comparison, B3LYP/MM optimization had only limited capability to alter the geometry of the zinc site, and in most cases was unable to remove artificial zinc coordinating water molecules (resulting in 6-coordinate zinc geometry) introduced during MM MD simulations. Computational expense precluded B3LYP/MM MD simulations. Our data show that B3LYP/MM optimization should be carried out from an appropriate starting geometry, and not one that contains significant distortions, such as those that may be introduced during preceding MM MD steps.

We found the GD3BJ dispersion correction to be important for accurate description of zinc coordination distances in simulations using B3LYP, with inclusion of the diffuse function important in systems containing sulfur atoms. In addition, only minor differences were observed when comparing structures obtained from a fully converged B3LYP-D3BJ simulation and after 250 steps of B3LYP-D3BJ treatment. Given the difficulty in reaching convergence when using DFT QM/MM optimization (which with our hardware infrastructure normally required around 750 steps and approximately 10 days of calculations), we then consider a 250-step optimization likely to be sufficient to obtain model geometries of acceptable levels of accuracy. Furthermore, the computational efficiency of the B3LYP-D3BJBJ/6-31+G(d) treatment was greatly improved when the structure was first optimized with B3LYP-D3BJ/6-31G*. A computationally efficient B3LYP-D3BJ-based QM/MM optimization approach could then be 250 steps of B3LYP-D3BJ/6-31G(d) first, followed by a further 250 steps with diffuse functions included. Of the three methods tested, the B3LYP-D3BJ approach then

provided the most accurate descriptions of the systems under test, but required previous geometry optimization of the starting structure by e.g. DFTB3 treatment.

Although MM MD may provide a robust and computationally efficient approach to obtaining a proper complex system for simulation (e.g., protonation state, stable conformation of the complex), especially when docked structures are used as initial structures, our results indicate that incorporation of an initial MM MD step is not essential when simulating ligand complexes of zinc containing proteins from crystal structures. In such cases reasonable descriptions of metal centers can be obtained by using QM/MM methods directly. In particular, although the semi-empirical DFTB3 method is less effective than DFT in predicting coordination bond lengths, and can present particular difficulties in dealing with sulfur atoms, DFTB3 consumes less computational resources than B3LYP-D3BJ, making possible QM/MM MD simulations on the nanosecond timescale. Moreover, DFTB3 QM/MM MD simulations have some ability to restore distorted zinc geometries, although caution should be taken when sulfur atoms are involved.

Conclusions

The extensive set of simulations described here for a range of zinc metalloprotein systems allows some general conclusions to be drawn regarding approaches that may be suitable to generate realistic models of zinc metalloprotein complexes. MM MD simulations using non-bonded models failed to maintain experimental zinc coordination and (as expected) showed strong tendencies towards octahedral geometries, due most likely to the treatment of zinc ions. DFTB3 MM/MD treatment retained experimental coordination when crystal structures were used as starting models, and was often able to remove water molecules added by prior MM MD treatment, but was less well able to describe zinc centers with coordinating waters exposed to bulk solvent and, in particular, zinc-sulfur interactions. B3LYP-D3BJ/MM geometry optimization well describes zinc centers and can restore distorted coordination distances close to experimental values, but does not result in large-scale changes to geometry and so requires an appropriate initial structure. Our data suggest that inclusion of empirical dispersion corrections in these simulations is beneficial, while use of diffuse functions in B3LYP-D3BJ/MM geometry optimization may improve the quality of the final model but does not always do so, and imposes a dramatic increase in computational cost. We conclude that a computational pipeline involving DFTB3 QM/MM MD simulations at the semi-empirical DFTB3 level of theory, followed by B3LYP-D3BJ/MM geometry optimization, should be sufficient to generate models of protein zinc centers, with preceding classical MM MD not essential unless there are requirements for significant optimization of the starting model, observation of the motions of the system over longer time scales, or examination of protonation states. In such cases LJ12-6-4 models provide a more accurate description of zinc than LJ12-6 models in terms of coordination distance, and the unrestrained LJ12-6 model is not recommended. While unrestrained simulations could be used to examine the potential movement of zinc binding residues, restrained simulations are more appropriate to simulate the dynamics of the zinc site while as far as possible retaining crystallographically observed geometry. Overall, the pipeline of MM MD, DFTB3/MM MD and B3LYP-D3BJ/MM geometry optimization provided good results in our tests. We suggest that this constitutes a robust and versatile approach suitable for modelling diverse zinc metalloproteins and their complexes.

Methods & Materials

Molecular Mechanics (MM) dynamics simulations. The PDB file for the Sfh-I inhibitor complex (PDB code: 7BJ9)[265] was obtained from Dr Philip Hinchliffe prior to release. Crystal structures of other protein-ligand complexes were obtained from the PDB database (https://www.rcsb.org/).[280,281] Crystallographic water molecules within 10 Å of the ZN ion were retained, other in the PDB files were removed. PDB headers and all lines other than 'ATOM', 'HEATM', 'TER' and 'END', were deleted. The protonation state of the protein was determined by PROPKA 3.0[282,283] through the PDB2PQR[284] server. Packages including tLEaP, Antechamber and ParmEd of AmberTools20[100] were used to parameterize the protein-ligand system. Hydrogens were added to the protein using tLEaP. Antechamber was used to generate parameterization files for the ligands: inhibitors were described by the general AMBER force field (GAFF)[285] and AM1-BCC charge method unless otherwise specified. The complex was solvated in a 12 Å water box using tLEaP. The Amber ff14SB forcefield and SPC/E water model were used to parameterize the system of protein and water molecules. After balancing the charge of the system using Na^+ and Cl^- counter ions, the topology and the coordinate files of a typical nonbonded LJ12-6 model of the system were generated (the LJ12-6 CM parameter set). The topology file editor ParmEd was then called to add C4 terms to the LJ12-6 topology file and coordinate file generating corresponding files for the LJ 12-6-4 model. After system preparation, the simulation used the pmemd engine of Amber18.[98]

The system first went through minimization of the hydrogen atoms followed by minimization of the water molecules, then minimization of side chains and finally of the whole protein-ligand complex. After the minimization, the system was slowly heated to 298K over 200 ps. After that, a 2-nanosecond equilibration process was carried out to optimize the system configuration. Distance restraints were implemented in the above steps to keep the zinc binding site stable. Positional restraints were applied to the protein backbone atoms and the zinc ion. The strength of positional restraints gradually decreased (starting from 25 kcal/mol) as the equilibrium progressed and these were totally removed in the last equilibration step (1 ns). Distance restraints were applied to help maintain the zinc coordination geometry. The values of the lower (r2) and upper (r3) bounds were respectively set to be ~0.15

Å from the crystallographically observed values. The constants rk2 and rk3 were set to be 50.0 kcal/mol. MD simulations (production runs) with or without (depending on model type) restraints were launched when the equilibration was completed. The non-bond cut off distance was set to 10 Å and the system temperature was set to 298K and regulated by Langevin dynamics with a collision frequency of 2. The Berendsen barostat was applied to control the system pressure at 1 atm. The SHAKE algorithm was applied to allow a larger time step. The simulation ran for 100 ns with a time step of 2 femtoseconds under the constant pressure periodic condition. MD trajectories were analyzed and RMSD values calculated by Cpptraj, a trajectory processing package included in AmberTools20. Records of the system status during MD simulations were analyzed by the Python script 'mdout_analyzer' of AmberTools19.[101] MD trajectories were visualized by VMD (Version 1.9.4)[286] and snapshot analysis (e.g., distance measurement between atoms) carried out using PyMoL (https://pymol.org/)[287]. 2.5 Å was set as the boundary distance for zinc coordination analysis.

QM/MM dynamics and geometry optimization. QM/MM calculations were carried out at two levels of QM theory: DFTB3 and DFT (B3LYP). In both DFTB3 and B3LYP QM/MM calculations, the QM region was defined as the inhibitor, zinc ions, any water molecule coordinated to the zinc ions, and the side chains of the zinc coordinating residues; while the rest of system was defined as the MM region. Link hydrogen atoms were automatically added to the system by Amber's QM/MM engine Sander. QM/MM MD was carried out from the last snapshot of preceding MM MD simulations to restore the coordination geometry of Zn^{++}, or directly from crystal structures. Calculations were performed using AMBER 18 which has built-in DFTB3 code and can handle both the QM and MM portions of the calculation. The QM region was modelled by DFTB3 theory, and the MM part was modelled using Amber forcefield ff14SB (the same profile as the MM production run). For the QM region, the nonbond cutoff distance was set to 8 Å and the SHAKE algorithm was used. QM/MM geometry optimization using DFT theory for the QM region was performed based on the result of DFTB3/MM MD or the last frame of MM MD simulation (depending on the system). The B3LYP/MM optimization was performed with AMBER 18 via its external QM program interface: Gaussian 16[119] was called to initiate single point calculations for the QM region and the MM portion was handled by AMBER 18 using the Amber ff14SB forcefield[272]. In the DFT optimization a QM cut-off distance of 8 Å was set. For the QM portion, we used B3LYP hybrid functionals with the 6-31G(d) or 6-31+G(d) basis sets. The D3 version of Grimme's dispersion correction with Becke-Johnson damping[137,138] was also applied, which is important for predicting accurate protein structures. In all QM/MM calculations, the particle mesh Ewald (PME) method was implemented to calculate long-range QM-QM and QM-MM electrostatic interactions. Covalent C-C bonds at the boundary of the QM and MM region were treated by adding hydrogen link-atoms, which were automatically placed by AMBER. SCF convergence thresholds was set to be SCF=(Conver=8). The outputs from QM/MM calculations were analyzed by Cpptraj in AmberTools20[101] and visualized using VMD (Version 1.9.4) and PyMoL (https://pymol.org/)[287]. As above, 2.5 Å was set as the boundary distance for zinc coordination analysis.

ASSOCIATED CONTENT

Supporting Information. Plots of zinc coordination distances during QM/MM simulations of Sfh-I:L-anti-1a complex, Figures showing RMSD plots and representative geometries for other simulations, Tables of coordination distances. This material is available free of charge via the Internet at https://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Professor Adrian J Mulholland, Email: Adrian J. Mulholland@bristol.ac.uk

Author Contributions

All authors designed the research. ZY conducted and analyzed the molecular simulations with guidance from RMT, CKC, SG, RS and AJM. ZY, JS and AJM wrote the manuscript, with all authors giving approval to the final version.

Funding Sources

We thank the China Scholarship Council and the U.K. Engineering and Physical Sciences Research Council (EPSRC) for funding studentships to ZY and RMT, respectively. This work was supported by the U.K. Biotechnology and Biological Sciences Research Council (BBSRC)-funded SouthWest Biosciences Doctoral Training Partnership (training grant reference BB/J014400/1). AJM thanks EPSRC and BBSRC for funding (grant numbers EP/M022609/1, EP/M013219/1, BB/M000354/1 and BB/L01386X/1). RS thanks RSC for Research Funds (Grants R19-3409 and R20-6912) and MCIN/ AEI/10.13039/501100011033 (Grant PID2020-113147GA-I00). This work was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (NIH) to JS under Award Number R01AI100560. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Notes

The authors declare no competing financial interest. Underlying computational data are available at the University of Bristol data repository, data.bris, at https://doi.org/10.5523/bris.10p78zgsappbz226bzrdagabq9.

ACKNOWLEDGEMENTS

This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bris.ac.uk/acrc/. We thank Pengfei Li (Loyola University) for advice regarding MD parameters

and Natalie Fey (School of Chemistry, University of Bristol) for her careful reading of the manuscript.

ABBREVIATIONS

DFT, density functional theory; MBL, metallo-beta-lactamase; MD, molecular dynamics; MM, molecular mechanics; MMTZ, thiol-based mercaptomethyl thiazolidine; QM, quantum mechanics; TGA, thioglycolic acid.

A Multiscale Workflow for Modelling Ligand Complexes of Zinc Metalloproteins

Zongfan Yang[a,b], Rebecca M. Twidale[a], Silvia Gervasoni[a,c,d], Reynier Suardíaz[a,e], Charlie Colenso[a,b], Eric J. M. Lang[a], James Spencer[b] and Adrian J. Mulholland[a]*

[a]Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, BS8 1TH, U.K.

[b]School of Cellular and Molecular Medicine, University of Bristol, Bristol, BS8 1TD, U.K.

[c]Department of Pharmaceutical Sciences, University of Milan, Via Mangiagalli, 25, I-20133 Milano, Italy.

[d]Department of Physics, University of Cagliari, S.P. Monserrato-Sestu km 0.700 I-09042, Monserrato (CA), Italy (present address)

[e]Departamento de Química Física, Facultad de Química, Universidad Complutense, 28040 Madrid, Spain (present address)

Supporting Information

Contents

Figure S13. Outcome of DFTB3/MM MD run directly from crystal structures.

Table S1. Zn coordination distance (Å) of four non-bonded MM models for Sfh-I: L-anti-1a complex.

Table S2. Coordination distances between the zinc ion and its coordinating atoms for QM/MM simulations of the Sfh-I:L-anti-1a complex.

Table S3. Atomic distances (Å) between the zinc ions and the zinc coordinating atoms for simulations of the ACE inhibitor complex model.

Table S4. Atomic distances (Å) between the zinc ions and the zinc coordinating atoms for simulations of the ACE2 inhibitor complex model.

Table S5. Zn coordination distance (Å) of four non-bonded MM models for the HDAC-2:SAHA complex.

Table S6. QM/MM optimized distances (Å) between the zinc ion and its coordinating atoms for simulations of the HDAC2:SAHA complex.

Table S7. Zn coordination distance (Å) of four non-bonded MM models for simulations of the IMP-1:TGA complex.

Table S8. Representative QM/MM optimized atomic distances (Å) between the zinc ions and the zinc coordinating atoms for simulations of the IMP-1:TGA complex model.

Table S9. Representative atomic distances (Å) between the zinc ions and the zinc coordinating atoms for simulations of the L1:Faropenem complex model.

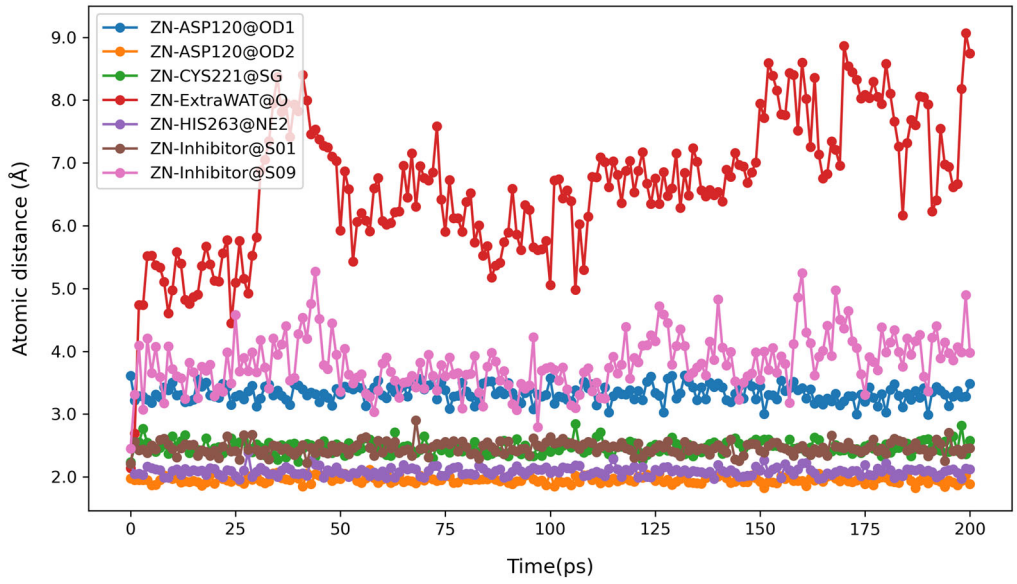


Figure S1. Zinc site coordination distances against simulation time for 200 ps DFTB3 QM/MM MD simulation of the Sfh-I:L-anti-1a complex. All Zn^{2+} ligation distances are recorded, including water coordination (ZN-ExtraWat@O) added by the MM MD. The major zinc geometry shift happens at the beginning of the QM/MM MD run. Data for each interaction was collected every picosecond, with dots indicating data collection points.

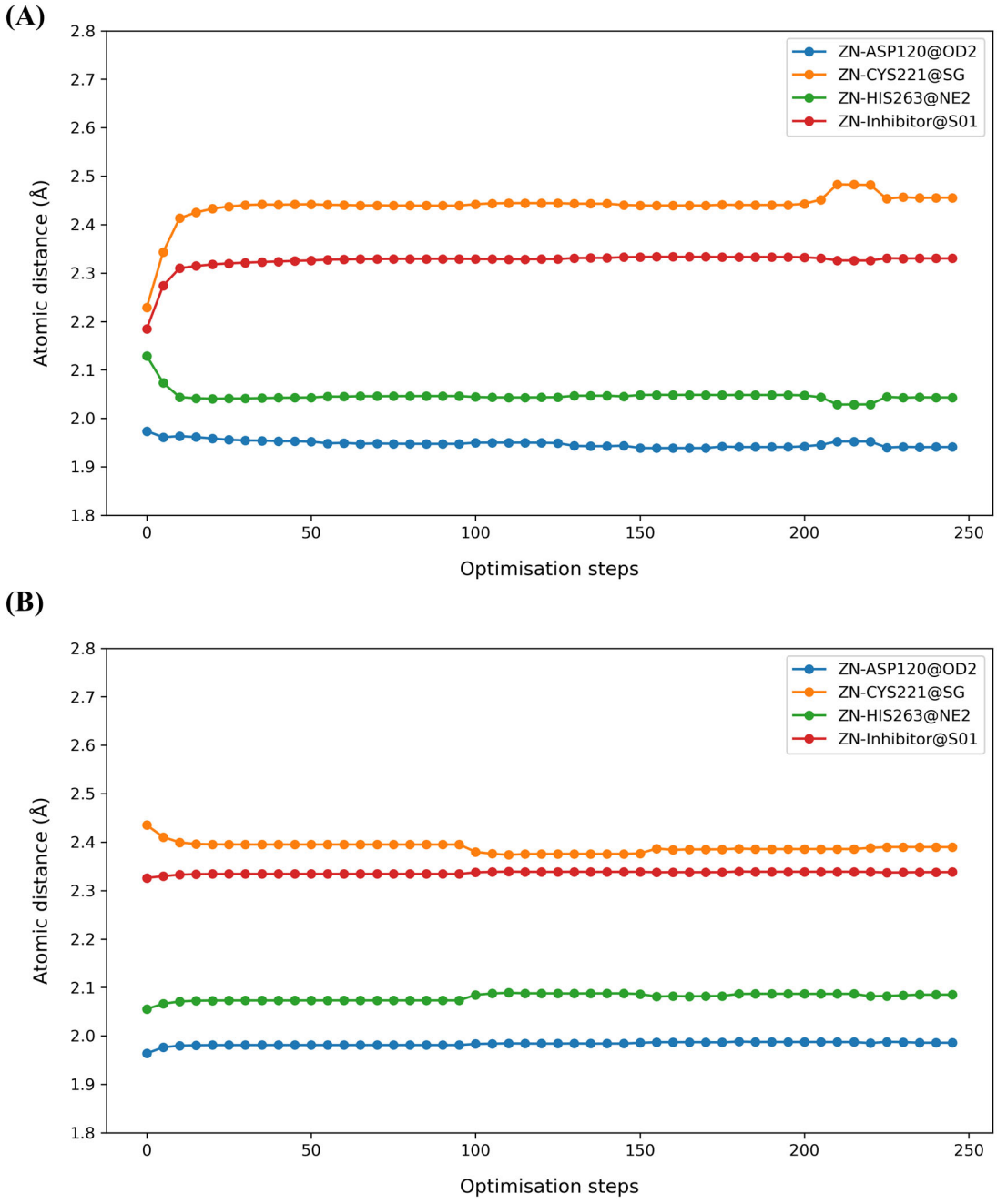


Figure S2. Zinc site coordination distances against simulation steps over 250-step B3LYP-D3BJ based QM/MM geometry optimization of the Sfh-I:L-anti-1a complex without (A) and with (B) diffuse function. The major zinc geometry shift happens at the beginning of the QM/MM optimization. Data for each interaction was collected every 5 steps, with dots indicating data collection points.

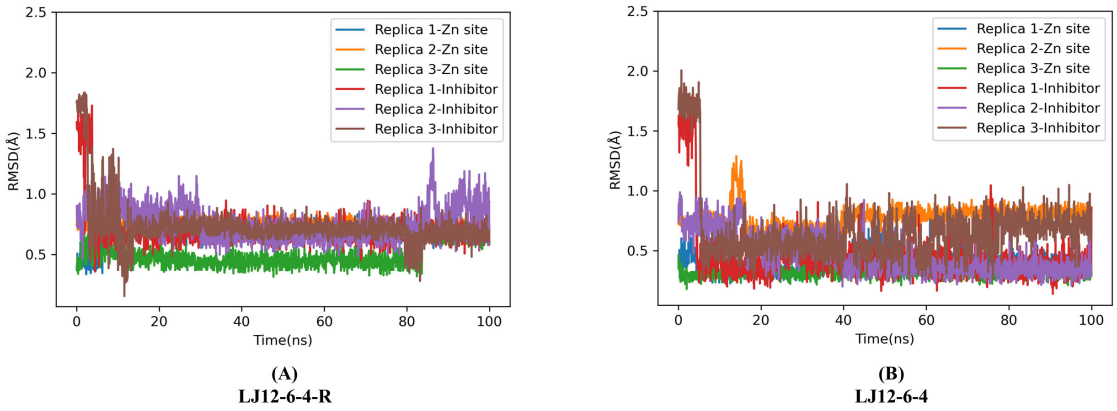


Figure S3. Time-dependence of RMSD values compared to the crystal structure for MM MD simulations of the ACE:L-captopril complex using LJ12-6-4 non-bonded models. Three replicate simulations were performed for each model and each replicate is 100 ns. The 'Zn site' refers to zinc ions and zinc coordinating residues and 'Inhibitor' refers to the zinc bound inhibitor L-captopril.

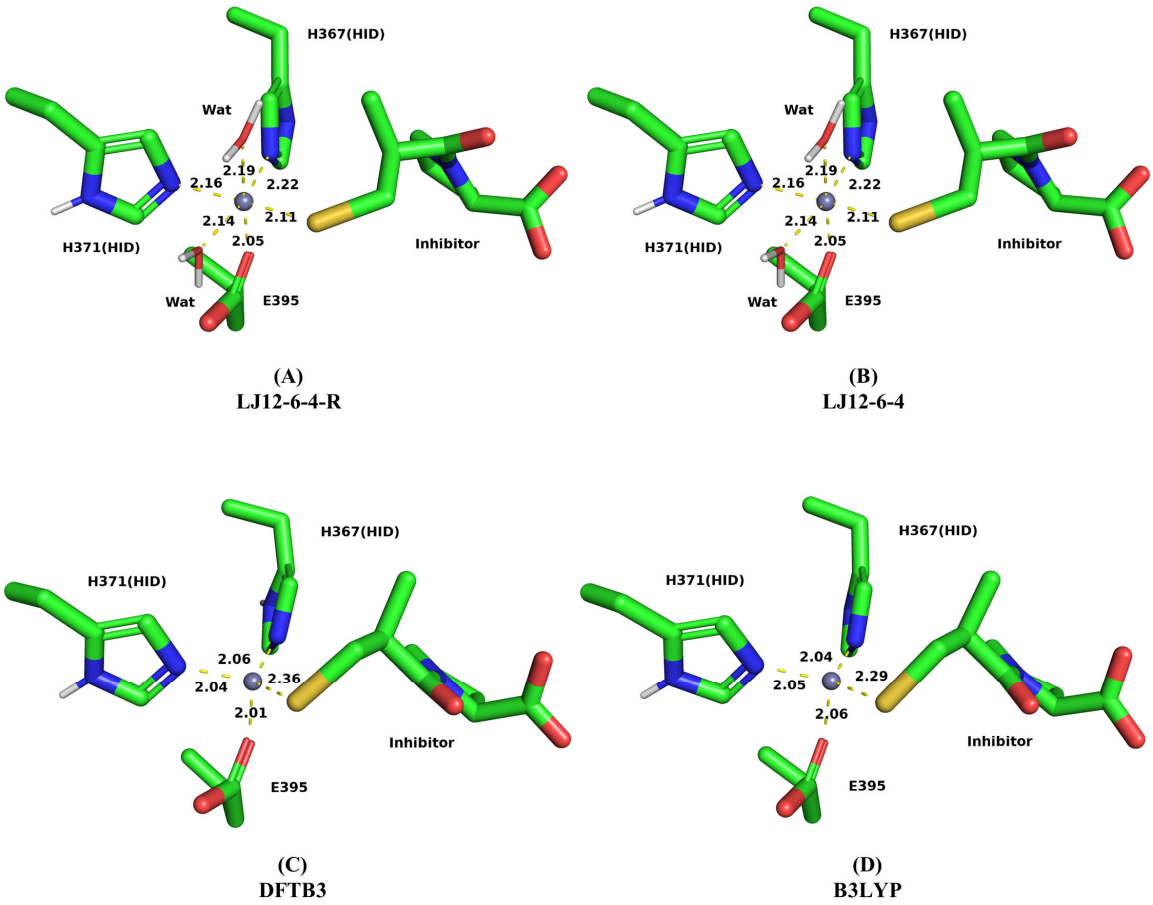


Figure S4. Representative zinc site geometries of the ACE:L-captopril complex after MM MD simulations and QM optimization. (A-B) Representative zinc site geometries after 100 ns MM MD simulations using LJ12-6-4 model with (A) and without (B) restraints. (C) Representative zinc geometry after 100 ps DFTB3/MM MD simulation. The extra water molecule was removed. (D) Representative zinc geometry after B3LYP-D3BJ based QM/MM geometry optimization.

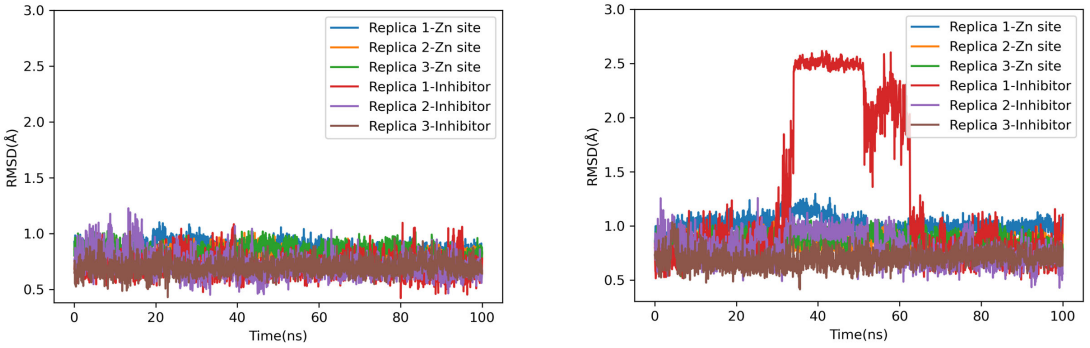


Figure S5. Time-dependence of RMSD values compared to the crystal structure for MM MD simulations of the ACE2:MLN-4760 complex using LJ12-6-4 non-bonded models. Three replicate simulations were performed for each model and each replicate is 100 ns. The 'Zn site' refers to zinc ions and zinc coordinating residues and 'Inhibitor' refers to the zinc bound inhibitor MLN-4760.

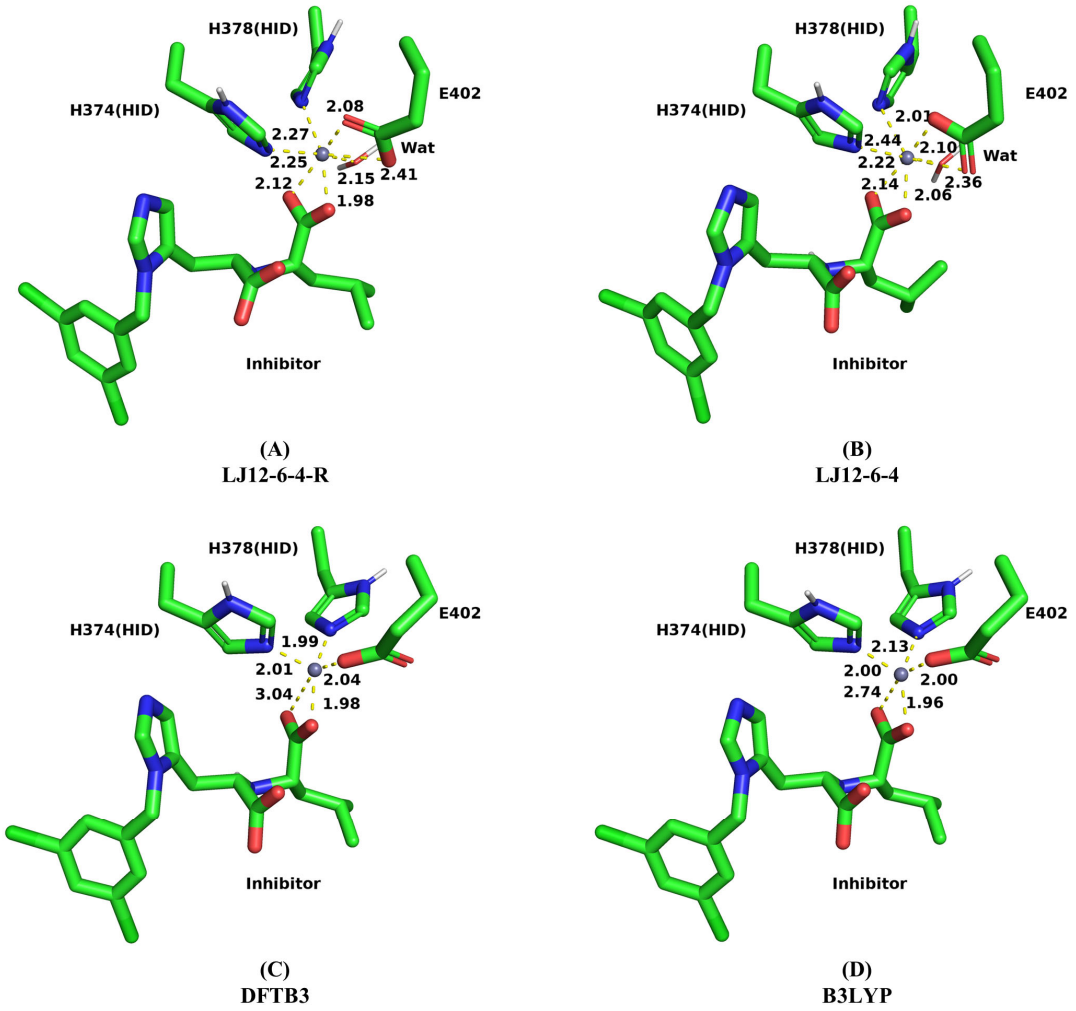


Figure S6. Representative zinc site geometries of the ACE2:MLN-4760 complexes after MM MD simulations and QM optimization. (A-B) geometries after 100 ns MM MD simulations using LJ12-6-4 model without (A) and with (B) restraints.

(C) Representative zinc geometry after 100 ps DFTB3/MM MD simulation. The extra water molecule was removed. (D) Representative zinc geometry after B3LYP-D3BJ based QM/MM geometry optimization.

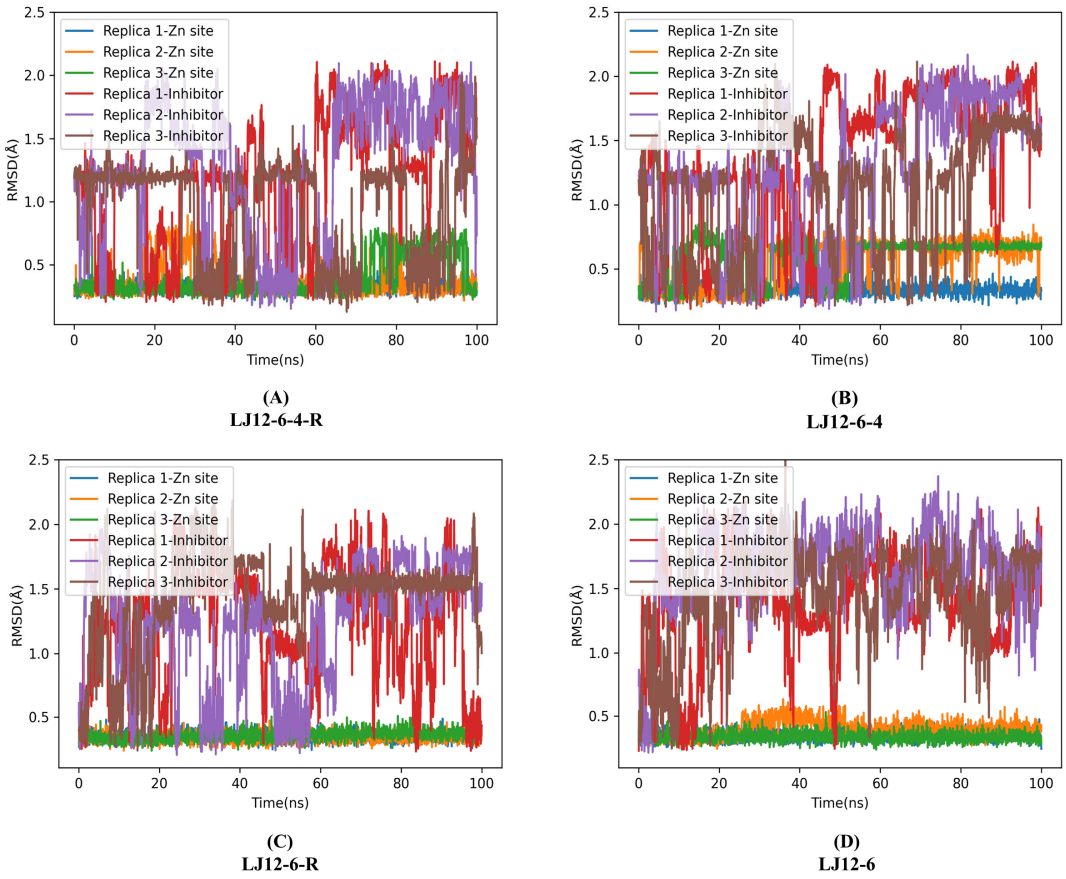


Figure S7. Time-dependence of RMSD values compared to the crystal structure for MM MD simulations of the HDAC2:SAHA complex using non-bonded models. Three replicate simulations were performed for each model and each replicate is 100 ns. The 'Zn site' refers to zinc ions and zinc coordinating residues and 'Inhibitor' refers to the zinc bound inhibitor SAHA.

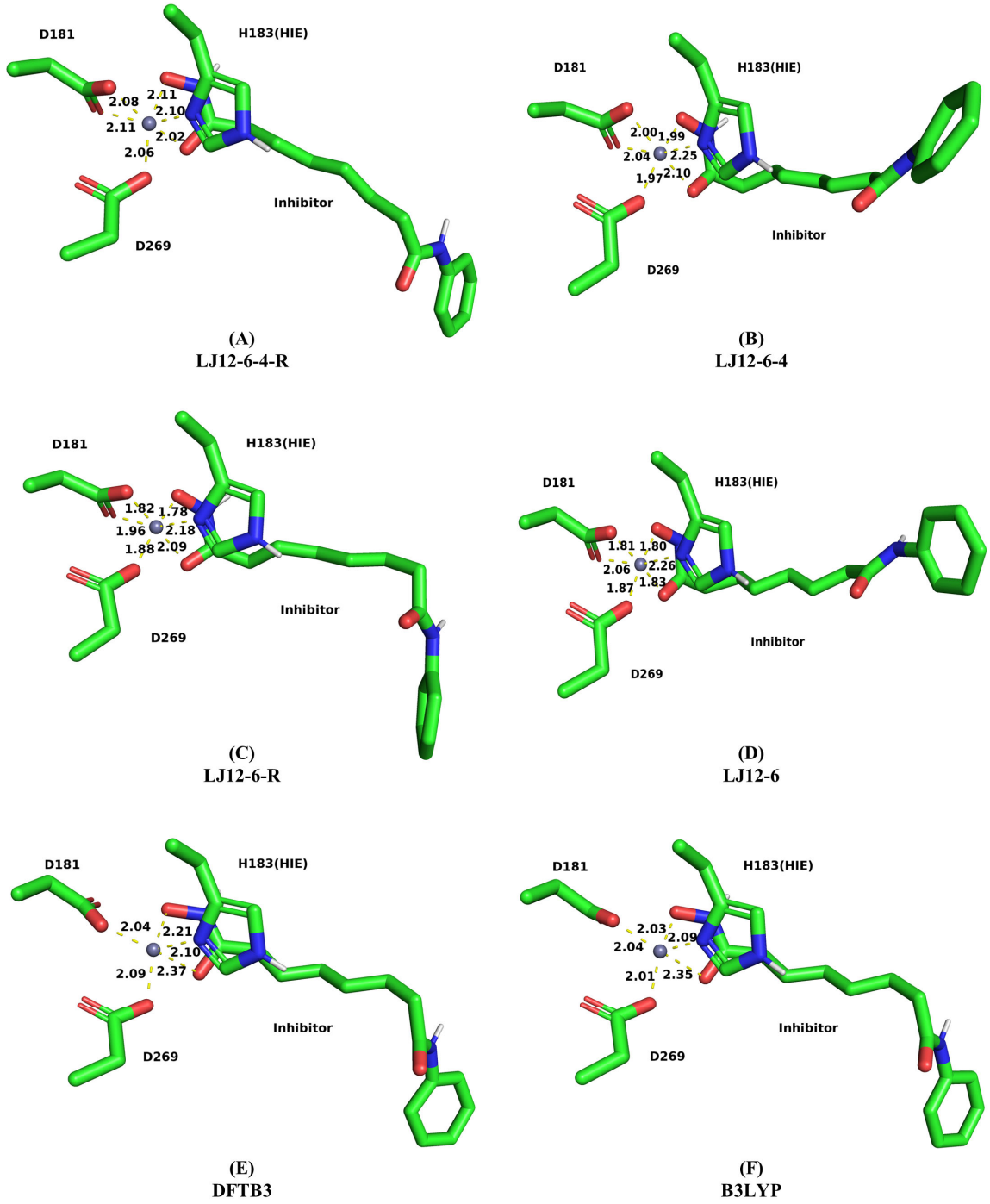


Figure S8. Representative zinc site geometries of the HDAC2:SAHA complex after MM MD simulations and QM optimization. (A-D) Representative zinc geometries of the four non-bonded models after 100ns MM MD simulation. (E) The representative zinc geometry after 100 ps DFTB3/MM MD simulation. Note restoration of mono-dentate zinc coordination by Asp 181. (F) Representative zinc geometry after B3LYP-D3BJ based QM/MM geometry optimization.

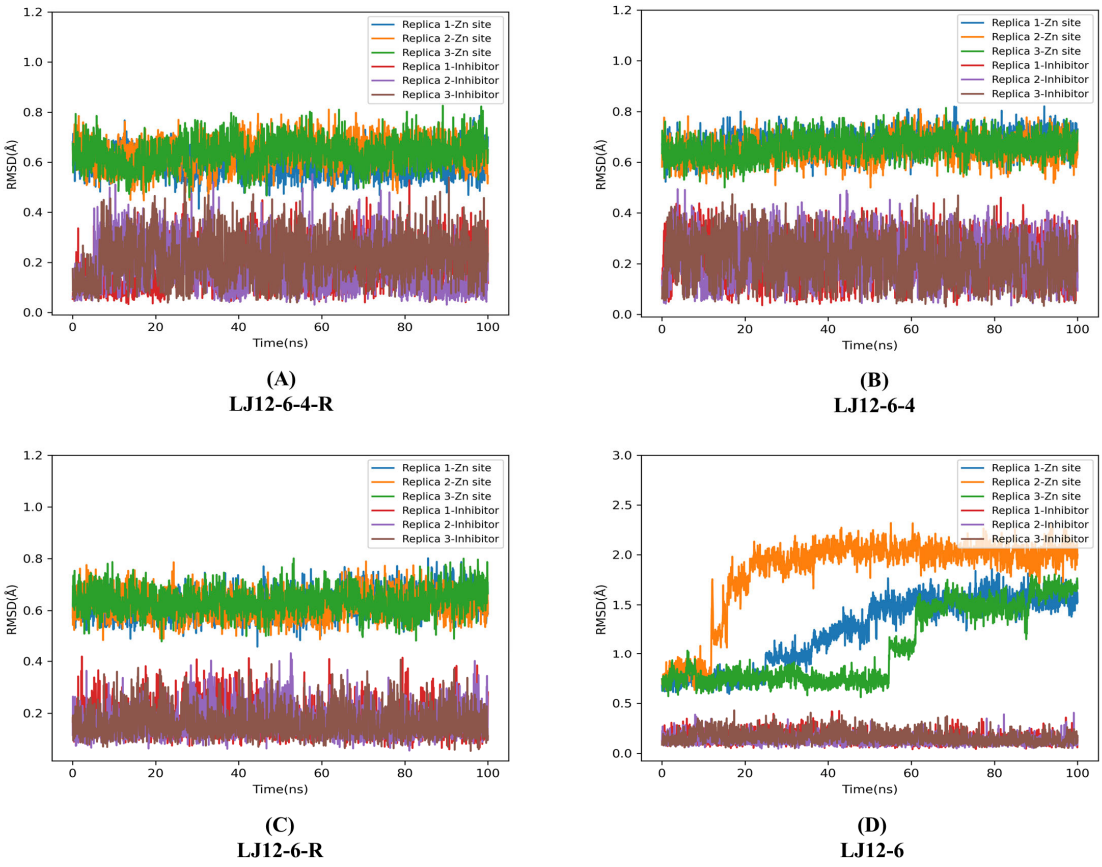


Figure S9. Time-dependence of RMSD values compared to the crystal structure for MM MD simulations of the IMP1:TGA complex using non-bonded models. Three replicate simulations were performed for each model and each replicate is 100 ns. The 'Zn site' refers to zinc ions and zinc coordinating residues and 'Inhibitor' refers to the zinc bound compound (TGA).

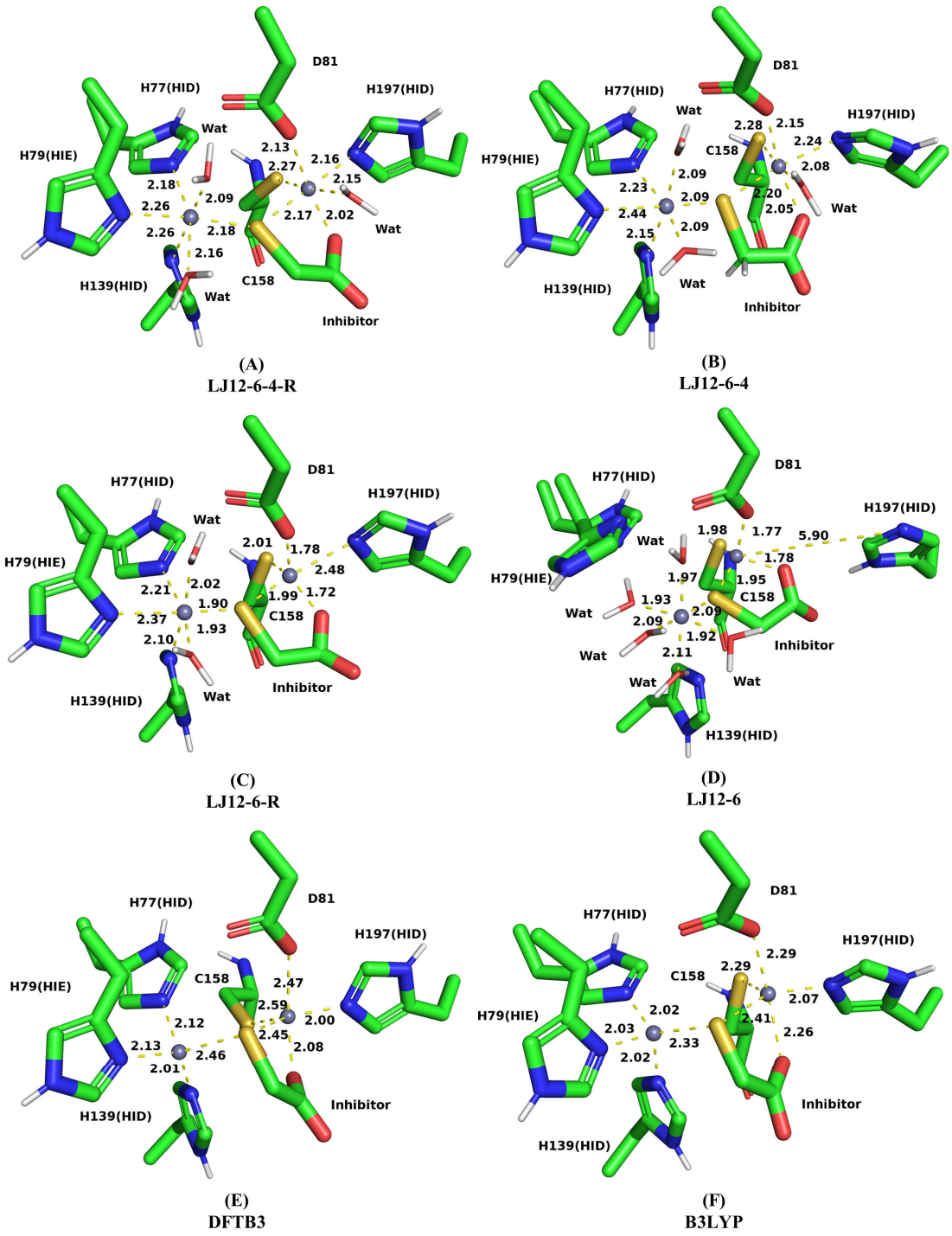


Figure S10. Representative zinc site geometries of the IMP1:TGA complex after MM MD simulations and QM optimization. (A - D) The representative zinc geometry of the four non-bonded models after 100ns MM MD simulation. (E) The representative zinc geometry after 100 ps DFTB3/MM MD simulation. The extra water molecule was removed. (F) The representative zinc geometry after B3LYP-D3BJ based QM/MM geometry optimization.

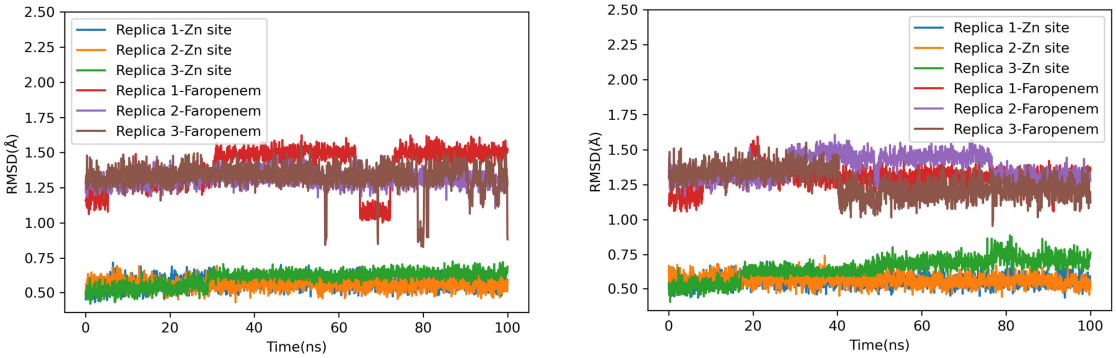


Figure S11. Time-dependence of RMSD values compared to the crystal structure for MM MD simulations of the L1:hydrolyzed faropenem complex using LJ12-6-4 non-bonded models. Three replicate simulations were performed for each model and each replicate is 100 ns. The 'Zn site' refers to zinc ions and zinc coordinating residues.

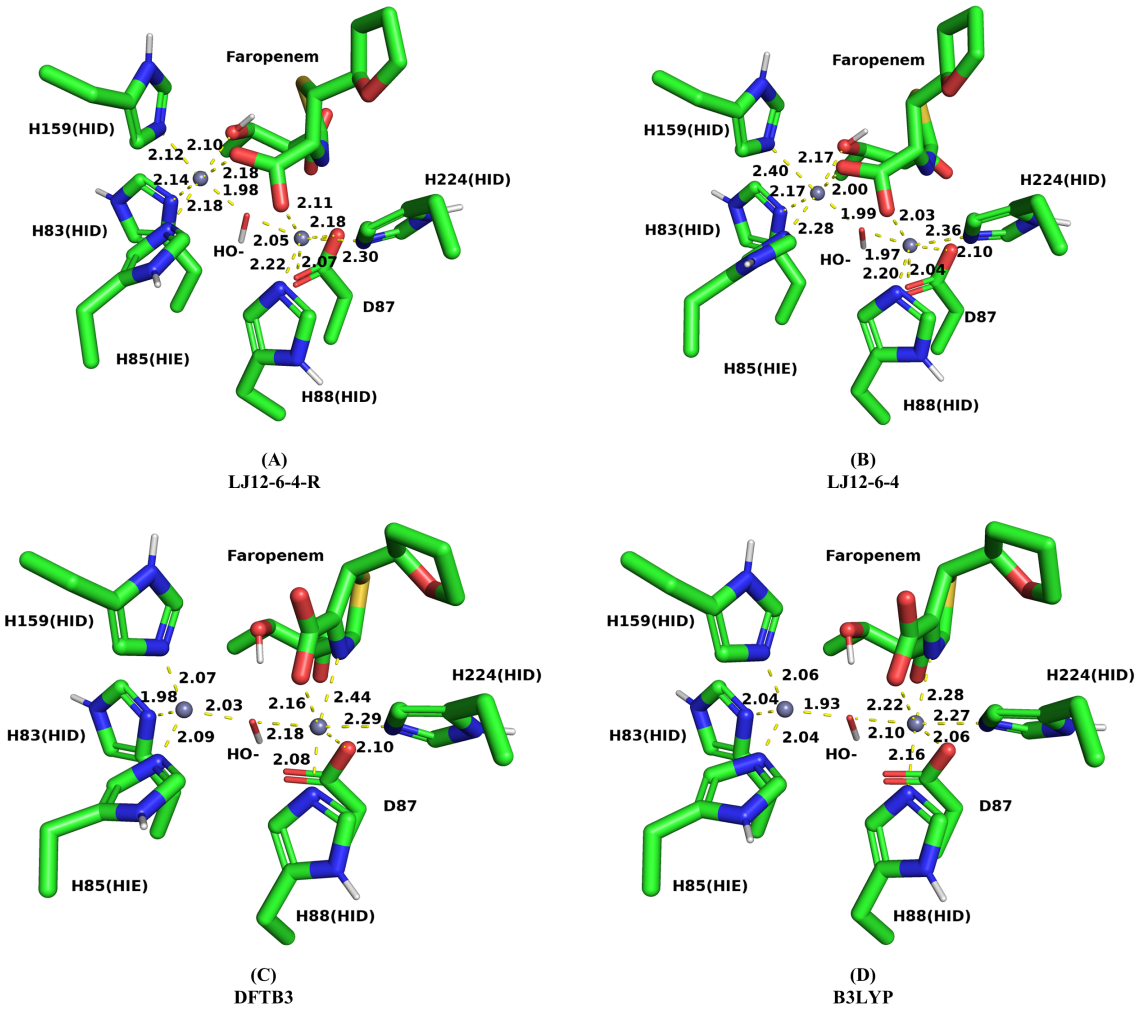


Figure S12. Representative zinc site geometries of the L1:hydrolyzed faropenem complex after MM MD simulations and QM optimization. (A-B) Representative zinc geometries for L1 complexes with hydrolyzed faropenem during MM MD simulations using LJ12-6-4 L1 non-bonded models with (A) and without (B) restraints, respectively. (C) Representative zinc geometry after 100 ps DFTB3/MM MD simulation. (D) The representative zinc geometry after B3LYP-D3BJ based QM/MM geometry optimization.

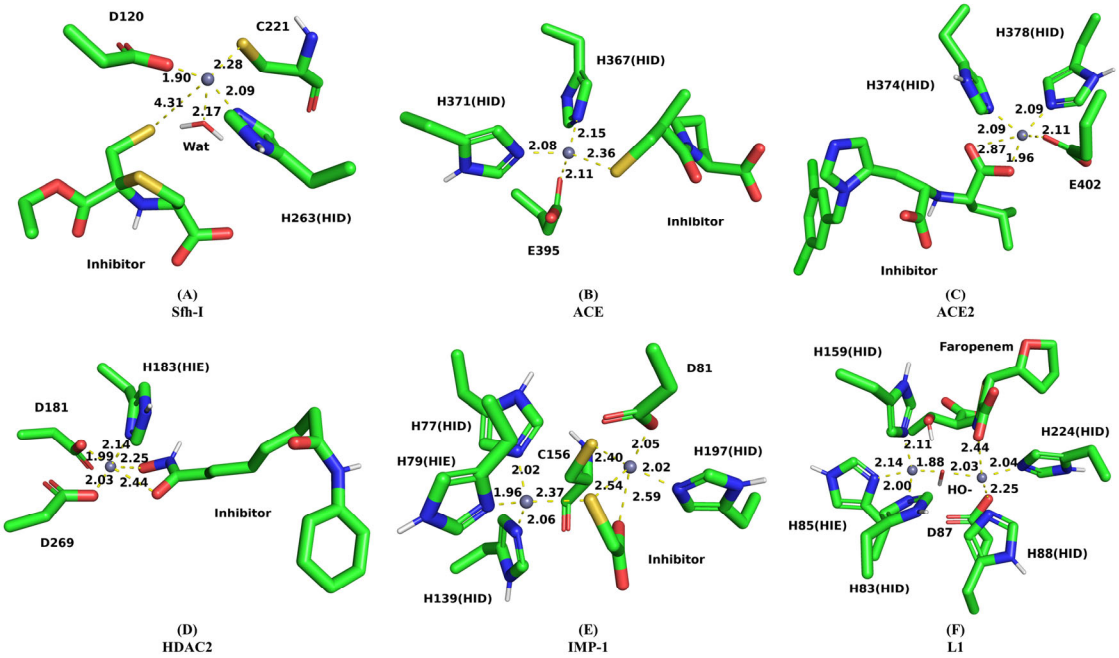


Figure S13. Outcome of DFTB3/MM MD run directly from crystal structures. Representative zinc site geometries of each protein-ligand complex after 2 ns DFTB3/MM MD simulation. Carbon atoms are in green, nitrogen atoms are in blue, oxygen atoms are in red and sulfur atoms are in dark yellow, zinc ions are represented as gray balls.

Table S1. Zn coordination distance (Å) of four non-bonded MM models for Sfh-I: L-anti-1a complex.

	Zn Coordinating Atom	LJ126	LJ126-R	LJ1264	LJ1264-R	Crystal
Mean Coordination Distance (Å)	Asp120-OD2	1.79	1.80	2.02	2.02	1.90
	Cys221-SG	1.99	2.03	2.20	2.28	2.37
	His263-NE2	3.73	2.44	2.29	2.23	2.10
	Inhibitor-S01	1.97	1.95	2.24	2.21	2.28
Mean CN		5.0	4.7	5.7	6.0	4.0
Mean Deviation	Asp120-OD2	-0.11	-0.10	0.12	0.12	
	Cys221-SG	-0.38	-0.34	-0.17	-0.09	
	His263-NE2	1.63	0.34	0.19	0.13	
	Inhibitor-S01	-0.31	-0.33	-0.04	-0.07	
TAD		2.43	1.11	0.53	0.41	
TAD per ligand		0.61	0.28	0.13	0.10	
Standard Error of TAD		1.16	1.24	0.06	0.47	
MPD	Inhibitor-S01	-5.6%	-5.4%	6.5%	6.3%	
	Asp120-OD2	-16.0%	-14.5%	-7.0%	-3.8%	
	Cys221-SG	77.6%	16.0%	9.2%	6.2%	
	His263-NE2	-13.6%	-14.5%	-1.9%	-3.2%	
Total Absolute MPD		112.9%	50.4%	24.6%	19.5%	

Deviation (Å) = simulation predicted value - crystal value; CN = Zn Coordination Number; TAD (Å) = Total Absolute Mean Deviation; MPD =Mean Percentage Deviation. The values presented are averages of the last frames of the three independent MM simulations.

Note: Additional Zn^{2+} coordinating atoms introduced during simulations that are not presented in the crystal structure are not reported in the table, but are included when calculating the mean coordination number.

Table S2. Coordination distances between the zinc ion and its coordinating atoms for QM/MM simulations of the Sfh-I:L-anti-1a complex.

		Action 1					DFTB3	
		Action 2		DFTB3	DFTB3	DFTB3	6-31G*/ GD3BJ	
	Zn Coordinating Atom	Action 3	DFTB3	6-31G*	6-31+G*	6-31G*/ GD3BJ	6-31+G*/ GD3BJ	Crystal
		Steps of Action 3		*250 steps*	*235 steps*[a]	*250 steps*	*250 steps*	
Coordination Distance (Å)	Inhibitor-S01		2.46	2.40	2.41	2.33	2.34	2.28
	Asp120-OD2		1.89	1.95	2.01	1.94	1.99	1.90
	Cys221-SG		2.58	2.49	2.43	2.46	2.39	2.37
	His263-NE2		2.13	2.10	2.12	2.04	2.09	2.10
Deviation	Inhibitor-S01		0.18	0.12	0.13	0.05	0.06	
	Asp120-OD2		-0.01	0.05	0.11	0.04	0.09	
	Cys221-SG		0.21	0.12	0.06	0.09	0.02	
	His263-NE2		0.03	0.00	0.02	-0.06	-0.01	
	Total Absolute		0.43	0.29	0.32	0.24	0.18	
Percentage Deviation	Inhibitor-S01		7.9%	5.3%	5.7%	2.2%	2.6%	
	Asp120-OD2		-0.5%	2.6%	5.8%	2.1%	4.7%	
	Cys221-SG		8.9%	5.1%	2.5%	3.8%	0.8%	
	His263-NE2		1.4%	0.0%	1.0%	-2.9%	-0.5%	
	Total Absolute		18.7%	13.0%	15.0%	11.0%	8.7%	
Total Absolute Deviation per ligand			0.11	0.07	0.08	0.06	0.05	
Total Absolute Deviation per ligand (%)			4.7%	3.2%	3.7%	2.7%	2.2%	

Deviation (Å) = simulation predicted value - crystal value; CN = Zn Coordination Number; TAD (Å) = Total Absolute Mean Deviation. DFTB3 refers to 100ps DFTB3/MM MD; 6-31G* and 6-31+G* refer to the basis set used in BL3YP based QM/MM geometry optimization; '+' in a basis set represents the diffuse function; GD3BJ is the empirical dispersion correction.

[a] The geometry optimization has reached the convergency criteria.

Table S3. Atomic distances (Å) between the zinc ions and the zinc coordinating atoms for simulations of the ACE inhibitor complex model.

	Zn Coordinating Atom	Action 1 Action 2 Action 3 Action 4	LJ1264[a]	LJ1264-R[a]	LJ1264 DFTB3 6-31G*	LJ1264 DFTB3 6-31+G*	LJ1264 DFTB3 6-31+G*	Crystal
Coordination Distance (Å)	His367-NE2		2.21	2.18	2.04	2.00	2.04	2.09
	His371-NE2		2.23	2.20	2.06	2.04	2.05	2.08
	Glu395-OE1		2.06	2.06	2.01	2.00	2.06	2.02
	Inhibitor-S		2.15	2.15	2.36	2.30	2.29	2.14
CN			6.0	6.0	4.0	4.0	4.0	4.0
Deviation	His367-NE2		0.12	0.09	-0.05	-0.09	-0.05	
	His371-NE2		0.15	0.12	-0.02	-0.04	-0.03	
	Glu395-OE1		0.04	0.04	-0.01	-0.02	0.04	
	Inhibitor-S		0.01	0.01	0.22	0.16	0.15	
TAD			0.32	0.26	0.30	0.31	0.27	
TAD per ligand			0.08	0.07	0.08	0.08	0.07	

Deviation (Å) = simulation predicted value - crystal value; CN = Zn Coordination Number; TAD (Å) = Total Absolute Mean Deviation. DFTB3 refers to 100ps DFTB3/MM MD; 6-31G* and 6-31+G* refer to the basis set used in BL3YP-D3BJ based QM/MM geometry optimization.

Note: Additional Zn^{2+} coordinating atoms introduced during simulations that are not presented in the crystal structure are not reported in the table, but are included when calculating the mean coordination number.

[a] The values presented are averages of the last frames of the three independent MM simulations.

Table S4. Atomic distances (Å) between the zinc ions and the zinc coordinating atoms for simulations of the ACE2 inhibitor complex model.

	Zn Coordinating Atom	Action 1 Action 2 Action 3 Action 4	LJ1264[a]	LJ1264-R[a]	LJ1264 DFTB3	LJ1264 DFTB3 6-31G*	LJ1264 DFTB3 6-31+G*	Crystal
Coordination Distance (Å)	His374-NE2		2.26	2.29	2.01	1.99	2.00	2.05
	His378-NE2		2.26	2.22	1.99	2.11	2.13	2.31
	Glu402-OE1		2.09	2.09	2.04	1.99	2.00	2.05
	Inhibitor-O1		2.03	2.01	1.98	1.94	1.96	2.04
CN			6.0	6.0	4.0	4.0	4.0	4.0
Deviation	His374-NE2		0.21	0.24	-0.04	-0.06	-0.05	
	His378-NE2		-0.05	-0.09	-0.32	-0.20	-0.18	
	Glu402-OE1		0.04	0.04	-0.01	-0.06	-0.05	
	Inhibitor-O1		-0.01	-0.03	-0.06	-0.10	-0.08	
TAD			0.31	0.40	0.43	0.42	0.36	
TAD per ligand			0.08	0.10	0.11	0.11	0.09	

Deviation (Å) = simulation predicted value - crystal value; CN = Zn Coordination Number; TAD (Å) = Total Absolute Mean Deviation. DFTB3 refers to 100ps DFTB3/MM MD; 6-31G* and 6-31+G* refer to the basis set used in BL3YP-D3BJ based QM/MM geometry optimization.

Note: Additional Zn^{2+} coordinating atoms introduced during simulations that are not presented in the crystal structure are not reported in the table, but are included when calculating the mean coordination number.

[a] The values presented are averages of the last frames of the three independent MM simulations.

Table S5. Zn coordination distance (Å) of four non-bonded MM models for the HDAC-2:SAHA complex.

	Zn Coordinating Atom	LJ126	LJ126-R	LJ1264	LJ1264-R	Crystal
Mean Coordination Distance (Å)	Asp181-OD2	2.01	1.85	2.05	2.04	2.00
	His183-ND1	2.18	2.16	2.25	2.18	2.07
	Asp269-OD2	1.85	1.84	2.01	2.00	1.90
	Inhibitor-O1	1.82	1.82	2.02	2.07	1.96
	Inhibitor-O2	1.94	2.04	2.10	2.10	2.33
Mean CN		6.0	6.0	6.0	6.0	5.0
Mean Deviation	Asp181-OD2	0.01	-0.15	0.05	0.04	
	His183-ND1	0.11	0.09	0.18	0.11	
	Asp269-OD2	-0.05	-0.06	0.11	0.10	
	Inhibitor-O1	-0.14	-0.14	0.06	0.11	
	Inhibitor-O2	-0.39	-0.29	-0.23	-0.23	
TAD		0.70	0.73	0.63	0.58	
TAD per ligand		0.14	0.15	0.13	0.12	
Standard Error of TAD		0.06	0.04	0.02	0.05	
MPD	Asp181-OD2	0.5%	-7.7%	2.7%	2.2%	
	His183-ND1	5.2%	4.3%	8.5%	5.2%	
	Asp269-OD2	-2.8%	-3.3%	5.6%	5.3%	
	Inhibitor-O1	-7.0%	-7.0%	3.1%	5.4%	
	Inhibitor-O2	-16.7%	-12.3%	-9.9%	-9.7%	
Total Absolute MPD		32.2%	34.6%	29.7%	27.8%	

Deviation (Å) = simulation predicted value - crystal value; Percentage Deviation = Deviation value / crystal value; CN = Zn Coordination Number; TAD (Å) = Total Absolute Mean Deviation; MPD =Mean Percentage Deviation.

Note: The values presented are averages of the last frames of the three independent MM simulations. Additional Zn^{2+} coordinating atoms introduced during simulations that are not presented in the crystal structure are not reported in the table, but are included when calculating the mean coordination number.

Table S6. QM/MM optimized distances (Å) between the zinc ion and its coordinating atoms for simulations of the HDAC2:SAHA complex.

		Action 1		DFTB3			
		Action 2		DFTB3	250 steps 6-31G*	DFTB3	
	Zn Coordinating Atoms	Action 3	DFTB3	6-31G*	6-31+G*	6-31G*	Crystal
		Steps of Action 3	100 ps	250 steps	250 steps	542 steps[a]	
Coordination Distance (Å)	Asp181-OD2		2.04	1.99	2.04	1.99	2.00
	His183-ND1		2.10	2.08	2.09	2.07	2.07
	Asp269-OD2		2.09	1.98	2.00	2.01	1.90
	Inhibitor-O1		2.21	1.99	2.03	1.99	1.96
	Inhibitor-O2		2.37	2.40	2.35	2.35	2.33
CN			5.0	5.0	5.0	5.0	5.0
Deviation	Asp181-OD2		0.04	-0.01	0.04	-0.01	
	His183-ND1		0.03	0.01	0.02	0.00	
	Asp269-OD2		0.19	0.08	0.10	0.11	
	Inhibitor-O1		0.25	0.03	0.07	0.03	
	Inhibitor-O2		0.04	0.07	0.02	0.02	
TAD			0.55	0.20	0.25	0.17	
TAD per ligand			0.11	0.04	0.05	0.03	

Deviation (Å) = simulation predicted value - crystal value; CN = Zn Coordination Number. Total Absolute Mean Deviation. DFTB3 refers to 100ps DFTB3/MM MD; 6-31G* and 6-31+G* refer to the basis set used in BL3YP-D3BJ based QM/MM geometry optimization.

[a] The geometry optimization has reached the convergency criteria.

Table S7. Zn coordination distance (Å) of four non-bonded MM models for simulations of the IMP-1:TGA complex.

	Atom-1	Atom-2	LJ126	LJ126-R	LJ1264	LJ1264-R	Crystal
Mean Atomic Distance (Å)	His77-NE2	Zn301	4.24	2.24	2.21	2.17	2.06
	His79-ND1	Zn301	4.31	2.30	2.29	2.24	2.05
	His139-NE2	Zn301	4.70	2.09	2.15	2.26	2.04
	Inhibitor-S2	Zn301	2.05	1.95	2.16	2.17	2.22
	Asp81-OD2	Zn302	1.73	1.81	2.17	2.04	2.03
	Cys158-SG	Zn302	2.00	2.04	2.24	2.24	2.28
	His197-NE2	Zn302	6.59	2.44	2.26	2.23	2.13
	Inhibitor-S2	Zn302	1.96	2.04	2.20	2.20	2.36
	Inhibitor-O2	Zn302	1.74	1.78	2.06	2.04	2.25
	Zn301	Zn302	3.88	3.82	4.21	4.02	3.55
Mean CN	Zn301		6.0	6.0	6.0	6.0	4.0
	Zn302		4.0	5.0	6.0	6.0	5.0
Mean Deviation	HIS77-NE2	Zn1	2.18	0.18	0.15	0.11	
	HIS79-ND1	Zn1	2.26	0.25	0.24	0.19	
	HIS139-NE2	Zn1	2.66	0.05	0.11	0.22	
	Inhibitor-S2	Zn1	-0.17	-0.27	-0.06	-0.05	
	ASP81-OD2	Zn2	-0.30	-0.22	0.14	0.01	
	CYS158-SG	Zn2	-0.28	-0.24	-0.04	-0.04	
	HIS197-NE2	Zn2	4.46	0.31	0.13	0.10	
	Inhibitor-S2	Zn2	-0.40	-0.32	-0.16	-0.16	
	Inhibitor-O2	Zn2	-0.51	-0.47	-0.19	-0.21	
Mean Deviation for Zn301 Coordination			7.27	0.74	0.56	0.57	
Mean Deviation for Zn302 Coordination			5.96	1.56	0.67	0.52	
TAD			13.23	2.30	1.23	1.09	
TAD per ligand			1.47	0.26	0.14	0.12	
Standard Error of TAD			1.80	0.27	0.17	0.01	

Deviation (Å) = simulation predicted value - crystal value; CN = Zn Coordination Number; TAD (Å) = Total Absolute Mean Deviation. The values presented are averages of the last frames of the three independent MM simulations.

Note: Additional Zn^{2+} coordinating atoms introduced during simulations that are not presented in the crystal structure are not reported in the table, but are included when calculating the mean coordination number.

Table S8. Representative QM/MM optimized atomic distances (Å) between the zinc ions and the zinc coordinating atoms for simulations of the IMP-1:TGA complex model.

			Action 1			**DFTB3**		**DFTB3**	
			Action 2		**DFTB3**	**250 steps 6-31G***	**DFTB3**	**449 step 6-31G***	
	Atom-1	**Atom-2**	**Action 3**	**DFTB3**	**6-31G***	**6-31+G***	**6-31G***	**6-31+G***	**Crystal**
			Steps of Action 3		250 steps	340 steps[a]	449 steps[a]	250 steps[a]	
Atomic Distance (Å)	His77-NE2	Zn301		2.08	2.02	2.05	2.02	2.04	2.06
	His79-ND1	Zn301		2.08	2.04	2.07	2.03	2.05	2.05
	His139-NE2	Zn301		2.08	2.03	2.06	2.02	2.06	2.04
	Inhibitor-S2	Zn301		2.20	2.32	2.32	2.33	2.34	2.22
	Asp81-OD2	Zn302		2.36	2.27	2.16	2.29	2.18	2.03
	Cys158-SG	Zn302		2.43	2.30	2.34	2.29	2.29	2.28
	His197-NE2	Zn302		2.01	2.06	2.08	2.07	2.10	2.13
	Inhibitor-S2	Zn302		2.43	2.41	2.43	2.41	2.41	2.36
	Inhibitor-O2	Zn302		2.37	2.28	2.48	2.26	2.44	2.25
Deviation	His77-NE2	Zn301		0.02	-0.04	-0.01	-0.04	-0.02	
	His79-ND1	Zn301		0.03	-0.01	0.02	-0.02	0.00	
	His139-NE2	Zn301		0.04	-0.01	0.02	-0.02	0.02	
	Inhibitor-S2	Zn301		-0.02	0.10	0.10	0.11	0.12	
	Asp81-OD2	Zn302		0.33	0.24	0.13	0.26	0.15	
	Cys158-SG	Zn302		0.15	0.02	0.06	0.01	0.01	
	His197-NE2	Zn302		-0.12	-0.07	-0.05	-0.06	-0.03	
	Inhibitor-S2	Zn302		0.07	0.05	0.07	0.05	0.05	
	Inhibitor-O2	Zn302		0.12	0.03	0.23	0.01	0.19	
TAD for Zn301 Coordination				0.11	0.16	0.15	0.19	0.16	
TAD for Zn302 Coordination				0.79	0.41	0.54	0.39	0.43	
TAD per ligand				0.10	0.06	0.08	0.06	0.07	

Deviation (Å) = simulation predicted value - crystal value; TAD (Å) = Total Absolute Mean Deviation. DFTB3 refers to 100ps DFTB3/MM MD; 6-31G* and 6-31+G* refer to the basis set used in BL3YP-D3BJ based QM/MM geometry optimisation.

[a] The geometry optimization has reached the convergency criteria.

Table S9. Representative atomic distances (Å) between the zinc ions and the zinc coordinating atoms for simulations of the L1:Faropenem complex model.

			Action 1					LJ1264-R	
			Action 2				LJ1264-R	DFTB3	
			Action 3			LJ1264-R	DFTB3	6-31G*	
	Atom-1	Atom-2	Action 4	LJ1264[a]	LJ1264-R[a]	DFTB3	6-31G*	6-31+G*	Crystal
Atomic Distance (Å)	HIS83-ND1	Zn267		2.20	2.17	1.98	2.02	2.04	2.06
	HIS85-NE2	Zn267		2.23	2.16	2.09	2.03	2.05	2.08
	HIS159-NE2	Zn267		2.29	2.15	2.07	2.04	2.06	2.05
	HO-	Zn267		1.99	1.97	2.03	1.92	1.93	1.92
	HIS88-NE2	Zn268		2.18	2.27	2.08	2.15	2.16	2.15
	ASP87-OD2	Zn268		2.07	2.12	2.10	2.04	2.06	2.12
	HIS224-NE2	Zn268		2.29	2.29	2.29	2.29	2.27	2.11
	HO-	Zn268		2.00	2.02	2.18	2.08	2.10	2.07
	Faropenem-O10	Zn268		2.53	2.48	2.16	2.12	2.22	2.22
	Faropenem-N06	Zn268		2.43	2.48	2.44	2.30	2.28	2.31
	Zn267	Zn268		3.68	3.74	3.95	3.87	3.90	3.57
CN	Zn267			6.0	6.0	4.0	4.0	4.0	4.0
	Zn268			6.3	6.0	6.0	6.0	6.0	6.0
Deviation	HIS83-ND1	Zn267		0.14	0.11	-0.08	-0.04	-0.02	
	HIS85-NE2	Zn267		0.15	0.08	0.01	-0.05	-0.03	
	HIS159-NE2	Zn267		0.24	0.10	0.02	-0.01	0.01	
	HO-	Zn267		0.07	0.05	0.11	0.00	0.01	
	HIS88-NE2	Zn268		0.03	0.12	-0.07	0.00	0.01	
	ASP87-OD2	Zn268		-0.05	0.00	-0.02	-0.08	-0.06	
	HIS224-NE2	Zn268		0.18	0.18	0.18	0.18	0.16	
	HO-	Zn268		-0.07	-0.05	0.11	0.01	0.03	
	Faropenem-O10	Zn268		0.31	0.26	-0.06	-0.10	0.00	
	Faropenem-N06	Zn268		0.12	0.17	0.13	-0.01	-0.03	
Deviation for Zn267 Coordination				0.60	0.34	0.22	0.10	0.07	
Deviation for Zn268 Coordination				0.76	0.78	0.57	0.38	0.29	
TAD				1.35	1.12	0.79	0.48	0.36	
TAD/Crystal ligand				0.14	0.11	0.08	0.05	0.04	
Standard Error of TAD				0.47	0.48	/	/	/	

Deviation (Å) = simulation predicted value - crystal value; CN = Zn Coordination Number; TAD (Å) = Total Absolute Mean Deviation. DFTB3 refers to 100ps DFTB3/MM MD; 6-31G* and 6-31+G* refer to the basis set used in BL3YP-D3BJ based QM/MM geometry optimisation.

Note: Additional Zn^{2+} coordinating atoms introduced during simulations that are not presented in the crystal structure are not reported in the table, but are included when calculating the mean coordination number.

[a] The values presented are averages of the last frames of the three independent MM simulations.

4.3 Computational investigation of MCR-like protein:ligand complexes

Statement.

The following material is not a part of the manuscript and is first reported in this PhD thesis.

4.3.1 Modelling the ICR^{Mc}:PEtN complex

Based on the methodology for modelling zinc metalloproteins established in section 4.2, we sought to apply the computational approach to MCR-1 like proteins. The non-covalent complex of the catalytic domain of the ICR^{Mc} Thr315Ala mutant and its substrate phosphoethanolamine (PEtN) (PDB code: 6BND[183]; Resolution: 1.66 Å,) was first modelled. The complex of the apo form of ICR^{Mc} is not available. In the substrate complex, the zinc ion is tetrahedrally coordinated by three protein ligands (Glu273, Asp497 and His499, equivalent to Glu246, Asp465 and His466 in MCR-1 structure) and the substrate PEtN (**Figure 4.1**). The zinc ion and all its coordinating ligands are designated as the 'Zn site'. Although His429 and His511 (equivalent to His395 and His478 in MCR-1 structure) are not coordinating the zinc ion, they are included in the extended zinc site (Ex Zn site) as they may play important roles in substrate binding. The protonation states of the two Histidine residues were set to be neutral according to the program PROPKA 3.1.

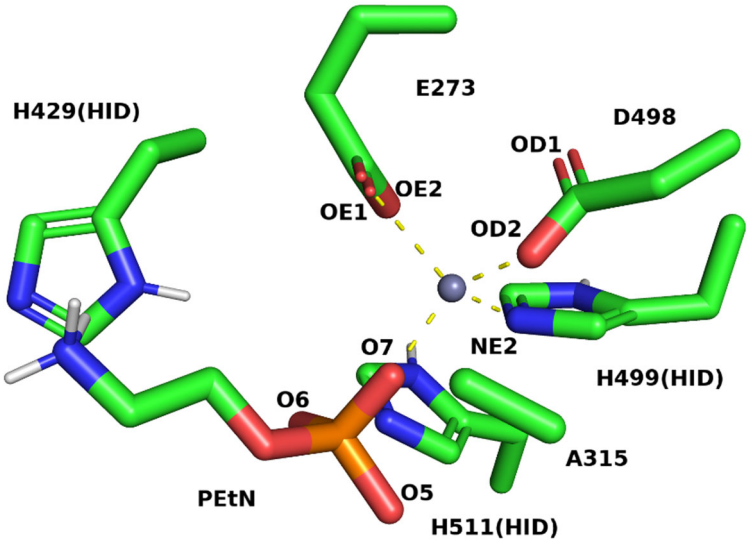


Figure 4-1. Zinc site geometry of the ICR^{Mc}:PEtN complex crystal structure.

ICR^{Mc} Thr315Ala mutant with substrate PEtN, coordination number (CN) = 4, PDB code: 6BND[183]. Carbon atoms are in green, nitrogen atoms are in blue, oxygen atoms are in red and sulfur atoms are in yellow, with zinc ions represented as gray balls. Dashed lines indicate zinc coordinating ligands.

4.3.1.1 MM MD simulations

Unrestrained LJ12-6-4 and restrained LJ12-6-4 MM MD simulations were carried out first to examine the extent of overall conformational changes and the effect on the system of different protonation states. Unrestrained MD simulations were used to examine the potential movement of zinc binding residues, while restrained simulations are used to simulate the dynamics of the zinc site retaining the crystallographically defined geometry as much as possible. RMSD plots (compared to the crystal structures) of the zinc sites of the individual models are presented in **Figure 4-2**.

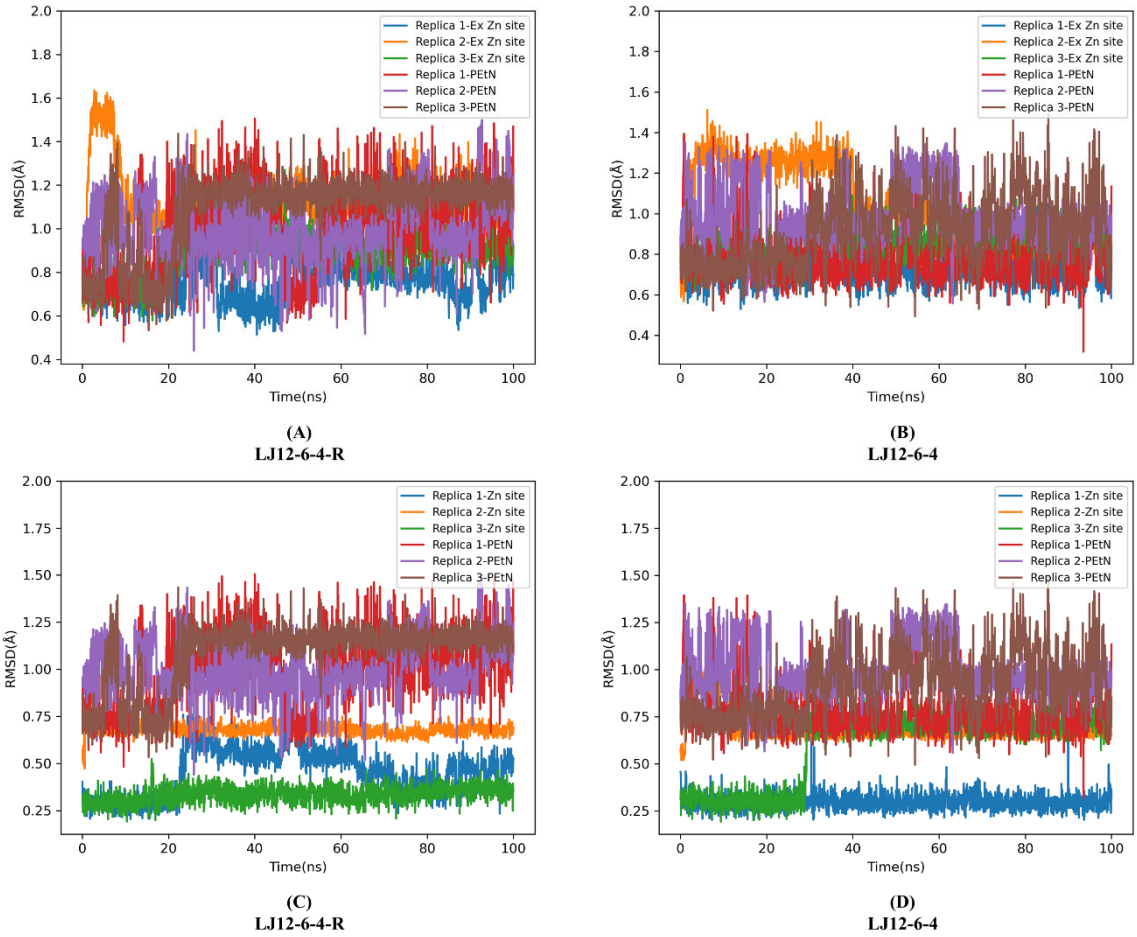


Figure 4-2. Time-dependence of RMSD values (compared to the crystal structure) for MM MD simulations of ICR^{Mc}:PEtN complex using LJ 12-6-4 non-bonded models (HID511).

Simulations were performed in triplicate for each model and each replicate is 100 ns. The 'Zn site' refers to zinc ions and zinc coordinating residues and 'Ex Zn site' refers to the extended zinc binding site including zinc ions, zinc coordinating residues, His429 and His511.

For both the restrained and unrestrained LJ12-6-4 simulations, the RMSD values of residues within the extended zinc binding site are consistently around 1.2 Å, suggesting that a stable zinc site geometry is maintained across the 100 ns simulation time (**Figure 4-2 A&B**). The RMSD value of the zinc bound substrate PEtN fluctuated between 0.9 and 1.2 Å in both the restrained model and the unrestrained model, compared to the crystal structure. The RMSD values of the protein backbone atoms from both models fluctuated and

converged at 1.0 -1.3 Å. In general, the unrestrained LJ12-6-4 model performed similarly to the restrained LJ12-6-4 model. The stable RMSDs (with only small fluctuations observed) suggests that on this time scale the system is not undergoing significant conformational changes from the experimental crystal structure. However, in both the unrestrained and restrained LJ12-6-4 types of MD simulations (each carried out in triplicate), the substrate PEtN lost its interaction with the zinc ion in one out of the three replicate simulations. One possible cause for this might be an incorrect protonation state of residues that interact with bound PEtN. RMSD plots of the zinc site excluding His429 and His511were between 0.2-0.8 Å (**Figure 4-2 C&D**), which were lower than the RMSD values with the inclusion of His429 and His511. This suggests that the movement of residues in the simulation was predominantly associated with the two Histidine residues. According to the work of Stogios et al.[183], the protonation state of His511 should be HIE511 (hydrogen on the epsilon nitrogen) while PROPKA 3.1 suggested His511 to be HID511 (hydrogen on the delta nitrogen). Unrestrained and restrained LJ12-6-4 MM MD simulations were performed in triplicate with His511 in the HIE511 protonation state.

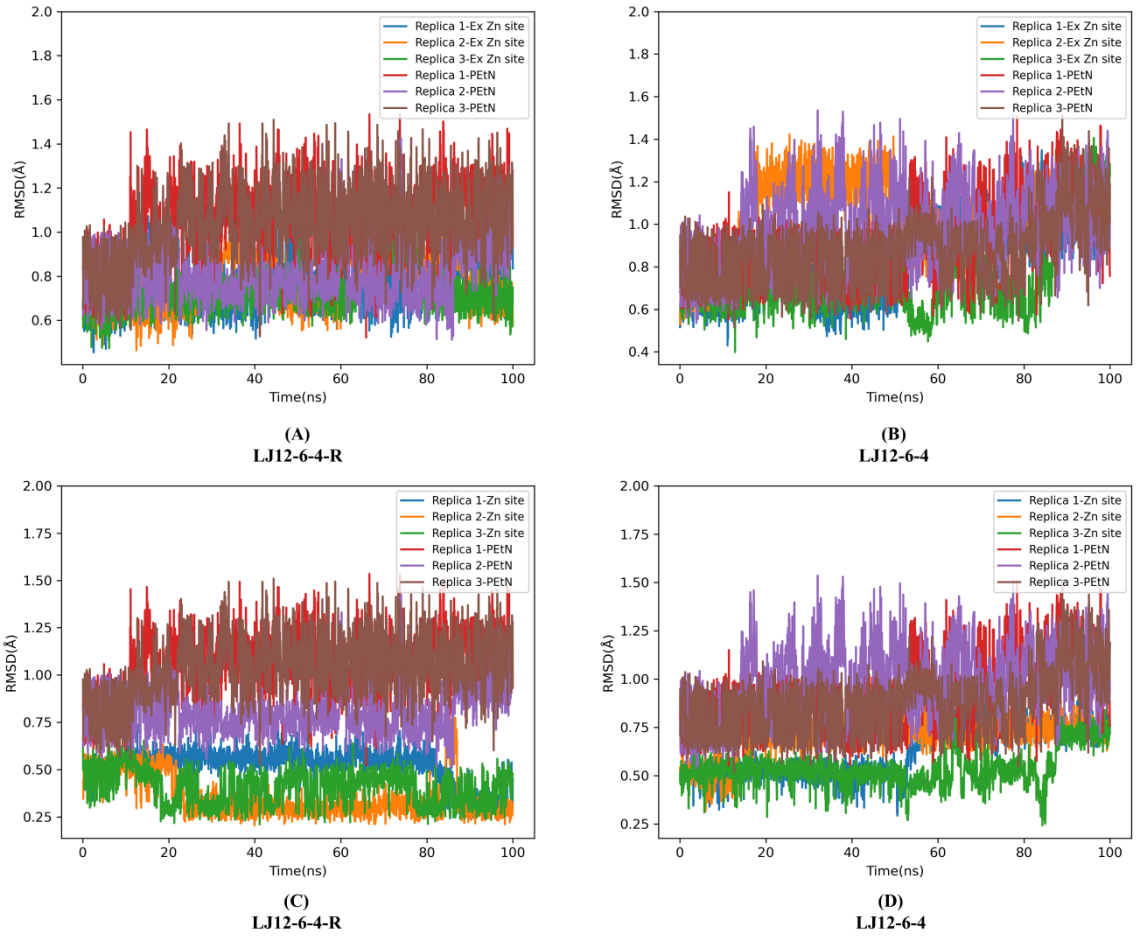


Figure 4-3. Time-dependence of RMSD values (compared to the crystal structure) for MM MD simulations of ICR^{Mc}:PEtN complex using LJ 12-6-4 non-bonded models (HIE511).

Simulations were performed in triplicate for each model and each replicate is 100 ns. The 'Zn site' refers to zinc ions and zinc coordinating residues and 'Ex Zn site' refers to the extended zinc binding site including zinc ions, zinc coordinating residues, His429 and His511.

With HIE511, RMSD values of the extended zinc binding site described by the restrained LJ12-6-4 model were between 0.8 Å and 1.0Å, 0.2 Å lower than the values from simulations with HID511 (**Figure 4-3A**). However, RMSD values of the extended zinc binding site described by the unrestrained LJ12-6-4 model fluctuated from 0.6 up to 1.4 Å (**Figure 4-3B**). RMSD plots of the zinc site excluding His429 and His511 showed that the RMSD values of restrained LJ12-6-4 simulations are 0.4 Å lower than those from simulations using the unrestrained LJ12-6-4 model (**Figure 4-3 C&D**). This suggested that the change from HID511 to HIE511 makes the zinc binding site geometry harder to maintain. Predictions of zinc binding site geometries from simulations using HIE511 were also worse than those obtained from simulations using HID511 in terms of preservation of PEtN binding. Specifically, the substrate PEtN lost its interaction with the zinc ion in 5 out of the total of 6 simulations. Altogether, the results of these simulations favoured assignment of His511 as the HID511, over the HIE511, form.

4.3.1.2 QM/MM refinement after the MM MD simulation

Subsequent DFTB3/MM MD simulations were carried out starting from the last frame of the 100ns MM MD simulation (a representative snapshot as the RSMD has converged by this point) and were followed by two 250-step B3LYP-D3BJ based QM/MM geometry optimizations, using the 6-31G(d) and the 6-31+G(d) basis sets, respectively. The DFTB3 MM MD step was trialed starting from the last frames of the restrained LJ12-6-4 MM MD simulations in which PEtN was preserved bound to the zinc ion. QM/MM MD simulations starting from the last frames of MM MD simulations using either the unrestrained or restrained LJ12-6-4 models should yield similar outcomes due to the negligible differences between the conformations predicted by the LJ12-6-4 or LJ12-6-4 models. The QM region includes the zinc ion, Glu273, Asp498, His499, PEtN and a zinc-coordinating water molecule that was introduced during the MM MD simulations. After 100ps of DFTB3/MM MD simulation, the extra zinc coordinating water molecule was removed from the zinc binding site and zinc coordination by Glu247 was shifted from bidentate to monodentate, restoring the zinc binding site to the tetrahedral geometry as observed in the starting crystal structure. Refinement using B3LYP-D3BJ QM/MM geometry optimization further optimized zinc coordination distances improving the accuracy of the model. In general, the QM/MM optimized structures were close to the conformations observed in the crystal structures. Zinc coordination geometries from the various treatments are reported in **Table 4-1.** Representative geometries for the ICR^{Mc} zinc centres at different stages of the modelling workflow are presented in **Figure 4-4**. Alignment of the ICR^{Mc} Zn sites to the crystal structures is shown in **Figure 4-15A.**

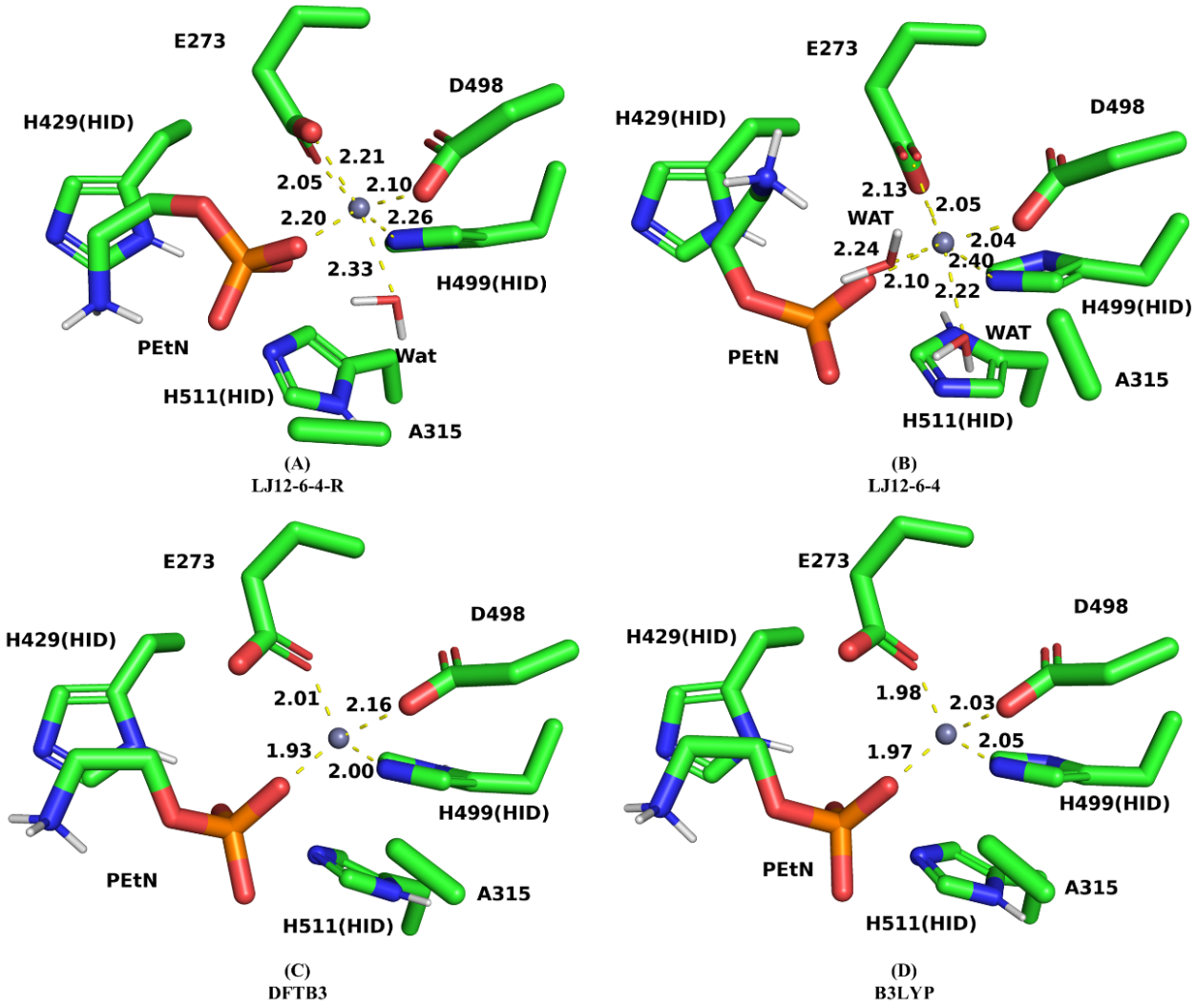


Figure 4-4. Zinc site geometries of ICR^{Mc}:PEtN complexes.

(A-B) Representative zinc geometries after MM MD simulations using LJ12-6-4 non-bonded models with (A) and without (B) restraints, respectively. (C) The representative zinc geometry after 100 ps DFTB3/MM MD simulation. (D) The representative zinc geometry after B3LYP-D3BJ based QM/MM geometry optimization.

Table 4-1. Atomic distances (Å) between the zinc ion and the zinc coordinating atoms for simulations of the ICR^{Mc}:PEtN complex.

	Zinc coordinating atom	Action 1 Action 2 Action 3 Action 4	LJ1264[a]	LJ1264-R[a]	LJ1264-R DFTB3 6-31G*	LJ1264-R DFTB3 6-31+G*	LJ1264-R DFTB3 6-31+G*	Crystal
Coordination Distance (Å)	Glu273-OE2		2.10	2.07	2.01	1.96	1.98	2.10
	Asp498-OD2		2.07	2.11	2.16	2.00	2.03	1.85
	His499-NE1		2.94	2.24	2.00	2.05	2.05	2.08
	PEtN-O7		2.68	12.36[b]	1.93	1.95	1.97	1.98
CN			6.0	6.0	4.0	4.0	4.0	4.0
Deviation	Glu273-OE2		0.00	-0.03	-0.09	-0.14	-0.12	
	Asp498-OD2		0.22	0.26	0.31	0.15	0.18	
	His499-NE1		0.86	0.16	-0.08	-0.03	-0.03	
	PEtN-O7		0.70	10.38	-0.05	-0.03	-0.01	
TAD			1.79	10.83	0.53	0.35	0.34	
TAD per ligand			0.45	2.71	0.13	0.09	0.09	

Deviation (Å) = simulation predicted value - crystal value; CN = Zn Coordination Number; TAD (Å) = Total Absolute Mean Deviation. DFTB3 refers to 100ps DFTB3/MM MD; 6-31G* and 6-31+G* refer to the basis set used in BL3YP-D3BJ based QM/MM geometry optimization.

Note: Additional Zn^{2+} coordinating atoms introduced during simulations that are not presented in the crystal structure are not reported in the table, but are included when calculating the mean coordination number.

[a] The values presented are averages of the last frames of three independent MM MD simulations.

[b] This value is the average of the three distance values: 32.90 Å, 2.20 Å and 1.98 Å. It is biased by the largest number 32.90 Å.

4.3.1.3 Direct DFTB3/MM MD simulations from the crystal structure

DFTB3/MM MD simulations directly starting from the crystal structure were also carried out. Both protonation states of His511 (HID and HIE) were tested. The simulations were run in triplicate with 2ns MD simulation for each replicate. Two sizes of the QM region were tested: the standard QM region and the extended QM region. The standard QM region includes the zinc ion, zinc coordinating protein residues (Glu273, Asp498 and His499) and the substrate PEtN. The extended QM region is based on the standard QM region but further includes residues His429 and His511, so covering the whole zinc binding site. RMSF (root mean square fluctuation) plots of the zinc site show that DFTB3 treatment is significantly better than MM treatment, but still cannot completely preserve the interaction between PEtN and the zinc ion **(Figure 4-5).** The predicted zinc geometry is similar to the representative geometry of 100 ps DFTB3/MM MD following the MM MD simulation (**Figure 4-4C**). In contrast to the results of MM MD simulations, no significant

difference was found between the two protonation states of His 511 using DFTB3/MM MD simulations starting from the crystal structure. However, the two sizes of QM region do not share a similar outcome. In DFTB3/MM MD simulations, the interaction between zinc and PEtN was more preserved in simulations with the standard QM region than in simulations with the extended QM region. When using the extended QM region, the substrate PEtN frequently lost its interaction with the zinc ion, with the dissociation time varying among simulations (**Table 4-2**). Longer dissociation times are observed in these simulations with HID511 (average: 750 ps) than with HIE511 (average: 320 ps) which may slightly favour the conclusion that HID511 is the preferred protonation state for His511 to stabilize the zinc-bound substrate. When PEtN is dissociating from coordinating zinc, solvent water molecules gradually approach the zinc ion and eventually occupy the free zinc-binding position(s). At the same time, the zinc coordinating Asp498 moves from monodentate to bidentate coordination. Together, these changes in zinc ligation result in 5- or 6-coordinated zinc binding geometries that are not consistent with the crystal structure.

Table 4-2. Interactions between PEtN and Zn^{2+} in simulations of the ICR^{Mc}:PEtN complex.

	The interaction between PEtN and Zn^{2+}	
	HIE511	**HID511**
Standard QM region		
Replicate-1	interacting	interacting
Replicate-2	interacting	lost at 1200ps
Replicate-3	interacting	interacting
Extended QM region		
Replicate-1	lost at 460ps	lost at 460ps
Replicate-2	lost at 150ps	lost at 1040ps
Replicate-3	lost at 350ps	interacting

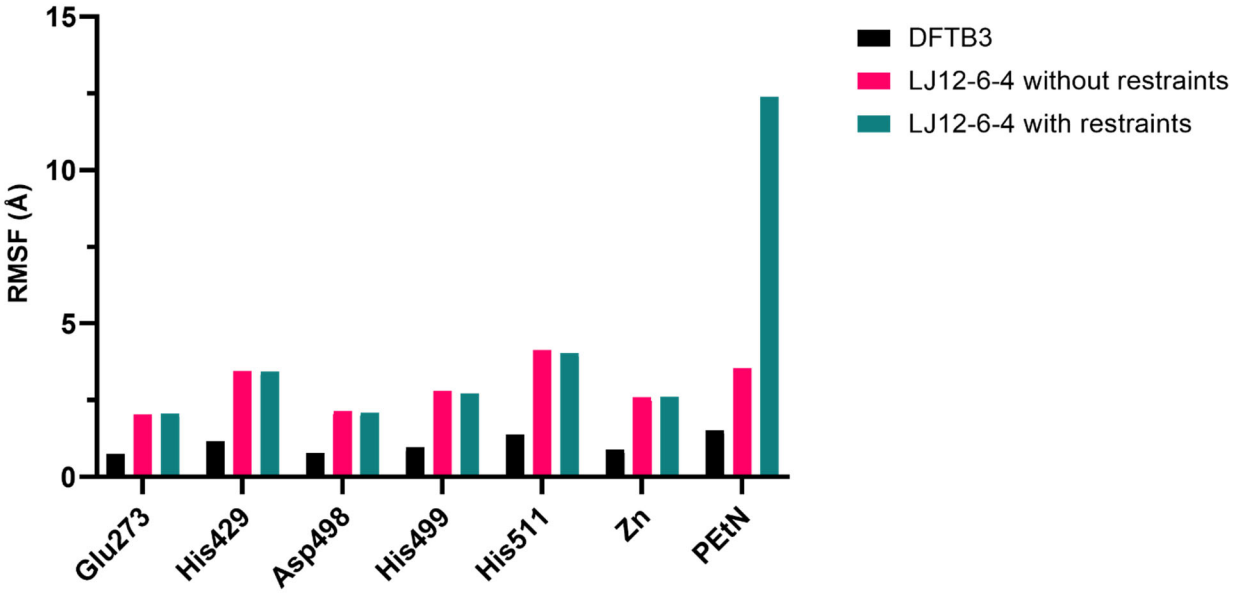


Figure 4-5. RMSF values (compared to the crystal structure) for the zinc ion, active site residues and PEtN in the zinc site of the ICR^{Mc}:PEtN complex (HID511) with MM and DFTB3 treatments.

4.3.2 Modelling the MCR-1:TGA complex

After modelling the substrate complex of MCR-1 like protein ICR^{Mc} using the pipeline, complexes of MCR-1 thioglycolic acid (TGA) were modelled. The crystal structure of di-zinc $MCR\text{-}1^{CD}$ (Resolution: 1.75 Å, PDB code: 5lrm[15]) was used for our study. In the crystal structure, the $Zn1^{2+}$ ion is tetrahedrally coordinated by four protein ligands (Glu246, Thr285, Asp465 and His466) and the $Zn2^{2+}$ is coordinated by His395 and His478 (**Figure 4-6B**). The two zinc ions, protein residues and small molecules coordinating the zinc are categorized as the zinc binding site. Due to the lack of consensus on the number of zinc ions in MCR-1, and the proposed mechanism of $MCR\text{-}1^{CD}$ substrate binding, both mono-zinc and di-zinc structures were modelled. Previous computational studies indicate that the mono-zinc form of $MCR\text{-}1^{CD}$ may be sufficient to transfer PEtN from the donor substrate to Thr285, while the second metal equivalent was necessary for transferring the phosphate group from phosphorylated Thr285 to membrane-bound lipid A.[142] The structure of mono-zinc $MCR\text{-}1^{CD}$ was obtained by removing the $Zn2^{2+}$ ion of the di-zinc $MCR\text{-}1^{CD}$ structure **(Figure 4-6A)**. Although His395 and His478 are not zinc coordinating residues in the mono-zinc structure, they are included in the zinc binding site as they play important roles in retaining MCR-1 activity.

Zinc is flexible with respect to the number of ligands that it can adopt in its first coordination shell. In aqueous solution, zinc is 6-coordinated, adopting an octahedral geometry. However, when binding to proteins, zinc is usually tetrahedrally coordinated, although zinc can be found to be 5-coordinated and, rarely in some catalytic sites, 6-coordinated.[243–246] Zinc in tetrahedral geometry is usually coordinated by 3 protein residues, while the fourth coordination position is often occupied by a water molecule when the protein is in apo form (e.g. as observed in Sfh-1, ACE and ACE2 (see above)). The coordination position of water is considered open and dynamic because the water molecule can exchange with the substrate or inhibitors when a protein-ligand complex is formed.[245,249] However, the $Zn1^{2+}$ ion of MCR-1 is tetrahedrally coordinated by four protein residues in the uncomplexed form, without leaving space for coordinating small molecules as are more usually found in tetrahedrally coordinated Zn^{2+} centres in proteins. The lack of an initial structure for the complex, the specificity of the zinc centre (i.e., 4 protein ligands coordinating zinc) and uncertainty surrounding the number of zinc ions in the active site make the modelling $MCR\text{-}1^{CD}$:TGA complex more challenging than the other systems investigated here.

4.3.2.1 Preparation of initial structures

The protonation state of TGA was assigned by reference to the TGA state in the previously described crystal structure 6JED[269] and the calculation by chemicalize web server (http://www.chemaxon.com). PLANTs[288] and AutoDock Vina[289] that have been validated in docking thiolates to metalloenzyme IMP-1[263], were used to dock TGA into the Zn^{2+} site of $MCR\text{-}1^{CD}$. For mono-zinc $MCR\text{-}1^{CD}$, the small molecule TGA was docked into the Zn^{2+} site of $MCR\text{-}1^{CD}$ by rigid docking with PLANTS[288] and AutoDock Vina[289] to generate initial structures

of the MCR-1^{CD}:TGA complex. One of the advantages of PLANTS is the easy application of distant restraints between the receptor and the ligand. Distance constraints between the zinc ions and the oxygen and sulfur atom with a weight of -7.5 kcal/mol were applied when using PLANTs, and the radius of the docking sphere was 10 Å centered at the Zn^{2+} ion. No distance restraints were used for AutoDock Vina, and the docking box dimension was 10 Å x 10 Å x 10 Å centering at the Zn^{2+} ion. The two docking programs yield a similar docking outcome for the mono-zinc MCR-1^{CD}:TGA complex. Although TGA could be docked into the zinc site of MCR-1^{CD}, the distance between TGA and the zinc ion was in both cases still over 3 Å (**Figure 4-6C**). MM minimizations with distance restraints in the modelling workflow subsequently resulted in TGA coordinating zinc. Eventually, TGA coordinates the Zn^{2+} ion as the fifth ligand via both thiol and carboxylate.

For the di-zinc MCR1^{CD}:TGA complex, crystal structures of IMP-1:TGA (Resolution: 1.57 Å, PDB code: 6JED[269]) and SMB-1:TGA (Resolution: 2.20 Å, PDB code: 3VQZ[290]) available in the PDB database were used as reference structures for the protonation state and binding pose of TGA. Autodock Vina was used for rigid docking without distance restraints and box dimension was 10 Å x 10 Å x 10 Å centring at the $Zn1^{2+}$ ion. In the docked di-zinc structure, TGA bridges the two Zn^{2+} ions *via* its sulphur atom, and coordinates the $Zn2^{2+}$ ion making the second zinc 4-coordinated as observed in the IMP-1:TGA complex (**Figure 4-6D**).

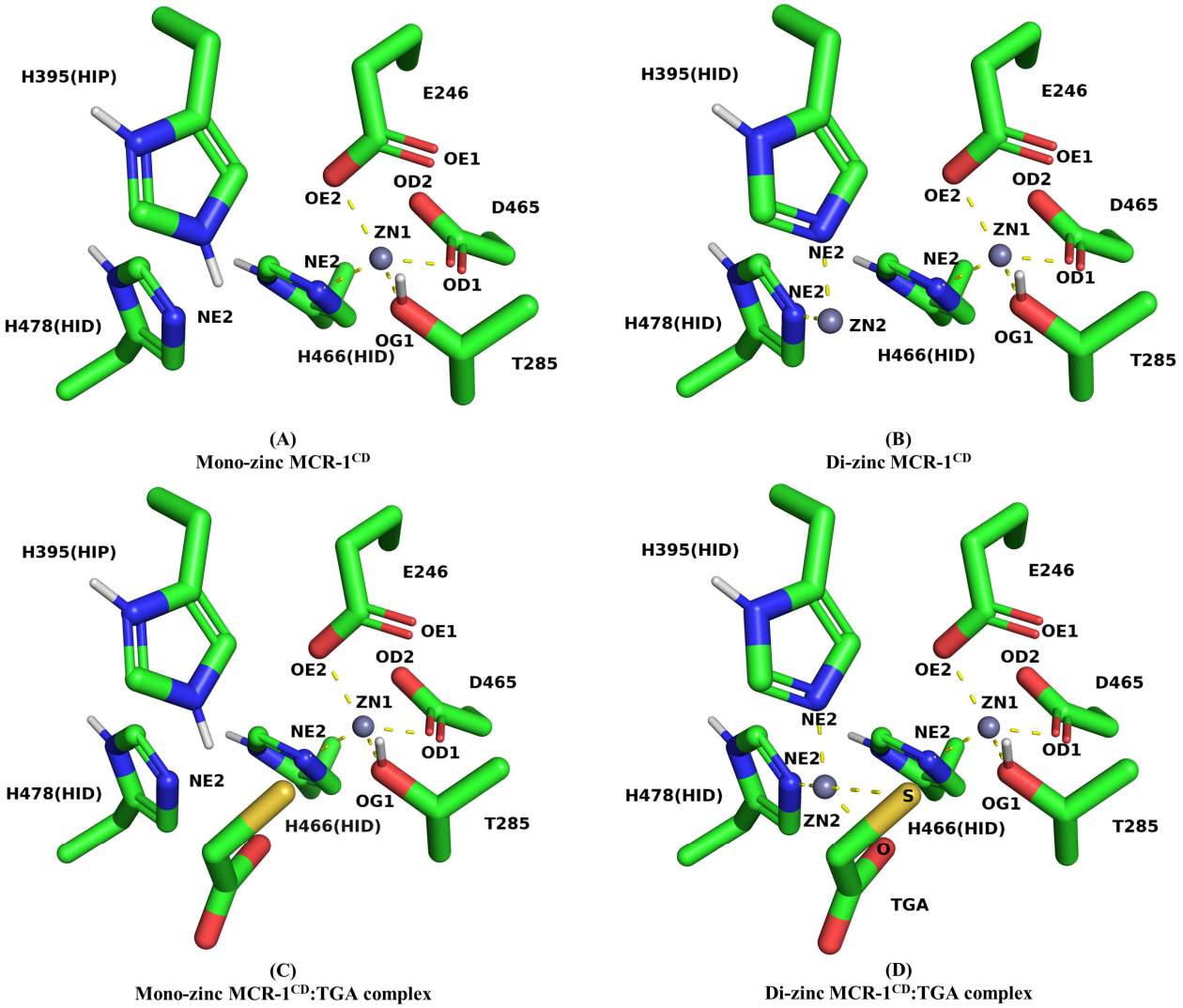


Figure 4-6. Zinc site geometries of MCR-1^{CD} and the MCR-1^{CD}:TGA complex initial structures generated by molecular docking.

MCR-1 catalytic domain docked with thiolate inhibitor TGA. Carbon atoms are in green, nitrogen atoms are in blue, oxygen atoms are in red and sulfur atoms are in yellow, with zinc ions represented as gray balls. Dashed lines indicate zinc coordinating ligands.

4.3.2.2 Implementing distance restraints

Distance restraints were used to maintain the zinc geometry of the MCR-1^{CD}:TGA complex during all steps of the MM simulations including system minimization, MD equilibration and MD production. The distance restraints used in the study are shown in **Table 4-3**. Because the initial structures of the MCR-1^{CD}:TGA complex were generated by molecular docking, the restraint parameters were carefully tuned to allow the zinc geometry to be flexible in the MM MD simulation. As Thr285 is observed to be modified by addition of phosphoethanolamine during the transfer reaction (i.e., phosphorylated), and may then lose its interaction with the zinc ion, the distance restraint between Thr285 and the zinc ion was removed in the production MM MD

simulations to allow free movement of Thr285. Any distance restraint that involves the TGA inhibitor was also removed during MD production runs.

Table 4-3. Distance restraints for MCR-1^{CD}:TGA complexes.

	Atom-1	Atom-2	Lower boundary(Å)	Upper boundary(Å)
Mono-zinc MCR-1^{CD}:TGA complex	ZN1	TGA-S	2.15	2.50
	ZN1	TGA-O	1.90	2.40
	ZN1	E246-OE2	1.90	2.30
	ZN1	D465-OD1	1.90	2.30
	ZN1	H466-NE2	1.90	2.30
	ZN1	T285-OG1	1.90	2.30
	Atom-1	Atom-2	Lower boundry	Upper boundry
Di-zinc MCR-1^{CD}:TGA complex	ZN1	TGA-S	2.15	2.50
	ZN1	E246-OE2	1.90	2.30
	ZN1	D465-OD1	1.90	2.30
	ZN1	H466-NE2	1.90	2.30
	ZN1	T285-OG1	1.90	2.30
	ZN2	TGA-S	2.15	2.50
	ZN2	E246-OE2	4.00	6.00
	ZN2	E246-OE1	4.00	6.00
	ZN2	H395-NE2	1.90	2.30
	ZN2	H478-NE2	1.90	2.30
	ZN2	ZN1	3.75	4.50

4.3.2.3 Mono-zinc MCR-1^{CD}:TGA complex

4.3.2.3.1 Similar predictions were made by MM MD simulations with/without restraints

The mono-zinc MCR-1^{CD}:TGA complex was first modelled. The protonation state of MCR-1^{CD} is assigned by the program PROPKA 3.1 with His395 manually set to be positive (i.e., HIP395) according to previous studies of MCR-1 substrate binding[142,273]. MM MD simulations were first performed followed by DFTB3/MM MD and B3LYP-D3BJ/MM geometry optimization. Unrestrained LJ12-6-4 and restrained LJ12-6-4 MM MD simulations were used to examine the conformational change and the protonation state of the complex system. Unrestrained simulations were used to examine the potential movement of zinc binding residues (e.g. the loss of the Thr285 - Zn^{2+} interaction), while restrained simulations were used to investigate the dynamics of the zinc site while retaining the crystallographically geometry as much as possible. The distance restraints used are listed in **Table 4-2.** MM MD simulations were run with 5, instead of 3, replicates due to the challenges of the MCR-1^{CD}:TGA complex and the length of each simulation is 100ns. The restrained LJ12-6-4 MM MD simulations showed similar results to the unrestrained simulations. For both the restrained and unrestrained LJ12-6-4 simulations, the RMSD values of the zinc binding site were consistently lower than 1.3 Å among all 5 replicates after 20 ns of simulations, suggesting that stable zinc site geometry is achieved across the

remaining 80 ns of simulations. The RMSD values of restrained LJ12-6-4 simulations and unrestrained LJ12-6-4 simulations were similar. The RMSD values, compared to the initial structure, of the inhibitor TGA fluctuate between 0.1 Å and 0.5 Å in both restrained and unrestrained models. The RMSD values of the protein backbone atoms from both models fluctuated and converged at 0.8-1.8 Å. These suggest that the predicted geometries of the zinc site of the MCR-1^{CD}:TGA complex are similar with or without restraints. MM MD simulations showed that the Zn^{2+} ion is 6-coordinated by Glu246, Asp465, His466, TGA (O and S atoms respectively) and a water molecule. The interaction between Thr285 and the Zn^{2+} ion was lost during MM MD simulations. RMSD plots (compared to the initial structures) of the zinc site are presented in **Figure 4-7**.

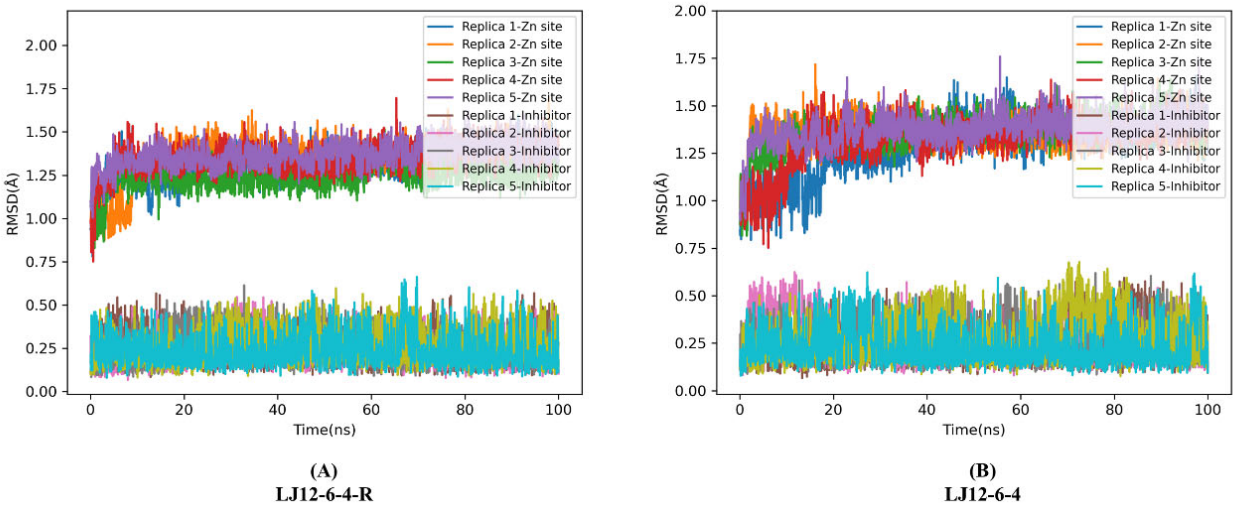


Figure 4-7. Time-dependence of RMSD values (compared to the initial structure) for MM MD simulations of mono-zinc MCR-1^{CD}:TGA complex using LJ 12-6-4 non-bonded models.

Simulations were performed in five replicates for each model and each replicate is 100 ns. The 'Zn site' refers to zinc ions and zinc coordinating residues and 'Inhibitor' refers to TGA.

4.3.2.3.2 QM/MM refinement of the MM MD predicted structure

Subsequent DFTB3/MM MD simulations were carried out starting from the last frame of the 100ns MM MD simulation (a representative snapshot as the RSMD has fully converged by this time) and were followed by two 250-step B3LYP-D3BJ based QM/MM geometry optimization runs using the 6-31G(d) and the 6-31+G(d) basis sets, respectively. Because the prediction results of restrained LJ12-6-4 and unrestrained LJ12-6-4 models are almost identical, using the last frame of unrestrained or restrained LJ12-6-4 MM MD simulations for DFTB3/MM MD should yield similar outcomes. The QM region includes the zinc ion, Glu246, Thr285, Asp465, His466, His395, His478, the inhibitor TGA, and the zinc coordinating water molecule added by the MM MD simulation. After 100ps of DFTB3/MM MD simulation, the added zinc-coordinating water molecule, as well as the interaction between the S atom of TGA and Zn^{2+}, were removed, turning the zinc binding geometry to tetrahedral. Refinement using B3LYP-D3BJ QM/MM geometry optimization further optimized zinc coordination distances. The result showed that, when TGA interacted with the zinc ion, Thr285 rotated

and lost its interaction with the zinc. Representative geometries for the MCR-1^{CD}:TGA complex zinc centres at different modelling conditions are presented in **Figure 4-8**. Zn coordination geometries from the various treatments are reported in **Table 4-4.** Alignment of the zinc binding site of MCR-1^{CD}:TGA complex to the initial structures is shown in **Figure 4-15B.**

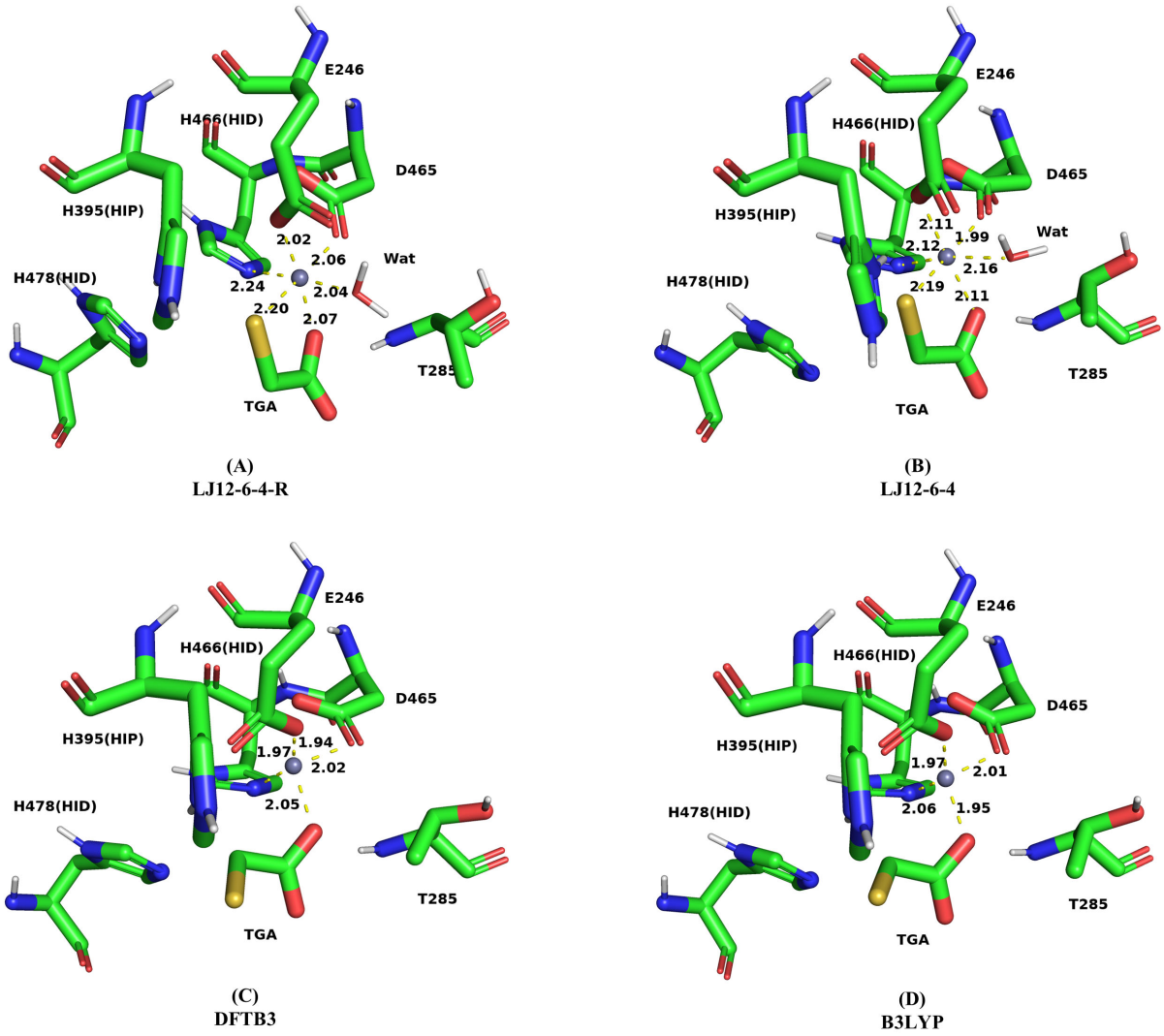


Figure 4-8. Zinc site geometries of mono-zinc MCR-1^{CD}:TGA complexes.

(A-B) Representative zinc geometries after MM MD simulations using LJ12-6-4 non-bonded models with (A) and without (B) restraints, respectively. (C) The representative zinc geometry after 100 ps DFTB3/MM MD simulation. (D) The representative zinc geometry after B3LYP-D3BJ based QM/MM geometry optimization.

Table 4-4. Atomic distances (Å) between the zinc ion and the zinc coordinating atoms for simulations of the mono-zinc MCR-1^{CD}:TGA complex.

	Zinc coordinating atom	Action 1 Action 2 Action 3 Action 4	LJ1264[a]	LJ1264-R[a]	LJ1264 DFTB3	LJ1264 DFTB3 6-31G*	LJ1264 DFTB3 6-31+G*	Initial
Coordination Distance (Å)	Glu246-OE2		2.09	2.07	1.94	1.96	1.97	2.04
	Thr285-OG1		5.39	5.22	5.56	5.48	5.49	2.10
	Asp465-OD1		2.04	2.04	2.02	1.99	2.01	1.99
	His466-NE2		2.23	2.25	1.97	2.05	2.06	2.10
	TGA-S		2.20	2.21	5.01	4.73	4.70	4.79
	TGA-O		2.09	2.07	2.05	1.93	1.95	3.91
CN			6.0	6.0	4.0	4.0	4.0	4.0
Deviation	Glu246-OE2		0.05	0.03	-0.10	-0.08	-0.07	
	Thr285-OG1		3.29	3.12	3.46	3.38	3.39	
	Asp465-OD1		0.05	0.05	0.03	0.00	0.02	
	His466-NE2		0.13	0.15	-0.13	-0.05	-0.04	
	TGA-S		-2.59	-2.58	0.22	-0.06	-0.09	
	TGA-O		-1.82	-1.84	-1.86	-1.98	-1.96	
TAD			3.52	3.35	3.72	3.51	3.52	

Deviation (Å) = simulation predicted value - initial value; CN = Zn Coordination Number; TAD (Å) = Total Absolute Mean Deviation. DFTB3 refers to 100ps DFTB3/MM MD; 6-31G* and 6-31+G* refer to the basis set used in BL3YP-D3BJ based QM/MM geometry optimization.

Note: Additional Zn^{2+} coordinating atoms introduced during simulations that are not presented in the crystal structure are not reported in the table, but are included when calculating the mean coordination number.

[a] The values presented are averages of the last frames of five independent MM MD simulations.

4.3.2.3.3 Direct DFTB3/MM MD simulations

DFTB3/MM MD simulations directly starting from three different initial structures were carried out. The detail of 3 initial structures are shown in **Figure 4-9**. The simulations were run in triplicate with 2ns MD simulation for each replicate. The QM region includes the zinc ion, Glu246, Thr285, Asp465, His466, His395, His478 and the inhibitor TGA.

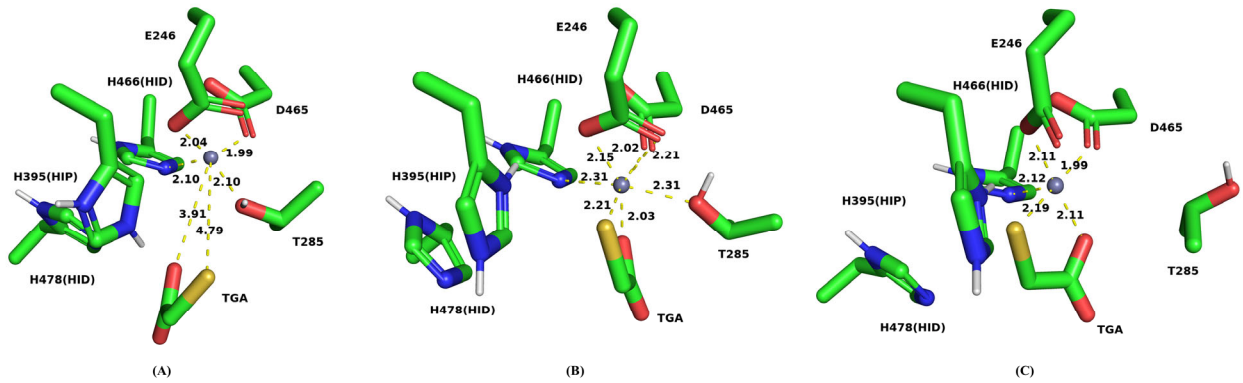


Figure 4-9. Initial zinc geometries of the MCR-1^{CD}:TGA complex for DFTB3 QM/MM treatment.

(A) MCR-1 catalytic domain docked with thiolate inhibitor TGA. (B) Snapshot of MCR-1 catalytic domain at MM equilibration with positional constraints. (C) Re-solvated MCR-1 catalytic domain at 100ns of MM MD production run, water molecules added by MM MD simulations were removed. Carbon atoms are in green, nitrogen atoms are in blue, oxygen atoms are in red and sulfur atoms are in yellow, with zinc ions represented as gray balls.

The inhibitor TGA did not stay in contact with the zinc ion in all 9 simulations using the three initial structures. Being more than 3.9 Å away from the zinc ion, TGA in the initial structure (A) never bound to the zinc ion in the 2ns DFTB3 QM/MM MD simulation. In simulations starting from the initial structures (B) and (C), TGA quickly lost its interaction with the zinc ion at the early stage of the simulation run. We mutated Thr285 to Ala285 *in silico* of the initial structure (B), then ran 2 ns DFTB3 QM/MM MD simulations in triplicate with the T285A mutant. In one of these three simulations, TGA remained bound to the zinc via its carboxylate oxygen for the full 2 ns duration, with a geometry similar to that shown in **Figure 4-8C**. However, in all three simulations, TGA bound to the zinc ion (via the carboxylate group) for more than the initial 100 ps, an outcome different to that obtained in simulations with T285.

Simulations above were carried out using His395 in positively charged state (i.e., HIP395). To investigate the effect of the protonation state of His395 on inhibitor binding, 200ps simulations were carried out in triplicate for the T285A mutant (i.e., with Ala285 present) with His 395 in the HID form (HID395) and initial structure (B) (i.e., with Thr285 present) respectively. The interaction between TGA and the zinc ion was still lost in simulations with the initial structure (B). Simulations using the T285A mutant as the initial structure with His 395 in the HID395 form gave similar results to those obtained in the simulations above (with His 395 in the HIP395 form). These results showed that in the presence of Thr285, the inhibitor TGA is not stably retained in the zinc binding site regardless of the protonation state of His395. However, the inhibitor may still leave the zinc binding site in simulations using the T285A mutant. In addition, the TGA thiol was found to dissociate from the zinc ion in the first 30 ps of almost all of these DFTB3/MM MD simulations (**Figure 4-10**). Therefore, all remaining contacts between the inhibitor TGA and the zinc ion were Zn-O, rather than Zn-S, interactions, which was interesting as it was expected that the thiol was likely to make the major contribution to zinc binding. However, this result may also reflect the limitations of the DFTB3 treatment, as described above (section 4.2).

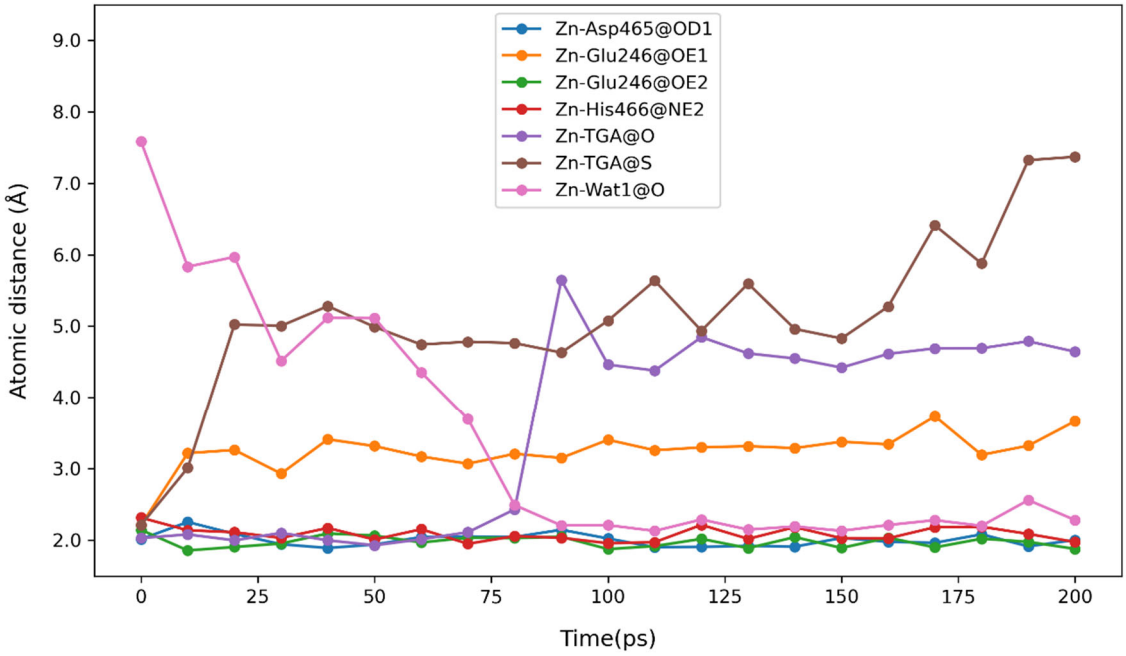


Figure 4-10. Zinc site coordination distances against simulation time for the beginning 200 ps of a 2 ns DFTB3/MM MD simulation of the mono-zinc MCR-1^{CD}:TGA complex.

All Zn^{2+} ligation distances are recorded, including water coordination (ZN-Wat1@O) added by later in the simulation. The interaction between the sulfur atom of TGA and Zn^{2+} is rapidly lost within10 ps. Data for each interaction was collected every picosecond, with dots indicating data collection points.

4.3.2.4 Di-zinc MCR-1^{CD}:TGA complex

The di-zinc form of the MCR-1^{CD} is proposed to play an important role in the second step of the catalytic reaction of MCR-1 (i.e., transferring the phosphate from phosphorylated Thr285 to lipid A). Although the second zinc site may be compatible with various divalent metal ions[142], a Zn^{2+} ion is more biologically possible. Referring to the binding pose of TGA complexed with the binuclear metallo-β-lactamases IMP-1 and SMB-1, TGA was expected to sit with its sulfur atom bridging the two zinc ions in the di-zinc MCR-1^{CD} models. Therefore, the distance restraints were set accordingly (**Table 4-3**). Unlike in the mono-zinc MCR-1^{CD} model, His395 is neutral in the di-zinc structure. The protonation states of side chains of protein residues in the di-zinc MCR-1^{CD} were predicted automatically via program PROPKA 3.1. MM MD simulations were run in 5 replicates and the length of each simulation is 100ns.

4.3.2.4.1 MM MD simulations with/without restraints

RMSD (compared to the initial structures) plots and representative geometries for the MCR-1^{CD} zinc centres at different stages of the modelling workflow are presented in **Figures 4-11** and **4-12**. The RMSD values of the inhibitor TGA are consistently small (0.1-0.6 Å) among simulations regardless of the use of the unrestrained or restrained LJ12-6-4 model. This suggested that the initial model allowed TGA to stably interact with the zinc binding site in all 10 simulations. Visual inspection showed that TGA remained bridging the two zinc ions by its sulfur atom. The RMSD of the zinc binding site fluctuated from 0.5 to 1.5 Å in simulations using the restrained LJ12-6-4 model and from 0.5 to 1.8 Å in simulations using the unrestrained LJ12-6-4 model. Although the RMSD of the zinc binding site of the di-zinc MCR-1^{CD}:TGA complex varies between simulation runs, the predicted geometries of the zinc binding site are consistent. However, we found that the Zn^{2+} ion modelled with the unrestrained LJ12-6-4 model may occasionally have 7 coordinating ligands (**Figure 4-12B**), which is unrealistic.

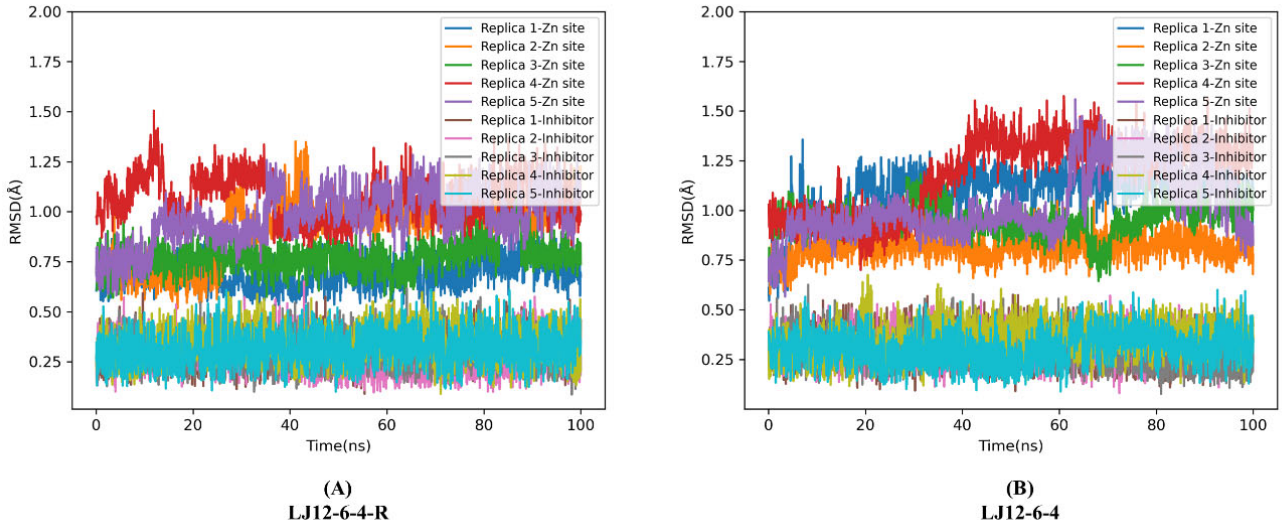


Figure 4-11. Time-dependence of RMSD values (compared to the initial structure) for MM MD simulations of the di-zinc MCR-1^{CD}:TGA complex using LJ 12-6-4 non-bonded models.

Simulations were performed in five replicates for each model with each replicate 100 ns. The 'Zn site' refers to zinc ions and zinc coordinating residues and 'Inhibitor' refers to TGA.

4.3.2.4.2 QM/MM refinement of the predicted structure

Subsequent DFTB3/MM MD simulations were carried out starting from the last frame of the restrained 100ns MM MD simulation, and were followed by two 250-step B3LYP-D3BJ based QM/MM geometry optimizations using the 6-31G(d) and the 6-31+G(d) basis set respectively. Because several water molecules (added by the MM MD simulation) are involved in the zinc binding site, and due to the known limitations of DFTB3 treatment on Zn-S interactions, two QM regions were used. The only difference between the two QM regions is the inclusion of zinc coordinating water molecules shown in **Figure 4-12A**. Both QM regions include the zinc ions, Glu246, Thr285, Asp465, His466, His395, His478 and the inhibitor TGA. The zinc-

coordinating water molecules were included in the water-included QM region and excluded in the water-excluded QM region. 100 ps DFTB3/MM MD simulations were performed in triplicate for each QM region.

Simulations with the two QM regions shared similar results for the zinc geometries. Interactions between the thiol of TGA and the $Zn1^{2+}$ and $Zn2^{2+}$ ions were lost or close to being lost (i.e., atomic distance of the Zn-S interaction is approaching or beyond 2.5 Å). The $Zn1^{2+}$ ion was often 4-coordinated by Glu246, Asp465, His466 and Thr285 or a water molecule. The $Zn2^{2+}$ ion was usually 5-coordinated by His395, His478, the TGA carboxyl and two water molecules. However, the distance between the two zinc ions in simulations with the water-excluded QM region (mean: 4.61 Å) was smaller than in simulations with the water-included QM region (mean: 5.51 Å). The shorter atomic distance between two zinc ions is also closer to the distance values in the reference structures (IMP-1: 3.55 Å, SMB-1: 3.60 Å). Moreover, in one of the three replicates using the water-excluded QM region the TGA thiol bridged the two zinc ions. This bridging geometry of TGA binding might be considered more realistic than other geometries because it is more consistent with the two reference structures (IMP-1 and SMB-1).

Refinement using B3LYP-D3BJ QM/MM geometry optimization was used to further optimize zinc coordination distances. Refinements started from the typical geometry obtained after DFTB3/MM MD treatment, as well as from the geometry described above where the TGA thiol bridged the two zinc ions. The predicted zinc binding geometry of the di-zinc MCR-1^{CD}:TGA complex is present in **Figure 4-12**. Zn coordination geometries from the various treatments are reported in **Table 4-5.** Alignment of the zinc binding site of the di-zinc MCR-1^{CD}:TGA complex to the initial structures is shown in **Figure 4-15C.**

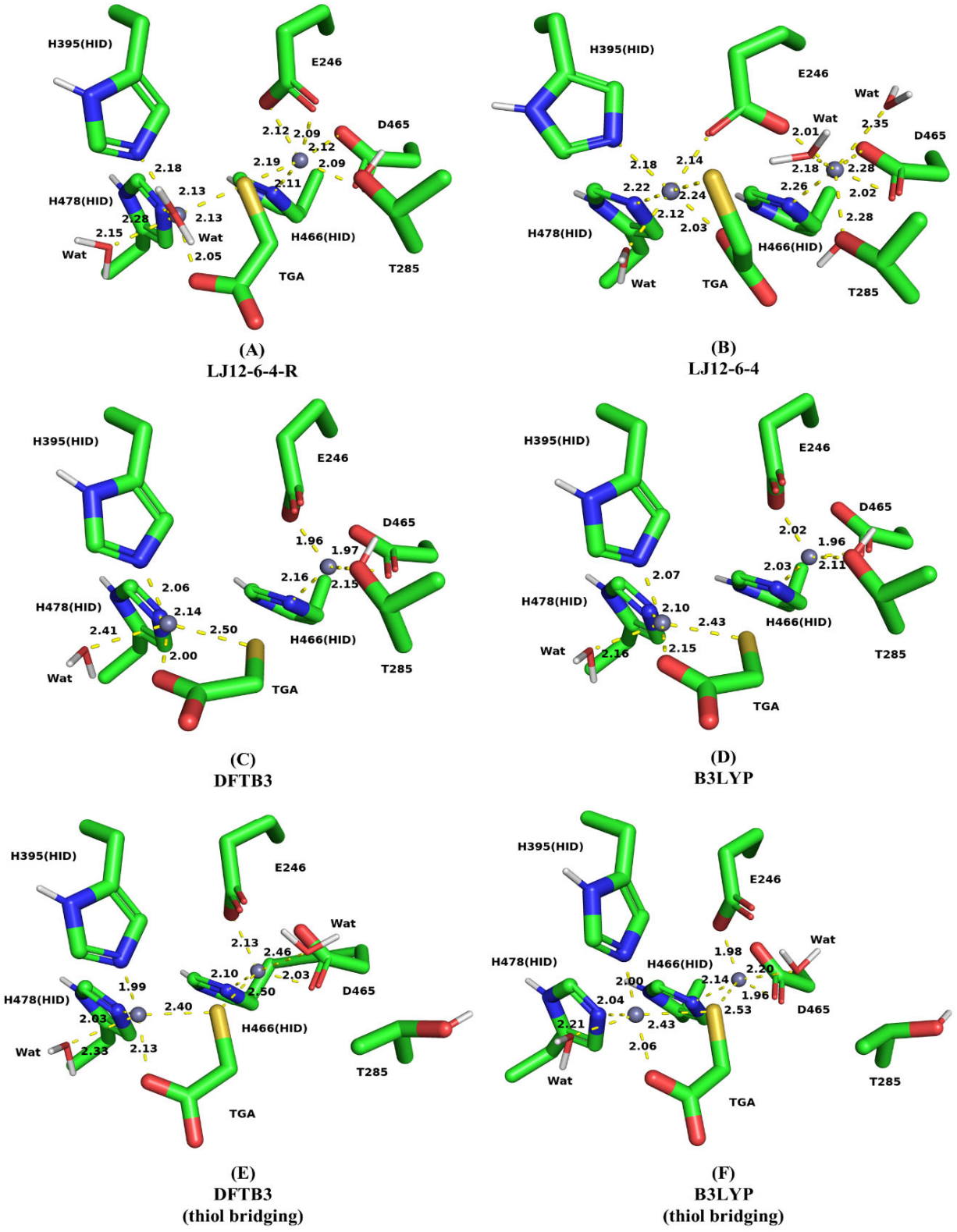


Figure 4-12. Zinc site geometries of di-zinc MCR-1^{CD}:TGA complexes.

(A-B) Representative zinc geometries after MM MD simulations using LJ12-6-4 non-bonded models with (A) and without (B) restraints, respectively. (C) The representative zinc geometry after 100 ps DFTB3/MM MD simulation. (D) The representative zinc geometry after B3LYP-D3BJ based QM/MM geometry optimization. (E-F) The zinc geometry that TGA was bridging the two zinc ions via its thiol after DFTB3/MM MD simulation (E) and B3LYP-D3BJ based QM/MM geometry optimization (F).

Table 4-5. Atomic distances (Å) between the zinc ion and the zinc coordinating atoms for simulations of the di-zinc MCR-1^{CD}:TGA complex.

			Action 1						
			Action 2			LJ1264-R	LJ1264-R DFTB3		
			Action 3		LJ1264-R	DFTB3	6-31G*		
	Atom-1	Atom-2	Action 4	LJ1264[a]	LJ1264-R[a]	DFTB3	6-31G*	6-31+G*	Initial
Mean Atomic Distance (Å)	Glu246-OE2	Zn1		2.08	2.08	1.96	1.97	2.02	2.04
	Thr285-OG1	Zn1		5.05	4.55	2.15	2.07	2.11	2.10
	Asp465-OD1	Zn1		2.05	2.05	1.97	1.95	1.96	1.99
	His466-NE2	Zn1		2.23	2.23	2.16	2.00	2.03	2.10
	TGA-S	Zn1		2.21	2.20	3.61	3.54	3.50	4.79
	HIS395-NE2	Zn2		2.22	2.16	2.06	2.04	2.07	1.87
	HIS478-NE2	Zn2		2.23	2.17	2.14	2.06	2.10	2.04
	TGA-S	Zn2		2.18	2.23	2.50	2.49	2.43	2.40
	TGA-O	Zn2		2.02	2.06	2.00	2.05	2.15	1.92
	Zn1	Zn2		4.30	4.30	5.16	5.15	5.03	4.50
Mean CN	Zn1			6.0	6.0	4.0	4.0	4.0	4.0
	Zn2			6.0	6.0	5.0	5.0	5.0	4.0
Mean Deviation	Glu246-OE2	Zn1		0.04	0.04	-0.08	-0.07		
	Thr285-OG1	Zn1		2.95	2.45	0.05	-0.03		
	Asp465-OD1	Zn1		0.06	0.06	-0.02	-0.04		
	His466-NE2	Zn1		0.13	0.13	0.06	-0.10		
	TGA-S	Zn1		-2.58	-2.59	-1.18	-1.25		
	HIS395-NE2	Zn2		0.35	0.29	0.19	0.17		
	HIS478-NE2	Zn2		0.19	0.13	0.10	0.02		
	TGA-S	Zn2		-0.22	-0.17	0.10	0.09		
	TGA-O	Zn2		0.10	0.14	0.08	0.13		
	Zn1	Zn2		-0.20	-0.20	0.66	0.65		
Mean Deviation for Zn1 Coordination				3.18	2.69	0.21	0.24		
Mean Deviation for Zn2 Coordination				1.04	0.94	1.13	1.06		
TAD				4.22	3.62	1.34	1.30		

Deviation (Å) = simulation predicted value - initial value; CN = Zn Coordination Number; TAD (Å) = Total Absolute Mean Deviation. DFTB3 refers to 100ps DFTB3/MM MD; 6-31G* and 6-31+G* refer to the basis set used in BL3YP-D3BJ based QM/MM geometry optimization.

Note: Additional Zn^{2+} coordinating atoms introduced during simulations that are not presented in the crystal structure are not reported in the table, but are included when calculating the mean coordination number.

[a] The values presented are averages of the last frames of five independent MM MD simulations.

4.3.2.4.3 Direct DFTB3/MM MD simulations

Inspired by the above 100 ps DFTB3/MM MD simulations using the water-excluded QM region, direct DFTB3/MM MD simulations without preceding MM MD simulations were performed from three different

initial structures. The first structure is that obtained after docking, which is the same structure as was used in previous MM MD simulations. The second initial structure is the re-solvated structure from a typical snapshot taken during MM MD simulations, where TGA is bound with the thiol bridging the two zinc ions. Water molecules added by MM MD simulations were removed. The main difference between the first and second initial structures is the pose of TGA. The last initial structure is a T285A mutant of the second initial structure. The 3 initial structures are shown in **Figure 3-13**. In each case the QM region used includes the zinc ions, Glu246, Thr285/Ala285, Asp465, His466, His395, His478 and the inhibitor TGA. The condition of whether to allow bond-breaking involving hydrogens in the QM region was also tested (QMshake=0 or QMshake=1) for each initial structure. The DFTB3/MM MD simulations were run in triplicate with 2ns MD simulation for each replicate. Details of the simulations are presented in **Table 3-6**.

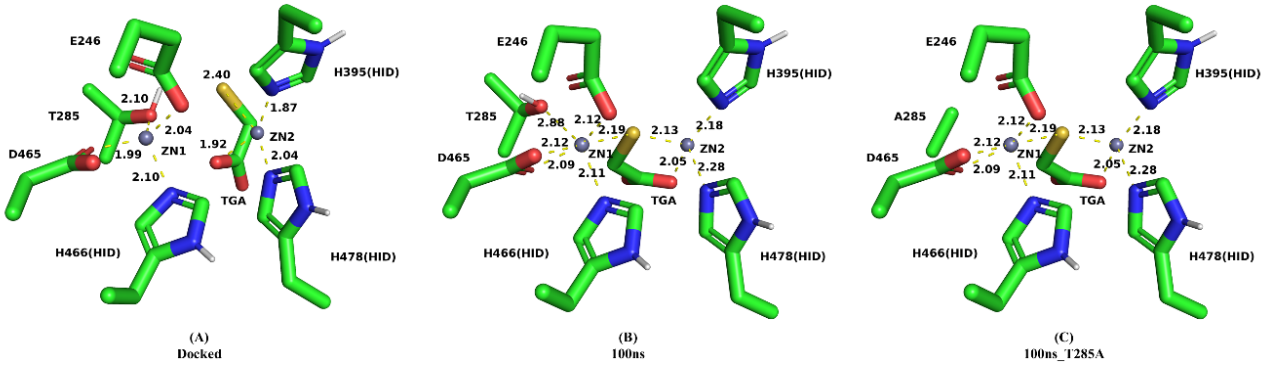


Figure 4-13. Zinc site geometries of the di-zinc MCR-1^{CD}:TGA complex initial structures.

MCR-1 catalytic domain docked with thiolate inhibitor TGA. Carbon atoms are in green, nitrogen atoms are in blue, oxygen atoms are in red and sulfur atoms are in yellow, with zinc ions represented as gray balls.

Simulations beginning from the three initial structures share similar results, with slight differences in the coordination number of the $Zn1^{2+}$ ion. When bond-breaking involving hydrogens was forbidden (QMshake=1), in simulations starting from the initial structure (B), the $Zn1^{2+}$ ion was more likely to be 4-coordinated, while for simulations using the initial structures (A) and (C), the $Zn1^{2+}$ ion was often 5 or 6-coordinated. Allowing bond-breaking involving hydrogens in the QM region (i.e., QMshake = 0) also yield similar results for each initial structure, but reduced the coordination number of the $Zn1^{2+}$ ion with initial structures (A) and (C). Simulations on the T285A mutant showed that, in the absence of Thr285, a water molecule, rather than the thiolate of TGA, entered the active site to coordinate the $Zn1^{2+}$ ion. The $Zn2^{2+}$ ion is usually 5 or 6-coordinated by His395, His478, TGA and water molecules among all simulations. The $Zn2^{2+}$ ion interacted with the inhibitor TGA via the carboxylate, rather than the thiol, group. TGA did not bridge the two zinc ions in any simulation and only interacted with the $Zn2^{2+}$ ion. Instead, a water molecule bridged the two zinc ions in 2 out of 18 simulations. This low frequency of obtaining the bridging water molecule is likely to be related to the distance between the two zinc ions. The distance between the two zinc ions is 4.5 Å in the crystal structure but usually increased to 5.5-6.0 Å in DFTB3/MM MD simulations. Due to this increase in the distance between the two zinc ions, one zinc ion usually has its coordination shell separated from the other, thus reducing the

opportunities for a single water molecule to bridge the two zinc ions. Representative zinc geometries with the residue Thr285 and Ala285 after the 2ns DFTB3/MM MD simulations are presented in **Figure 4-14**.

Table 4-6. Zinc geometries of the di-zinc MCR-1^{CD}:TGA complex predicted by direct DFTB3/MM MD simulations.

	Initial structure	CN of Zn1	Number of water molecules	CN of Zn2	Number of water molecules
Qmshake=1	A	5.3	1.0	5.3	2.3
	B	4.3	0.7	5.7	2.7
	C	5.7	2.7	5.7	2.7
Qmshake=0	A	4.7	0.7	5.7	2.7
	B	4.3	0.3	5.7	2.7
	C	4.3	1.0	5.3	2.3

The values presented are averages of the last frames of three independent DFTB3/MM MD simulations.

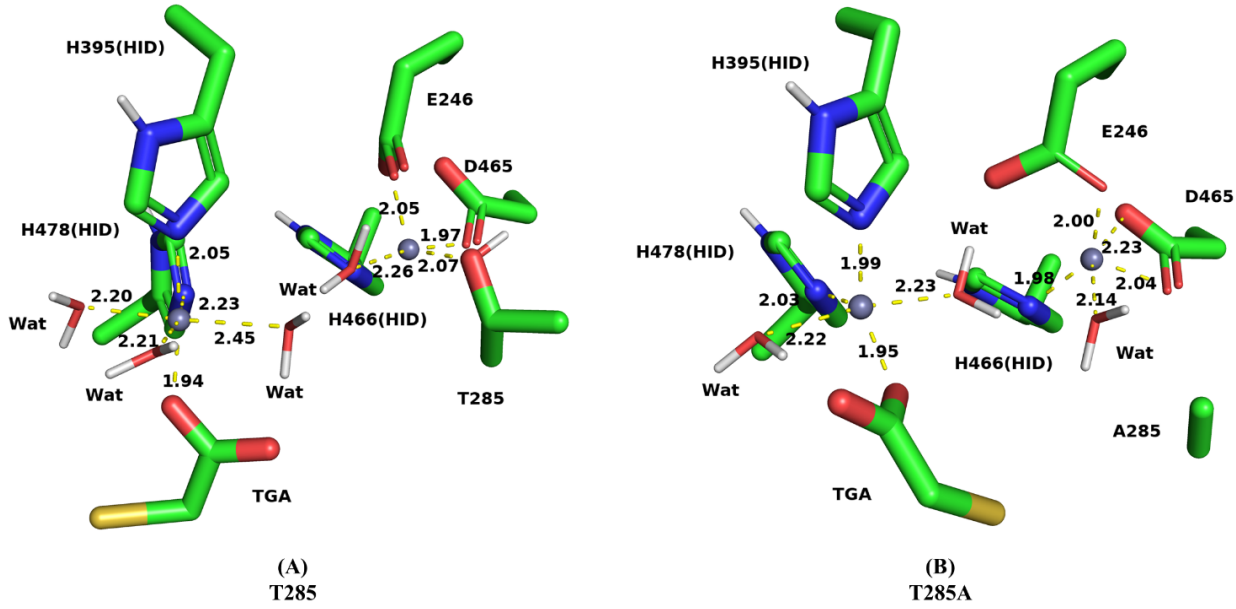


Figure 4-14. Outcome of DFTB3/MM MD run directly from initial structures.

Representative zinc site geometries of each protein-ligand complex after 2 ns DFTB3/MM MD simulation. Carbon atoms are in green, nitrogen atoms are in blue, oxygen atoms are in red and sulfur atoms are in dark yellow, zinc ions are represented as grey balls.

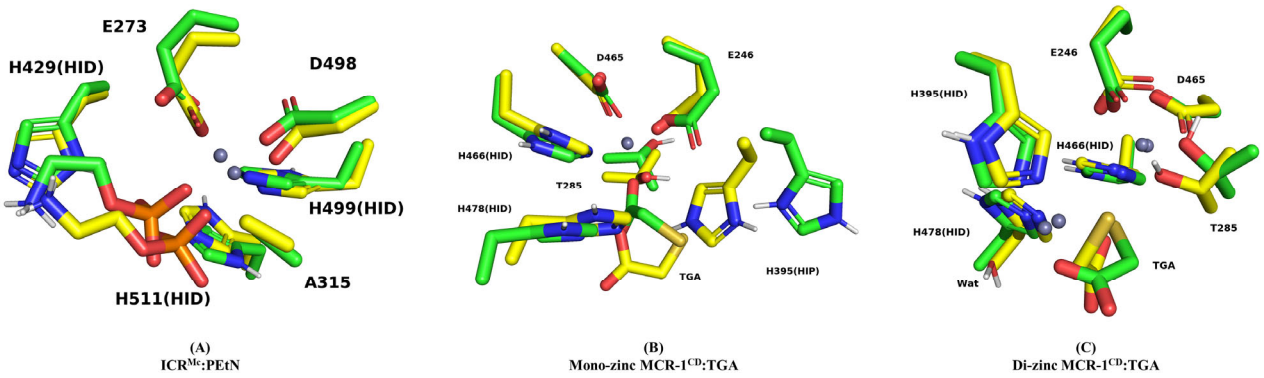


Figure 4-15. Zinc binding sites of the 3 MCR-like protein-ligand systems included in the study after simulations compared to the initial structures.

Carbon atoms are in green/yellow/purple depending on the model: yellow represents the crystal structure; green represents a representative structure after DFTB3/MM MD and B3LYP-D3BJ based QM/MM geometry optimization. Nitrogen atoms are in blue, oxygen atoms are in red and sulfur atoms are in dark yellow, zinc ions are represented as gray balls.

4.3.3 Discussion

Based on the methodology established in section 4.2, we are interested to apply the computational protocol to MCR-1 like PEtN transferases. Due to the lack of experimental complex structures of MCR-1 and our interests in searching for inhibitors of MCR-1, we sought to investigate the behaviour of a weak inhibitor of MCR-1^{CD} identified in Chapter 2, thioglycolic acid (TGA), to MCR-1^{CD}. Before simulating the MCR-1^{CD}:TGA complex, we first applied the protocol to the MCR-1 like the ICR^{Mc}:PEtN complex where a crystal structure is available. The modelling of the substrate complex of ICR^{Mc} catalytic domain suggests that the substrate PEtN cannot stably bind to the zinc ion during MD simulations (**Figure 4-4**). Although the standard 3OB parameter set used for DFTB3 treatment has some known disadvantages when used with phosphorus-containing systems[291], this should have a minor impact on the simulations as they do not involve a change in the coordination number of the phosphorus. For investigation of structural properties involving phosphoryl transfer reactions, the specific 3OB/OPhyd parameter set is available[291]. In the crystal structure, the PEtN molecule mainly interacts with the zinc ion, His511 and Tyr338. His511 is not bound to zinc, but its equivalent in MCR is important (MIC drops when it is mutated)[15]. The protonation state and role of this Histidine remain uncertain. The simulations here show that the protonation state of the non-zinc bound residue His511 heavily impacts the binding of the substrate PEtN to the zinc ion of the catalytic domain of ICR^{Mc}. These simulations, in which both protonation states of His511 were tested, indicate that His511 is likely to be protonated on the delta nitrogen (i.e., HID511).

Another important non-zinc bound residue in the zinc binding site is His429, which is predicted to be in the HID429 state according to PROPKA 3.1. Tuning the protonation state of His429 might help to stabilize substrate binding to some extent. Stogios *et al* also claimed that Tyr338 from an adjacent subunit plays an important role for positioning PEtN, and showed that the full length Y338A mutant ICR^{Mc} loses colistin

resistance in the experiment. They suggested that dimerization of the catalytic domain may be required to facilitate the coordination of PEtN. Since the crystal structure of the dimeric catalytic domain of the protein is not available, only the monomeric ICR^{Mc} catalytic domain was modelled in the MD simulations. No interactions between Tyr 338 and PEtN were observed in the simulations. Simulations using dimeric forms of the ICR^{Mc} catalytic domain may facilitate the binding of PEtN. Another possible explanation for the instability of PEtN binding is that the protonation state of PEtN is not appropriate. In particular, the amine tail of PEtN is usually facing in a different orientation after MD simulations, compared to the crystal structure. The amine tail might be neutral rather than being positively protonated in the ICR^{Mc} microenvironment (PEtN was identified as having two pKa values at 5.61 and 10.39[292]). The neutral ammonia tail reduces the overall charge of the substrate molecule which may promote PEtN binding to Zn^{2+}.

Lastly, the binding of PEtN to the ICR^{Mc} catalytic domain might be a weak interaction which only lasts for a short time. Simulations using the DFTB3/MM MD treatment (giving higher accuracy than the MM MD treatment) show that the coordination of the PEtN molecule to Zn^{2+} is not stable for ICR^{Mc} T315A mutant. The zinc ion of T315A mutant is coordinated by three protein residues which leaves space within the zinc coordination shell for coordination by PEtN. It can be reasonably predicted that the binding of the phosphate of PEtN to Zn^{2+} will be harder for the WT ICR^{Mc} protein as Thr315 occupies the fourth coordination position of Zn^{2+}, making the zinc shell more crowded. A hypothesis for binding of the substrate PEtN could be that the phosphate group of PEtN is transferred to Thr315 via temporary coordination to Zn^{2+} in the presence of a Zn^{2+}-Thr315 coordinating interaction (i.e., a 5 coordinate intermediate species), rather than by substituting for the coordination of Thr315. This hypothesis is consistent with the findings in the study of Stogios *et al.*: PEtN is only present in the crystal structure of the T315A mutant while only a phosphate group (likely to be added during bacterial expression) is detected in the crystal structure of the WT protein (i.e., the electron density of PEtN was not observed).

After applying the multiscale computational approach on the the ICR^{Mc}:PEtN complex, then we used the workflow to investigate complexes of MCR-1 with the small thiol TGA. Simulations of the mono-zinc $MCR\text{-}1^{CD}$:TGA model suggested that the presence of the Thr285-Zn interaction prevents TGA from binding to the zinc. In simulations using both MM MD and QM/MM refinement, Thr285 rotated almost 180 degrees opposite to the position of the zinc ion when TGA is coordinating to zinc. For DFTB3/MM MD simulations starting directly from initial structures obtained by docking, TGA only binds to zinc in simulations of the T285A mutant, which is consistent with the findings described above for the model of the ICR^{Mc}:PEtN complex. These results indicate that absence of zinc coordination by Thr285 may be a prerequisite for TGA to bind to the zinc because this provides the opportunity (e.g., empty ligand position and space near the zinc ion) for TGA to coordinate zinc. These results also suggest that the zinc ion is more likely to be 4-coordinated rather than to be 5- or 6-coordinated. Because the zinc is already 4-coordinated in the uncomplexed form of MCR-1, it is difficult for the external TGA to interact with the zinc. The zinc ion (i.e., $Zn1^{2+}$) is mostly 4-coordinated in the catalytic mechanism study of MCR-1 carried out by Suardiaz et al[142]. The zinc ion ($Zn1^{2+}$) is temporarily

5-coordinated when the substrate PEtN first enters the zinc coordination shell. The entry of PEtN increased the distance between His466 and the zinc ion from 2.1 (crystal value in 5LRN[15]) to 2.4 Å approaching the boundary 2.5 Å. PEtN was then added to Thr285 forming phosphorylated Thr285 which finally coordinate zinc as single ligand. During the second step of the reaction, a second zinc (i.e., $Zn2^{2+}$) is recruited and His466 changed from coordinating the primary metal ion ($Zn1^{2+}$) to coordinating the second zinc ($Zn2^{2+}$). The coordinating shift of His466 provided the opportunity for PEtN and Thr285 stably coordinate the $Zn^{2+}{}_1$ ion without pushing away other existing zinc ligands. Take together the results of this section and outcome from Suardiaz et al, it is reasonable to assume that the entry of an external ligand to the $Zn^{2+}{}_1$ coordination shell involves the substitution/leaving of an existing ligand (e.g., Thr285 and His466). Loss of coordination by two different ligands may reflect the differing computational treatments between the two studies. TGA is not highly capable of substituting the existing ligands of $Zn1^{2+}$ and it cannot be added to one of the existing ligands, which makes it hard to bind to the $Zn1^{2+}$ ion. This may be a possible reason for the weak inhibitory effect of TGA against MCR-1^{CD}. The protonation state of His395 is not as crucial as the presence of Thr285 for the mono-zinc MCR-1^{CD}:TGA model. Presence of a positively charged HIP395 did not significantly contribute to the positioning of the negatively charged TGA. His395 is also conserved in the structure of the ICR^{Mc}:PEtN complex (His429), but it is unlikely to be positively charged in this complex structure due to its close distance to the positive amino group of PEtN.

For the di-zinc MCR-1^{CD}:TGA model, the results showed that TGA usually only coordinated with the $Zn2^{2+}$ ion via its carboxylate group, instead of bridging both zinc ions with its thiol. Therefore, the binding pose of TGA predicted by the simulations is not fully consistent with the pose of TGA in the two reference structures (of its complexes with the zinc-dependent β-lactamases IMP-1 and SMB-1). This may be due to the significant increase in the $Zn1^{2+}$ - $Zn2^{2+}$ atomic distance during MD simulations. The increased distance between two zinc ions (5.5-6.0 Å) makes it impossible for either a water molecule or TGA to bridge the two zinc ions with coordination distance to both zinc ions remaining within 2.5 Å. Taken together, the predicted structure of the complex of MCR-1^{CD}:TGA using the workflow developed here suggests that the inhibitor may not bind strongly to the mono-zinc form of MCR-1^{CD} via the $Zn1^{2+}$ ion, but can bind to the $Zn2^{2+}$ ion in the possible dizinc form of MCR-1^{CD}. Considering the catalytic mechanism of MCR-1[142], TGA may come into the $Zn2^{2+}$ site after the first step of MCR-1 catalysis. Then it may disrupt the second step of MCR-1's catalytic reaction (i.e., transferring phosphate from phosphorylated Thr285 to lipid A) by binding to the second Zn^{2+} ion required for this reaction step. The simulation results of MCR-1^{CD}:TGA here also suggests that it is possible for a second zinc to be present in the MCR-1 structure. The MCR-1 protein in solution may consists of both mono-zinc and dizinc forms. However, considering that simulations predict that TGA can stably bind to the $Zn2^{2+}$ ion in a di-zinc MCR-1^{CD} structure, but that TGA only weakly inhibits the activity of MCR-1^{CD} in the chromogenic assay described in Chapter 2, it is likely that the dizinc form is not present in significant quantity in solution.

Although zinc geometries of the MCR-1^{CD}:TGA complex have been successfully investigated using the workflow presented here, caution needs to be taken in interpreting the results, and the accuracy and reliability of the predictions needs to be further improved. The results show that DFTB3 treatment has limitations in dealing with Zn - S interactions, especially in the presence of Zn - O interactions. This disadvantage may have an important impact on the prediction of interactions between proteins and thiolates, resulting in predicted binding poses of thiolates being inconsistent with crystal structures. In the simulations here, the sulfur atom of a thiolate group often tends to leave the zinc ion when water molecules or carboxylate groups are close. This may lead to, for example, TGA predicted as binding to the zinc ion through its carboxylate, rather than its thiolate part (e.g. MCR-1^{CD}:TGA complex, IMP-1:TGA complex) or the MMTZ inhibitor being replaced by a water molecule (e.g. in the Sfh-I:L-anti-1a complex) in DFTB3/MM MD simulations. D3BJ corrections are important for modelling protein-ligand interactions[291] and were found to improve the prediction accuracy of the B3LYP method during workflow development. The DFTB3-D3BJ method may improve the accuracy of prediction as well as helping to overcome the limitations of the workflow when dealing with Zn-S interactions. It will be interesting to investigate the performance of simulations using DFTB3-D3BJ in the future.

The protonation state of the interacting ligand is another issue to be cautious about. The protonation state of TGA used in our model is the fully deprotonated state, mainly because TGA in the two reference structures bridges the two zinc ions via its thiolate, not the carboxylate. However, according to the pK_a plot from Chemicalize (https://chemicalize.com), the dominating protonation state of TGA in the range of pH 6 to pH 8 has a deprotonated carboxylate and protonated thiolate group. It is nevertheless possible that the microenvironment of cells fully deprotonates TGA. Further simulations using the current simulation approach for the partially deprotonated TGA:MCR-1 complex may be useful.

In this chapter, a multiscale modelling workflow is developed and extensively tested to generate accurate models of Zn metalloprotein complexes. The workflow is then applied to complex structures of ICR^{Mc} and MCR-1^{CD}. The simulation results are consistent with the limited experimental data suggesting that stable complexes can be obtained with Thr285 Ala mutants and support the relative instability of the ICR^{Mc} noncovalent PEtN complex. The simulation outcome of MCR-1^{CD}:TGA complex is consistent with the idea that TGA is not a strong inhibitor of mono-zinc MCR-1^{CD}.

Chapter 5. General discussion and prospects

The overall aim of this PhD project is to investigate potential inhibitors of the mobile colistin resistance determinant MCR-1 via experimental and computational approaches. Virtual screening using BUDE was performed to identity potential inhibitors of MCR-1. Representatives of the top-ranked compounds were examined in biological experiments. The effects of several compounds that have been reported to interact with zinc containing proteins in the literature, as well as some empirical suggestions (e.g., thiolates, boronates, MCR-1 substrate similarities, etc.) were also investigated in experimental assays. Two biochemical assays were developed to investigate interactions between MCR-1^{CD} (the MCR-1 catalytic domain) and ligands. TGA showed weak binding and inhibitory effects on MCR-1 in these assays. To investigate the binding mode of TGA to MCR-1^{CD}, for which the complex crystal structure was not available, MD simulations were performed. A multiscale modelling approach combining MM and QM/MM MD calculations was developed. The multiscale approach was first used to model protein:ligand complexes of zinc containing proteins for which crystal structures were available. After method validation with these structures, the multiscale modelling approach was applied to an MCR-like system that has an available crystal structure, the ICR^{Mc}:PEtN complex. The computational approach was lastly used to predict binding interactions in the MCR-1^{CD}:TGA complex.

Colistin susceptibility assays were used to validate the effects of the above compounds on MCR-1 expressing strains. Compound C4 was found to potentiate colistin activity against bacterial strains with MCR-1 and/or ArnT-mediated colistin resistance. Compound C4 potentiated activity of the carbapenem meropenem against *E.coli* strains carrying the metallo-β-lactamase IMP-1. In experiments compound C4 did not permeabilize the outer cell membrane nor compromise the cell membrane integrity. Compound C4 did not potentiate colistin activity against the colistin sensitive ATCC-25922 *E.coli* and Ecl8 *K. pneumoniae* strains. Although the mechanism of action of compound C4 was not fully uncovered, these findings provide some clues towards this and suggest the direction of further investigations.

5.1 The zinc geometry in the catalytic site of MCR-1

The multiscale simulation approach was used to predict binding interactions in the MCR-1^{CD}:TGA complex without a crystal structure for reference. Although the workflow is validated, due to the complexity of zinc ions, interpretation of the simulation results should be cautious. The $Zn1^{2+}$ ion in the MCR-1^{CD} protein has four coordinating residues (Glu246, Thr285, Asp465 and His466) which differs from the other zinc proteins included in our simulations. Except for the MCR-like ICR^{Mc} protein, the Zn^{2+} ion in mono-zinc metalloproteins tested here, including Sfh-1, HDAC-2, ACE and ACE-2, is bound only to three protein ligands leaving the fourth coordination position open for non-protein ligands. A geometry in which the zinc ion is bound to three amino acid residues with the fourth position usually occupied by a water molecule is also the most common geometry in catalytic sites[249]. The 4-coordinated primary zinc centre of MCR-1 ($Zn1^{2+}$)

complicates the prediction of the zinc coordination pose of the small thiol TGA. Results of QM/MM MD simulations suggest that TGA cannot stably coordinate the $Zn1^{2+}$ ion in the mono-zinc form of MCR-1, either detaching entirely from the metal centre or retaining coordination only via its carboxylate, rather than the thiol, group. Although unexpected, this is consistent with findings that the carboxylate, rather than thiol, groups of ligands with both functionals can coordinate to the catalytic metal ion in mono-zinc metallo-β-lactamases[59]. On the other hand, TGA can stably coordinate the possible second $Zn2^{2+}$ ion when present in the MCR-1 active site, mostly via its carboxylate group. This contrasts with the mode of binding observed when TGA intercalates between the two Zn^{2+} ions and occupies a coordination site for both metal ions in crystallized SMB-1:TGA and IMP-1:TGA complexes.

As described in Chapter 2, half of the top-scoring ligands identified by BUDE are predicted to bind far from the $Zn1^{2+}$ ion in MCR-1^{CD}, while most of these ligands could easily interact with the $Zn2^{2+}$ ion when present. In the MD simulations in Chapter 4, it was hard to retain the binding of TGA to the Zn^{2+} ion in the mono-zinc MCR-1^{CD} structure, but binding was stable with the di-zinc MCR-1^{CD} structure. Recent QM cluster studies of the catalytic mechanism[142,273] of MCR-1 showed that the transfer of PEtN from a membrane phospholipid to Thr285 of MCR-1 can proceed in the presence of only a single $Zn1^{2+}$ ion. However, these calculations suggested that the second $Zn2^{2+}$ ion in a dinuclear site may be necessary to support the transfer of PEtN from the phosphorylated Thr285 to the final lipid A acceptor. All these calculations support the possible recruitment of a second zinc ion to the catalytic site of MCR-1 during the second stage of the PEtN transfer reaction. Experimental determination of the number of zinc ions in MCR-1, and the dependence of activity on zinc content, would be needed to validate this hypothesis.

5.2 BUDE screening and computational modelling

During the BUDE virtual screening, thiolate compounds in the ZINC8 ligand library were not included. These ligands are excluded because thiolates are known ZBGs that chelate zinc ions and abolish the activity of zinc proteins. The selectivity of such zinc chelators is typically poor as thiolates may not only bind to zinc ions in MCR-1, but may also react with other zinc-containing proteins in the human body, thus causing off-target effects and reducing the likely safety of any compound used *in vivo*[167]. Although we did not directly obtain an inhibitor of MCR-1 from the top-scoring ligands given by BUDE, one derivate of these candidate ligands, compound C4, successfully potentiated colistin activity against MCR-1 producing strains. This demonstrates a need to increase the reliability with which ligands for metalloproteins can be identified in virtual screening approaches, and justifies development of methods for better filtering screening hits, for instance, application of the multiscale pipeline described here.

The multiscale pipeline may be applied directly after virtual screening to postprocess the conformations of the top ligands[293]. Because standard MD methods are usually dependent on the initial structure[294], the current

multiscale modelling method described in Chapter 4 is not fully ready for such an application. The prediction accuracy of Zn-ligand interactions needs to be further improved to minimize the need for manual modification of the initial structures based on individual knowledge and experience. To improve the accuracy of MM MD predictions, modification of the forcefield and water models may be starting points. During the validation of the multiscale modelling approach, the effect of using different water models in MM MD non-bonded model was found not to be a dominating factor for obtaining the final zinc geometry. Recently, new LJ12-6-4 parameters for the OPC, OPC3, and TIP4P water models have been published[295]. We have tested the OPC3 and OPC water models with the state-of-the-art Amber ff19SB forcefield[296]; the problem of an extra water adding to a non-octahedral zinc geometry remained, and no significant improvement was found compared to the currently used combination of ff14SB and the SPC/E water model. Using more detailed water models (e.g. OPC3 and OPC water models) does not stop formation of a zinc octahedral binding geometry. Therefore, the SPC/E water model was used for majority of these simulations, because it was found to perform well in combination with the LJ12-6-4 model[252].

The major difference in the Amber ff19SB, compared to the ff14SB forcefield, is the updated parameters for protein backbone atoms developed using machine learning techniques. However, the description of protein side chains is the same for ff14SB and ff19SB, and polarization effects are not well modelled. The development of MM forcefields that have representations of polarization effects, or are fully derived by machine learning with high-quality training set, is thriving recently[105,297]. These novel MM forcefields may provide better calculation of zinc-ligand interactions compared to the currently used ff14SB. It will be worthwhile to look at these forcefields in further work. For QM/MM MD calculations, the addition of the empirical dispersion corrections D3BJ to the DFTB3 method is of interest, and improves the accuracy with which experimental structures are predicted, as described in Chapter 4. It may also be interesting to evaluate the performance of the recently developed semi-empirical GFN2-xTB method[298] on zinc-ligand interactions in zinc metalloprotein complexes in the future.

5.3 Directly investigation of molecular interactions with SPR

The mechanism of action of compound C4 remains unexplained, but the data presented here do not support C4 directly compromising cell membranes. The possibility that C4 binds to zinc ions to inactivate the protein cannot be ruled out. This is because compound C4 potentiates colistin activity against both MCR-1 producing *E. coli* and MCR-1 and ArnT mediated colistin resistant *K. pneumoniae* strains. Compound C4 also potentiates meropenem against IMP-1 mediated carbapenem resistant cells but not KPC-2 mediated carbapenem resistant cells. MCR-1, ArnT and IMP-1 are metalloproteins while KPC-2 is a serine β-lactamase. As explained in Chapter 1, the potential activities of the C4 scaffold remain to be fully explored.

Investigation of the mechanism of C4 and the structure of its putative complex with MCR-1 may provide useful information for further screening studies. Current biochemical and biophysical assays developed in

Chapter 2 are not optimal as they produce low absorbance/fluorescence signal and require high amount of MCR-1 protein. To directly investigate molecular interactions between the compound (e.g., C4) and MCR-1 protein, developing an assay platform using SPR (surface plasmon resonance) technique is an option. Completing binding activity between compounds and protein with SPR platform should also provide useful information for understanding of the mechanism of compound C4. Although His-tagged MCR-1^{CD} could be immobilized to the Ni-NTA chip for use in SPR experiments, the binding affinity is expected to be low because the interaction between Ni-NTA and His-tag is not very strong. Weak binding to the SPR chip via His-tag may cause MCR-1^{CD} proteins to be washed off throughout the assay leading to unreliable results. To immobilize the protein on the chip of SPR equipment, the biotin-streptavidin reaction can be utilized. The binding between biotin and streptavidin is one of the strongest known non-covalent biological interactions. To biotinylate MCR-1^{CD}, the AviTag-peptide $(GLNDIFEAQKIEWHE)^{299}$ recognition site may be added in the sequence of MCR-1^{CD}. The 15 amino acid AviTag-peptide may be introduced at the C-terminus of MCR-1^{CD} to maximize the exposure of the zinc binding site of MCR-1^{CD}. In preliminary work the *mcr-1* gene fused to a C-terminal avi-tag coding sequence has been introduced into the pOPINF T7 plasmid and transformed into the SoluBL21 (DE3) *E.coli* cells for protein production. The method described in Chapter 2 can then be applied to the production and purification of MCR-1^{CD} proteins. The next step will then be to test whether MCR-1^{CD} labelled with the AviTag-peptide can be biotinylated by mixing with *E. coli* biotin ligase (BirA) for immobilization on streptavidin chips (series S SA). Once this has been achieved, it will be possible to apply the SPR method to known and candidate ligands of MCR-1. In addition, recent studies have reported that cyclic boronates interact with zinc ions in MBLs which inactivated these enzymes[59]. The interactions of boronates and MCR-1^{CD} is also worthy to investigate with SPR technique at the molecular level.

In this PhD project, a combination of computational and experimental approaches was used to seek and investigate potential inhibitors of MCR-1. Virtual ligand screening was carried out with BUDE on the zinc site of MCR-1 and laboratory experiments are developed to investigate interactions between MCR-1 and possible ligands. Top ranked ligands from BUDE, reported inhibitors of zinc metalloproteins and compounds proposed based on our knowledge are tested using colistin susceptibility assays. Compound C4 is identified as an inhibitor of MCR-1. It potentiates colistin activity against both MCR-1 and ArnT mediated colistin resistant cells and also potentiates meropenem activity against IMP-1 producing strains. Investigation of the mechanism of action C4 has been attempted and found that C4 does not affect the cell membrane integrity/permeability. To investigate molecular interactions between MCR-1 and its ligands, a multiscale modelling workflow has been developed and validated on a variety of complex structures of zinc metalloproteins. This was then applied to known and modelled complexes of MCR-1 and related proteins, and obtains results that are consistent with current experimental findings. The work presented here provides the basis for a toolkit that can be used to explore inhibitor discovery for MCR-1 mediated colistin resistance, as well as computational tools with more general application to zinc metalloproteins.

References

(1) Landers, T. F.; Cohen, B.; Wittum, T. E.; Larson, E. L. A Review of Antibiotic Use in Food Animals: Perspective, Policy, and Potential. *Public Health Rep.* **2012**, *127* (1), 4–22. https://doi.org/10.1177/003335491212700103.

(2) Butler, M. S.; Paterson, D. L. Antibiotics in the Clinical Pipeline in October 2019. *J. Antibiot. (Tokyo).* **2020**, *73* (6), 329–364. https://doi.org/10.1038/s41429-020-0291-8.

(3) Jernigan, J. A.; Hatfield, K. M.; Wolford, H.; Nelson, R. E.; Olubajo, B.; Reddy, S. C.; McCarthy, N.; Paul, P.; McDonald, L. C.; Kallen, A.; Fiore, A.; Craig, M.; Baggs, J. Multidrug-Resistant Bacterial Infections in U.S. Hospitalized Patients, 2012–2017. *N. Engl. J. Med.* **2020**, *382* (14), 1309–1319. https://doi.org/10.1056/NEJMoa1914433.

(4) Laxminarayan, R.; Amábile-Cuevas, C. F.; Cars, O.; Evans, T.; Heymann, D. L.; Hoffman, S.; Holmes, A.; Mendelson, M.; Sridhar, D.; Woolhouse, M.; Røttingen, J.-A. UN High-Level Meeting on Antimicrobials—What Do We Need? *Lancet* **2016**, *388* (10041), 218–220. https://doi.org/10.1016/S0140-6736(16)31079-0.

(5) Holmes, A. H.; Moore, L. S. P.; Sundsfjord, A.; Steinbakk, M.; Regmi, S.; Karkey, A.; Guerin, P. J.; Piddock, L. J. V. Understanding the Mechanisms and Drivers of Antimicrobial Resistance. *Lancet* **2016**. https://doi.org/10.1016/S0140-6736(15)00473-0.

(6) Magiorakos, A. P.; Srinivasan, A.; Carey, R. B.; Carmeli, Y.; Falagas, M. E.; Giske, C. G.; Harbarth, S.; Hindler, J. F.; Kahlmeter, G.; Olsson-Liljequist, B.; Paterson, D. L.; Rice, L. B.; Stelling, J.; Struelens, M. J.; Vatopoulos, A.; Weber, J. T.; Monnet, D. L. Multidrug-Resistant, Extensively Drug-Resistant and Pandrug-Resistant Bacteria: An International Expert Proposal for Interim Standard Definitions for Acquired Resistance. *Clin. Microbiol. Infect.* **2012**, *18* (3), 268–281. https://doi.org/10.1111/j.1469-0691.2011.03570.x.

(7) Lazarus, B.; Paterson, D. L.; Mollinger, J. L.; Rogers, B. A. Do Human Extraintestinal Escherichia Coli Infections Resistant to Expanded-Spectrum Cephalosporins Originate from Food-Producing Animals? A Systematic Review. *Clinical Infectious Diseases*. 2015. https://doi.org/10.1093/cid/ciu785.

(8) Schwarz, S.; Johnson, A. P. Transferable Resistance to Colistin: A New but Old Threat. *J. Antimicrob. Chemother.* **2016**, *71* (8), 2066–2070. https://doi.org/10.1093/jac/dkw274.

(9) Xu, Y.; Wei, W.; Lei, S.; Lin, J.; Srinivas, S.; Feng, Y. An Evolutionarily Conserved Mechanism for Intrinsic and Transferable Polymyxin Resistance. *MBio* **2018**, *9* (2), 1–18. https://doi.org/10.1128/mBio.02317-17.

(10) Sun, J.; Zhang, H.; Liu, Y. H.; Feng, Y. Towards Understanding MCR-like Colistin Resistance. *Trends Microbiol.* **2018**, *xx*, 1–15. https://doi.org/10.1016/j.tim.2018.02.006.

(11) Tacconelli, E.; Sifakis, F.; Harbarth, S.; Schrijver, R.; van Mourik, M.; Voss, A.; Sharland, M.; Rajendran, N. B.; Rodri´guez-Baño, J. Surveillance for Control of Antimicrobial Resistance. *Lancet Infect. Dis.* **2017**, 99–106. https://doi.org/10.1016/S1473-3099(17)30485-1.

(12) Peleg, A. Y.; Hooper, D. C. Hospital-Acquired Infections Due to Gram-Negative Bacteria. *N. Engl. J. Med.* **2010**, *362* (19), 1804–1813. https://doi.org/10.1056/NEJMra0904124.

(13) Jim O'Neill. Tackling Drug-Resistant Infections Globally: Final Report and Recommendations. Government of the United Kingdom 2016. https://doi.org/https://doi.org/APO-63983.

(14) de Kraker, M. E. A.; Stewardson, A. J.; Harbarth, S. Will 10 Million People Die a Year Due to Antimicrobial Resistance by 2050? *PLoS Med.* **2016**. https://doi.org/10.1371/journal.pmed.1002184.

(15) Hinchliffe, P.; Yang, Q. E.; Portal, E.; Young, T.; Li, H.; Tooke, C. L.; Carvalho, M. J.; Paterson, N. G.; Brem, J.; Niumsup, P. R.; Tansawai, U.; Lei, L.; Li, M.; Shen, Z.; Wang, Y.; Schofield, C. J.; Mulholland, A. J.; Shen, J.; Fey, N.; Walsh, T. R.; Spencer, J. Insights into the Mechanistic Basis of Plasmid-Mediated Colistin Resistance from Crystal Structures of the Catalytic Domain of MCR-1. *Sci. Rep.* **2017**, *7* (January), 1–10. https://doi.org/10.1038/srep39392.

(16) Vila, J.; Sáez-López, E.; Johnson, J. R.; Römling, U.; Dobrindt, U.; Cantón, R.; Giske, C. G.; Naas, T.; Carattoli, A.; Martínez-Medina, M.; Bosch, J.; Retamar, P.; Rodríguez-Baño, J.; Baquero, F.; Soto, S. M. Escherichia Coli: An Old Friend with New Tidings. *FEMS Microbiol. Rev.* **2016**, *40* (4), 437–463. https://doi.org/10.1093/femsre/fuw005.

(17) Thelwall, S.; Nsonwu, O.; Rooney, G.; Chudasama, D.; Wasti, S.; Anselmo, J.; Hope, R. *Annual Epidemiological Commentary: Gram-Negative Bacteraemia, MRSA Bacteraemia, MSSA Bacteraemia and C. Difficile Infections, up to and Including Financial Year April 2017 to March 2018*; 2018.

(18) *Critically Important Antimicrobials for Human Medicine, 6th Revision*; Geneva:World Health Organization, 2019.

(19) Liu, Y. Y.; Chandler, C. E.; Leung, L. M.; McElheny, C. L.; Mettus, R. T.; Shanks, R. M. Q.; Liu, J. H.; Goodlett, D. R.; Ernst, R. K.; Doi, Y. Structural Modification of Lipopolysaccharide Conferred by Mcr-1 in Gram-Negative ESKAPE Pathogens. *Antimicrob. Agents Chemother.* **2017**, *61* (6). https://doi.org/10.1128/AAC.00580-17.

(20) Liu, Y. Y.; Wang, Y.; Walsh, T. R.; Yi, L. X.; Zhang, R.; Spencer, J.; Doi, Y.; Tian, G.; Dong, B.; Huang, X.; Yu, L. F.; Gu, D.; Ren, H.; Chen, X.; Lv, L.; He, D.; Zhou, H.; Liang, Z.; Liu, J. H.; Shen, J. Emergence of Plasmid-Mediated Colistin Resistance Mechanism MCR-1 in Animals and Human Beings in China: A Microbiological and Molecular Biological Study. *Lancet Infect. Dis.* **2016**, *16* (2), 161–168. https://doi.org/10.1016/S1473-3099(15)00424-7.

(21) Baron, S.; Hadjadj, L.; Rolain, J. M.; Olaitan, A. O. Molecular Mechanisms of Polymyxin Resistance: Knowns and Unknowns. *Int. J. Antimicrob. Agents* **2016**, *48* (6), 583–591. https://doi.org/10.1016/j.ijantimicag.2016.06.023.

(22) Poirel, L. J. N. Polymyxins: Antibacterial Activity, Susceptibility Testing, and Resistance Mechanisms Encoded by Plasmids or Chromosomes. **2017**, *30* (2), 557–596.

(23) Davies, M.; Walsh, T. R. A Colistin Crisis in India. *Lancet Infect. Dis.* **2018**, *18* (3), 256–257. https://doi.org/10.1016/S1473-3099(18)30072-0.

(24) Olaitan, A. O.; Morand, S.; Rolain, J.-M. Mechanisms of Polymyxin Resistance: Acquired and Intrinsic Resistance in Bacteria. *Frontiers in Microbiology*. 2014, p 643.

(25) Laurent, P.; Aurélie, J.; Patrice, N. Polymyxins: Antibacterial Activity, Susceptibility Testing, and Resistance Mechanisms Encoded by Plasmids or Chromosomes. *Clin. Microbiol. Rev.* **2017**, *30* (2), 557–596. https://doi.org/10.1128/CMR.00064-16.

(26) Brown, P.; Dawson, M. J. Development of New Polymyxin Derivatives for Multi-Drug Resistant Gram-Negative Infections. *J. Antibiot. (Tokyo).* **2017**, *70* (4), 386–394. https://doi.org/10.1038/ja.2016.146.

(27) Pérez Chavarría, R. Understanding Colistin Resistance in Gram-Negative Bacteria.PhD Thesis.The University of Bristol., The University of Bristol, 2021.

(28) Wanty, C.; Anandan, A.; Piek, S.; Walshe, J.; Ganguly, J.; Carlson, R. W.; Stubbs, K. A.; Kahler, C. M.; Vrielink, A. The Structure of the Neisserial Lipooligosaccharide Phosphoethanolamine Transferase A (LptA) Required for Resistance to Polymyxin. *J. Mol. Biol.* **2013**, *425* (18), 3389–3402. https://doi.org/10.1016/j.jmb.2013.06.029.

(29) Sabnis, A.; Hagart, K. L. H.; Klöckner, A.; Becce, M.; Evans, L. E.; Furniss, R. C. D.; Mavridou, D. A. I.; Murphy, R.; Stevens, M. M.; Davies, J. C.; Larrouy-Maumus, G. J.; Clarke, T. B.; Edwards, A. M. Colistin Kills Bacteria by Targeting Lipopolysaccharide in the Cytoplasmic Membrane. *Elife* **2021**, *10*, e65836. https://doi.org/10.7554/eLife.65836.

(30) Olaitan, A. O.; Morand, S.; Rolain, J. M. Mechanisms of Polymyxin Resistance: Acquired and Intrinsic Resistance in Bacteria. *Frontiers in Microbiology*. 2014. https://doi.org/10.3389/fmicb.2014.00643.

(31) Moffatt, J. H.; Harper, M.; Boyce, J. D. Mechanisms of Polymyxin Resistance BT - Polymyxin Antibiotics: From Laboratory Bench to Bedside; Li, J., Nation, R. L., Kaye, K. S., Eds.; Springer International Publishing: Cham, 2019; pp 55–71. https://doi.org/10.1007/978-3-030-16373-0_5.

(32) Cheung, C. H. P.; Dulyayangkul, P.; Heesom, K. J.; Avison, M. B. Proteomic Investigation of the Signal Transduction Pathways Controlling Colistin Resistance in Klebsiella Pneumoniae. *Antimicrob. Agents Chemother.* **2020**, *64* (April), 1–10. https://doi.org/10.1101/2020.05.05.078428.

(33) A., G. E. The Pleiotropic Two-Component Regulatory System PhoP-PhoQ. *J. Bacteriol.* **2001**, *183* (6), 1835–1842. https://doi.org/10.1128/JB.183.6.1835-1842.2001.

(34) Wösten, M. M. S. M.; Kox, L. F. F.; Chamnongpol, S.; Soncini, F. C.; Groisman, E. A. A Signal Transduction System That Responds to Extracellular Iron. *Cell* **2000**, *103* (1), 113–125. https://doi.org/10.1016/S0092-8674(00)00092-1.

(35) Mitrophanov, A. Y.; Jewett, M. W.; Hadley, T. J.; Groisman, E. A. Evolution and Dynamics of Regulatory Architectures Controlling Polymyxin B Resistance in Enteric Bacteria. *PLOS Genet.* **2008**, *4* (10), e1000233.

(36) Gogry, F. A.; Siddiqui, M. T.; Sultan, I.; Haq, Q. M. R. Current Update on Intrinsic and Acquired Colistin Resistance Mechanisms in Bacteria. *Frontiers in Medicine*. 2021.

(37) Cox, A. D.; Wright, J. C.; Li, J.; Hood, D. W.; Moxon, E. R.; Richards, J. C. Phosphorylation of the Lipid A Region of Meningococcal Lipopolysaccharide: Identification of a Family of Transferases That Add Phosphoethanolamine to Lipopolysaccharide. *J. Bacteriol.* **2003**, *185* (11), 3270–3277. https://doi.org/10.1128/JB.185.11.3270-3277.2003.

(38) Anandan, A.; Evans, G. L.; Condic-Jurkic, K.; O'Mara, M. L.; John, C. M.; Phillips, N. J.; Jarvis, G. A.; Wills, S. S.; Stubbs, K. A.; Moraes, I.; Kahler, C. M.; Vrielink, A. *Structure of a Lipid A Phosphoethanolamine Transferase Suggests How Conformational Changes Govern Substrate Binding*; 2017; Vol. 114. https://doi.org/10.1073/pnas.1612927114.

(39) Zhi, C.; Lv, L.; Yu, L. F.; Doi, Y.; Liu, J. H. Dissemination of the Mcr-1 Colistin Resistance Gene. *Lancet Infect. Dis.* **2016**, *16* (3), 292–293. https://doi.org/10.1016/S1473-3099(16)00063-3.

(40) WANG, X.; ZHANG, H.; Sun, J.; LIU, Y.-H.; Feng, Y. *The MCR-1 Colistin Resistance: A New Challenge to Global Public Health*; 2017; Vol. 62. https://doi.org/10.1360/N972016-01084.

(41) Xavier, B. B.; Lammens, C.; Ruhal, R.; Malhotra-Kumar, S.; Butaye, P.; Goossens, H.; Malhotra-Kumar, S. Identification of a Novel Plasmid-Mediated Colistinresistance Gene, Mcr-2, in Escherichia Coli, Belgium, June 2016. *Eurosurveillance* **2016**, *21* (27), 6–11. https://doi.org/10.2807/1560-7917.ES.2016.21.27.30280.

(42) Ling, Z.; Yin, W.; Shen, Z.; Wang, Y.; Shen, J.; Walsh, T. R. Epidemiology of Mobile Colistin Resistance Genes Mcr-1 to Mcr-9. *J. Antimicrob. Chemother.* **2020**, *75* (11), 3087–3095. https://doi.org/10.1093/jac/dkaa205.

(43) Wei, W.; Srinivas, S.; Lin, J.; Tang, Z.; Wang, S.; Ullah, S.; Goutham Kota, V.; Feng, Y. Defining ICR-Mo, an Intrinsic Colistin Resistance Determinant from Moraxella Osloensis. **2018**. https://doi.org/10.1371/journal.pgen.1007389.

(44) Li, R.; Yu, H.; Xie, M.; Chen, K.; Dong, N.; Lin, D.; Chan, E. W. C.; Chen, S. Genetic Basis of Chromosomally-Encoded Mcr-1 Gene. *Int. J. Antimicrob. Agents* **2018**, *51* (4), 578–585. https://doi.org/10.1016/j.ijantimicag.2017.11.015.

(45) Du, H.; Chen, L.; Tang, Y. W.; Kreiswirth, B. N. Emergence of the Mcr-1 Colistin Resistance Gene in Carbapenem-Resistant Enterobacteriaceae. *Lancet Infect. Dis.* **2016**, *16* (3), 287–288. https://doi.org/10.1016/S1473-3099(16)00056-6.

(46) Falgenhauer, L.; Waezsada, S. E.; Yao, Y.; Imirzalioglu, C.; Käsbohrer, A.; Roesler, U.; Michael, G. B.; Schwarz, S.; Werner, G.; Kreienbrock, L.; Chakraborty, T. Colistin Resistance Gene Mcr-1 in Extended-Spectrum β-Lactamase-Producing and Carbapenemase-Producing Gram-Negative Bacteria in Germany. *Lancet Infect. Dis.* **2016**, *16* (3), 282–283. https://doi.org/10.1016/S1473-3099(16)00009-8.

(47) Quan, J.; Li, X.; Chen, Y.; Jiang, Y.; Zhou, Z.; Zhang, H.; Sun, L.; Ruan, Z.; Feng, Y.; Akova, M.; Yu, Y. Prevalence of Mcr-1 in Escherichia Coli and Klebsiella Pneumoniae Recovered from Bloodstream Infections in China: A Multicentre Longitudinal Study. *Lancet Infect. Dis.* **2017**, *17* (4), 400–410. https://doi.org/10.1016/S1473-3099(16)30528-X.

(48) Haenni, M.; Poirel, L.; Kieffer, N.; Châtre, P.; Saras, E.; Métayer, V.; Dumoulin, R.; Nordmann, P.; Madec, J. Y. Co-Occurrence of Extended Spectrum β Lactamase and MCR-1 Encoding Genes on Plasmids. *Lancet Infect. Dis.* **2016**, *16* (3), 281–282. https://doi.org/10.1016/S1473-3099(16)00007-4.

(49) Poirel, L.; Kieffer, N.; Liassine, N.; Thanh, D.; Nordmann, P. Plasmid-Mediated Carbapenem and Colistin Resistance in a Clinical Isolate of Escherichia Coli. *Lancet Infect. Dis.* **2016**, *16* (3), 281. https://doi.org/10.1016/S1473-3099(16)00006-2.

(50) Xu, Y.; Lin, J.; Cui, T.; Srinivas, S.; Feng, Y. Mechanistic Insights into Transferable Polymyxin Resistance among Gut Bacteria. *J. Biol. Chem.* **2018**, jbc.RA117.000924. https://doi.org/10.1074/jbc.RA117.000924.

(51) Wei, P.; Song, G.; Shi, M.; Zhou, Y.; Liu, Y.; Lei, J.; Chen, P.; Yin, L. Substrate Analog Interaction with MCR-1 Offers Insight into the Rising Threat of the Plasmid-Mediated Transferable Colistin Resistance. *J* **2018**, *32*, 1085–1098. https://doi.org/10.1096/fj.201700705R.

(52) Stojanoski, V.; Sankaran, B.; Prasad, B. V. V.; Poirel, L.; Nordmann, P.; Palzkill, T. Structure of the Catalytic Domain of the Colistin Resistance Enzyme MCR-1. *BMC Biol.* **2016**. https://doi.org/10.1186/s12915-016-0303-0.

(53) Coates, K.; Walsh, T. R.; Spencer, J.; Hinchliffe, P. 1.12 Å Resolution Crystal Structure of the Catalytic Domain of the Plasmid-Mediated Colistin Resistance Determinant MCR-2. *Acta Crystallogr. Sect. Struct. Biol. Commun.* **2017**. https://doi.org/10.1107/S2053230X17009669.

(54) Jian Sun, Yongchang Xu, Rongsui Gao, Jingxia Lin, Wenhui Wei, Swaminath Srinivas, Defeng Li, Run-Shi Yang, Xing-Ping Li, Xiao-Ping Liao, Ya-Hong Liu, Y. F. Deciphering MCR-2 Colistin Resistance. *MBio* **2017**, *8* (3).

(55) Sun, J.; Xu, Y.; Gao, R.; Lin, J.; Wei, W.; Srinivas, S.; Li, D.; Yang, R. S.; Li, X. P.; Liao, X. P.; Liu, Y. H.; Feng, Y. Deciphering MCR-2 Colistin Resistance. *MBio* **2017**. https://doi.org/10.1128/mBio.00625-17.

(56) Millán, J. L. Alkaline Phosphatases. *Purinergic Signal.* **2006**, *2* (2), 335. https://doi.org/10.1007/s11302-005-5435-6.

(57) Fage, C. D.; Brown, D. B.; Boll, J. M.; Keatinge-Clay, A. T.; Trent, M. S. Crystallographic Study of the Phosphoethanolamine Transferase EptC Required for Polymyxin Resistance and Motility in Campylobacter Jejuni. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2014**. https://doi.org/10.1107/S1399004714017623.

(58) Antonoplis, A.; Zang, X.; Wegner, T.; Wender, P. A.; Cegelski, L. Vancomycin–Arginine Conjugate Inhibits Growth of Carbapenem-Resistant E. Coli and Targets Cell-Wall Synthesis. *ACS Chem. Biol.* **2019**, *14* (9), 2065–2070. https://doi.org/10.1021/acschembio.9b00565.

(59) González-Bello, C.; Rodríguez, D.; Pernas, M.; Rodríguez, Á.; Colchón, E. β-Lactamase Inhibitors To Restore the Efficacy of Antibiotics against Superbugs. *J. Med. Chem.* **2020**, *63* (5), 1859–1881. https://doi.org/10.1021/acs.jmedchem.9b01279.

(60) Tooke, C. L.; Hinchliffe, P.; Bragginton, E. C.; Colenso, C. K.; Hirvonen, V. H. A.; Takebayashi, Y.; Spencer, J. β-Lactamases and β-Lactamase Inhibitors in the 21st Century. *J. Mol. Biol.* **2019**, *431* (18), 3472–3500. https://doi.org/https://doi.org/10.1016/j.jmb.2019.04.002.

(61) Tuon, F. F.; Rocha, J. L.; Formigoni-Pinto, M. R. Pharmacological Aspects and Spectrum of Action of Ceftazidime–Avibactam: A Systematic Review. *Infection* **2018**, *46* (2), 165–181. https://doi.org/10.1007/s15010-017-1096-y.

(62) Wu, G.; Cheon, E. Meropenem-Vaborbactam for the Treatment of Complicated Urinary Tract Infections Including Acute Pyelonephritis. *Expert Opin. Pharmacother.* **2018**, *19* (13), 1495–1502. https://doi.org/10.1080/14656566.2018.1512586.

(63) Klingler, F.-M.; Wichelhaus, T. A.; Frank, D.; Cuesta-Bernal, J.; El-Delik, J.; Müller, H. F.; Sjuts, H.; Göttig, S.; Koenigs, A.; Pos, K. M.; Pogoryelov, D.; Proschak, E. Approved Drugs Containing Thiols as Inhibitors of Metallo-β-Lactamases: Strategy To Combat Multidrug-Resistant Bacteria. *J. Med. Chem.* **2015**, *58* (8), 3626–3630. https://doi.org/10.1021/jm501844d.

(64) Liu, B.; Trout, R. E. L.; Chu, G.-H.; McGarry, D.; Jackson, R. W.; Hamrick, J. C.; Daigle, D. M.; Cusick, S. M.; Pozzi, C.; De Luca, F.; Benvenuti, M.; Mangani, S.; Docquier, J.-D.; Weiss, W. J.; Pevear, D. C.; Xerri, L.; Burns, C. J. Discovery of Taniborbactam (VNRX-5133): A Broad-Spectrum Serine- and Metallo-β-Lactamase Inhibitor for Carbapenem-Resistant Bacterial Infections. *J. Med. Chem.* **2020**, *63* (6), 2789–2801. https://doi.org/10.1021/acs.jmedchem.9b01518.

(65) Gogry, F. A.; Siddiqui, M. T.; Sultan, I.; Haq, Q. M. R. Current Update on Intrinsic and Acquired Colistin Resistance Mechanisms in Bacteria . *Frontiers in Medicine* . 2021.

(66) Mousavi, S. M.; Babakhani, S.; Moradi, L.; Karami, S.; Shahbandeh, M.; Mirshekar, M.; Mohebi, S.; Moghadam, M. T. Bacteriophage as a Novel Therapeutic Weapon for Killing Colistin-Resistant Multi-Drug-Resistant and Extensively Drug-Resistant Gram-Negative Bacteria. *Curr. Microbiol.* **2021**, *78* (12), 4023–4036. https://doi.org/10.1007/s00284-021-02662-y.

(67) Dong, H.; Xiang, H.; Mu, D.; Wang, D.; Wang, T. Exploiting a Conjugative CRISPR/Cas9 System to Eliminate Plasmid Harbouring the Mcr-1 Gene from Escherichia Coli. *Int. J. Antimicrob. Agents* **2019**, *53* (1), 1–8. https://doi.org/https://doi.org/10.1016/j.ijantimicag.2018.09.017.

(68) Yonglin, Z.; Shui, L.; Tingting, W.; Hui, L.; Shusheng, T.; Jianfeng, W.; Yang, W.; Xuming, D. Pterostilbene, a Potential MCR-1 Inhibitor That Enhances the Efficacy of Polymyxin B. *Antimicrob. Agents Chemother.* **2021**, *62* (4), e02146-17. https://doi.org/10.1128/AAC.02146-17.

(69) Zhou, Y.; Wang, T.; Guo, Y.; Liu, S.; Wang, J.; Shen, Y.; Tang, S.; Wang, Y.; Deng, X. In Vitro/Vivo Activity of Potential MCR-1 Inhibitor in Combination with Colistin Againsts Mcr-1-Positive Klebsiella Pneumonia. *Front. Microbiol.* **2018**, *9* (JUL), 1–10. https://doi.org/10.3389/fmicb.2018.01615.

(70) Zhou, Y.; Wang, J.; Guo, Y.; Liu, X.; Liu, S.; Niu, X.; Wang, Y.; Deng, X. Discovery of a Potential MCR-1 Inhibitor That Reverses Polymyxin Activity against Clinical Mcr-1-Positive Enterobacteriaceae. *J. Infect.* **2019**, *78* (5), 364–372. https://doi.org/10.1016/j.jinf.2019.03.004.

(71) Lengauer, T.; Rarey, M. Computational Methods for Biomolecular Docking. *Current Opinion in Structural Biology*. 1996, pp 402–406. https://doi.org/10.1016/S0959-440X(96)80061-3.

(72) Kim, J.; Kwon, J.; Kim, M.; Do, J.; Lee, D.; Han, H. Computational Protein–Ligand Docking and Virtual Drug Screening with the AutoDock Suite. *Nat. Protoc.* **2016**, *48* (7), 829–834. https://doi.org/10.1038/pj.2016.37.

(73) Jorgensen, W. L. The Many Roles of Computation in Drug Discovery. *Science (80-.).* **2004**, *303* (5665), 1813–1818. https://doi.org/10.1126/science.1096361.

(74) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nature Reviews Drug Discovery*. 2004, pp 935–949. https://doi.org/10.1038/nrd1549.

(75) DesJarlais, R. L.; Seibel, G. L.; Kuntz, I. D.; Furth, P. S.; Alvarez, J. C.; Ortiz de Montellano, P. R.; DeCamp, D. L.; Babé, L. M.; Craik, C. S. Structure-Based Design of Nonpeptide Inhibitors Specific for the Human Immunodeficiency Virus 1 Protease. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87* (17), 6644–6648. https://doi.org/10.1073/pnas.87.17.6644.

(76) Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Frontiers in Pharmacology*. 2018, p 923.

(77) De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* **2016**, *59* (9), 4035–4061. https://doi.org/10.1021/acs.jmedchem.5b01684.

(78) Trosset, J. Y.; Scheraga, H. a. Reaching the Global Minimum in Docking Simulations: A Monte Carlo Energy Minimization Approach Using Bezier Splines. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (14), 8011–8015. https://doi.org/10.1073/pnas.95.14.8011.

(79) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.* **1997**, *267* (3), 727–748. https://doi.org/10.1006/jmbi.1996.0897.

(80) Gohlke, H.; Klebe, G. Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors. *Angewandte Chemie - International Edition*. 2002, pp 2644–2676. https://doi.org/10.1002/1521-3773(20020802)41:15<2644::AID-ANIE2644>3.0.CO;2-O.

(81) Leach, A. R. Molecular Modelling : Principles and Applications. *Computers*. 2001, p 784. https://doi.org/10.1016/S0097-8485(96)00029-0.

(82) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785–2791. https://doi.org/https://doi.org/10.1002/jcc.21256.

(83) Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr. Comput. Aided. Drug Des.* **2011**, *7* (2), 146–157. https://doi.org/10.1016/j.biotechadv.2011.08.021.Secreted.

(84) Korb, O.; Stützle, T.; Exner, T. E. PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design BT - Ant Colony Optimization and Swarm Intelligence; Dorigo, M., Gambardella, L. M., Birattari, M., Martinoli, A., Poli, R., Stützle, T., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2006; pp 247–258.

(85) Korb, O.; Stützle, T.; Exner, T. E. Empirical Scoring Functions for Advanced Protein-Ligand Docking with PLANTS. *J. Chem. Inf. Model.* **2009**, *49* (1), 84–96. https://doi.org/10.1021/ci800298z.

(86) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739–1749. https://doi.org/10.1021/jm0306430.

(87) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *J. Med. Chem.* **2006**, *49* (21), 6177–6196. https://doi.org/10.1021/jm051256o.

(88) McIntosh-Smith, S.; Wilson, T.; Ibarra, A. Á.; Crisp, J.; Sessions, R. B. Benchmarking Energy Efficiency, Power Costs and Carbon Emissions on Heterogeneous Systems. *Comput. J.* **2012**, *55* (2), 192–205. https://doi.org/10.1093/comjnl/bxr091.

(89) Neudert, G.; Klebe, G. DSX: A Knowledge-Based Scoring Function for the Assessment of Protein-Ligand Complexes. *J. Chem. Inf. Model.* **2011**, *51* (10), 2731–2745. https://doi.org/10.1021/ci200274q.

(90) Velec, H. F. G.; Gohlke, H.; Klebe, G. DrugScoreCSD-Knowledge-Based Scoring Function Derived from Small Molecule Crystal Data with Superior Recognition Rate of near-Native Ligand Poses and Better Affinity Prediction. *J. Med. Chem.* **2005**, *48* (20), 6296–6303. https://doi.org/10.1021/jm050436v.

(91) Trudy McKee; James R. McKee. *Biochemistry: The Molecular Basis of Life, International Edition*, Third Edit.; McGraw-Hill, 2003.

(92) Uzman, A.; Lodish, H.; Berk, A.; Zipursky, L.; Baltimore, D. *Molecular Cell Biology (4th Edition) New York, NY, 2000, ISBN 0-7167-3136-3*; 2000; Vol. 29. https://doi.org/10.1016/S1470-8175(01)00023-6.

(93) George A. Jeffrey. *An Introduction to Hydrogen Bonding*; Oxford University Press, 1997.

(94) Hospital, A.; Goñi, J. R.; Orozco, M.; Gelpí, J. L. Molecular Dynamics Simulations: Advances and Applications. *Adv. Appl. Bioinform. Chem.* **2015**, *8*, 37–47. https://doi.org/10.2147/AABC.S70333.

(95) Hollingsworth, S. A.; Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99* (6), 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011.

(96) De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* **2016**, *59* (9), 4035–4061. https://doi.org/10.1021/acs.jmedchem.5b01684.

(97) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. A Practical Guide to Modelling Enzyme-Catalysed Reactions. *Chem. Soc. Rev.* **2012**, *41* (8), 3025–3038. https://doi.org/10.1039/c2cs15297e.

(98) Case DA, I. Y. B-S, Brozell SR, Cerutti DS, Cheatham TE III, Cruzeiro WD V., Darden TA,; Duke RE, Gilson MK, Gohlke H, Goetz AW, Greene D, Harris R, Homeyer N, Huang Y,; Izadi S, A. K., Kurtzman T, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, M, Y. Dm. K. P. Amber 2018. University of California, San Francisco: 2018.

(99) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688. https://doi.org/10.1002/jcc.20290.

(100) D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, G. Giambasu, T. Giese, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, D. M. Y. and P. A. K. AMBER 2019. University of California, San Francisco:2019.

(101) D.A. Case, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R Harris, S. Izadi, K. Kasava- jhala, A. Kovalenko, R. Krasny, T. Kurtzman, T., D. M. Y. and P. A. K. AMBER 2020. University of California, San Francisco: 2020.

(102) Brooks, B. R.; Brooks III, C. L.; Mackerell Jr., A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30* (10), 1545–1614. https://doi.org/https://doi.org/10.1002/jcc.21287.

(103) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103* (19), 3596–3607. https://doi.org/10.1021/jp984217f.

(104) Durrant, J. D.; McCammon, J. A. Molecular Dynamics Simulations and Drug Discovery. *BMC Biol.* **2011**, *9* (1), 71. https://doi.org/10.1186/1741-7007-9-71.

(105) Xu, M.; Zhu, T.; Zhang, J. Z. H. Molecular Dynamics Simulation of Zinc Ion in Water with an Ab Initio Based Neural Network Potential. *J. Phys. Chem. A* **2019**, *123* (30), 6587–6595. https://doi.org/10.1021/acs.jpca.9b04087.

(106) Li, R.; Obc, /; Lence, E.; Van Der Kamp, M. W.; González-Bello, C.; Mulholland, A. J. Organic & Biomolecular Chemistry Organic & Biomolecular Chemistry QM/MM Simulations Identify the Determinants of Catalytic Activity Differences between Type II Dehydroquinase Enzymes †. *Org. Biomol. Chem.* **2018**, *16*, 4443. https://doi.org/10.1039/c8ob00066b.

(107) Gong, W.; Wu, R.; Zhang, Y. Thiol versus Hydroxamate as Zinc Binding Group in HDAC Inhibition: An Ab Initio QM/MM Molecular Dynamics Study. *J. Comput. Chem.* **2015**, *36* (30), 2228–2235. https://doi.org/10.1002/jcc.24203.

(108) Wang, X.; Li, R.; Cui, W.; Li, Q.; Yao, J. QM/MM Free Energy Simulations of an Efficient Gluten Hydrolase (Kuma030) Implicate for a Reactant-State Based Protein-Design Strategy for General Acid/ Base Catalysis. *Sci. RepoRtS* | **2018**, *8*, 7042. https://doi.org/10.1038/s41598-018-25471-z.

(109) Wu, S.; Xu, D.; Guo, H. QM/MM Studies of Monozinc β-Lactamase CphA Suggest That the Crystal Structure of an Enzyme-Intermediate Complex Represents a Minor Pathway. *J. Am. Chem. Soc.* **2010**, *132* (51), 17986–17988. https://doi.org/10.1021/ja104241g.

(110) Ahmadi, S.; Barrios Herrera, L.; Chehelamirani, M.; Hostaš, J.; Jalife, S.; Salahub, D. R. Multiscale Modeling of Enzymes: QM-Cluster, QM/MM, and QM/MM/MD: A Tutorial Review. *Int. J. Quantum Chem.* **2018**, *118* (9), 1–34. https://doi.org/10.1002/qua.25558.

(111) Vidossich, P.; Magistrato, A. QM/MM Molecular Dynamics Studies of Metal Binding Proteins. *Biomolecules* **2014**, *4* (3), 616–645. https://doi.org/10.3390/biom4030616.

(112) Pokorná, P.; Kruse, H.; Krepl, M.; Šponer, J. QM/MM Calculations on Protein–RNA Complexes: Understanding Limitations of Classical MD Simulations and Search for Reliable Cost-Effective QM Methods. *J. Chem. Theory Comput.* **2018**, *14* (10), 5419–5433. https://doi.org/10.1021/acs.jctc.8b00670.

(113) Hou, G.; Cui, Q. Stabilization of Different Types of Transition States in a Single Enzyme Active Site: QM/MM Analysis of Enzymes in the Alkaline Phosphatase Superfamily. *J. Am. Chem. Soc.* **2013**, *135* (28), 10457–10469. https://doi.org/10.1021/ja403293d.

(114) Hou, G.; Cui, Q. QM/MM Analysis Suggests That Alkaline Phosphatase (AP) and Nucleotide Pyrophosphatase/Phosphodiesterase Slightly Tighten the Transition State for Phosphate Diester Hydrolysis Relative to Solution: Implication for Catalytic Promiscuity in the AP Superfamil. *J. Am. Chem. Soc.* **2012**, *134* (1), 229–246. https://doi.org/10.1021/ja205226d.

(115) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to

Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/https://doi.org/10.1016/j.softx.2015.06.001.

(116) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* **2020**, *153* (4), 44130. https://doi.org/10.1063/5.0014475.

(117) Metz, S.; Kästner, J.; Sokol, A. A.; Keal, T. W.; Sherwood, P. ChemShell—a Modular Software Package for QM/MM Simulations. *WIREs Comput. Mol. Sci.* **2014**, *4* (2), 101–110. https://doi.org/https://doi.org/10.1002/wcms.1163.

(118) Lu, Y.; Farrow, M. R.; Fayon, P.; Logsdail, A. J.; Sokol, A. A.; Catlow, C. R. A.; Sherwood, P.; Keal, T. W. Open-Source, Python-Based Redevelopment of the ChemShell Multiscale QM/MM Environment. *J. Chem. Theory Comput.* **2019**, *15* (2), 1317–1328. https://doi.org/10.1021/acs.jctc.8b01036.

(119) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16, Revision A.03. 2016.

(120) McCall, K. A.; Huang, C.; Fierke, C. A. Function and Mechanism of Zinc Metalloenzymes. *J. Nutr.* **2000**, *130* (5), 1437S-1446S. https://doi.org/10.1093/jn/130.5.1437S.

(121) Anzellotti, A. I.; Farrell, N. P. Zinc Metalloproteins as Medicinal Targets. *Chem. Soc. Rev.* **2008**, *37* (8), 1629–1651. https://doi.org/10.1039/B617121B.

(122) Andreini, C.; Banci, L.; Bertini, I.; Rosato, A. Counting the Zinc-Proteins Encoded in the Human Genome. *J. Proteome Res.* **2006**, *5* (1), 196–201. https://doi.org/10.1021/pr050361j.

(123) Peters, M. B.; Yang, Y.; Wang, B.; Füsti-Molnár, L.; Weaver, M. N.; Merz, K. M. Structural Survey of Zinc-Containing Proteins and Development of the Zinc AMBER Force Field (ZAFF). *J. Chem. Theory Comput.* **2010**, *6* (9), 2935–2947. https://doi.org/10.1021/ct1002626.

(124) Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.* **2013**. https://doi.org/10.1021/ct400146w.

(125) Pang, Y. P. Successful Molecular Dynamics Simulation of Two Zinc Complexes Bridged by a Hydroxide in Phosphotriesterase Using the Cationic Dummy Atom Method. *Proteins Struct. Funct. Genet.* **2001**, *45* (3), 183–189. https://doi.org/10.1002/prot.1138.

(126) Li, P.; Merz, K. M. Metal Ion Modeling Using Classical Mechanics. *Chem. Rev.* **2017**, *117* (3), 1564–1686. https://doi.org/10.1021/acs.chemrev.6b00440.

(127) Li, P.; Merz, K. M. Taking into Account the Ion-Induced Dipole Interaction in the Nonbonded Model of Ions. *J. Chem. Theory Comput.* **2014**. https://doi.org/10.1021/ct400751u.

(128) Li, P.; Song, L. F.; Merz, K. M. Systematic Parameterization of Monovalent Ions Employing the Nonbonded Model. *J. Chem. Theory Comput.* **2015**. https://doi.org/10.1021/ct500918t.

(129) Li, P.; Song, L. F.; Merz, K. M. Parameterization of Highly Charged Metal Ions Using the 12-6-4 LJType Nonbonded Model in Explicit Water. *J. Phys. Chem. B* **2015**, *119*, 29. https://doi.org/10.1021/jp505875v.

(130) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7* (4), 931–948. https://doi.org/10.1021/ct100684s.

(131) Lu, X.; Gaus, M.; Elstner, M.; Cui, Q. Parametrization of DFTB3/3OB for Magnesium and Zinc for Chemical and Biological Applications. *J. Phys. Chem. B* **2015**, *119* (3), 1062–1082. https://doi.org/10.1021/jp506557r.

(132) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. General Performance of Density Functionals. *J. Phys. Chem. A* **2007**, *111* (42), 10439–10452. https://doi.org/10.1021/jp0734474.

(133) Douglas-Gallardo, O. A.; Shepherd, I.; Bennie, S. J.; Ranaghan, K. E.; Mulholland, A. J.; Vöhringer-Martinez, E. Electronic Structure Benchmark Calculations of CO2 Fixing Elementary Chemical Steps in RuBisCO Using the Projector-Based Embedding Approach. *J. Comput. Chem.* **2020**, *41* (24), 2151–2157. https://doi.org/10.1002/jcc.26380.

(134) Claeyssens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schütz, M.; Thiel, S.; Thiel, W.; Werner, H. J. High-Accuracy Computation of Reaction Barriers in Enzymes. *Angew. Chemie - Int. Ed.* **2006**, *45* (41), 6856–6859. https://doi.org/10.1002/anie.200602711.

(135) Zheng, M.; Xu, D. New Delhi Metallo-β-Lactamase I: Substrate Binding and Catalytic Mechanism. *J. Phys. Chem. B* **2013**, *117* (39), 11596–11607. https://doi.org/10.1021/jp4065906.

(136) Wang, D.; Helquist, P.; Wiest, O. Zinc Binding in HDAC Inhibitors: A DFT Study. *J. Org. Chem.* **2007**, *72* (14), 5446–5449. https://doi.org/10.1021/jo070739s.

(137) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15). https://doi.org/10.1063/1.3382344.

(138) Stefan Grimme, S. E. and L. G. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465. https://doi.org/10.1002/jcc.21759.

(139) Gervasoni, S.; Spencer, J.; Hinchliffe, P.; Pedretti, A.; Vairoletti, F.; Mahler, G.; Mulholland, A. J. A Multiscale Approach to Predict the Binding Mode of Metallo Beta-Lactamase Inhibitors. *Proteins Struct. Funct. Bioinforma.* **2021**. https://doi.org/10.1002/prot.26227.

(140) Twidale, R. M.; Hinchliffe, P.; Spencer, J.; Mulholland, A. J. Crystallography and QM/MM Simulations Identify Preferential Binding of Hydrolyzed Carbapenem and Penem Antibiotics to the L1 Metallo-β-Lactamase in the Imine Form. *J. Chem. Inf. Model.* **2021**, *61* (12), 5988–5999. https://doi.org/10.1021/acs.jcim.1c00663.

(141) Chen, K.; Xu, L.; Wiest, O. Computational Exploration of Zinc Binding Groups for HDAC Inhibition. *J. Org. Chem.* **2013**, *78* (10), 5051–5055. https://doi.org/10.1021/jo400406g.

(142) Suardíaz, R.; Lythell, E.; Hinchliffe, P.; Van Der Kamp, M.; Spencer, J.; Fey, N.; Mulholland, A. J. Catalytic Mechanism of the Colistin Resistance Protein MCR-1. *Org. Biomol. Chem.* **2021**, *19* (17), 3813–3819. https://doi.org/10.1039/d0ob02566f.

(143) Brem, J.; Panduwawala, T.; Hansen, J. U.; Hewitt, J.; Liepins, E.; Donets, P.; Espina, L.; Farley, A. J. M.; Shubin, K.; Campillos, G. G.; Kiuru, P.; Shishodia, S.; Krahn, D.; Leśniak, R. K.; Schmidt (Adrian), J.; Calvopiña, K.; Turrientes, M.-C.; Kavanagh, M. E.; Lubriks, D.; Hinchliffe, P.; Langley, G. W.; Aboklaish, A. F.; Eneroth, A.; Backlund, M.; Baran, A. G.; Nielsen, E. I.; Speake, M.; Kuka, J.; Robinson, J.; Grinberga, S.; Robinson, L.; McDonough, M. A.; Rydzik, A. M.; Leissing, T. M.; Jimenez-Castellanos, J. C.; Avison, M. B.; Da Silva Pinto, S.; Pannifer, A. D.; Martjuga, M.;

Widlake, E.; Priede, M.; Hopkins Navratilova, I.; Gniadkowski, M.; Belfrage, A. K.; Brandt, P.; Yli-Kauhaluoma, J.; Bacque, E.; Page, M. G. P.; Björkling, F.; Tyrrell, J. M.; Spencer, J.; Lang, P. A.; Baranczewski, P.; Cantón, R.; McElroy, S. P.; Jones, P. S.; Baquero, F.; Suna, E.; Morrison, A.; Walsh, T. R.; Schofield, C. J. Imitation of β-Lactam Binding Enables Broad-Spectrum Metallo-β-Lactamase Inhibitors. *Nat. Chem.* **2022**, *14* (1), 15–24. https://doi.org/10.1038/s41557-021-00831-x.

(144) Brem, J.; Cain, R.; Cahill, S.; McDonough, M. A.; Clifton, I. J.; Jiménez-Castellanos, J.-C.; Avison, M. B.; Spencer, J.; Fishwick, C. W. G.; Schofield, C. J. Structural Basis of Metallo-β-Lactamase, Serine-β-Lactamase and Penicillin-Binding Protein Inhibition by Cyclic Boronates. *Nat. Commun.* **2016**, *7* (1), 12406. https://doi.org/10.1038/ncomms12406.

(145) H., O. C.; A., M.; M., K. S.; H., S. A. Novel Method for Detection of β-Lactamases by Using a Chromogenic Cephalosporin Substrate. *Antimicrob. Agents Chemother.* **1972**, *1* (4), 283–288. https://doi.org/10.1128/AAC.1.4.283.

(146) Guo, Y.; Wang, J.; Niu, G.; Shui, W.; Sun, Y.; Zhou, H.; Zhang, Y.; Yang, C.; Lou, Z.; Rao, Z. A Structural View of the Antibiotic Degradation Enzyme NDM-1 from a Superbug. *Protein Cell* **2011**, *2* (5), 384–394. https://doi.org/10.1007/s13238-011-1055-9.

(147) King, D. T.; Worrall, L. J.; Gruninger, R.; Strynadka, N. C. J. New Delhi Metallo-β-Lactamase: Structural Insights into β-Lactam Recognition and Inhibition. *J. Am. Chem. Soc.* **2012**, *134* (28), 11362–11365. https://doi.org/10.1021/ja303579d.

(148) McIntosh-Smith, S.; Price, J.; Sessions, R. B.; Ibarra, A. A. High Performance in Silico Virtual Drug Screening on Many-Core Processors. *Int. J. High Perform. Comput. Appl.* **2015**, *29* (2), 119–134. https://doi.org/10.1177/1094342014528252.

(149) Parker, M. J.; Sessions, R. B.; Badcoe, I. G.; Clarke, A. R. The Development of Tertiary Interactions during the Folding of a Large Protein. *Fold. Des.* **1996**, *1* (2), 145–156. https://doi.org/10.1016/S1359-0278(96)00023-5.

(150) Irwin, J. J.; Shoichet, B. K. ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45* (1), 177–182. https://doi.org/10.1021/ci049714+.

(151) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52* (7), 1757–1768. https://doi.org/10.1021/ci3001277.

(152) Sterling, T.; Irwin, J. J. ZINC 15 - Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **2015**, *55* (11), 2324–2337. https://doi.org/10.1021/acs.jcim.5b00559.

(153) Multi-million pound supercomputer speeds up research at the University of Bristol http://www.bris.ac.uk/news/2017/may/blue-crystal4.html (accessed 2018 -02 -24).

(154) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; 2006. https://doi.org/10.1007/978-0-387-46312-4.

(155) Jackson, D. A.; Mabury, S. A. Enzymatic Kinetic Parameters for Polyfluorinated Alkyl Phosphate Hydrolysis by Alkaline Phosphatase. *Environ. Toxicol. Chem.* **2012**, *31* (9), 1966–1971. https://doi.org/10.1002/etc.1922.

(156) Lorenz, U. Protein Tyrosine Phosphatase Assays. *Curr Protoc Immunol.* **2011**, No. 434, 1–14. https://doi.org/10.1002/0471142735.im1107s93.Protein.

(157) Bahta, M.; Liu, F.; Kim, S. E.; Stephen, A. G.; Fisher, R. J.; Burke, T. R. Oxime-Based Linker Libraries as a General Approach for the Rapid Generation and Screening of Multidentate Inhibitors. *Nat. Protoc.* **2012**, *7* (4), 686–702. https://doi.org/10.1038/nprot.2012.007.

(158) Lountos, G. T.; Cherry, S.; Tropea, J. E.; Waugh, D. S. Structural Analysis of Human Dual-Specificity Phosphatase 22 Complexed with a Phosphotyrosine-like Substrate. *Acta Crystallogr. Sect. FStructural Biol. Commun.* **2015**, *71*, 199–205. https://doi.org/10.1107/S2053230X15000217.

(159) Coppi, M.; Cannatelli, A.; Antonelli, A.; Baccani, I.; Di Pilato, V.; Sennati, S.; Giani, T.; Rossolini, G. M. A Simple Phenotypic Method for Screening of MCR-1-Mediated Colistin Resistance. *Clin. Microbiol. Infect.* **2018**, *24* (2), 201.e1-201.e3. https://doi.org/10.1016/j.cmi.2017.08.011.

(160) Karvanen, M.; Malmberg, C.; Lagerbäck, P.; Friberg, L. E.; Cars, O. Colistin Is Extensively Lost during Standard In Vitro Experimental Conditions. *Antimicrob. Agents Chemother.* **2017**, *61* (11), e00857-17. https://doi.org/10.1128/AAC.00857-17.

(161) Prim, N.; Rivera, A.; Coll, P.; Mirelis, B. Is Colistin Susceptibility Testing Finally on the Right Track? *Antimicrob. Agents Chemother.* **2018**, *62* (4), 1–2. https://doi.org/10.1128/AAC.02067-17.

(162) Matuschek, E.; Åhman, J.; Webster, C.; Kahlmeter, G. Antimicrobial Susceptibility Testing of Colistin – Evaluation of Seven Commercial MIC Products against Standard Broth Microdilution for Escherichia Coli, Klebsiella Pneumoniae, Pseudomonas Aeruginosa, and Acinetobacter Spp. *Clin. Microbiol. Infect.* **2018**, *24* (8), 865–870. https://doi.org/10.1016/j.cmi.2017.11.020.

(163) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. The Maximal Affinity of Ligands. *Proc. Natl. Acad. Sci.* **1999**, *96* (18), 9997–10002. https://doi.org/10.1073/pnas.96.18.9997.

(164) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand Efficiency: A Useful Metric for Lead Selection. *Drug Discov. Today* **2004**, *9* (10), 430–431. https://doi.org/10.1016/S1359-6446(04)03069-7.

(165) Shultz, M. D. Setting Expectations in Molecular Optimizations: Strengths and Limitations of Commonly Used Composite Parameters. *Bioorganic Med. Chem. Lett.* **2013**, *23* (21), 5980–5991. https://doi.org/10.1016/j.bmcl.2013.08.029.

(166) Berrow, N. S.; Alderton, D.; Sainsbury, S.; Nettleship, J.; Assenberg, R.; Rahman, N.; Stuart, D. I.; Owens, R. J. A Versatile Ligation-Independent Cloning Method Suitable for High-Throughput Expression Screening Applications. *Nucleic Acids Res.* **2007**, *35* (6), e45–e45. https://doi.org/10.1093/nar/gkm047.

(167) Zhang, L.; Zhang, J.; Jiang, Q.; Zhang, L.; Song, W.; Zhang Ã, L.; Zhang Ã, J. Zinc Binding Groups for Histone Deacetylase Inhibitors. *J. Enzyme Inhib. Med. Chem.* **2018**, *33* (1), 714–721. https://doi.org/10.1080/14756366.2017.1417274.

(168) Tuccinardi, T.; Bertini, S.; Granchi, C.; Ortore, G.; Macchia, M.; Minutolo, F.; Martinelli, A.; Supuran, C. T. Salicylaldoxime Derivatives as New Leads for the Development of Carbonic Anhydrase Inhibitors. **2013**. https://doi.org/10.1016/j.bmc.2012.08.057.

(169) Sixto-López, Y.; Bello, M.; Correa-Basurto, J. Exploring the Inhibitory Activity of Valproic Acid against the HDAC Family Using an MMGBSA Approach. *J. Comput. Aided. Mol. Des.* **2020**, No. I. https://doi.org/10.1007/s10822-020-00304-2.

(170) King, D. T.; Worrall, L. J.; Gruninger, R.; Strynadka, N. C. J. New Delhi Metallo-β-Lactamase: Structural Insights into β-Lactam Recognition and Inhibition. *J. Am. Chem. Soc.* **2012**, *134* (28), 11362–11365. https://doi.org/10.1021/ja303579d.

(171) Wang, X.; Wu, S.; Xu, D.; Xie, D.; Guo, H. Inhibitor and Substrate Binding by Angiotensin-Converting Enzyme: Quantum Mechanical/Molecular Mechanical Molecular Dynamics Studies. *J. Chem. Inf. Model.* **2011**, *51* (5), 1074–1082. https://doi.org/10.1021/ci200083f.

(172) Watermeyer, J. M.; Kröger, W. L.; O'Neill, H. G.; Sewell, B. T.; Sturrock, E. D. Characterization of Domain-Selective Inhibitor Binding in Angiotensin-Converting Enzyme Using a Novel Derivative of Lisinopril. *Biochem. J.* **2010**, *428* (1), 67–74. https://doi.org/10.1042/BJ20100056.

(173) Vandenbroucke, R. E.; Libert, C. Is There New Hope for Therapeutic Matrix Metalloproteinase Inhibition? *Nat. Rev. Drug Discov.* **2014**, *13* (12), 904–927. https://doi.org/10.1038/nrd4390.

(174) Chen, A. Y.; Adamek, R. N.; Dick, B. L.; Credille, C. V.; Morrison, C. N.; Cohen, S. M. Targeting Metalloenzymes for Therapeutic Intervention. *Chem. Rev.* **2018**. https://doi.org/10.1021/acs.chemrev.8b00201.

(175) Dormán, G.; Cseh, S.; Hajdú, I.; Barna, L.; Kónya, D.; Kupai, K.; Kovács, L.; Ferdinandy, P. Matrix Metalloproteinase Inhibitors. *Drugs* **2010**, *70* (8), 949–964. https://doi.org/10.2165/11318390-000000000-00000.

(176) Betts, J. W.; Phee, L. M.; Momin, M. H. F. A.; Umland, K.-D.; Brem, J.; Schofield, C. J.; Wareham, D. W.; Abdul Momin, M. H. F.; Umland, K.-D.; Brem, J.; Schofield, C. J.; Wareham, D. W. In Vitro and in Vivo Activity of ML302F: A Thioenolate Inhibitor of VIM-Subfamily Metallo β-Lactamases. *Medchemcomm* **2016**, *7* (1), 190–193. https://doi.org/10.1039/C5MD00380F.

(177) Salimraj, R.; Hinchliffe, P.; Kosmopoulou, M.; Tyrrell, J. M.; Brem, J.; Berkel, S. S.; Verma, A.; Owens, R. J.; McDonough, M. A.; Walsh, T. R.; Schofield, C. J.; Spencer, J. Crystal Structures of VIM-1 Complexes Explain Active Site Heterogeneity in VIM-Class Metallo-β-Lactamases. *FEBS J.* **2018**, *0* (0). https://doi.org/10.1111/febs.14695.

(178) Siemann, S.; Brewer, D.; Clarke, A. J.; Dmitrienko, G. I.; Lajoie, G.; Viswanatha, T. IMP-1 Metallo-β-Lactamase: Effect of Chelators and Assessment of Metal Requirement by Electrospray Mass Spectrometry. *Biochim. Biophys. Acta - Gen. Subj.* **2002**, *1571* (3), 190–200. https://doi.org/https://doi.org/10.1016/S0304-4165(02)00258-1.

(179) Mikata, Y.; Nodomi, Y.; Kizu, A.; Konno, H. Quinoline-Attached Triazacyclononane (TACN) Derivatives as Fluorescent Zinc Sensors. **2014**, *43*, 1684. https://doi.org/10.1039/c3dt52107a.

(180) Somboro, A. M.; Amoako, D. G.; Sekyere, J. O.; Kumalo, H. M.; Khan, R.; Bester, L. A.; Essack, S. Y.; Nojiri, H. *,4,7-Triazacyclononane Restores the Activity of-Lactam Antibiotics against Metallo-Lactamase-Producing Enterobacteriaceae: Exploration of Potential Metallo-Lactamase Inhibitors*; 2019; Vol. 85.

(181) Jones, H. B. L.; Wells, S. A.; Prentice, E. J.; Kwok, A.; Liang, L. L.; Arcus, V. L.; Pudney, C. R. A Complete Thermodynamic Analysis of Enzyme Turnover Links the Free Energy Landscape to Enzyme Catalysis. *FEBS J.* **2017**, *284* (17), 2829–2842. https://doi.org/10.1111/febs.14152.

(182) Wanty, C.; Anandan, A.; Piek, S.; Walshe, J.; Ganguly, J.; Carlson, R. W.; Stubbs, K. A.; Kahler, C. M.; Vrielink, A. The Structure of the Neisserial Lipooligosaccharide Phosphoethanolamine Transferase A (LptA) Required for Resistance to Polymyxin. *J. Mol. Biol.* **2013**, *425* (18), 3389–3402. https://doi.org/10.1016/j.jmb.2013.06.029.

(183) Stogios, P. J.; Cox, G.; Zubyk, H. L.; Evdokimova, E.; Wawrzak, Z.; Wright, G. D.; Savchenko, A. Substrate Recognition by a Colistin Resistance Enzyme from Moraxella Catarrhalis. *ACS Chem. Biol.* **2018**, *13* (5), 1322–1332. https://doi.org/10.1021/acschembio.8b00116.

(184) Wei, P.; Song, G.; Shi, M.; Zhou, Y.; Liu, Y.; Lei, J.; Chen, P.; Yin, L. Substrate Analog Interaction with MCR-1 Offers Insight into the Rising Threat of the Plasmid-Mediated Transferable Colistin Resistance. *J* **2018**, *32*, 1085–1098. https://doi.org/10.1096/fj.201700705R.

(185) EUCAST. Recommendations for MIC Determination of Colistin (Polymyxin E) As Recommended by the Joint CLSI-EUCAST Polymyxin Breakpoints Working Group. *Http://Www.Eucast.Org* **2016**, No. March, 22, 2016.

(186) Anderson, A. C.; Burnett, A. J. N.; Hiscock, L.; Maly, K. E.; Weadge, J. T. The Escherichia Coli Cellulose Synthase Subunit G (BcsG) Is a Zn2+-Dependent Phosphoethanolamine Transferase. *J. Biol. Chem.* **2020**, *295* (18), 6225–6235. https://doi.org/10.1074/jbc.RA119.011668.

(187) Hernández, I.; Fernández, J. A.; Niell, F. X. A Comparative Study of Alkaline Phosphatase Activity in Two Species of Gelidium (Gelidiales, Rhodophyta). *Eur. J. Phycol.* **1995**, *30* (1), 69–77. https://doi.org/10.1080/09670269500650811.

(188) Matsumoto, Y.; Hayama, K.; Sakakihara, S.; Nishino, K.; Noji, H.; Iino, R.; Yamaguchi, A. Evaluation of Multidrug Efflux Pump Inhibitors by a New Method Using Microfluidic Channels. *PLoS One* **2011**, *6* (4), e18547.

(189) Berrow, N. S.; Alderton, D.; Sainsbury, S.; Nettleship, J.; Assenberg, R.; Rahman, N.; Stuart, D. I.; Owens, R. J. A Versatile Ligation-Independent Cloning Method Suitable for High-Throughput Expression Screening Applications. *Nucleic Acids Res.* **2007**, *35* (6). https://doi.org/10.1093/nar/gkm047.

(190) Challand, M. R.; Salvadori, E.; Driesener, R. C.; Kay, C. W. M.; Roach, P. L.; Spencer, J. Cysteine Methylation Controls Radical Generation in the Cfr Radical AdoMet RRNA Methyltransferase. *PLoS One* **2013**, *8* (7). https://doi.org/10.1371/journal.pone.0067979.

(191) Jeannot, K.; Bolard, A.; Plésiat, P. Resistance to Polymyxins in Gram-Negative Organisms. *Int. J. Antimicrob. Agents* **2017**, *49* (5), 526–535. https://doi.org/https://doi.org/10.1016/j.ijantimicag.2016.11.029.

(192) Stokes, J. M.; French, S.; Ovchinnikova, O. G.; Bouwman, C.; Whitfield, C.; Brown, E. D. Cold Stress Makes Escherichia Coli Susceptible to Glycopeptide Antibiotics by Altering Outer Membrane Integrity. *Cell Chem. Biol.* **2016**, *23* (2), 267–277. https://doi.org/10.1016/j.chembiol.2015.12.011.

(193) Petrou, V. I.; Herrera, C. M.; Schultz, K. M.; Clarke, O. B.; Vendome, J.; Tomasek, D.; Banerjee, S.; Rajashankar, K. R.; Dufrisne, M. B.; Kloss, B.; Kloppmann, E.; Rost, B.; Klug, C. S.; Trent, M. S.; Shapiro, L.; Mancia, F. Structures of Aminoarabinose Transferase ArnT Suggest a Molecular Basis for Lipid A Glycosylation. *Science (80-.).* **2016**, *351* (6273), 608–612. https://doi.org/10.1126/science.aad1172.

(194) Poirel, L.; Jayol, A.; Bontron, S.; Villegas, M.-V.; Ozdamar, M.; Türkoglu, S.; Nordmann, P. The MgrB Gene as a Key Target for Acquired Resistance to Colistin in Klebsiella Pneumoniae. *J. Antimicrob. Chemother.* **2015**, *70* (1), 75–80. https://doi.org/10.1093/jac/dku323.

(195) Maria, F.; Jing, Y.; Shyamasree, D. M.; Nicholas, T.; Thamarai, S. Genome Sequence of Klebsiella Pneumoniae Ecl8, a Reference Strain for Targeted Genetic Manipulation. *Genome Announc.* **2013**, *1* (1), e00027-12. https://doi.org/10.1128/genomeA.00027-12.

(196) Muheim, C.; Götzke, H.; Eriksson, A. U.; Lindberg, S.; Lauritsen, I.; Nørholm, M. H. H.; Daley, D. O. Increasing the Permeability of Escherichia Coli Using MAC13243. *Sci. Rep.* **2017**, *7* (1), 17629. https://doi.org/10.1038/s41598-017-17772-6.

(197) Helander, I. M.; Mattila-Sandholm, T. Fluorometric Assessment of Gram-Negative Bacterial Permeabilization. *J. Appl. Microbiol.* **2000**, *88* (2), 213–219. https://doi.org/https://doi.org/10.1046/j.1365-2672.2000.00971.x.

(198) Maisuria, V. B.; Okshevsky, M.; Déziel, E.; Tufenkji, N. Proanthocyanidin Interferes with Intrinsic Antibiotic Resistance Mechanisms of Gram-Negative Bacteria. *Adv. Sci.* **2019**, *6* (15), 1802333. https://doi.org/https://doi.org/10.1002/advs.201802333.

(199) Ramkumar, I.; Annette, F.; R., R.; L., E. A. A Fluorescent Microplate Assay Quantifies Bacterial Efflux and Demonstrates Two Distinct Compound Binding Sites in AcrB. *Antimicrob. Agents Chemother.* **2021**, *59* (4), 2388–2397. https://doi.org/10.1128/AAC.05112-14.

(200) Riccardi, C.; Nicoletti, I. Analysis of Apoptosis by Propidium Iodide Staining and Flow Cytometry. *Nat. Protoc.* **2006**, *1* (3), 1458–1461. https://doi.org/10.1038/nprot.2006.238.

(201) Huang, E.; Yousef, A. E. The Lipopeptide Antibiotic Paenibacterin Binds to the Bacterial Outer Membrane and Exerts Bactericidal Activity through Cytoplasmic Membrane Damage. *Appl. Environ. Microbiol.* **2014**, *80* (9), 2700–2704. https://doi.org/10.1128/AEM.03775-13.

(202) Stogios, P. J.; Savchenko, A. Molecular Mechanisms of Vancomycin Resistance. *Protein Sci.* **2020**, *29* (3), 654–669. https://doi.org/https://doi.org/10.1002/pro.3819.

(203) Antonoplis, A.; Zang, X.; Wegner, T.; Wender, P. A.; Cegelski, L. Vancomycin-Arginine Conjugate Inhibits Growth of Carbapenem-Resistant E. Coli and Targets Cell-Wall Synthesis. *ACS Chem. Biol.* **2019**, *14* (9), 2065–2070. https://doi.org/10.1021/acschembio.9b00565.

(204) Muheim, C.; Götzke, H.; Eriksson, A. U.; Lindberg, S.; Lauritsen, I.; Nørholm, M. H. H.; Daley, D. O. Increasing the Permeability of Escherichia Coli Using MAC13243. *Sci. Rep.* **2017**, *7* (1), 1–11. https://doi.org/10.1038/s41598-017-17772-6.

(205) Lamers, R. P.; Cavallari, J. F.; Burrows, L. L. The Efflux Inhibitor Phenylalanine-Arginine Beta-Naphthylamide (PAβN) Permeabilizes the Outer Membrane of Gram-Negative Bacteria. *PLoS One* **2013**, *8* (3), 1–7. https://doi.org/10.1371/journal.pone.0060666.

(206) Renau, T. E.; Léger, R.; Flamme, E. M.; Sangalang, J.; She, M. W.; Yen, R.; Gannon, C. L.; Griffith, D.; Chamberland, S.; Lomovskaya, O.; Hecker, S. J.; Lee, V. J.; Ohta, T.; Nakayama, K. Inhibitors of Efflux Pumps in Pseudomonas Aeruginosa Potentiate the Activity of the Fluoroquinolone Antibacterial Levofloxacin. *J. Med. Chem.* **1999**, *42* (24), 4928–4931. https://doi.org/10.1021/jm9904598.

(207) Olga, L.; S., W. M.; Angela, L.; Jorge, G.; Richard, F.; May, L.; Johanne, B.; Deidre, C.; Suzanne, C.; Tom, R.; Roger, L.; Scott, H.; Will, W.; Kazuki, H.; Hiroko, I.; J., L. V. Identification and Characterization of Inhibitors of Multidrug Resistance Efflux Pumps in Pseudomonas Aeruginosa: Novel Agents for Combination Therapy. *Antimicrob. Agents Chemother.* **2001**, *45* (1), 105–116. https://doi.org/10.1128/AAC.45.1.105-116.2001.

(208) Schumacher, A.; Steinke, P.; Bohnert, J. A.; Akova, M.; Jonas, D.; Kern, W. V. Effect of 1-(1-Naphthylmethyl)-Piperazine, a Novel Putative Efflux Pump Inhibitor, on Antimicrobial Drug Susceptibility in Clinical Isolates of Enterobacteriaceae Other than Escherichia Coli. *J. Antimicrob. Chemother.* **2006**, *57* (2), 344–348. https://doi.org/10.1093/jac/dki446.

(209) Liu, S.; Zhang, J.; Zhou, Y.; Hu, N.; Li, J.; Wang, Y.; Niu, X.; Deng, X.; Wang, J. Pterostilbene Restores Carbapenem Susceptibility in New Delhi Metallo-β-Lactamase-Producing Isolates by Inhibiting the Activity of New Delhi Metallo-β-Lactamases. *Br. J. Pharmacol.* **2019**, *176* (23), 4548–4557. https://doi.org/https://doi.org/10.1111/bph.14818.

(210) Wayne PA: Clinical and Laboratory Standards Institute. 30th Ed. CLSI Supplement M100. *Performance Standards for Antimicrobial Susceptibility.* 2020.

(211) The European Committee on Antimicrobial Susceptibility Testing. Broth microdilution - EUCAST reading guide, version 3.0, 2021 https://www.eucast.org/ast_of_bacteria/mic_determination/.

(212) Albiger, B.; Brolund, A.; Byfors, S.; Giske, C. G.; Kahlmeter, G.; Kohlenberg, A.; Struelens, M.; Tegmark Wisell, K.; European Centre for Disease Prevention and Control. Laboratory Manual for Carbapenem and Colistin Resistance Detection and Characterisation for the Survey of Carbapenem- and/or Colistin-Resistant Enterobacteriaceae : Version 2.0. *Eur. Cent. Dis. Prev. Control* **2019**, 16.

(213) Miroux, B.; Walker, J. E. Over-Production of Proteins InEscherichia Coli: Mutant Hosts That Allow Synthesis of Some Membrane Proteins and Globular Proteins at High Levels. *J. Mol. Biol.* **1996**, *260* (3), 289–298. https://doi.org/https://doi.org/10.1006/jmbi.1996.0399.

(214) Zhou, A.; Kang, T. M.; Yuan, J.; Beppler, C.; Nguyen, C.; Mao, Z.; Nguyen, M. Q.; Yeh, P.; Miller, J. H. Synergistic Interactions of Vancomycin with Different Antibiotics against Escherichia Coli:

Trimethoprim and Nitrofurantoin Display Strong Synergies with Vancomycin against Wild-Type E. Coli. *Antimicrob. Agents Chemother.* **2015**, *59* (1), 276–281. https://doi.org/10.1128/AAC.03502-14.

(215) Stokes, J. M.; French, S.; Ovchinnikova, O. G.; Bouwman, C.; Whitfield, C.; Brown, E. D. Cold Stress Makes Escherichia Coli Susceptible to Glycopeptide Antibiotics by Altering Outer Membrane Integrity. *Cell Chem. Biol.* **2016**, *23* (2), 267–277. https://doi.org/10.1016/j.chembiol.2015.12.011.

(216) A., H. J.; M., H. R. Colistin MIC Variability by Method for Contemporary Clinical Isolates of Multidrug-Resistant Gram-Negative Bacilli. *J. Clin. Microbiol.* **2013**, *51* (6), 1678–1684. https://doi.org/10.1128/JCM.03385-12.

(217) Navon-Venezia, S.; Kondratyeva, K.; Carattoli, A. Klebsiella Pneumoniae: A Major Worldwide Source and Shuttle for Antibiotic Resistance. *FEMS Microbiol. Rev.* **2017**, *41* (3), 252–275. https://doi.org/10.1093/femsre/fux013.

(218) Garneau-Tsodikova, S.; Labby, K. J. Mechanisms of Resistance to Aminoglycoside Antibiotics: Overview and Perspectives. *Medchemcomm* **2016**, *7* (1), 11–27. https://doi.org/10.1039/C5MD00344J.

(219) Phillips, S. R.; Wilson, L. J.; Borkman, R. F. Acrylamide and Iodide Fluorescence Quenching as a Structural Probe of Tryptophan Microenvironment in Bovine Lens Crystallins. *Curr. Eye Res.* **1986**, *5* (8), 611–620. https://doi.org/10.3109/02713688609015126.

(220) Chmyrov, A.; Sandén, T.; Widengren, J. Iodide as a Fluorescence Quencher and Promoter—Mechanisms and Possible Implications. *J. Phys. Chem. B* **2010**, *114* (34), 11282–11291. https://doi.org/10.1021/jp103837f.

(221) El-Halfawy, O. M.; Valvano, M. A. Antimicrobial Heteroresistance: An Emerging Field in Need of Clarity. *Clin. Microbiol. Rev.* **2015**, *28* (1), 191–207. https://doi.org/10.1128/CMR.00058-14.

(222) Halaby, T.; Kucukkose, E.; Janssen, A. B.; Rogers, M. R. C.; Doorduijn, D. J.; van der Zanden, A. G. M.; al Naiemi, N.; Vandenbroucke-Grauls, C. M. J. E.; van Schaik, W. Genomic Characterization of Colistin Heteroresistance in Klebsiella Pneumoniae during a Nosocomial Outbreak. *Antimicrob. Agents Chemother.* **2016**, *60* (11), 6837–6843. https://doi.org/10.1128/AAC.01344-16.

(223) Cheong, H. S.; Kim, S. Y.; Wi, Y. M.; Peck, K. R.; Ko, K. S. Colistin Heteroresistance in Klebsiella Pneumoniae Isolates and Diverse Mutations of PmrAB and PhoPQ in Resistant Subpopulations. *J. Clin. Med.* **2019**, *8* (9). https://doi.org/10.3390/jcm8091444.

(224) Aurélie, J.; Patrice, N.; Adrian, B.; Laurent, P. Heteroresistance to Colistin in Klebsiella Pneumoniae Associated with Alterations in the PhoPQ Regulatory System. *Antimicrob. Agents Chemother.* **2021**, *59* (5), 2780–2784. https://doi.org/10.1128/AAC.05055-14.

(225) El-Halfawy, O. M.; Valvano, M. A. Chemical Communication of Antibiotic Resistance by a Highly Resistant Subpopulation of Bacterial Cells. *PLoS One* **2013**, *8* (7), e68874.

(226) Silva, A.; Sousa, A. M.; Alves, D.; LourenÃ Â§o, A.; Pereira, M. O. Heteroresistance to Colistin in Klebsiella Pneumoniae Is Triggered by Small Colony Variants Sub-Populations within Biofilms. *Pathog. Dis.* **2016**, *74* (5), ftw036. https://doi.org/10.1093/femspd/ftw036.

(227) Wang, G.; Zhao, G.; Chao, X.; Xie, L.; Wang, H. The Characteristic of Virulence, Biofilm and Antibiotic Resistance of Klebsiella Pneumoniae. *International Journal of Environmental Research and Public Health* . 2020. https://doi.org/10.3390/ijerph17176278.

(228) Nirwati, H.; Sinanjung, K.; Fahrunissa, F.; Wijaya, F.; Napitupulu, S.; Hati, V. P.; Hakim, M. S.; Meliala, A.; Aman, A. T.; Nuryastuti, T. Biofilm Formation and Antibiotic Resistance of Klebsiella Pneumoniae Isolated from Clinical Samples in a Tertiary Care Hospital, Klaten, Indonesia. *BMC Proc.* **2019**, *13* (11), 20. https://doi.org/10.1186/s12919-019-0176-7.

(229) Vuotto, C.; Longo, F.; Pascolini, C.; Donelli, G.; Balice, M. P.; Libori, M. F.; Tiracchia, V.; Salvia, A.; Varaldo, P. E. Biofilm Formation and Antibiotic Resistance in Klebsiella Pneumoniae Urinary Strains. *J. Appl. Microbiol.* **2017**, *123* (4), 1003–1018. https://doi.org/https://doi.org/10.1111/jam.13533.

(230) Naparstek, L.; Carmeli, Y.; Navon-Venezia, S.; Banin, E. Biofilm Formation and Susceptibility to Gentamicin and Colistin of Extremely Drug-Resistant KPC-Producing Klebsiella Pneumoniae. *J. Antimicrob. Chemother.* **2014**, *69* (4), 1027–1034. https://doi.org/10.1093/jac/dkt487.

(231) Shein, A. M. S.; Wannigama, D. L.; Higgins, P. G.; Hurst, C.; Abe, S.; Hongsing, P.; Chantaravisoot, N.; Saethang, T.; Luk-in, S.; Liao, T.; Nilgate, S.; Rirerm, U.; Kueakulpattana, N.; Laowansiri, M.; Srisakul, S.; Muhummudaree, N.; Techawiwattanaboon, T.; Gan, L.; Xu, C.; Kupwiwat, R.; Phattharapornjaroen, P.; Rojanathanes, R.; Leelahavanichkul, A.; Chatsuwan, T. Novel Colistin-EDTA Combination for Successful Eradication of Colistin-Resistant Klebsiella Pneumoniae Catheter-Related Biofilm Infections. *Sci. Rep.* **2021**, *11* (1), 21676. https://doi.org/10.1038/s41598-021-01052-5.

(232) Riquelme, S. A.; Ahn, D.; Prince, A. ***Pseudomonas Aeruginosa*** and ***Klebsiella Pneumoniae*** Adaptation to Innate Immune Clearance Mechanisms in the Lung. *J. Innate Immun.* **2018**, *10* (5–6), 442–454. https://doi.org/10.1159/000487515.

(233) Sahoo, A.; Swain, S. S.; Behera, A.; Sahoo, G.; Mahapatra, P. K.; Panda, S. K. Antimicrobial Peptides Derived From Insects Offer a Novel Therapeutic Option to Combat Biofilm: A Review. *Frontiers in Microbiology*. 2021, p 1077.

(234) Campos, M. A.; Vargas, M. A.; Regueiro, V.; Llompart, C. M.; Albertí, S.; Bengoechea, J. A. Capsule Polysaccharide Mediates Bacterial Resistance to Antimicrobial Peptides. *Infect. Immun.* **2004**, *72* (12), 7107–7114. https://doi.org/10.1128/IAI.72.12.7107-7114.2004.

(235) Kell, D. B.; Heyden, E. L.; Pretorius, E. The Biology of Lactoferrin, an Iron-Binding Protein That Can Help Defend Against Viruses and Bacteria. *Frontiers in Immunology*. 2020, p 1221.

(236) Hinchliffe, P.; Tanner, C. A.; Krismanich, A. P.; Labbé, G.; Goodfellow, V. J.; Marrone, L.; Desoky, A. Y.; Calvopiña, K.; Whittle, E. E.; Zeng, F.; Avison, M. B.; Bols, N. C.; Siemann, S.; Spencer, J.; Dmitrienko, G. I. Structural and Kinetic Studies of the Potent Inhibition of Metallo-β-Lactamases by 6-Phosphonomethylpyridine-2-Carboxylates. *Biochemistry* **2018**, *57* (12), 1880–1892. https://doi.org/10.1021/acs.biochem.7b01299.

(237) Takebayashi, Y.; Wan Nur Ismah, W. A. K.; Findlay, J.; Heesom, K. J.; Zhang, J.; Williams, O. M.; MacGowan, A. P.; Avison, M. B. Prediction of Cephalosporin and Carbapenem Susceptibility in Multi-Drug Resistant Gram-Negative Bacteria Using Liquid Chromatography-Tandem Mass Spectrometry. *bioRxiv* **2017**, 138594. https://doi.org/10.1101/138594.

(238) Tooke, C. L.; Hinchliffe, P.; Lang, P. A.; Mulholland, A. J.; Brem, J.; Schofield, C. J.; Spencer, J. Molecular Basis of Class A β-Lactamase Inhibition by Relebactam. *Antimicrob. Agents Chemother.* **2019**, *63* (10), e00564-19. https://doi.org/10.1128/AAC.00564-19.

(239) Andreini, C.; Banci, L.; Bertini, I.; Rosato, A. Zinc through the Three Domains of Life. *J. Proteome Res.* **2006**, *5* (11), 3173–3178. https://doi.org/10.1021/pr0603699.

(240) Rasia, R. M.; Vila, A. J. Exploring the Role and the Binding Affinity of a Second Zinc Equivalent in B. Cereus Metallo-β-Lactamase. *Biochemistry* **2002**, *41* (6), 1853–1860. https://doi.org/10.1021/bi010933n.

(241) Amaro, R. E.; Mulholland, A. J. Multiscale Methods in Drug Design Bridge Chemical and Biological Complexity in the Search for Cures. *Nat. Rev. Chem.* **2018**, *2* (4). https://doi.org/10.1038/s41570-018-0148.

(242) Huggins, D. J.; Biggin, P. C.; Dämgen, M. A.; Essex, J. W.; Harris, S. A.; Henchman, R. H.; Khalid, S.; Kuzmanic, A.; Laughton, C. A.; Michel, J.; Mulholland, A. J.; Rosta, E.; Sansom, M. S. P.; van der Kamp, M. W. Biomolecular Simulations: From Dynamics and Mechanisms to Computational Assays of Biological Activity. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2019**, *9* (3), 1–23. https://doi.org/10.1002/wcms.1393.

(243) Dudev, T.; Lin, Y. lin; Dudev, M.; Lim, C. First-Second Shell Interactions in Metal Binding Sites in Proteins: A PDB Survey and DFT/CDM Calculations. *J. Am. Chem. Soc.* **2003**, *125* (10), 3168–3180. https://doi.org/10.1021/ja0209722.

(244) Zhang, J.; Yang, W.; Piquemal, J. P.; Ren, P. Modeling Structural Coordination and Ligand Binding in Zinc Proteins with a Polarizable Potential. *J. Chem. Theory Comput.* **2012**, *8* (4), 1314–1324. https://doi.org/10.1021/ct200812y.

(245) Dudev, T.; Lim, C. Tetrahedral vs Octahedral Zinc Complexes with Ligands of Biological Interest: A DFT/CDM Study. *J. Am. Chem. Soc.* **2000**, *122* (45), 11146–11153. https://doi.org/10.1021/ja0010296.

(246) Laitaoja, M.; Valjakka, J.; Jänis, J. Zinc Coordination Spheres in Protein Structures. *Inorg. Chem.* **2013**, *52* (19), 10983–10991. https://doi.org/10.1021/ic401072d.

(247) Dudev, T.; Lim, C. Principles Governing Mg, Ca, and Zn Binding and Selectivity in Proteins. *Chem. Rev.* **2003**, *103* (3), 773–787. https://doi.org/10.1021/cr020467n.

(248) Dudev, T.; Lim, C. Competition among Metal Ions for Protein Binding Sites: Determinants of Metal Ion Selectivity in Proteins. *Chem. Rev.* **2014**, *114* (1), 538–556. https://doi.org/10.1021/cr4004665.

(249) Daniel, A. G.; Farrell, N. P. The Dynamics of Zinc Sites in Proteins: Electronic Basis for Coordination Sphere Expansion at Structural Sites. *Metallomics* **2014**, *6* (12), 2230–2241. https://doi.org/10.1039/c4mt00213j.

(250) Marek, M.; Shaik, T. B.; Heimburg, T.; Chakrabarti, A.; Lancelot, J.; Ramos-Morales, E.; Da Veiga, C.; Kalinin, D.; Melesina, J.; Robaa, D.; Schmidtkunz, K.; Suzuki, T.; Holl, R.; Ennifar, E.; Pierce, R. J.; Jung, M.; Sippl, W.; Romier, C. Characterization of Histone Deacetylase 8 (HDAC8) Selective Inhibition Reveals Specific Active Site Structural and Functional Determinants. *J. Med. Chem.* **2018**, *61* (22), 10000–10016. https://doi.org/10.1021/acs.jmedchem.8b01087.

(251) Macdermott-Opeskin, H.; Mcdevitt, C. A.; O'mara, M. L. Comparing Nonbonded Metal Ion Models in the Divalent Cation Binding Protein PsaA. *J. Chem. Theory Comput.* **2020**, *16* (3), 1913–1923. https://doi.org/10.1021/acs.jctc.9b01180.

(252) Panteva, M. T.; Giambaşu, G. M.; York, D. M. Comparison of Structural, Thermodynamic, Kinetic and Mass Transport Properties of Mg2+ Ion Models Commonly Used in Biomolecular Simulations. *J. Comput. Chem.* **2015**, *36* (13), 970–982. https://doi.org/10.1002/jcc.23881.

(253) Chen, H.; Giese, T. J.; Golden, B. L.; York, D. M. Divalent Metal Ion Activation of a Guanine General Base in the Hammerhead Ribozyme: Insights from Molecular Simulations. *Biochemistry* **2017**, *56* (24), 2985–2994. https://doi.org/10.1021/acs.biochem.6b01192.

(254) Pu, Z.; Zhao, M.; Zhang, Y.; Sun, W.; Bao, Y. Dynamic Description of the Catalytic Cycle of Malate Enzyme: Stereoselective Recognition of Substrate, Chemical Reaction, and Ligand Release. *J. Phys. Chem. B* **2018**, *122* (51), 12241–12250. https://doi.org/10.1021/acs.jpcb.8b05135.

(255) Xu, M.; He, X.; Zhu, T.; Zhang, J. Z. H. A Fragment Quantum Mechanical Method for Metalloproteins. *J. Chem. Theory Comput.* **2019**, No. Mm. https://doi.org/10.1021/acs.jctc.8b00966.

(256) Duarte, F.; Amrein, A.; Caroline, S.; Kamerlin, L. Modeling Catalytic Promiscuity in the Alkaline Phosphatase Superfamily. *Phys. Chem. Chem. Phys* **2013**, *15*, 11160. https://doi.org/10.1039/c3cp51179k.

(257) Xu, D.; Cui, Q.; Guo, H. Quantum Mechanical/Molecular Mechanical Studies of Zinc Hydrolases. *Int. Rev. Phys. Chem.* **2014**, *33* (1), 1–41. https://doi.org/10.1080/0144235X.2014.889378.

(258) Kulik, H. J.; Zhang, J.; Klinman, J. P.; Martínez, T. J. How Large Should the QM Region Be in QM/MM Calculations? The Case of Catechol O-Methyltransferase. *J. Phys. Chem. B* **2016**, *120* (44), 11381–11394. https://doi.org/10.1021/acs.jpcb.6b07814.

(259) Christensen, A. S.; Kubař, T.; Cui, Q.; Elstner, M. Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications. *Chem. Rev.* **2016**, *116* (9), 5301–5337. https://doi.org/10.1021/acs.chemrev.5b00584.

(260) Wang, C.; Guo, H. Inhibitor Binding by Metallo-β-Lactamase IMP-1 from Pseudomonas Aeruginosa: Quantum Mechanical/Molecular Mechanical Simulations. *J. Phys. Chem. B* **2007**, *111* (33), 9986–9992. https://doi.org/10.1021/jp073864g.

(261) Wang, Y. T.; Cheng, T. L. Refined Models of New Delhi Metallo-Beta-Lactamase-1 with Inhibitors: An QM/MM Modeling Study. *J. Biomol. Struct. Dyn.* **2016**, *34* (10), 2214–2223. https://doi.org/10.1080/07391102.2015.1110834.

(262) Xu, D.; Guo, H.; Cui, Q. Antibiotic Binding to Dizinc β-Lactamase L1 from Stenotrophomonas Maltophilia: SCC-DFTB/CHARMM and DFT Studies. *J. Phys. Chem. A* **2007**, *111* (26), 5630–5636. https://doi.org/10.1021/jp068746s.

(263) Silvia Gervasoni, James Spencer, Philip Hinchliffe, Alessandro Pedretti, Franco Vairoletti, Graciela Mahler, A. J. M. A Multiscale Approach to Predict the Binding Mode of Metallo Beta-Lactamase Inhibitors. *Proteins* **2021**. https://doi.org/10.1002/prot.26227.

(264) Rossi, M. A.; Martinez, V.; Hinchliffe, P.; Mojica, M. F.; Castillo, V.; Moreno, D. M.; Smith, R.; Spellberg, B.; Drusano, G. L.; Banchio, C.; Bonomo, R. A.; Spencer, J.; Vila, A. J.; Mahler, G. 2-Mercaptomethyl-Thiazolidines Use Conserved Aromatic-S Interactions to Achieve Broad-Range Inhibition of Metallo-β-Lactamases. *Chem. Sci.* **2021**, *12* (8), 2898–2908. https://doi.org/10.1039/d0sc05172a.

(265) Hinchliffe, P.; Moreno, D. M.; Rossi, M.-A.; Mojica, M. F.; Martinez, V.; Villamil, V.; Spellberg, B.; Drusano, G. L.; Banchio, C.; Mahler, G.; Bonomo, R. A.; Vila, A. J.; Spencer, J. 2-Mercaptomethyl Thiazolidines (MMTZs) Inhibit All Metallo-β-Lactamase Classes by Maintaining a Conserved Binding Mode. *ACS Infect. Dis.* **2021**. https://doi.org/10.1021/acsinfecdis.1c00194.

(266) Akif, M.; Georgiadis, D.; Mahajan, A.; Dive, V.; Sturrock, E. D.; Isaac, R. E.; Acharya, K. R. High-Resolution Crystal Structures of Drosophila Melanogaster Angiotensin-Converting Enzyme in Complex with Novel Inhibitors and Antihypertensive Drugs. *J. Mol. Biol.* **2010**, *400* (3), 502–517. https://doi.org/10.1016/j.jmb.2010.05.024.

(267) Towler, P.; Staker, B.; Prasad, S. G.; Menon, S.; Tang, J.; Parsons, T.; Ryan, D.; Fisher, M.; Williams, D.; Dales, N. A.; Patane, M. A.; Pantoliano, M. W. ACE2 X-Ray Structures Reveal a Large Hinge-Bending Motion Important for Inhibitor Binding and Catalysis. *J. Biol. Chem.* **2004**, *279* (17), 17996–18007. https://doi.org/10.1074/jbc.M311191200.

(268) Lauffer, B. E. L.; Mintzer, R.; Fong, R.; Mukund, S.; Tam, C.; Zilberleyb, I.; Flicke, B.; Ritscher, A.; Fedorowicz, G.; Vallero, R.; Ortwine, D. F.; Gunzner, J.; Modrusan, Z.; Neumann, L.; Koth, C. M.; Kaminker, J. S.; Heise, C. E.; Steiner, P. Histone Deacetylase (HDAC) Inhibitor Kinetic Rate Constants Correlate with Cellular Histone Acetylation but Not Transcription and Cell Viability. *J. Biol. Chem.* **2013**, *288* (37), 26926–26943. https://doi.org/10.1074/jbc.M113.490706.

(269) Jun-ichi Wachino, a Reo Kanechi, a Erina Nishino, a Marie Mochizuki, a Wanchun Jin, a Kouji Kimura, a Hiromasa Kurosaki, B.; A, Y. A. 4-Amino-2-Sulfanylbenzoic Acid as a Potent Subclass B3 Metallo-β-Lactamase-Specific Inhibitor Applicable for Distinguishing Metallo-β-Lactamase Subclasses. **2019**, No. June, 1–7.

(270) Lucic, A.; Hinchliffe, P.; Malla, T. R.; Tooke, C. L.; Brem, J.; Calvopiña, K.; Lohans, C. T.; Rabe, P.; McDonough, M. A.; Armistead, T.; Orville, A. M.; Spencer, J.; Schofield, C. J. Faropenem Reacts with Serine and Metallo-β-Lactamases to Give Multiple Products. *Eur. J. Med. Chem.* **2021**, *215*. https://doi.org/10.1016/j.ejmech.2021.113257.

(271) Saavedra, M. J.; Peixe, L.; Sousa, J. C.; Henriques, I.; Alves, A.; Correia, A. Sfh-I, a Subclass B2 Metallo-β-Lactamase from a Serratia Fonticola Environmental Isolate. *Antimicrob. Agents Chemother.* **2003**, *47* (7), 2330–2333. https://doi.org/10.1128/AAC.47.7.2330-2333.2003.

(272) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput* **2015**, *11*, 50. https://doi.org/10.1021/acs.jctc.5b00255.

(273) Lythell, E.; Suardíaz, R.; Hinchliffe, P.; Hanpaibool, C.; Visitsatthawong, S.; Oliveira, A. S. F.; Lang, E. J. M.; Surawatanawong, P.; Lee, V. S.; Rungrotmongkol, T.; Fey, N.; Spencer, J.; Mulholland, A. J. Resistance to the "Last Resort" Antibiotic Colistin: A Single-Zinc Mechanism for Phosphointermediate Formation in MCR Enzymes. *Chem. Commun.* **2020**, *56* (50), 6874–6877. https://doi.org/10.1039/d0cc02520h.

(274) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. Effects of Dispersion in Density Functional Based Quantum Mechanical/Molecular Mechanical Calculations on Cytochrome P450 Catalyzed Reactions. *J. Chem. Theory Comput.* **2012**, *8* (11), 4637–4645. https://doi.org/10.1021/ct300329h.

(275) Zheng, H.; Cooper, D. R.; Porebski, P. J.; Shabalin, I. G.; Handing, K. B.; Minor, W. CheckMyMetal: A Macromolecular Metal-Binding Validation Tool. *Acta Crystallogr. Sect. D Struct. Biol.* **2017**, *73*, 223–233. https://doi.org/10.1107/S2059798317001061.

(276) Zheng, H.; Chordia, M. D.; Cooper, D. R.; Chruszcz, M.; Müller, P.; Sheldrick, G. M.; Minor, W. Validation of Metal-Binding Sites in Macromolecular Structures with the CheckMyMetal Web Server. *Nat. Protoc.* **2014**, *9* (1), 156–170. https://doi.org/10.1038/nprot.2013.172.

(277) Wilcox, D. E. Binuclear Metallohydrolases. *Chem. Rev.* **1996**, *96* (7), 2435–2458. https://doi.org/10.1021/cr950043b.

(278) Schenk, G.; Mitić, N.; Gahan, L. R.; Ollis, D. L.; McGeary, R. P.; Guddat, L. W. Binuclear Metallohydrolases: Complex Mechanistic Strategies for a Simple Chemical Reaction. *Acc. Chem. Res.* **2012**, *45* (9), 1593–1603. https://doi.org/10.1021/ar300067g.

(279) E., B. S.; M., L. D.; C., H. D.; W., H. W. Metallo-β-Lactamases: Structure, Function, Epidemiology, Treatment Options, and the Development Pipeline. *Antimicrob. Agents Chemother.* **2021**, *64* (10), e00397-20. https://doi.org/10.1128/AAC.00397-20.

(280) Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T.N.Bhat, Helge Weissig, I. N. S. and P. E. B. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242. https://doi.org/10.1093/nar/28.1.235.

(281) Burley, S. K.; Bhikadiya, C.; Bi, C.; Bittrich, S.; Chen, L.; Crichlow, G. V.; Christie, C. H.; Dalenberg, K.; Di Costanzo, L.; Duarte, J. M.; Dutta, S.; Feng, Z.; Ganesan, S.; Goodsell, D. S.; Ghosh, S.; Green, R. K.; Guranovic, V.; Guzenko, D.; Hudson, B. P.; Lawson, C. L.; Liang, Y.; Lowe, R.; Namkoong, H.; Peisach, E.; Persikova, I.; Randle, C.; Rose, A.; Rose, Y.; Sali, A.; Segura, J.; Sekharan, M.; Shao, C.; Tao, Y. P.; Voigt, M.; Westbrook, J. D.; Young, J. Y.; Zardecki, C.; Zhuravleva, M. RCSB Protein Data Bank: Powerful New Tools for Exploring 3D Structures of Biological Macromolecules for Basic and Applied Research and Education in Fundamental Biology, Biomedicine, Biotechnology, Bioengineering and Energy Sciences. *Nucleic Acids Res.* **2021**, *49* (1), D437–D451. https://doi.org/10.1093/nar/gkaa1038.

(282) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of p K a Values. *J. Chem. Theory Comput.* **2011**, *7* (7), 2284–2295. https://doi.org/10.1021/ct200133y.

(283) Olsson, M. H. M.; SØndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical p K a Predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525–537. https://doi.org/10.1021/ct100578z.

(284) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: An Automated Pipeline for the Setup of Poisson-Boltzmann Electrostatics Calculations. *Nucleic Acids Res.* **2004**, *32* (WEB SERVER ISS.), 665–667. https://doi.org/10.1093/nar/gkh381.

(285) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174. https://doi.org/10.1002/jcc.20035.

(286) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

(287) Schrodinger LLC. The PyMOL Molecular Graphics System, Version 2.3. 2019.

(288) Korb, O.; Stützle, T.; Exner, T. E. An Ant Colony Optimization Approach to Flexible Protein–Ligand Docking. *Swarm Intell.* **2007**, *1* (2), 115–134. https://doi.org/10.1007/s11721-007-0006-9.

(289) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–461. https://doi.org/https://doi.org/10.1002/jcc.21334.

(290) Jun-ichi, W.; Yoshihiro, Y.; Shigetarou, M.; Hiromasa, K.; Yoshichika, A.; Keigo, S. Structural Insights into the Subclass B3 Metallo-β-Lactamase SMB-1 and the Mode of Inhibition by the Common Metallo-β-Lactamase Inhibitor Mercaptoacetate. *Antimicrob. Agents Chemother.* **2013**, *57* (1), 101–109. https://doi.org/10.1128/AAC.01264-12.

(291) Gaus, M.; Lu, X.; Elstner, M.; Cui, Q. Parameterization of DFTB3/3OB for Sulfur and Phosphorus for Chemical and Biological Applications. *J. Chem. Theory Comput.* **2014**, *10* (4), 1518–1537. https://doi.org/10.1021/ct401002w.

(292) Myller, A. T.; Karhe, J. J.; Pakkanen, T. T. Preparation of Aminofunctionalized TiO2 Surfaces by Binding of Organophosphates. *Appl. Surf. Sci.* **2010**, *257* (5), 1616–1622. https://doi.org/https://doi.org/10.1016/j.apsusc.2010.08.109.

(293) Rehman, M.; AlAjmi, M.; Hussain, A.; Rather, G.; Khan, M. High-Throughput Virtual Screening, Molecular Dynamics Simulation, and Enzyme Kinetics Identified ZINC84525623 as a Potential Inhibitor of NDM-1. *Int. J. Mol. Sci.* **2019**, *20* (4), 819. https://doi.org/10.3390/ijms20040819.

(294) Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W. Computational Methods in Drug Discovery. *Pharmacol. Rev.* **2014**, *66* (1), 334 LP – 395. https://doi.org/10.1124/pr.112.007336.

(295) Li, Z.; Song, L. F.; Li, P.; Merz, K. M. Systematic Parameterization of Divalent Metal Ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB Water Models. *J. Chem. Theory Comput.* **2020**. https://doi.org/10.1021/acs.jctc.0c00194.

(296) Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migues, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. Ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **2020**, *16* (1), 528–552. https://doi.org/10.1021/acs.jctc.9b00591.

(297) Jing, Z.; Liu, C.; Cheng, S. Y.; Qi, R.; Walker, B. D.; Piquemal, J. P.; Ren, P. Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications. *Annu. Rev. Biophys.* **2019**, *48*, 371–394. https://doi.org/10.1146/annurev-biophys-070317-033349.

(298) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-XTB - An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-

Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15* (3), 1652–1671. https://doi.org/10.1021/acs.jctc.8b01176.

(299) Fairhead, M.; Howarth, M. Site-Specific Biotinylation of Purified Proteins Using BirA BT - Site-Specific Protein Labeling: Methods and Protocols; Gautier, A., Hinner, M. J., Eds.; Springer New York: New York, NY, 2015; pp 171–184. https://doi.org/10.1007/978-1-4939-2272-7_12.